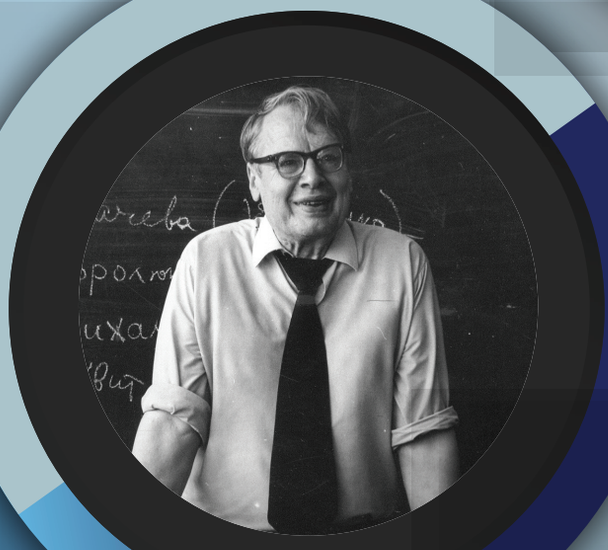
RTA

ISSN 1932-2321

JOURNAL IS REGISTERED IN THE LIBRARY OF THE U.S. CONGRESS

RELIABILITY: THEORY & APPLICATIONS

INTERNATIONAL GROUP ON RELIABILITY



RELIABILITY
RISK ANALYSIS
MAINTENANCE
SAFETY

GNEDENKO FORUM PUBLICATIONS

#4

(76) VOL.18

DECEMBER

2023

SAN DIEGO

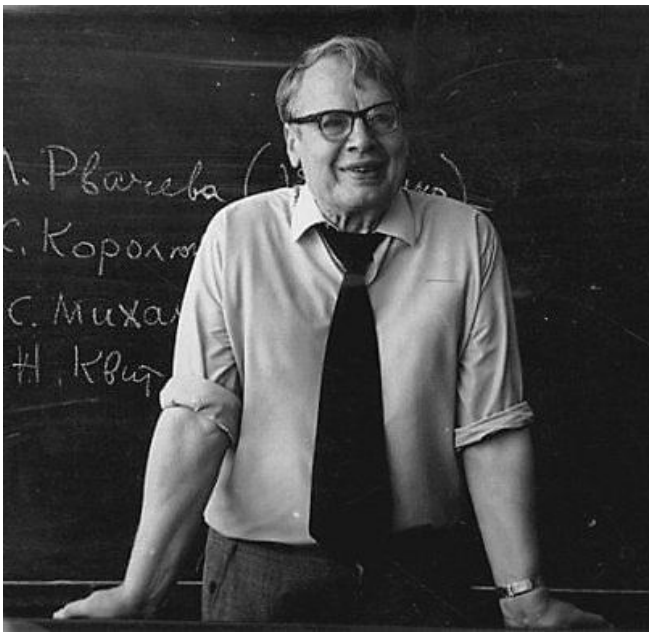
ISSN 1932-2321

© "Reliability: Theory & Applications", 2006, 2007, 2009-2023
© " Reliability & Risk Analysis: Theory & Applications", 2008

© I.A. Ushakov
© A.V. Bochkov, 2006-2023
© Kristina Ushakov, Cover Design, 2023
http://www.gnedenko.net/Journal/index.htm

All rights are reserved

The reference to the magazine "Reliability: Theory & Applications"
at partial use of materials is obligatory.



RELIABILITY:

THEORY & APPLICATIONS

Vol.18 No.4 (76),
December 2023

San Diego
2023

Editorial Board

Editor-in-Chief

Rykov, Vladimir (Russia)
Doctor of Sci, Professor, Department of Applied Mathematics & Computer Modeling, Gubkin Russian State Oil & Gas University, Leninsky Prospect, 65, 119991 Moscow, Russia.
e-mail: vladimir_rykov@mail.ru

Managing Editors

Bochkov, Alexander (Russia)
Doctor of Technical Sciences, Scientific Secretary JSC NIIAS, Scientific-Research and Design Institute Informatization, Automation and Communication in Railway Transport, Moscow, Russia, 107078, Orlikov pereulok, 5, building 1
e-mail: a.bochkov@gmail.com

Gnedenko, Ekaterina (USA)
PhD, Lecturer Department of Economics Boston University, Boston 02215, USA
e-mail: gnedenko@bu.edu

Bushinskaya, Anna (Russia)
Candidate of Tech. Sci. Leading Research Fellow of the Sci & Engng Center of the Russian Academy of Sciences, Ekaterinburg
e-mail: bushinskaya@gmail.com

Sazonov, Aleksey (Russia)
Leading Specialist of the Standardization Department, JSC NIIAS (Joint Stock Company "Design &Research Institute for Information "Technology, Signaling and Telecommunications on Railway Transport), Bild 1, 5 Orlikov Pereulok, Moscow, Russia, 107078
e-mail: sazonoff2007@gmail.com

Deputy Editors

Dimitrov, Boyan (USA)
Ph.D., Dr. of Math. Sci., Professor of Probability and Statistics, Associate Professor of Mathematics (Probability and Statistics), GMI Engineering and Management Inst. (now Kettering)
e-mail: bdimitro@kettering.edu

Gnedenko, Dmitry (Russia)
Doctor of Sci., Assos. Professor, Department of Probability, Faculty of Mechanics and Mathematics, Moscow State University, Moscow, 119899, Russia
e-mail: dmitry@gnedenko.com

Kashtanov, Victor A. (Russia)
PhD, M. Sc (Physics and Mathematics), Professor of Moscow Institute of Applied Mathematics, National Research University "Higher School of Economics" (Moscow, Russia)
e-mail: VAKashtan@yandex.ru

Krishnamoorthy, Achyutha (India)
M.Sc. (Mathematics), PhD (Probability, Stochastic Processes & Operations Research), Professor Emeritus, Department of Mathematics, Cochin University of Science & Technology, Kochi-682022, INDIA.
e-mail: achyuthacusat@gmail.com

Recchia, Charles H. (USA)
PhD, Senior Member IEEE Chair, Boston IEEE Reliability Chapter A Joint Chapter with New Hampshire and Providence, Advisory Committee, IEEE Reliability Society
e-mail: charles.recchia@macom.com

Shybinsky Igor (Russia)
Doctor of Sci., Professor, Division manager, VNIIAS (Russian Scientific and Research Institute of Informatics, Automatics and Communications), expert of the Scientific Council under Security Council of the Russia
e-mail: igor-shubinsky@yandex.ru

Yastrebenetsky, Mikhail (Ukraine)
Doctor of Sci., Professor. State Scientific and Technical Center for Nuclear and Radiation Safety (SSTC NRS), 53, Chernishevska str., of.2, 61002, Kharkov, Ukraine
e-mail: ma_yastrebenetsky@sstc.com.ua

Associate Editors

Aliyev, Vugar (Azerbaijan)
Doctor of Sci., Professor, Chief Researcher of the Institute of Physics of the National Academy of Sciences of Azerbaijan, Director of the AMIR Technical Services Company
e-mail: prof.vugar.aliyev@gmail.com

Balakrishnan, Narayanaswamy (Canada)
Professor of Statistics, Department of Mathematics and Statistics, McMaster University
e-mail: bala@mcmaster.ca

Carrión García, Andrés (Spain)
Professor Titular de Universidad, Director of the Center for Quality and Change Management, Universidad Politécnica de Valencia, Spain
e-mail: acarrion@eio.upv.es

Chakravarthy, Srinivas (USA)
Ph.D., Professor of Industrial Engineering & Statistics, Departments of Industrial and Manufacturing Engineering & Mathematics, Kettering University (formerly GMI-EMI) 1700, University Avenue, Flint, MI48504
e-mail: schakrav@kettering.edu

Cui, Lirong (China)
PhD, Professor, School of Management & Economics, Beijing Institute of Technology, Beijing, P. R. China (Zip:100081)
e-mail: lirongcui@bit.edu.cn

Finkelstein, Maxim (SAR)
Doctor of Sci., Distinguished Professor in Statistics/Mathematical Statistics at the UFS. Visiting researcher at Max Planck Institute for Demographic Research, Rostock, Germany and Visiting research professor (from 2014) at the ITMO University, St Petersburg, Russia
e-mail: FinkelM@ufs.ac.za

Kaminsky, Mark (USA)
PhD, principal reliability engineer at the NASA Goddard Space Flight Center
e-mail: mkaminskiy@hotmail.com

Krivtsov, Vasiliy (USA)
PhD. Director of Reliability Analytics at the Ford Motor Company. Associate Professor of Reliability Engineering at the University of Maryland (USA)
e-mail: VKrivtso@Ford.com, krivtsov@umd.edu

Lemeshko Boris (Russia)
Doctor of Sci., Professor, Novosibirsk State Technical University, Professor of Theoretical and Applied Informatics Department
e-mail: Lemeshko@ami.nstu.ru

Lesnykh, Valery (Russia)
Professor, Doctor of Sci., Adviser to Director General, LLC Gazprom gaznadzor, Novocheryo-mushkinskaya Street, 65, Moscow, 117418, Russia
e-mail: vvlesnykh@gmail.com

Levitin, Gregory (Israel)
PhD, The Israel Electric Corporation Ltd. Planning, Development & Technology Division. Reliability & Equipment Department, Engineer-Expert; OR and Artificial Intelligence applications in Power Engineering, Reliability.
e-mail: levitin@iec.co.il

Limnios, Nikolaos (France)
Professor, Université de Technologie de Compiègne, Laboratoire de Mathématiques, Appliquées Centre de Recherches de Royallieu, BP 20529, 60205 COMPIEGNE CEDEX, France
e-mail: Nikolaos.Limnios@utc.fr

Papic, Ljubisha (Serbia)
PhD, Professor, Head of the Department of Industrial and Systems Engineering Faculty of Technical Sciences Cacak, University of Kragujevac, Director and Founder the Research Center of Dependability and Quality Management (DQM Research Center), Prijevor, Serbia
e-mail: dqmcenter@mts.rs

Ram, Mangey (India)
Professor, Department of Mathematics, Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun, India. Visiting Professor, Institute of Advanced Manufacturing Technologies, Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia.
e-mail: mangeyram@gmail.comq

Timashev, Sviatoslav (Russia)
Doctor of Sci., Professor, Director and principal scientist the Sci & Engng Center of the Russian Academy of Sciences, Ekaterinburg
e-mail: timashevs@cox.net

Zio, Enrico (Italy)
PhD, Full Professor, Direttore della Scuola di Dottorato del Politecnico di Milano, Italy.
e-mail: Enrico.Zio@polimi.it

e-Journal ***Reliability: Theory & Applications*** publishes papers, reviews, memoirs, and bibliographical materials on Reliability, Quality Control, Safety, Survivability and Maintenance.

Theoretical papers must contain new problems, finger practical applications and should not be overloaded with clumsy formal solutions.

Priority is given to descriptions of case studies.
General requirements for presented papers.

1. Papers must be presented in English in MS Word or LaTeX format.
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper must be spell-checked.
4. For those whose language is not English, we kindly recommend using professional linguistic proofs before sending a paper to the journal.

The manuscripts complying with the scope of journal and accepted by the Editor are registered and sent for external review. The reviewed articles are emailed back to the authors for revision and improvement.

The decision to accept or reject a manuscript is made by the Editor considering the referees' opinion and considering scientific importance and novelty of the presented materials. Manuscripts are published in the author's edition. The Editorial Board are not responsible for possible typos in the original text. The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Publication in this e-Journal is equal to publication in other International scientific journals.

Papers directed by Members of the Editorial Boards are accepted without referring.
The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Send your papers to Alexander Bochkov, e-mail: a.bochkov@gmail.com

Table of Contents

EXPLORING NOVEL EXTENSION OF SUJA DISTRIBUTION: UNVEILING PROPERTIES AND DIVERSE APPLICATIONS 32

C. Subramanian, M. Subhashree, Aafaq A. Rather

This research article introduces and explores the area biased techniques of the Suja distribution, presenting novel derivations and insights. The estimation of the one-parameter area biased Suja distribution is accomplished using maximum likelihood, providing a robust framework for modeling real-world data. A comprehensive study of several statistical properties is conducted to unveil the characteristics and behaviors of this new model. To demonstrate its practical applicability, the proposed distribution is applied to real data of weather temperature. The analysis showcases the distribution's effectiveness in capturing the intricacies of temperature patterns, revealing its potential utility in weather modeling and related applications. The research contributes to the advancement of statistical modeling techniques and enriches our understanding of the Suja distribution's versatility in handling diverse datasets.

STATISTICAL MODELS FOR FORECASTING EMERGENCY SITUATIONS OF A BIOLOGICAL AND SOCIAL CHARACTER..... 41

Valery Akimov, Ekaterina Ivanova, Irina Oltyan

The article considers a statistical model for predicting emergency situations of a biological and social nature. Particular attention is paid to the calculation of indicators of resource provision of the medical care system and mortality rates during the spread of the epidemic.

A NEW EXTENDED EXPONENTIATED DISTRIBUTION WITH PROPERTIES AND APPLICATIONS..... 46

K. Selvakumar, Jameel A. Ansari, K. Kavitha, C. Subramanian, D. Vedavathi Saraja, Arshad Ahmad Khan, Rashid A. Ganaie, Aafaq A. Rather, Maryam Mohiuddin

This manuscript focuses on the statistical properties and estimation methods of the exponentiated Suja distribution, which is characterized by two parameters: scale and shape. From a frequentist perspective, our primary emphasis is on estimation techniques. Additionally, we derive statistical and reliability characteristics for the model. We explore various estimation procedures, including order statistics, entropies, reliability analysis, and the maximum likelihood method. To assess the model's superiority and practical utility, we analyze real lifetime data sets. Overall, this study provides a comprehensive analysis of the exponentiated Suja distribution, offering insights into its statistical properties, estimation techniques, and real-life applications.

LEVERAGING AUXILIARY VARIABLES: ADVANCING MEAN ESTIMATION THROUGH CONDITIONAL AND UNCONDITIONAL POST-STRATIFICATION 57

G.R.V. Triveni, Faizan Danish

This article presents a novel class of estimators designed for post-stratification to estimate the mean of a study variable using information from auxiliary variables. Through a rigorous examination of bias and Mean Square Error (MSE), we demonstrate the potential to improve estimation accuracy up to the first order of approximation. We also thoroughly explore both Conditional and Unconditional post-stratification properties, enhancing our understanding of the estimator's performance. To assess the effectiveness of our proposed estimator, we conduct a comprehensive numerical illustration. The results affirm its superiority over existing estimators in both Conditional and Unconditional Poststratification scenarios, exhibiting the highest Percentage Relative Efficiency. Additionally, graphical analysis reveals that Conditional post- stratification outperforms Unconditional post-stratification. These findings underscore the significant practical value of our proposed

estimator in enhancing the accuracy of mean estimation in post-stratification studies. By accurately estimating population parameters, our novel class of estimators contributes to more informed decision-making in various fields of study. The utilization of auxiliary variables allows for better utilization of available information and leads to more reliable and robust conclusions. Overall, the novel class of estimators introduced in this article represents a valuable contribution to the field of post-stratification. As researchers continue to explore and apply these estimators, they have the potential to revolutionize data analysis methods, becoming indispensable tools for survey and research design. The improvements in estimation accuracy brought about by these estimators are particularly crucial in situations where reliable data is scarce or challenging to obtain, making them invaluable for decision-makers and researchers alike. With the increased accuracy and efficiency of our proposed estimators, they provide a pathway for better resource allocation, cost-effective decision-making, and improved policy formulation. Policymakers and researchers can confidently rely on these estimators to produce more accurate results and achieve better outcomes in various domains. In conclusion, the novel class of estimators for post-stratification presented in this article opens up new avenues for advancing statistical estimation methods. The fusion of auxiliary variables with traditional poststratification techniques represents a powerful approach to enhance estimation accuracy. Embracing and incorporating these estimators into research practices will undoubtedly bring us closer to making data-driven decisions that have a meaningful impact on society.

ASSESSMENT OF GENERALIZED LIFETIME PERFORMANCE INDEX FOR LINDLEY DISTRIBUTION USING PROGRESSIVE TYPE-II SAMPLES 69

Abhimanyu S Yadav, Mahendra Saha, Amartya Bhattacharya, Arindam Gupta

A meaningful subject of discourse in manufacturing industries is the assessment of the lifetime performance index. In manufacturing industries, the lifetime performance index is used to measure the performance of the product. A generalized lifetime performance index (GLPI) is defined by taking into consideration the median of the process measurement when the lifetime of products follows a parametric distribution may serve better the need of quality engineers and scientists in industry. The present study constructs various point estimators of the GLPI based on progressive type II right censored data for the Lindley distributed lifetime in both classical and Bayesian setup. We perform Monte Carlo simulations to compare the performances of the maximum likelihood and Bayes estimates with a gamma prior of Cy (L) under progressive type-II right censoring scheme. Finally, the validity of the model is adjudged through analysis of a data set.

EXPLORING THE LENGTH BIASED TORNUMONKPE DISTRIBUTION: PROPERTIES, ESTIMATIONS AND PRACTICAL APPLICATIONS 87

D. Vedavathi Saraja, B. Jayakumar, Madhulika Mishra, Priya Deshpande, C. Subramanian, Rashid A. Ganaie, Aafaq A. Rather, Bilal Ahmad Bhat

In this study, we introduce a novel extension of the Tornumonkpe distribution, known as the length biased Tornumonkpe distribution. This distribution holds particular significance as it belongs to the family of weighted distributions, specifically the length biased variant. Through an in-depth analysis, we explore the mathematical and statistical properties of this distribution, shedding light on its unique characteristics. To estimate the model parameters of this new distribution, we employ the well-established technique of maximum likelihood estimation. This allows us to accurately determine the parameters and enhance our understanding of the distribution's behavior. To demonstrate the practical applicability and advantages of the length biased Tornumonkpe distribution, we showcase its performance using a real-life time data set. Through this empirical examination, we investigate the distribution's superiority and flexibility, providing valuable insights into its potential use in various domains.

ENHANCING ENGINEERING SCIENCES WITH UMA DISTRIBUTION: A PERFECT FIT AND VALUABLE CONTRIBUTIONS 99

R. A. Ganaie, C. Subramanian, V. P. Soumya, R. Shenbagaraja, Mahfooz Alam, D. Vedavathi Saraja, Rushika Kinjawadekar, Aafaq A. Rather, Showkat A. Dar

In this study, we introduce a novel class of distributions called the length biased Uma distribution. This

distribution is a specific instance of the broader weighted distribution family, known for its versatility in various applications. We explore the structural properties of the length biased Uma distribution and propose a robust parameter estimation technique based on maximum likelihood estimation. To assess its efficacy, we apply the newly introduced distribution to two real-world datasets, evaluating its flexibility and performance in comparison to existing models. The results obtained demonstrate the potential of the length biased Uma distribution as a valuable addition to the repertoire of statistical distributions, offering valuable insights for a wide range of practical applications.

A NOVEL EXTENSION OF INVERSE EXPONENTIAL DISTRIBUTIONS: A HEAVY-TAILED MODEL WITH UPSIDE DOWN BATHTUB SHAPED HAZARD RATE 112

Jabir Bengalath, Bindu Punathumparambath

Heavy-tailed distributions have garnered interest due to their advantageous statistical and reliability characteristics, rendering them valuable in applied fields such as economics, finance, and risk management. Such distributions offer robust properties, making them pertinent to studies in various areas like econometrics, statistics, and insurance. Thus, the primary objective of this paper is to propose a Two parameter right skewed-upside down bathtub type, heavy tailed distribution, which is a generalisation of Inverse Exponential distribution and is referred to as Modi Inverse Exponential distribution. We derive several mathematical and statistical features, including quantile function, mode, median, skewness, kurtosis, and mean deviation. Additionally, the reliability and hazard rate functions are also derived. Stochastic ordering and order statistics of the proposed distribution were derived. We also investigate the tail behaviour of the proposed model. Furthermore, estimation methods such as maximum likelihood estimation and its asymptotic confidence bound, percentile method, and Cramer-von-Mises method were examined. To demonstrate the appropriateness of the suggested model, we have considered two distinct real datasets along with three distinct models and concluded that the proposed model is more adaptable.

A NOVEL METHODOLOGY IN DEVELOPING STRESS-STRENGTH RELIABILITY MODEL FOR WEIBULL DISTRIBUTION: A COMPARISON OF ARTIFICIAL NEURAL NETWORK (ANN) AND RESPONSE SURFACE ANALYSIS (RSA) 128

Dr. Saurabh L. Raikar, Prof. Rajesh S. Prabhu Gaonkar

Stress strength interference theory is widely used in evaluating the reliability of mechanical components. Various interference models have been developed when stress and strength follow a wide range of distributions. But when stress and strength follow Weibull distribution, a closed form of interference model is not available. This paper deals with developing a methodology for obtaining a closed form interference model for a given application when the stress and strength follow Weibull distribution. The method of artificial neural network (ANN) and response surface analysis (RSA) are used in modelling and analysis. The validation experiment has been conducted and the error obtained shows that the proposed methodology performs reasonably well.

RELIABILITY MODELLING OF UTENSILS MANUFACTURING SYSTEM WITH TEMPERATURE DEPENDENT MAINTENANCE 141

Manisha Gaba, Dalip Singh, Sheetal, Kajal Sachdeva

In this paper, a stochastic model for utensils manufacturing system with preventive maintenance (PM) is analysed in detail. The operation is affected by variation in the temperature dependent maintenance. The entire manufacturing process of utensils goes through four subsystems viz., Circle cutting subsystem 1, Pressing subsystem 2, Spinning subsystem 3 and Polishing & Packing 4. The system has series structure of all the subsystems. The system is put under PM on the winter time and after PM it operates as new. The PM time distributions are considered as arbitrary and the time to failure as well as repair of each subsystem follows a negative exponential distribution. All random variables are statistically independent. Several measures for evaluating the effectiveness of a system, including mean time to system failure (MTSF), system availability (in summer and winter), busy period of repairman and expected number of repairs (in summer and winter) are

derived using a regenerative point technique and Markov process. The system is also analysed for particular values of the parameters.

OPTIMIZING MULTI-OBJECTIVE MULTI-INDEX TRANSPORTATION PROBLEMS: A SMART ALGORITHMIC SOLUTION WITH LINDO SOFTWARE 154

Ajjaz Maqbool Dar, K. Selvakumar, S. Ramki, K. M. Karuppasamy, Jameel A. Ansari, Aafaq A. Rather

In the present paper, we create an algorithm to address the transportation problem with numerous objectives and indexes. The transportation problem exists when there are more supply points, more demand points, and various means of transportation are used to meet demand or when moving certain types of goods. The transportation problem may frequently be more complex than the typical form of transportation problem. We create a model that blends fuzzy multi-objective programming and the multiindex transportation problem' by using LINDO software to resolve all related problems. Additionally, the decision-maker may present a variety of data and it may be further improved. The new algorithmfor addressing transport problems in fuzzy environments is demonstrated numerically.

CHARACTERIZATION OF NEW QUASI LINDLEY DISTRIBUTION BY TRUNCATED MOMENTS AND CONDITIONAL EXPECTATION OF ORDER STATISTICS 168

Mohd. Amir, Mohammad Faizan, Rafiqullah Khan

Characterization of a probability distribution plays an important role in probability and statistics. Before a particular probability distribution model is applied to fit the real-world data, it is necessary to confirm whether the given probability distribution satisfies the underlying requirements by its characterization. The aim of this paper is to find characterization results New Quasi Lindley distribution. These results are established using the relation between truncated moments and failure rate functions and conditional expectation of adjacent order statistics. The first characterization result is based on relation between left truncation moment and failure rate function while the second result is based on relation between right truncated moment and reverse failure rate function. In the third characterization result we used conditional expectation of order statistics when the conditioned one is adjacent order statistics. Further, some of its important deductions are also discussed.

ANALYSIS OF AN M/M/1/K FEEDBACK WORKING VACATION QUEUE WITH RENEGING 178

Krishan, Neetu Gupta

The analysis of an M/M/1/N feedback working vacation queueing system with reneging is presented in this paper Customers may become impatient and even disappointed when they see a long line. In the literature on queueing, customer dissatisfaction caused on by unsatisfactory service is referred as feedback. In the case of feedback, customers retry services after receiving unsatisfactory or incomplete. First, we create the equations for the steady-state probabilities using the Markov process method. The steady-state probabilities are then solved by the matrix method. We then provide some system performance measures. We create a cost model using performance analysis. Finally, we give some numerical examples to show how the various model parameters affect the system's behaviour

PERFORMANCE CHARACTERIZATION OF TWO-SERVER BATCH SERVICE QUEUE WITH SECOND OPTIONAL SERVICE 189

Andwilile Abrahamu George, P. Vijaya Laxmi

In this paper, we analyze the performance of a finite capacity two-server Markovian batch service queueing model with second optional service. The servers provide two kinds of services, the first essential service (FES), which is provided to all incoming customers and the second optional service (SOS) to those who demand it after completing FES. The service times of the two servers are identical and are exponentially distributed. Matrix-

decomposition method is used to obtain the steady-state probabilities of the model. Numerical results and discussion are presented to demonstrate the impact of the model parameters on the system behavior. Furthermore, the cost model optimization is developed to determine the optimal service rates using the Quasi-Newton method to minimize the expected cost. Finally, the findings of this work show that the blocking probability is monotonically decreases as finite buffer size increases and approaches its minimum value of zero when finite buffer is sufficiently large.

PROPERTIES OF QUADRASOPHIC FUZZY SET AND ITS APPLICATIONS 208

G. Aruna, J. Jesintha Rosline

Fuzzy set theory is a distinctive way of approaching ambiguous information. In this artifact, we introduce a new extension of fuzzy set known as Quadrasophic Fuzzy set and its properties. The Quadrasophic Fuzzy set has four parameters. The attributes and operations of the Quadrasophic Fuzzy sets are defined with pertinent examples. The arithmetic aggregator operators with a redefined level of 0.5 are introduced. The theorems of aggregator operators of Quadrasophic Fuzzy sets are explained using mathematical formulations. Suitable results and examples are provided to enlighten the proposed method. The arithmetic aggregator operators of the proposed method have been used in decision- making to get the optimal solution with supplementary statistics. Additionally, the selection of appropriate fertilizer in farming is demonstrated using the operators of the suggested model. A decision-making approach is also used to develop the proposed method in order to identify the ideal solution. An illustration is provided to examine the unique feature of the proposed method to resolve the decision-making problems with a perfect solution.

A PENTAGONAL FUZZY-BASED SOLUTION OF MULTIPLE OBJECTIVE LPP 221

Junaid Basha M, Nandhini S, Nur Aisyah Abdul Fataf

In this paper, the researchers compare proposed approach and Excel solver. In this proposed technique, the researchers converted fuzzy multiple objective linear programming problems (FMOLPP) into multiple objective linear programming problems (MOLPP) with the help of Defuzzified mean of maxima method. Before that, the researchers changed the pentagonal fuzzy numerical valuation to a triangular fuzzy value by using the proposed theorem. Further, the crisp value of MOLPP is solved using standard simplex algorithms. Then the outcomes of the optimal solutions are compared with both the results.

DIFFERENT ESTIMATION METHODS FOR THE PARAMETER OF XGAMMA DISTRIBUTION AND THEIR COMPARISON 229

Sukanta Pramanik, Sandipan Maiti

The xgamma distribution is vital in reliability/survival analysis and biomedical research. In this article, different estimation methods are proposed for the parameter of this distribution. The distribution is a unique finite mixture of exponential distribution and gamma distribution. Some further properties of the distribution that are not available in the earlier literature are studied. We consider the maximum likelihood estimator, least squares estimator, weighted least squares estimator, percentile estimator, the maximum product spacing estimator, the minimum spacing absolute distance estimator, the minimum spacing absolute log-distance estimator, Cram er von Mises estimator, Anderson Darling estimator, right-tailed Anderson Darling estimator, and compare them using a comprehensive simulation study. For comparison purposes, the estimators' bias and mean squared error are considered. A real data example is also a part of this work. Some model selection techniques are used to choose the best fitting of the distribution to the data.

CONFIDENCE INTERVAL USING MAXIMUM LIKELIHOOD ESTIMATION FOR THE PARAMETERS OF POISSON TYPE LENGTH BIASED EXPONENTIAL CLASS MODEL 242

Rajesh Singh, Preeti A. Badge, Pritee Singh

In this research paper, Confidence interval using Maximum likelihood estimation is obtained for Poisson type Length biased exponential class for the parameters. Failure intensity, mean time to failure and likelihood function

for the parameter is obtained. Confidence interval has been derived for parameters using maximum likelihood estimation. To study the performance of confidence interval, average length and coverage probability are calculated using Monte Carlo simulation technique. From the obtained intervals it is concluded that Confidence interval for the parameter perform better for appropriate choice of execution time and certain values of parameters.

A TWO NON IDENTICAL UNITS COLD STANDBY SYSTEM WITH CORRELATED FAILURE TIME AND REPAIR MACHINE FAILURE 252

Alka Chaudhary, Shivali Sharma and Anika Sharma

The paper deals with a system composed of two-non identical units (unit-1 and unit-2). Initially, unit-1 is operative and unit-2 is kept in cold standby. The cold standby unit can't fail in its standby mode. Each unit of the system has two possible modes: Normal (N) and total failure (F). When the unit-1 fails the cold standby (unit-2) becomes operative instantaneously with the help of a perfect and instantaneous switching device. A single repairman is always available with the system to repair a failed unit and failed RM. Unit-1 gets priority in operation and repair over unit-2. However, the RM gets priority in repair over any of the units. The RM machine is good initially and can't fail unless it becomes operative. The system failure occurs when both the units are in total failure mode. The joint distribution of failure and repair times for each unit is taken bivariate exponential distribution. Each repaired unit works as good as new. Using regenerative point technique, various important measures of system effectiveness have been obtained.

DECISION MAKING THROUGH FUZZY LINEAR PROGRAMMING APPROACH 263

Pandit U. Chopade, Mahesh M. Janolkar, Kirankumar L. Bondar

In this study a real world industrial MPS problem is addressed using the SMF approach. A decision maker, analyst and implementer, all play significant roles in making judgements in an uncertain environment, which is where this difficulty arises in the chocolate manufacturing business. As analysts our goal is to identify a solution with a higher LOS that will enable the decision maker to reach a conclusion. Because all the coefficients including the goals, technical and resource factors are well defined. The MPS problem is taken into consideration. With 24 constraints and 6 variables, this is regarded as one of the sufficiently large problem, which LOV is appropriate for getting satisfactory OR can be determined by a decision maker. To increase the satisfactory income, the decision maker can also advice to the analyst some feasible modification to FI. This collaborative process between the analyst, decision maker and implementer must continue until the best possible solution is found and put into action.

DENSITY BY MODULI AND LACUNARY STATISTICAL CONVERGENCE OF DOUBLE SEQUENCES 276

A. G. K. Ali, A. M. Brono, A. Masha

In this paper, we introduced and studied the concept of lacunary statistical convergence of double sequence with respect to modulus function where the modulus function is an unbounded double sequence. We also introduced the concept of lacunary strong convergence of double sequence via modulus function. We further characterized those lacunary convergence of double sequence for which the lacunary statistically convergent of double sequence with respect to modulus function equals statistically convergent of double sequence with respect to modulus function. Finally, we established some inclusion relations between these two lacunary methods and proved some essential analogue for double sequence.

ESTIMATION OF STRESS STRENGTH RELIABILITY USING PRANAV DISTRIBUTION 288

Ankitha Lukose, Chacko V M

This paper deals with the estimation of stress strength reliability parameter R, which is the probability of Y less than X when X and Y are two independent distribution with different scale parameter and same shape parameter The maximum likelihood method is used to find an estimator for R. We also obtain the asymptotic distribution of

the maximum likelihood estimator ofR. Based on this asymptotic distribution, the asymptotic confidence interval can be obtained. We also propose bootstrap confidence interval for the parameter R. Analysis of a simulated data and a real life data have been presented for illustrative purposes.

FAST AND ROBUST BIVARIATE CONTROL CHARTS FOR INDIVIDUAL OBSERVATIONS..... 296

Sajesh T A

There are various circumstances where it is important to simultaneously monitor or control two or more related quality characteristics. Independently tracking these quality characteristics might be quite deceptive. Hotelling's T2 chart, in which the T2 statistics are generated using the classical estimates of location and scatter, is the most well-known multivariate process monitoring and control approach. It is well known that the existence of outliers in a dataset has a significant impact on classical estimators. Any statistic that is computed using the classical estimates will be distorted by even a single outlier. The non-robustness issue is investigated in this study, which also suggests four robust bivariate control charts based on the robust Gnandesikan-Kettenring estimator. This study employs four highly robust scale estimators, with the best breakdown point, namely the Qn estimator, Sn estimator, MAD estimator, and t estimator, in order to robustify the Gnandesikan- Kettenring estimator. Through the use of a Monte Carlo simulation and a real-life data, the performance of the suggested control charts is assessed. The four techniques all outperform the traditional method and provide greater computing efficiency.

STATISTICAL MODELS FOR FORECASTING EMERGENCY SITUATIONS OF MAN-CAUSED CHARACTER..... 309

Valery Akimov, Ekaterina Ivanova, Yuri Shishkov

The aim of the study is to develop predictive and analytical solutions for technogenic threats for urban areas, the mathematical basis of which is Bayesian classifiers. The result of the work is a formalized description of models for predicting the consequences of a heat supply shutdown; consequences of a power outage; consequences of oil and oil products spills; the consequences of the discharge of liquid technological waste into the hydrosphere; the consequences of the release of hazardous chemicals into the environment.

CHARACTERIZATION OF CHANDBHAS-P DISTRIBUTION AND ITS APPLICATIONS IN MEDICAL SCIENCE..... 314

Praseeja C B, Prasanth C B, C Subramanian, Unnikrishnan T

The current research attempts the length biased version of new two-parameter Sujatha distribution, which is referred as ChandBhas-P Distribution (CBPD). Its different structural properties are discussed and the model parameters of this novel distribution are predicted by using the Maximum Likelihood Estimation. The distribution was examined with two real lifetime sets of data. The first set of data is birth weight of new born babies, randomly selected from a hospital at Thrissur, kerala and the second set is the weight of children of age range between three months to four years-collected from a few babysitting centres and play schools across Thrissur, and both are employed in order to discuss the goodness of fit.

FORECASTING OF EXTREME RISK USING MARKOV-SWITCHING GARCH MODELS: EVIDENCE FROM GLOBAL ENERGY MARKETS 325

S. Kavitha, G. Mokesh Rayalu, D. Pachiyappan, P. Manigandan

This paper investigates the Markov-Switching GARCH and Single-Regime (SR) GARCH models for the extreme-risk prediction of the global energy markets. Using daily data from Jan. 2020 to July. 2022, we find the MS-GARCH-types models are appropriate for both developed and emerging energy markets because they efficiently measure the extreme risk of energy commodities in various cases. Meanwhile, the regime-switching model's capture-dynamic structures in the financial markets and this model is only better than single-regime models in terms of long position risk predicting, rather than short position risk forecasting. That is, on the

downside risk predicting, it just outperforms the single regime. Through competitive models, this study examines the risk forecast of energy commodities in different conditions. The findings have strong implications for investors and policymakers in selecting the appropriate model to predict the extreme risk of energy commodities when facing asset allocation, portfolio selection, and risk management.

PRICE RISK ANALYSIS USING GARCH FAMILY MODELS: EVIDENCE FROM INDIAN NATIONAL STOCK EXCHANGE FUTURE MARKET 338

M. Valavan, Mohammed Ahmar Uddin, S. Rita

The prediction of time-varying volatility plays an important role in financial data. In the paper, a comprehensive analysis of the mean return and conditional variance of NSE index is performed to use GARCH, EGARCH and TGARCH models with Normal innovation and Student's t innovation. Conducting a bootstrap simulation study which shows the Model Confidence Set (MCS) captures the superior models across a range of significance levels. The experimental results show that, under various loss functions, the GARCH using Student's t innovation model is the best model for volatility predictions of NSE among the six models.

IDZ DISTRIBUTION: PROPERTIES AND APPLICATION 346

Idzhar A. Lakibul

This paper introduces a novel two - parameter continuous distribution. This distribution is derived from the mixture of the Exponential, Weibull and Ailamujia distributions. The derived distribution is named as "Idz distribution". The probability density function of the Idz distribution is derived and some of its plots are presented. It can be observed that the Idz distribution can generate right tailed unimodal, non-monotonic decreasing and exponential shapes. Further, survival and hazard functions of the Idz distribution are derived. It reveals that the hazard function of the Idz distribution can accommodate three types of failure rate behaviors, namely, non-monotonic constant, right tailed unimodal and nonmonotonic decreasing. Moreover, some properties of Idz distribution such as moments, mean, variance, moment generating function, order statistics and maximum likelihood estimates are derived. In addition, the proposed distribution is applied into a Breast Cancer data and compare with the Exponentiated Generalized Inverse Rayleigh distribution, the Ailamujia Inverted Weibull distribution and the New Extended Exponentiated Weibull distribution. Result shows that the Idz distribution gives better estimates as compared with the said distributions for a given dataset.

A LITERATURE REVIEW ON DISCRETE-TIME QUEUEING MODELS **355**

Harini R, Indhira K

In this paper, a quantitative research survey is carried out on discrete-time queueing models. In real-life scenarios, the idea of discrete-time queues has taken on a new meaning. This survey mainly focuses on the unfolding of discrete-time queueing models in recent decades, challenges implied on them and their influence in various fields. The ultimate goal of this paper is to provide enough information to all the researchers and analysts who toil in this field and wish to know more about these models. A few open issues and intriguing future research paths has been discussed.

QUANTILE RESIDUAL ENTROPY FOR SOME LIFE TIME DISTRIBUTIONS 372

Javid Gani Dar, Mohammad Younus Bhat, Shahid Tamboli, Shaikh Sarfaraj, Aafaq A. Rather, Maryam Mohiuddin, Showkat Ahmad Dar

This study explores the concept of residual entropy as an alternative approach to traditional entropy measures. The field of information theory, built upon Shannon's entropy, has been instrumental in understanding the dynamics of systems. However, existing literature has recognized the limitations of applying traditional entropy measures to systems that have already been in existence for a certain duration. This study delves into the concept of residual entropy, acknowledging the need for a more suitable approach for such systems. Specifically, we investigate the characteristics of residual entropy using a quantile-based framework. By deriving the quantile

residual entropy function for various lifetime models, we gain insights into the reordering and ageing phenomena captured by the quantile version of the residual entropy equation. Our findings contribute to an enhanced understanding of residual entropy and provide a novel perspective on analyzing and interpreting the behavior of established systems.

SAMPLING PLANS BASED ON TRUNCATED LIFE TEST FOR LOGISTIC FAMILY OF DISTRIBUTIONS 382

Sriramachandran G V

In this article, we develops optimal sample size for acceptance number (zero and one) for single and double sampling plans by fixing consumer's risk and test completion time, with the assumption that, the life of the item follows logistic family of distributions (i.e. Logistic Rayleigh distribution/Logistic exponential distribution/Logistic Weibull distribution). The optimal size obtained for single and double sampling plans for logistic family of distributions are compared with baseline distributions and the results are discussed.

ECONOMIC ORDER QUANTITY MODEL FOR IMPERFECT ITEMS WITH SHORTAGE BACKORDERING 391

Priyanka Singh, A. R. Nigwal, U. K. Khedlekar

This study presents the development of an economic order quantity model (EOQ) specifically designed for imperfect quality items. The model takes into consideration three distinct scenarios: (a) Model I trigger a reorder when the inventory level reaches zero; (b) Model II initiates a reorder when the backordered quantity equals the imperfect quantity; (c) Model III initiates a reorder when the shortage persists. To distinguish between perfect and imperfect quality products, a screening process is implemented for each product lot. Upon product delivery from the supplier to the vendor, all received products undergo immediate inspection through the screening process. Following the EOQ ordering policy, the vendor sells imperfect products to customers at a reduced cost at the end of the cycle, rather than returning them to the supplier. To fulfil the remaining demand for high-quality products, the vendor procures such products from a local vendor at a higher price. This study optimizes the duration of positive inventory, selling price, and total profit per unit time. Model I, which exhibits the longest duration of positive inventory, demonstrates greater business stability compared to the other two models. The concavity property is analytically and numerically demonstrated, and a sensitivity analysis is provided to explore the impact of model parameters on outputs.

DESIGNING A HYBRID SINGLE SAMPLING PLAN FOR LIFE-TIME ASSESSMENT USING THE EXPONENTIAL-RAYLEIGH DISTRIBUTION 410

Radhika A, Nandhini M, Jeslin J

The approach of statistical quality control known as "product control" deals with the steps involved in making judgments on one or more batches of completed goods produced by production processes. One of the main categories of product control is sampling inspection by variables, which includes processes for selecting numerous individual units based on sample measurements for a quality characteristic under investigation. These approaches are predicated on the knowledge of the functional form of the probability distribution and the presumption that the quality feature is measured on a continuous scale. The literature on product control contains inspection techniques that were created with the implicit presumption that the quality characteristic is distributed normally with the associated attributes. In this study, a single variable sampling plan is developed and assessed under the assumption that the quality characteristic will be distributed using an Exponential-Rayleigh distribution. This article discusses the development of reliability sampling plans for intermittent test batches using type-I and type-II censoring data. To build a sampling strategy using the Exponential-Rayleigh distribution, this work offers a two-parameter continuous probability distribution. One of the main categories of acceptance sampling is sampling inspection by variables, which involves processes for making decisions regarding the disposition of numerous individual units based on sample measurements of those units for a quality feature under investigation. Assume that the sample inspection's number of defective items follows the Poisson distribution. The suggested SSP's ideal parameters are determined using a multi-objective genetic algorithm, which is concerned with concurrently minimizing the average number of samples and inspection costs a maximizing the likelihood of the acceptance sampling plan. The

Rayleigh distribution is an appropriate model for life-testing studies, and the Exponential Rayleigh Distribution is studied as a model for a lifetime random variable. The paper also analyses the effectiveness of reliable single sampling plans designed using the median lifetime of products. The efficiency of these sampling plans is evaluated in terms of sample size and sampling risks. Poisson probabilities are used to determine the parameters of the sampling plans, to protect both producers and consumers from risks. For manufacturing enterprises to analyze the viability of the sample plan, necessary tables and procedures are constructed with acceptable examples.

REDESCENDING M-ESTIMATOR BASED LASSO FOR FEATURE SELECTION 419

R. Muthukrishnan, C. K. James

Aim: Regression analysis is one of the statistical methods which helps to model the data and helps in prediction, a large data set with higher number of variables will often create problem due to its dimensionality and hence create difficulties to gather important information from the data, so it is a need of a method which can simultaneously choose important variables which contains most of the information and hence helps to fit the model. Least absolute shrinkage and selection operator (LASSO) is a popular choice for shrinkage estimation and variable selection. But LASSO uses the conventional least squares technique for feature selection which is very sensitive to outliers. As a result, when the data set is contaminated with bad observations (Outliers), the LASSO technique gives unreliable results, so in this paper the focus is to create a method which can resist to outliers in the data and helps in giving a meaningful result. Method: proposed a new procedure, a LASSO method by adding weights which uses the concept of redescending M-estimator, which can resist outliers in both dependent and independent variables. The observation with greater importance receives a higher weight and less weight to the least important observation. Findings: The efficiency of the proposed method has been studied in the real and simulation environment and compared with other existing procedures with measures like Median Absolute Error (MDAE), False Positive Rate (FPR), False Negative Rate (FNR), Mean Absolute Percentage Error (MAPE). The proposed method with the redescending M-estimator shows a higher resistance to outliers compared to conventional LASSO and other robust existing procedures. Conclusion: The study reveals that the proposed method outperforms other existing procedures in terms of MDAE, FPR, FNR and MAPE, indicating its superior performance in variables selection under outlier contaminated datasets.

PARAMETER ESTIMATION OF SCALE MUTH DISTRIBUTION(SMD) UNDER TYPE-1 CENSORING USING CLASSICAL AND BAYESIAN APPROACHES 429

Agni Saroj, Prashant K. Sonker, Shalini Kumari, Rakesh Ranjan, Mukesh Kumar

The lifetime distributions are used to understand and explain the real life circumstances in various fields (medical, engineering, etc.). Many times it is very tough task to complete an experiment with complete data due to lack of time, money or some other factors and get the data in incomplete form. to draw the information from such type of data (incomplete data), we use some censoring techniques. In the field of statistics, there are several censoring techniques available where type-I censoring is most commonly used due to its simplicity. In this article, the scale Muth distribution (SMD) is considered as a lifetime distribution under type-I censoring scheme. The parameter estimation has been done by classical as well as Bayesian approach. Under the classical paradigm, two most popular methods were used maximum likelihood estimation (MLE) and the maximum product of spacing estimation (MPSE). And under the Bayesian paradigm, we used the informative priors for each parameter and obtained the estimates by considering the squared error loss function using an approximation method, Metropolis Hasting (MH) algorithm. The performance of each estimator is evaluated by their mean squared error or simulated risk. Also, a real data set is used to illustrate the real phenomena and to estimate the parameter using above-mentioned techniques under type-I censoring scheme.

ROBUST MAHALANOBIS DEPTH BASED ON MINIMUM REGULARIZED COVARIANCE DETERMINANT ESTIMATOR FOR HIGHDIMENSIONAL DATA 446

R Muthukrishnan, Surabhi S Nair

Handling of high-dimensional data is an important issue in robust literature. For analyzing data, location measure plays a vital role in almost all statistical methods. The location parameter of a distribution is used to find the central value. Many computational methods are used to find the measure of location for analyzing data. The data depth procedure is one approach to finding the true representative of the entire data and it is one of the key concepts in multivariate data analysis. Data depth is a term used to describe how deep a particular point is inside the broad multivariate data cloud. Instead of the

typical smallest to biggest rank, the sample points can be ordered from the center outward. Mahalanobis depth is one of the popular depth procedures. The traditional approach used to find Mahalanobis depth is based on Mahalanobis distance, it is based on the classical sample mean vector and covariance matrix. So the conventional Mahalanobis depth is sensitive to outliers and may fail when the data is contaminated. To solve this problem, the Minimum Covariance Determinant (MCD) estimators are used instead of classical estimators. However, the MCD estimators cannot be calculated in high dimensional data sets, in which the variable number is higher than the subset size. To calculate Mahalanobis depth values in high dimensional data, propose a new depth function namely the Robust Regularized Mahalanobis Depth (RRMD), which can be calculated in high dimensional data sets. The proposed procedure is based on Minimum Regularized Covariance Determinants (MRCD) estimators, this study shows that the proposed depth function is successful in finding the deepest point in high dimensional data sets with real and simulation studies up to a certain level of contamination.

DISCRETE INVERSE GAMMA DISTRIBUTION BASED LOAD-SHARE MODEL WITH APPLICATION 453

Rachna Srivastava, Pramendra Singh Pundir

In reliability engineering, the multi-component load-sharing models are being used to amplify system's reliability. This study consists of the k-component load-sharing parallel system model considering each component's failure time distribution as discrete inverse gamma. The classical and Bayesian analysis for this model is performed. The maximum likelihood estimates along with their standard errors for the parameters, system's reliability function, hazard rate function and reversed second rate of failure function are obtained. The asymptotic confidence intervals as well as two bootstrap intervals like bootstrap-p and bootstrap-t confidence intervals are constructed. Further, Bayes estimates along with their posterior standard errors and highest posterior density credible intervals for the parameters and system's reliability characteristics are obtained by using Markov Chain Monte Carlo techniques. A detailed simulation table is formed to demonstrate the effectiveness of the theory developed. Finally, a real data set is used to illustrate the applicability of the model.

SURVIVAL ANALYSIS OF A MULTI-STATE SEMI-MARKOV MODEL ON INFECTIOUS DISEASE CONSIDERING VARIOUS LEVELS OF SEVERITY 466

Sujata Sukhija, Rajeev Kumar

The aim of the paper is to carry out survival analysis of a novel multi-state model on infectious disease considering various levels of severity using semi-Markov processes. Various levels of severity of the disease over time and transitions between these severity levels have been considered. Transition probabilities and expected waiting times are derived. Expressions for mean survival time, expected total time in home isolation, and expected total time in hospital are obtained. The analysis of the proposed model is carried out through numerical computation and plotting several graphs. Important conclusions are drawn. The modelling framework proposed here can be used to model any infectious disease irrespective of disease states. The study will be helpful in designing effective measures to control the infectious disease and selecting the appropriate intervention policies.

ON AN IMPATIENT CONSUMER QUEUE WITH SECONDARY SERVICE, MULTIPLE VACATIONS AND SERVER BREAKDOWNS 481

K. Jyothsna, P. Vijaya Kumar, P. Vijaya Laxmi

This study presents a limited buffer secondary service queue with multiple vacations and server breakdowns. The model under consideration includes two types of impatient policies: balking and reneging. After the completion of the essential primary service, only few consumers choose to proceed with secondary service with a certain probability. During the active period of the server, it is subject to breakdown and the broken down server is immediately sent for repair. Further, the server will go on vacation as soon as there are no waiting consumers in the queue. On returning from a vacation, if the system is still empty the server leaves for another vacation and continues to do so until atleast one consumer is found at a vacation termination epoch. The model is analyzed under steady-state conditions and the explicit expressions of various performance indices are evaluated. A few numerical results illustrate how the model parameters have an effect on the performance metrics.

MODELING OF RELIABILITY AND SURVIVAL DATA WITH EXPONENTIATED GENERALIZED INVERSE LOMAX DISTRIBUTION 493

Sule Omeiza Bashiru, Ibrahim Ismaila Itopa

In this paper, a new four-parameter distribution is developed and studied by combining the properties of the exponentiated generalized-G family of distributions and the features of the Inverse Lomax distribution. The newly developed distribution is called the exponentiated generalized inverse Lomax distribution that extends the classical inverse Lomax distribution. The shape of the hazard rate function is very flexible because it possesses increasing, decreasing, and inverted (upside-down) bathtub shapes. Some important characteristics of the exponentiated generalized inverse Lomax distribution are derived, including moments, moment generating function, survival function, hazard function and order statistics. The method of maximum likelihood estimation is used to obtain estimates of the unknown parameters of the new model. The application of the new model is based on two real-life data sets used to show the modeling potential of the proposed distribution. The exponentiated generalized inverse Lomax distribution turns out to be the best by capturing important details in the structure of the data sets considered.

RELIABILITY ANALYSIS OF THE SHAFT SUBJECTED TO TWISTING MOMENT AND BENDING MOMENT FOR NORMALLY DISTRIBUTED STRENGTH AND STRESS 502

Md. Yakoob Pasha, M. Tirumala Devi, T. Sumathi Uma Maheswari

Shaft is the rotating component that transmits power from one place to another. The shafts are commonly subject to torsional and bending moments and combinations of these moments. In general, shafts are subjected to a combination of torsional and bending stresses. The design of a shaft is essential, subject to its strength and stress. This paper presents the reliability analysis of the shaft subjected to (a) twisting moment, (b) bending moment and (c) combined twisting and bending moment for which stress and strength follow the normal distribution.

GROUP RUNS AND MODIFIED GROUP RUNS CONTROL CHARTS FOR MONITORING LINEAR REGRESSION PROFILES 513

Onkar Ghadge, Vikas Ghute

Profile monitoring is a critical tool for manufacturing industries to evaluate and maintain quality performance, as well as detect faults. The process of profile monitoring involves observing how variables interact with one another throughout a given period. This enables the understanding of any changes in their functional relationship over time. Generally, control charts are used for monitoring profiles. This paper proposes two new methods to enhance the monitoring of simple and multiple linear regression profiles in Phase II. The proposed methods are based on group runs (GR) and modified group runs (MGR) control charting schemes. The procedure to obtain optimal design parameters for the proposed methods is discussed in detail. The effectiveness of the suggested techniques is assessed through the ARL standard. The study found that the proposed GR and MGR monitoring methods displayed superior performance compared to other available monitoring methods in the literature. A real-life example is illustrated using proposed GR and MGR charting schemes.

ON A CLASS OF LORENTZIAN PARA-KENMOTSU MANIFOLDS ADMITTING QUARTER-SYMMETRIC METRIC CONNECTION **525**

S. Sunitha Devi, K. L. Sai Prasad

In this present paper, a class of Lorentzian almost paracontact metric manifolds known as the LP-Kenmotsu (Lorentzian para-Kenmotsu) is considered that accepts a connection of quarter-symmetric. In this work, it was found that an LP-Kenmotsu manifold is either <p-symmetric or concircular <p-symmetric with respect to quarter-symmetric metric connection if and only if it is symmetric with respect to the Riemannian connection, provided the scalar curvature ofRiemannian connection is constant.

EVALUATION OF SAMPLE SIZE AND EFFICIENT FIELD SAMPLING PLAN IN HDP APPLE ORCHARDS 533

Tabasum Mushtaq, Mushtaq A. Lone, S. A. Mir, Sonali Kedar Powar, Aafaq A. Rather, Adil H. Khan, Faizan Danish

An essential stage in research is choosing an adequate sample size and sampling strategy. In order to obtain the most accurate estimates possible when surveying high density apple orchards, this paper provides the proper procedure for selecting the sample and an effective sampling strategy. For this study, primary information gathered during a two-year period from the SKUAST-Kashmir exotic apple block Plate I was employed. This investigation was conducted using the TCSA of exotic apple trees of the Gala and Fuji types. The sample was obtained using a variety of sampling techniques in order to find the parameters of population. Findings revealed that using proportional allocation of a stratified sample technique, in both the varieties, produces the most efficient population parameter estimates.

ESTIMATION OF STRESS-STRENGTH RELIABILITY BASED ON KME MODEL..... 539

Kavya P., Manoharan M.

In reliability theory the estimation of stress-strength reliability is an important problem. It has many applications in engineering and physics areas. In many practical situations, the assumption of identical strength distributions may not be quite realistic because components of a system are of different structure. Here we establish the estimation of stress-strength reliability of the KM-Exponential (KME) distribution. In this article, we consider the case that the stress-strength variables are independent. KME distribution is parsimonious in parameter and has decreasing failure rate. The stress-strength reliability based on KME model is estabished and using maximum likelihood estimation method, the estimation of the stress-strength reliability is derived and also the asymptotic distribution. Simulation method is used to show the performance of the parameters and the 95% confidence interval is also calculated. With the help of simulated data, we depict the application of the stress-strength reliability of KME distribution.

REDUNDANCY OPTIMIZATION FOR A SYSTEM COMPRISING ONE OPERATIVE UNIT AND N HOT STANDBY UNITS 547

Parveen, Dalip Singh, Anil Kumar Taneja

In many industries and applications, downtime or failure can have serious consequences, such as financial losses, safety hazards, or reputational damage. A hot standby unit can help minimize the impact of such events by providing a backup that can quickly and seamlessly take over in the event of a failure. Further, the question of as to how many hot standby units should be used also needs to be addressed. So, an N+1-Unit-system is investigated wherein N units are on hot standby, whereas one unit is operational and the system is such that the hot standby units can take over seamlessly if the single operative unit fails. The system breaks down completely when all the units fail. It is assumed that the failure rates of all the operational units and the redundant units will vary exponentially. To get different performability measurements, the regenerative point technique has been applied to optimize the value ofN.

FUZZY CONTROL CHARTS BASED ON RANKING OF PENTAGONAL FUZZY NUMBERS.... 563

Mohammad Ahmad, Weihu Cheng, Zhao Xu, Abdul Kalam, Ahteshamul Haq

A Control Chart is a fundamental approach in Statistical Process Control. When uncommon causes of variability are present, sample averages will plot beyond the control boundaries, making the control chart a particularly effective process monitoring approach. Uncertainties are caused by the measuring system, including the gauges operators and ambient circumstances. In this paper, the concept of fuzzy set theory is used for dealing with uncertainty. The control limits are to converted into fuzzy control limits using the membership function. The fuzzy X-R and X-S control chart is developed by using the ranking of the pentagonal fuzzy number system. An illustrative example is done with the discussed technique to make fuzzy X-R and X-R control charts and increase the flexibility of the control limit.

N-POWER HALF LOGISTIC-G FAMILY: APPLICATIONS TO MEDICAL AND TRAFFIC DATA 575

Pankaj Kumar, Laxmi Prasad Sapkota, Vijay Kumar

This research article introduces a novel family of distributions achieved through the methodology of the n-power transformation technique. The study focuses on one specific member that is inverse Weibull distribution within this family, which showcases a hazard function exhibiting distinct J, reverse-J, bathtub, or monotonically increasing shapes. The article explores the essential characteristics of this distribution and employs the maximum likelihood estimation (MLE) method to estimate its associated parameters. To evaluate the accuracy of the estimation procedure, a simulation experiment is conducted, revealing a decrease in biases and mean square errors as sample sizes increase, even when working with small samples. Furthermore, the practical application of the proposed distribution is demonstrated by analyzing two real medical and traffic datasets. By employing model selection criteria and conducting goodness-of-fit test statistics, the article establishes that the proposed model surpasses existing models in performance. The application of this research work can be significant in various fields where modeling and analyzing hazard functions or survival data are essential, while also making contributions to probability theory and statistical inferences.

POWER KOMAL DISTRIBUTION WITH PROPERTIES AND APPLICATION IN RELIABILITY ENGINEERING 591

Rama Shanker, Mousumi Ray, Hosenur Rahman Prodhani

The statistical analysis and modeling of reliability data from engineering is really a challenge for statistician because the reliability data from engineering are stochastic in nature. Recently one parameter Komal distribution was introduced in statistics literature for the analysis and modeling of failure time data from engineering. Komal distribution, being one parameter distribution, does not provide good fit to some engineering data due to its theoretical or applied point of view. In this article we propose a two-parameter power Komal distribution, which includes Komal distribution as particular case, for the analysis and modeling of data from reliability engineering. Its statistical properties including behavior of probability density function and cumulative distribution function for varying values of parameters have been presented. The first four raw moments and the variance of the proposed distribution has been derived and given. The expressions for hazard rate function and mean residual life function have been obtained and their behaviors for varying values of parameters have been presented. The stochastic ordering which is very much useful comparing the stochastic nature has also been discussed. Method of maximum likelihood has been discussed for estimating the parameters. Application of the distribution has been investigated using a real lifetime dataset from engineering. The goodness of fit of power Komal distribution has been tested using Akaike Information criterion and Kolmogorov-Smirnov statistic. The goodness of fit of power Komal distribution shows that it gives much closure fit over two-parameter power Garima distribution, Power Shanker distribution and Weibull distribution and one parameter exponential distribution, Shanker distribution, Garima distribution and Komal distribution. As the power Komal distribution gives much better fit over Weibull distribution, which is very much useful for modeling and analysis of data from reliability engineering, the final recommendation is that the power Komal distribution should be preferred over the considered distributions including Weibull for modeling data from reliability engineering.

PROFIT ANALYSIS OF REPAIRABLE COLD STANDBY SYSTEM UNDER REFRESHMENTS 604

Ajay Kumar, Ashish Sharma

In the generation of science and technology, every company wants to increase the reliability of their products. So, they used the concept of cold standby redundancy, timely repair of the failed unit and providing limited refreshments to the available technician when required. This paper aims to explore the system of two identical units where the primary unit is operative and the secondary unit is in cold standby mode. When the primary unit fails due to any fault then secondary unit starts working immediately. Here, times of failure of unit and technician refreshment request follow the general distribution whereas times of repair of unit and refreshment follow the exponential distributions. Such types of systems are used in industries and education systems to prevent losses. The system's performance is calculated by using concepts of mean time

to system failure, availability, busy period of the server, expected number of visits made by the server and profit function using the semi-Markov process and regenerative point technique. Tables are used to explore the performance of the system.

POWER WEIBULL QUANTILE FUNCTION AND IT'S RELIABILITY ANALYSIS 613

Jeena Joseph, Sonitta Tony

In this article, we propose a new class of distributions defined by a quantile function, which is the sum of the quantile functions of the Power and Weibull distributions. Various distributional properties and reliability characteristics of the class are discussed. To examine the usefulness of the model, the model is applied to a real life datasets. Parameters are estimated using maximum likelihood estimation technique.

GENERALISED EXPONENTIAL RATIO-CUM-PRODUCT ESTIMATOR FOR ESTIMATING POPULATION VARIANCE IN SIMPLE RANDOM SAMPLING 625

Rafia Jan, T. R. Jan, Faizan Danish

This study presents a comprehensive investigation into the estimation of population variance for a study variable using Simple Random Sampling, with the incorporation of two auxiliary variables. To address the challenges of variance estimation in complex scenarios, a novel approach termed the "Proposed Generalized Exponential Ratio-cum-Product Estimator" is introduced. This innovative estimator belongs to a class of estimators that rely on exponential functions of the auxiliary variables, providing enhanced precision and efficiency in variance estimation. To thoroughly assess the performance of the proposed estimators, the research develops equations for both Mean Square Errors and Biases, unveiling their statistical properties. The study systematically explores the conditions under which these estimators demonstrate superior efficiency compared to traditional alternative estimators, thereby enabling researchers to identify contexts where their utilization is most beneficial. The empirical aspect of the research constitutes a significant contribution to the study's validity. Through empirical analysis, the proposed estimators are directly compared against the conventional Unbiased Sample Variance Estimator, showcasing their clear superiority in terms of efficiency. Furthermore, Mean Square Errors and Percent Relative Efficiency are calculated for all estimators and subjected to theoretical and empirical comparisons with existing estimation methods. These findings corroborate the advantageous attributes of the proposed estimator in real-world scenarios, reinforcing its practicality and reliability in various research domains. Beyond methodological developments, this study also delves into the real-world implications and applications of the proposed estimators. It highlights the potential benefits of utilizing these estimators in situations where study variables exhibit intricate relationships with auxiliary variables, offering valuable insights into multifaceted data sets and multidimensional factors. Additionally, a comprehensive sensitivity analysis is undertaken to assess the robustness of the proposed estimators under varying assumptions and sampling schemes. The researchers' meticulous evaluation enhances the credibility of the proposed estimators and ensures their adaptability across diverse practical scenarios. Overall, this study's significance extends beyond statistical theory, presenting valuable practical implications for researchers and practitioners across different fields. Improved population variance estimation leads to enhanced decision-making, optimal resource allocation, and deeper insights into underlying phenomena. By introducing the proposed estimator and thoroughly examining its performance through rigorous theoretical and empirical analyses, this research lays a solid foundation for more robust and efficient variance estimation techniques. The insights gained from this study can reshape statistical practices, paving the way for advancements in diverse scientific disciplines and inspiring further knowledge exploration.

EXPONENTIAL-PARETO MIXTURE DISTRIBUTION 632

Irina Peshkova

In this paper we introduce the Exponential-Pareto mixture distribution. This distribution is associated as mixture of light and heavy-tailed data which arise in a wide class applications including risk analysis. Characteristic function, failure rate function, mean excess, conditional excess distribution are derived. It is proved that the limiting distribution of maxima among n values of rv's with Exponential-Pareto distribution has Frechet-type form. The maximal likelihood estimation of parameters is discussed. The upper bound of uniform distance between Exponential-Pareto mixture and Pareto distributions is derived.

RELIABILITY ASSESSMENTS USING STOCHASTIC DEGRADATION PROCESS FOR CURRENT TIME ANALYSIS CUMULATIVE DAMAGE MODELS 646

G. Sathya Priyanka, S. Rita, M. Iyappan

The reliability study of such incredibly reliable items is inappropriate for the use of failure time data analysis and testing methodologies. More trustworthy information can be obtained from degradation data than from standard censored failure-time data, especially in cases where few or no failures are anticipated. The market for lighting has given a lot of attention to high-power white light emitting diodes (HPWLEDs). But as one of the more dependable electronic goods, it may not be expected to fail in either a traditional or even an accelerated life test. DDDM, or data-driven degradation methodology, is used in this research. Using data on lumen maintenance gathered from the IES LM-80-08 lumen maintenance test standard and based on the general degradation path model, the dependability of HPWLED was predicted. Testing such devices in typical working situations, and occasionally even under worse conditions, is difficult enough without trying to collect an adequate amount of time-to-failure data. Modern items are made with superb quality and high reliability in mind. Some safety-critical parts and systems are even made to last for an incredibly long time in order to prevent the disastrous effects of probable breakdowns. A cumulative damage model based on stochastic degradation processes has been developed in this paper. A suitable numerical representation is used to support the analytical findings. As a result, the degradation analysis approach has been developed to address dependability modeling issues using data on product degradation gleaned from historical records or degradation testing.

SOFTWARE QUALITY ANALYSIS BASED ON SELECTIVE PARAMETERS USING ENHANCED ENSEMBLE MODEL 657

Rakhi Singh, Mamta Bansal, Surabhi Pandey

Software Quality Analysis refers to the process of evaluating and assessing the quality of software products or applications. It involves analyzing various aspects of the software to determine its level of quality, identify potential issues or defects, and make informed decisions to expand the software's overall quality. There are investigated different software quality models based on machine learning algorithms. Nevertheless, the explored approaches have an inconsistent understanding of the software product quality and high complexity. This research presents an enhanced ensemble model (EEM) that involves Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Extreme Learning Machine (ELM) to assess the optimal outcomes. This model performance is computed based on multiple selective parameters namely functional suitability and maintainability of the software and compared along with several algorithms namely Decision Tree, Random Forest, AdaBoost, and Naive Bayes. The outcome of this ensemble model demonstrates that it offers highly accurate results for software validation, verification, and overall product development process to analyze the functional suitability as well as software maintainability. The measured accuracy on Decision Tree, Random Forest, Naive Bayes, AdaBoost, and EEM is found 92.08%, 93.35%, 94.50%, 95.60%, and 99.14%, respectively

OPTIMIZATION OF SYSTEM PARAMETERS OF 2: 3 GOOD SERIAL SYSTEM USING DEEP LEARNING 670

Shakuntla Singla, Shilpa Rani, Umar Muhammad Modibbo, Irfan Ali

In this paper, Optimization of System Parameters of 2:3 Good Serial System controlled with the help of a controlling unit Using Deep Learning Optimization with packing unit in series with priority in repair and single server which never fails is carried out. There are three units of different capacities working in parallel in which if three/two units are working then the system is working at full/reduced capacity. Working of these online parallel and offline units is managed by the controlling unit, which also manages the preventive maintenance of all type of units, together with a 24/7 repair facility is modeled for reliability performance measurements. Taking exponential failure and repair rates of units and facilities a steady state transition diagram (depicting transition rates and states) is drawn using Markov process. The system parameters are modelled using Regenerative Point graphical Technique (RPGT) and optimized using Deep learning methods such as Adam, SGD, RMS prop. The results of the optimization may be used to validate and challenge existing models and assumptions about the systems.

RAP AND AVAILABILITY ANALYSIS OF MANUFACTURING SYSTEM: SMO AND PSO..... 680

Sarita Devi, Nakul Vashishth, Deepika Garg

In today's scenario manufacturing industries are highly complex and prone to failure. That's why redundancy allocation problem (RAP) and time dependent availability analysis plays a major role for the successful life cycle of a manufacturing industry. RAP is a Np-hard problem which is very difficult to solve by traditional methods. Therefore in this paper, RAP for the Manufacturing system is solved by Spider monkey optimization. SMO is recent meta-heuristic technique. Till now it is not used to solve RAP. Further results are compared with the Particle swarm optimization algorithm and comparison validates the better performance of SMO in this problem. As mentioned above, for avoiding the complete breakdown of the manufacturing system time- dependent availability is analyzed in this study. Firstly failure and repair data is collected from the manufacturing system then with the help of this information transition diagram is developed. Further equations are developed from transition diagram by using Markov birth death process then equations are solved with the use of Runge-Kutta method. This methodology is implemented in MATLAB.

INTEGRATING TESTING COVERAGE, EFFORT AND CHANGE POINT IN A SOFTWARE RELIABILITY GROWTH MODEL: A COMPREHENSIVE ANALYSIS 692

Sudeep Kumar, Anu G Aggarwal

Software reliability growth models (SRGMs) are essential for forecasting and controlling the reliability of software systems. In present article, we propose an enhanced SRGM that incorporates three important factors: testing coverage, testing effort, and change point detection. We introduce a novel testing coverage function that captures the delayed S-shaped behaviour commonly observed in software reliability growth. Weibull distribution is utilized to model the testing effort. Finally, we address the impact of change points in software reliability. To assess how well our suggested model works, we conducted experiments using real-world software failure data provide by Tandem computers. The results demonstrate that our model outperforms existing SRGMs by providing more accurate predictions and a better understanding of the interplay between testing coverage, testing effort, and change points.

PREDICTION OF RELIABILITY CHARACTERISTICS OF THRESHER PLANT BASIS ON GENERAL AND COPULA DISTRIBUTION..... 701

Urvashi, Shikha Bansal

In the agriculture field industry, farming tools play an important role. Any type of machinery's performance is influenced by factors including dependability, accessibility, and operating conditions. Different types of modern machinery are being used in today's modern world, so the farming system has become very easy. Thresher plants are essential equipment in the agriculture field industry, and these plants have many uses. A transition diagram for the system is used to develop a mathematical model of the thresher plant. Partial differential equations are created associated with the help of a transition diagram and solved using Laplace transforms and the supplementary variables approach to assessing the system's reliability. The copula approach was used to design the experiment, and the same methodology was used to assess the outcomes. The main aim of the present article is to evaluate the reliability factors, Profit, and sensitive analysis of a threshers plant. It is also possible to compute the dependability factor with the aid of general distributions and compare it to that copula distribution.

APPLICATION OF MACHINE LEARNING ALGORITHMS IN THE PROBLEMS OF IMPROVING MODE RELIABILITY OF MODERN POWER SYSTEMS..... 716

Viktor Kurbatsky, Huseyngulu Guliyev, Nikita Tomin, Famil Ibrahimov, Nijat Huseynov

In order to increase the regime reliability of energy systems, the experience of applying machine learning algorithms and models for various issues of operative-dispatching and counter-accident management was reviewed. It is indicated that an effective solution to this problem is the use of machine learning algorithms and models that are able to learn to predict and control the operating modes of the power system, taking into account many changing influencing factors. The experience of using machine learning technology in the tasks of operational dispatch and emergency control of EPS is presented, which clearly shows the prospects of such studies for subsequent practical implementation in the work of various automated control systems for electric power networks of power systems. Until recently, models based on neural network structures

have remained the most popular among machine approaches in predictive problems. The advantages of using this structure are shown, first of all, by the fact that the neural network structure makes it possible to obtain models with good approximating abilities. A comparative analysis of the effectiveness of various models in predicting electricity consumption is given. The issues of voltage and reactive power regulation in the electrical network of power systems using an artificial neural network are considered and the effectiveness of this approach is shown. A model and algorithm for estimating voltage stability in power system nodes under various influencing factors is proposed, as well as results are presented that confirm the reliability of the results obtained.

OPTIMIZATION OF PRIORITY SERVICE WITH EFFICIENT COORDINATION OF ADMISSION CONTROL, EMERGENCY VACATION OF AN UNRELIABLE SERVER 729

G. Ayyappan, S. Nithya

In this paper, we establish a single server retrial queueing system with two types of customers, admission control, balking, emergency vacation, differentiate breakdown, and restoration. There are two distinct factors which must be considered when classifying priority and ordinary customers. The non-preemptive priority discipline proposed by this concept. Ordinary and priority customers arrive in accordance with Poisson processes. For both priority and ordinary customers, the server continuously offers a single service that is distributed arbitrarily. In this study, we compute the Laplace transforms of the time-dependent probabilities of system states using a probability generating function and the supplementary variable technique. The sensitivity analysis of system descriptions is assisted by study of numerical findings.

PERFORMANCE ANALYSIS OF BULK ARRIVAL GENERAL SERVICE QUEUE WITH FEEDBACK, IMPATIENT CUSTOMERS AND SECOND OPTIONAL SERVICE 744

P. Vijaya Laxm, Hasan A. Qrewi, E. Girija Bhavani

This paper analyzes the steady state behavior of batch arrival non-Markovian service queue with feedback, balking, reneging, and second optional service (SOS). The steady-state probabilities are computed using the probability generating function. After completing the first essential service (FES), if a customer is unsatisfied with it, he may choose to rejoin the system (feedback), opt for the SOS, or depart from the system with specific probabilities. Once a customer arrives, he decides immediately to join the queue or refuses to join (balking). Furthermore, after joining the queue if a customer does not get service within a specific time, may become impatient, and decide to leave the line without getting any service (reneging). Reneging time follows exponential distribution while service time (FES and SOS) follow general distribution. Also, the cost model was presented to determine the optimal service rates to minimize the expected cost. Finally, various performance measures and numerical illustrations are provided.

RELIABILITY MODELLING OF A PARALLEL-COLD STANDBY SYSTEM WITH REPAIR PRIORITY 760

Puran Rathi, Anuradha, S.C. Malik

This paper deals with the reliability modelling of a parallel cold standby system of four units. The units operate in two phases; phase-I and phase-II. In phase-I, two identical units (called main units) work in parallel and the other two identical units (called duplicate units) have been taken as spare in cold standby. The units of phase-I and of the phase-II are not identical. The priority to repair the units of phase-I has been given over the repair of the units of the phase-II. However, no priority is given for operation of the units of both phases. There is a single repair facility which tackles all types of faults whenever occurred in the system. After repair each unit works as new and the switches devices are considered as perfect. The repair time of the units follows arbitrary probability distribution while the failure time of the units is assumed as constant. The behaviour of mean sojourn time (MST), transition probabilities, mean time to system failures (MTSF), availability, expected number of repairs for both phase-I and phase-II units, expected number of visits of the server, busy period of the server and finally the profit function are obtained in steady state by making use of well-known semi-Markov process (SMP) and Regenerative Point Technique (RPT) for arbitrary values of the parameters in steady state. Novelty and Application: A four-unit system is configured in two phases namely; phase-I and phase-II under some novel assumptions with a practical visualization in metallic bush manufacturing industries.

ANALYSIS OF M, MAP/PH\, PH2/1 NON-PREEMPTIVE PRIORITY QUEUEING MODEL WITH DELAYED WORKING VACATIONS, IMMEDIATE FEEDBACK,IMPATIENT CUSTOMER, DIFFERENTIATE BREAKDOWN AND PHASE TYPE REPAIR 771

G. Ayyappan, N. Arulmozhi

The arrival of high priority customers is governed by the Poisson process while that of low priority customers is governed by the Markovian Arrival Process, and the service times are determined by a distinct Phase-type distribution. When the service is finished and the system is empty, the server stays idle for a random period (delay time). If a customer arrives within the delayed period, the server resumes normal service to the customer immediately. Otherwise, at the end of the delayed period, the server will take a working vacation and will instantly provide slow service to customers (high priority customers only). The Matrix analytic method is used to investigate the system. We also discussed the steady-state vector and busy period for our concept. The estimated and visually displayed performance measures of the system

ANALYSIS OF MAP/ PH1, PH2, PH3/1 QUEUEING-INVENTORY SYSTEM WITH TWO COMMODITIES 791

S. Meena, N. Arulmozhi, G. Ayyappan, K. Jeganathan

In this work, a single server implements a two-commodity inventory queueing system. We assume that both commodities have a finite capacity. Customers arrive by a Markovian Arrival Process, there is a need for a single item, and either or both types of commodities are required, and this requirement is modeled using certain probabilities. The lead times are exponentially distributed, and the service times have a PH distribution. We use matrix analytical techniques to investigate the queueing inventory system and adopt an (s, S)-type replenishment policy that is dependent on the type of commodity. In the steady state, the joint and individual probability distribution of the Esystem, inventory level, and server status is obtained. A few significant performance measures are attained. Our mathematical concept is then illustrated with a few numerical examples.

EXPLORING THE ADAPTABILITY OF THE UNIT INVERSE WEIBULL DISTRIBUTION FOR MODELING DATA ON THE UNIT INTERVAL 806

Shameera T, BlNDU P.P

This paper derives a new lifetime distribution called the unit inverse Weibull distribution (UIWD) from inverse weibull distribution. Various statistical properties such as the survival function, hazard rate function, revised hazard rate function, cumulative hazard rate function, moments, and quartiles have been discussed. Additionally, we have explored other properties like skewness, kurtosis, order statistics, and the quantile function. Various methods of estimation, including maximum likelihood, moments, percentiles, and the Cramer-Von Mises, have been discussed. Simulation studies were conducted to assess the accuracy and precision of the parameters. Comparative analyses were performed to highlight the effectiveness and utility of the proposed model in comparison to other existing models, using two real-life applications. Finally, real life data analysis reveals that derived distribution can provide a better fit than several well-known distributions.

On MIXTURE OF BURR XII AND NAKAGAMI DISTRIBUTIONS: PROPERTIES AND APPLICATIONS 821

Hemani Sharma, Parmil Kumar

The Burr XII and Nakagami distributions hold significant importance in both lifetime distribution and wealth distribution analyses. The Burr XII distribution serves as a valuable tool for understanding the distribution of wealth and wages within specific societies, while the Nakagami distribution finds its application in the realm of communication engineering. The incorporation of finite mixture distributions, aimed at accounting for unobserved variations, has gained substantial traction, particularly in the estimation of dynamic discrete choice models. This research delves into the fundamental characteristics of the mixture Burr XII and Nakagami distributions. The study introduces parameter estimation techniques and explores various aspects, including the cumulative distribution function, hazard rate, failure rate, inverse hazard function, odd function, cumulative hazard function, rth moment, moment generating function, characteristic function,

moments, mean and variance, Renyi and Beta entropies, mean deviation from mean, and mean time between failures (MTBF). The paper also addresses the estimation of the mixing parameter through a Bayesian approach. To illustrate the effectiveness of the proposed model, two real-life datasets are examined.

STARTING MODE OF SYNCHRONOUS MACHINES WITH MASSIVE ROTORS 842

Laman Hasanova, Nurali Yusifbayli

It is known that in synchronous machines with massive rotors, it is required to take into account the change in the equivalent rotor of active resistance depending on the frequency of the current in the rotor at starting of these machines. A three-phase mathematical model of these machines has been compiled, the equations of which are written in axes rotating at the speed of the machine rotor. Study of the start-up modes of these machines and operation in synchronous mode with a load surge (dynamic mode) has been carried out on this model. The studies have allowed for making the following conclusions. When starting synchronous machines with massive rotors, it is most preferable to take into account the change in the equivalent resistance in the form of a linear sliding function. It was found out that in the synchronous mode of operation of these machines, including in stable dynamic modes, there is no need to take into account changes in the rotor resistance-sliding, since the sliding in these modes oscillatory damps around zero, thereby not affecting changes in the value of the equivalent resistance. Therefore, in these modes it is recommended to consider the value of this resistance constant and definite at slip equal to zero.

DESIGNING OF PHASE II ARL-UNBIASED 52-CHART WITH IMPROVED PERFORMANCE USING REPETITIVE SAMPLING 850

Sonam Jaiswal, Nirpeksh Kumar

In this paper, we consider a two-sided Phase II S2-chart with probability limits because the surveillance of both an increase and decrease in the process variance plays a decisive role in a continuous quality improvement program. We propose a two-sided average run length unbiased S2-chart under the repetitive sampling with probability limits for a fixed in-control average run length and average sample size to eliminate the average run length biasedness. It is well established that the Shewhart-type charts are less sensitive to detect small to moderate changes in the process parameters. Therefore, a repetitive sampling scheme is taken into consideration to improve the S2-chart's ability to detect changes in the process variance. Under the repetitive sampling methodology, the detection ability of the chart is improved by using more than one samples if the first sample does not provide sufficient evidence to decide whether the process is in-control or out-of-control. The proposed chart is compared with the existing charts, such as the equal tailed standard Shewhart S2-chart and unequal tailed S2-chart under repetitive sampling. Results show that the proposed chart is more efficient than the existing chart. Finally, an illustration has been provided in the favor of the proposed chart with the help of a published dataset.

EVALUATION OF PERFORMANCE MEASURES FOR RELIABLE AND SECURE PHISHING DETECTION SYSTEM 861

Pratikkumar A Barot, Sunil A Patel, H B Jethva

Phishing is an illegal act and security breach which acquires a user's confidential information without consent. Anti-phishing techniques used to detect and prevent such malicious acts to provide data safety to the end user. Researchers proposed an anti-phishing solution with the help of techniques like the blacklist record, heuristic function, visual similarity, and machine learning algorithm. In recent times many researchers proposed machine learning techniques for phishing detection and achieve more than 90% accuracy. However, there is reliability issue in the accuracy measures used by the researchers. In real life, the phishing dataset is unbalanced. Most of the researchers ignore this data quality during their research work design. In the case of unbalanced data, traditional accuracy measure does not give proper performance evaluation. It shows biased performance evaluations. In this paper, we experimented with an unbalanced dataset of phishing detection and did detailed result analysis to highlight the reliability issue of traditional performance evaluation measures for unbalanced data classification. We experiment with four classification algorithms and found that more than 90% of accuracy does not entitle any classifier as secure and safe if the dataset is unbalanced. Our work highlights the data factors and algorithmic limitations that compromise the system security and data safety.

DEVELOPING A NEW HUNTSBERGER TYPE SHRINKAGE ESTIMATOR FOR THE ENTROPY OF EXPONENTIAL DISTRIBUTION UNDER DIFFERENT LOSS FUNCTIONS 871

Priyanka Sahni, Rajeev Kumar

The aim of the paper is to develop a new Huntsberger type shrinkage estimator for the entropy function of the exponential distribution. The present paper proposes a Huntsberger type shrinkage estimator for the entropy function of the exponential distribution. This Huntsberger type shrinkage entropy estimator is based on test statistic, which eliminates arbitrariness of choice of shrinkage factor. For the developed estimator risk expressions under LINEX loss function and squared error loss function have been calculated. To assess the efficacy of the proposed estimator, numerical computations are performed, and graphical analysis is carried out for risk and relative risks for the proposed estimator. It is also compared with the existing best estimator for distinct degrees of asymmetry and different levels of significance. Based on the criteria of relative risk, it is found that the proposed Huntsberger type shrinkage estimator is better than the existing estimator for the entropy function of the exponential distribution for smaller values of level of significance and degrees of freedom.

THE SABUR DISTRIBUTION: PROPERTIES AND APPLICATION RELATED TO ENGINEERING DATA 882

Aijaz Ahmad, Afaq Ahmad, Aafaq A. Rather

This paper introduces a novel probability distribution called the Sabur distribution (SD), characterized by two parameters. It offers a comprehensive analysis of this distribution, encompassing various properties such as moments, moment-generating functions, deviations from the mean and median, mode and median, Bonferroni and Lorenz curves, Renyi entropy, order statistics, hazard rate functions, and mean residual functions. Furthermore, the paper delves into the graphical representation of the probability density function, cumulative distribution function and hazard rate function to provide a visual understanding of their behavior. The distribution's parameters are estimated using the well-known method of maximum likelihood estimation. The paper also showcases the practical applicability of the Sabur distribution through real-world examples, underscoring its performance and relevance in various scenarios.

COST & PROFIT ANALYSIS OF TWO-DIMENSIONAL STATE M/M/2 QUEUING MODEL WITH CORRELATED SERVERS, MULTIPLE VACATION, BALKING AND CATASTROPHES 893

Sharvan Kumar, Indra

The present study obtains the time-dependent solution of a two-dimensional state Markovian queuing model with infinite capacity, correlated servers, multiple vacation, balking and catastrophes. Inter arrival times follow an exponential distribution with parameters A and service times follow Bivariate exponential distribution BVE (p, p, v) where p is the service time parameter and v is the correlation parameter. Both the servers go on vacation with probability one when there are no units in the system and the servers keeps on taking a sequence of vacations of random length each time the system becomes empty, till it finds at least one unit in the system to start each busy period referred as multiple vacation. The unit finds a long queue and decides not to join it; may be considered as balking. All the units are ejected from the system when catastrophes occur and the system becomes temporarily unavailable. The system reactivates when new units arrive. Occurrence of catastrophes follow Poisson distribution with rate £,. Laplace transform approach has been used to find the time-dependent solution. By using differential-difference equations, the recursive expressions for probabilities of exactly i arrivals and j departures by time t are obtained. The probabilities of this model are consistent to the results of "Pegden & Rosenshine". The model estimates the total expected cost, total expected profit and obtained the optimal values by varying time t for cost and profit. These important key measures give a greater understanding of the model behaviour. Numerical analysis and graphical representations have been done by using Maple software.

INNOVATIVE METHODS OF ENSURING THE FUNCTIONAL SAFETY OF TRAIN CONTROL SYSTEMS 909

I.B. Shubinsky, E.N. Rozenberg, H. Schabe

The paper examines the specificity of the modern intelligent control systems. The Big Data technology and Data Science algorithms open up great potential in train traffic management based on hazard prevention. An example is given of high reliability and acceptable accuracy of hazardous railway infrastructure failure prediction using methods based on artificial intelligence. A great deal of attention is given to economical methods of ensuring the required levels of functional safety of train control systems. For that purpose, the efficiency of the digital twin-based method was evaluated. It is shown that, under certain conditions, this method allows significantly reducing the cost of a control system while achieving an acceptably high level of functional safety. The method of virtual second channels is based on the same principle of using information redundancy rather than hardware redundancy. The paper presents and analyses the method of virtual second channels in respect to an axle counter-based train control system. It is established that it is possible to ensure a safety integrity level of an entire control system with a virtual second channel at least as high as SIL3. The above methods ensure, on the one hand, a reduction of the amount of equipment and significantly lower cost of the systems and, on the other hand, requires the creation of additional software and substantiation of the acceptability of the achieved level of functional safety. This matter is within the competence of the developer of the control system.

A FUZZY INNOVATIVE ORDERING PLAN USING STOCK DEPENDENT HOLDING COST OF INSPECTION WITH SHORTAGES IN TIME RELIABILITY DEMAND USING TFN 921

Sivan V, Thirugnanasambandam K, Sivasankar N, Sanidari

The considerations in this paper are, the demand is consistent with time deterioration, the holding cost is dependent based on the quantity of stock available in the system, and the ordering cost is linear and time-dependent. This system should be considered in terms of fuzziness. It is assumed that the shortages are permitted partially, the order is inspected, defective items are identified, by using penalty cost, the defective items should be minimized. Under the classical model and fuzzy environment, the mathematical equation is arrived at to find the optimal solution of total relevant cost with optimal order quantity and time using triangular fuzzy numbers. Defuzzification has been accomplished through the use of the signed distance method of integration. The solutions have been arrived and the model numerical problem of three levels of values (lower, medium and upper) in parametric changes has been verified. Using Sensitivity analysis, the solution is used to validate the changes in different parameter values of the system. To demonstrate the convexity of the TRC function over time, it has used a three-dimensional mesh graph.

STOCHASTIC ANALYSIS OF JUICE PLANT SUBJECT TO REPAIR FACILITY 940

Amit Kumar, Pinki Kumari

The performance of a juice plant is analyzed by using the base state and the regenerative point graphical technique. The juice plant under consideration consists of three distinct units. It is considered that units A and B may be in a complete failed state through partial failure mode but unit C is in only partially failed state. If one of the units A or B or C partially fails then the system works in a reduced state. When any unit is completely failed then the system is in failed state and no unit can fail further when the system is in a failed state. A technician is always available to repair the failed unit. In this paper, the failure time and repair time follow general distributions. Tables are used to describe the reliability measures such as mean time to system failure, availability and profit values of juice plant.

STOCHASTIC MODEL ON EARLY-STAGE BREAST CANCER WITH TWO TYPES OF TREATMENTS 948

Suman, Rajeev Kumar

The aim of the paper is to study effectiveness of the different treatments of early-stage breast cancer through analysis of a stochastic model. The early-stage breast cancer is a term used to describe breast cancer that is detected at an early stage of development, typically before it has spread to other parts of the body. Early detection of breast cancer is critical as it greatly increases the chances of successful treatment and saves lives. At early-stage breast cancer of the patient, two types of treatment namely, tamoxifen and tamoxifen combined with radiation therapy are commonly used. As it is essential to

consider innovative and cost-effective strategies for early detection and treatment. Investigations through analysis of the stochastic model on early-stage breast cancer with these two types of treatments may help the stakeholders. Keeping this in view, in the present paper, a stochastic model is developed for the early-stage breast cancer considering two treatment types, namely tamoxifen and tamoxifen combined with radiation therapy. The model is analyzed by Markov process and regenerative point technique. Mean sojourn time refers to the average amount of time spent by a patient in a particular state before transitioning to another state and mean survival time refers to the average time a patient survives after diagnosis of breast cancer. Mean sojourn time and mean survival time have been calculated. Sensitivity analysis is a technique to understand how changes in input variables or parameters affect the output or outcome of a model and it helps assess the robustness, reliability, and stability of a model by quantifying the impact of variations in input factors. The paper also includes sensitivity and relative sensitivity analyses of the model which explore the impact of different parameters on the survivability of the patient. The MATLAB software has been used for numerical computing and plotting various graphs. The investigation through our analysis of the stochastic model shows that the mean survival time lessens with the rise in the rates of transition and mean survival time from the treatment, tamoxifen plus radiation is higher than the treatment, tamoxifen only. It is concluded that tamoxifen plus radiation is more effective and useful than only tamoxifen for treatment of early-stage breast cancer.

SYNTHETIC RELIABILITY MODELING AND PERFORMANCE ENHANCEMENT FOR MULTI-UNIT SERIAL SYSTEMS: UNVEILING INSIGHTS VIA GUMBEL-HOUGARD FAMILY COPULA APPROACH 964

Ismail Muhammad Musa, Ibrahim Yusuf

This paper presents a comprehensive study of a series-parallel system comprising four interconnected subsystems: subsystem-1, subsystem-2, subsystem-3, and subsystem-4. Subsystem-1 stands as a single unit, subsystem-2 consists of three identical units in active parallel, subsystem-3 involves two identical units in series, while subsystem-4 incorporates two identical units in parallel. The system operates under good conditions, considering various failure rates and repair rates. The investigation employs Laplace transforms and Supplementary variable techniques to analyze the system's performance. Key reliability parameters, including Availability, Reliability, Mean Time to Failure (MTTF), Sensitivity, and Cost, are evaluated for specific values of failure and repair rates. The paper delves into the intricate analysis of a multi-unit series system, focusing on its reliability and performance evaluation. The study employs the Gumbel-Hougard Family Copula approach, a sophisticated and robust methodology to capture the interdependencies among system units. By utilizing this advanced technique, the paper provides a comprehensive understanding of the system's behavior under varying operating conditions. Various reliability and performance metrics, including Availability, Mean Time to Failure (MTTF), and Component Importance Measures, are thoroughly examined, offering valuable insights for optimizing the system's reliability and performance. The results are presented in a clear and visually appealing manner, utilizing tables and figures to aid in the comprehension of the findings.

A BAYESIAN PREDICTION FOR THE TOTAL FERTILITY RATE OF AFGHANISTAN USING THE AUTO-REGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) MODEL 980

Sayed Rahmi Khuda Haqbin and Athar Ali Khan

In this article, a simple methodology to predict the total fertility rate of Afghanistan via a Bayesian statistical analysis method has been applied. R- statistical analysis tool is used for data analysis. To forecast, the "bayesforecast" package is needed. It is a substitute package in Rfor the "forecast" package in the traditional (frequentest) statistical method. The Bayesian data analysis using the specific case of the general auto-regressive integrated moving average model (ARIMA) is processed as follows; As the first step, the stationarity of the given data-set is assessed, the time series has been made stationary by taking differences. After fitting several models, as the most appropriate fitted model, the ARIMA (1, 2,1) model has been fitted to the data. The accuracy of the fitted model is examined, and thereafter, the developed model is analyzed. The posterior computation is done, using the Markov Chain Monte Carlo (MCMC) simulation method. The method ultimately focuses on drawing relevant inferences including the 16 years prediction, and the results are; in general, found to be satisfactory.

EXPONENTIATED DISCRETE HYPO EXPONENTIAL DISTRIBUTION AND ITS GENERALIZATIONS 998

Krishnakumari.K, Dais George

Generalizations of standard probability distributions is a thought-provoking concept in statistical literature and was inspired by many researchers in recent days. This is because the addition of parameters may increase the flexibility of the new models. Now a days various generalization techniques are available in literature. In this work, we proposed a generalization of discrete hypo exponential distribution and studied its various properties. A real data analysis is carried out and check the flexibility of the new model by comparing it with other standard distributions. Two generalizations of the proposed distribution are introduced.

AN INNOVATIVE APPROACH FOR RELIABILITY MODELING OF HVDC CONVERTER STATION 1011

Aditya Tiwary, R. S. Mandloi

Assessment of reliability indices is important when availability and unavailability of the system or systems or components or group of components are to be assessed. There are various reliability indexes which are very important for overall performance of any complex engineering system. Reliability block diagram modeling is required to be formulated for evaluating different essential and important reliability parameters of any complex engineering system. In view of above, in this paper, reliability block diagram modeling of HVDC converter station is represented and formulated. The schematic diagram of the HVDC converter station is available in literature and based on that schematic diagram the modeling of HVDC converter station is formulated in this paper. After the reliability block diagram modeling of HVDC converter station, the mean time to failure (MTTF) of each and every components of HVDC converter station are also evaluated and represented in the result and discussion section. The reliability of each and every component of the HVDC converter station is evaluated and expressed in result section. Assessment of unavailability is also obtained and shown in result section.

TIME DEPENDENT BEHAVIOUR OF A SINGLE SERVER QUEUEING SYSTEM WITH DIFFERENTIATED WORKING VACATIONS SUBJECT TO SYSTEM DISASTER 1019

V Karthick, V Suvitha

This study investigates the time dependent behaviour of the single server queue with differentiated working vacations. The model also takes into account the possibility of a disaster happening during busy periods and working vacations, with the repair procedure starting right away. The time-dependent probabilities of system size are described in terms of modified Bessel functions in the paper using explicit equations that were generated using generating functions. Numeric instances have been added to support the theoretical findings even more.

ON THE PROPERTIES AND APPLICATIONS OF TOPP-LEONE GOMPERTZ INVERSE RAYLEIGH DISTRIBUTION 1032

Sule Omeiza, Bashirub O.Y., Halid Anyigba, Kogi State,

In this study, we introduce a new four-parameter continuous probability distribution known as the Topp-Loene Gompertz Inverse Rayleigh (TLGoIRa) distribution. This novel model extends the Gompertz Inverse Rayleigh distribution. We present various mathematical properties of the distribution, including moments, moment generating functions, quantile functions, survival functions, hazard functions, reversed hazard functions, and odd functions. We also derive the distribution of order statistics, yielding both the maximum and minimum order statistics. This process of parameter estimation using the maximum likelihood estimation method is discussed. Furthermore, we present two real-life applications that illustrate the effectiveness and robustness of the TLGoIRa distribution when compared to several considered lifetime models. Our analysis reveals that the TLGoIRa distribution demonstrates superior robustness in comparison to the competing lifetime models. Additionally, the study highlights the distribution's efficacy in fitting biomedical datasets.

SOME APPLICATIONS OF TRANSMUTED LOG-UNIFORM DISTRIBUTION 1046

Ashin K Shaji, Rani Sebastian

As a generalization of the Log Uniform distribution, Transmuted Log - Uniform distribution is introduced and its properties are studied. We obtained graphical representations of its pdf, cdf, hazard rate function and survival function. We have derived statistical properties such as moments, mean deviations, and the quantile function for the Transmuted Log-Uniform distribution. We also obtained the order statistics of the new distribution. Method of maximum likelihood is used for estimating the parameters. Estimation of stress strength parameters is also done. We applied the Transmuted Log-Uniform distribution to a real data set and compared it with Transmuted Weibull distribution and Transmuted Quasi-Akash distribution. It was found that the Transmuted Log-Uniform distribution was a better fit than the Transmuted Weibull distribution and Transmuted Quasi-Akash distribution distributions based on the values of the AIC, CAIC, BIC, HQIC, the Kolmogorov-Smirnov (K-S) goodness-of-fit statistic and the p-values.

MEASUREMENTS OF BRIDGE STRUCTURES USING NON-DESTRUCTIVE TESTING METHODS AND THEIR STABILITY IN WIND GUSTS 1056

Alena Rotaru

As bridge structures become older and older, they are subject to wear and tear due to ageing, weather conditions or environmental effects, as well as due to surprise structural modifications substantially affecting the condition of the structures. Therefore, the condition assessment of bridge structures is a must for the safety and absence of risk. The condition assessment of bridge structures is also necessary for the maintenance and repair of existing structures having been in service for more than 30 years, in order to avoid breakdowns and save human lives. This paper states the condition assessment performed with the use of various nondestructive test methods.

STATISTICAL MODELS FOR FORECASTING NATURAL EMERGENCIES..... 1067

Valery Akimov, Maxim Bedilo, Olga Derendiaeva

The article considers predictive-analytical solutions for natural hazards for urbanized areas, the mathematical basis of which is Bayesian classifiers. The result of the work is a formalized description of models for predicting forest fires, the consequences of earthquakes and floods resulting from floods.

EXPLORING NOVEL EXTENSION OF SUJA DISTRIBUTION: UNVEILING PROPERTIES AND DIVERSE APPLICATIONS

C. Subramanian[1], M. Subhashree[2], Aafaq A. Rather[3,*]

•

[1,2] Department of Statistics, Annamalai University, Annamalai Nagar, Tamil Nadu, India
[3,*]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[1]manistat@yahoo.co.in, [2]Subhashreem17@gmail.com, [3,*]aafaq7741@gmail.com

Abstract

This research article introduces and explores the area biased techniques of the Suja distribution, presenting novel derivations and insights. The estimation of the one-parameter area biased Suja distribution is accomplished using maximum likelihood, providing a robust framework for modeling real-world data. A comprehensive study of several statistical properties is conducted to unveil the characteristics and behaviors of this new model. To demonstrate its practical applicability, the proposed distribution is applied to real data of weather temperature. The analysis showcases the distribution's effectiveness in capturing the intricacies of temperature patterns, revealing its potential utility in weather modeling and related applications. The research contributes to the advancement of statistical modeling techniques and enriches our understanding of the Suja distribution's versatility in handling diverse datasets.

Keywords: Area-biased distribution, Suja distribution, Maximum likelihood Estimator

1. Introduction

The concept of area- biased distribution was explained earlier by Fisher [5]. He first introduced weighted distribution which is a combination of model specification and data interpretation. Also, stated that "the estimation of frequencies based on the effective methods of ascertainment". The size-biased are the special cases of the weighted distribution. Later, Rao [16] explained that the weighted distributions as many applications such as medicine, reliability, ecology, behavioral science, finance, insurance, etc. A discrete Poisson area-biased Lindley distribution was purposed by Bashir and Rasul [3] and explained its properties. Also, applied in biological data and compared with Poisson distribution. Zahida et.al. [18] Purposed a new extension of Weibull distribution called area- biased weighted Weibull distribution. The characterization of this model was derived and shown how the model fitted to the problem of ball bearing data. Rama Shanker [10] introduced a new one- parameter Suja distribution and estimated the parameter using method of moments and maximum likelihood. The important properties were explained and finally, compared this model with other one-parameter distributions by applying a real lifetime data.

Ayesha Fazal [2] introduced an area-biased Poisson exponential distribution with its moments and other properties, Also, the goodness of fit for this model has been discussed to show how it fit in real

data sets. Many studies on length biased distribution case has been published, for example; Rather and Subramanian [11], Rather and Subramanian [12], Rather and Ozel [13], Rather and Subramanian [19], Rather et al. [18]. A new generalized area- biased Aradhana distribution was introduced by Elangovan and Mohanasundari [4] and estimated the parameters by maximum likelihood estimator. Finally, applied a real lifetime data set in the model to show how it works. The new extension of Suja distribution called length-biased Suja distribution was given by Ibrahim Al-omari and Islam Khaled [6]. The various properties of this model were explained and shown the usefulness of the model in the real data set. Later, Ibrahim Al-omari et.al. [7] extended the length-biased Suja distribution to power length-biased Suja distribution with its characteristics and estimated the two parameters of this proposed distribution by maximum likelihood method. Finally, they illustrated a real data to show the performance of this model. Arun Kumar Rao and Himanshu Pandey [1] studied the parameter estimation of area-biased Rayleigh distribution by using maximum likelihood estimation and Bayesian estimation with quasi and inverted gamma priors. The weighted Suja distribution as an extension of most important Suja distribution was discussed by Islam Khaled and Ibrbhim Al-Omari [8]. Also, explained its statistical properties and application in the ball bearings data. The new generalization method as area-biased was introduced by Nuri Celik [9] for beta, Rayleigh and log-normal distributions. Also, the main statistical properties and parameters estimation were studied. At last, some of the real data examples were used. Shanker, Upadhyay and Shukla [17] were given a new two parameter quasi Suja distribution with its characteristics and parameters estimation. Also, they illustrated with real data to show its performance.

2. Area-biased Suja Distribution (ABSD)

The probability density function (p.d.f.) and cumulative distribution function (c.d.f.) of the Suja distribution is given by

$$f(y,\alpha) = \frac{\alpha^5}{\alpha^4+24}(1+y^4)e^{-\alpha z}\,; y>0, \alpha>0 \tag{1}$$

$$F(y,\alpha) = 1-\left[1+\frac{(\alpha^4y^4+4\alpha^3y^3+120\alpha^2y^2+24\alpha y)}{\alpha^4+24}\right]\,; y>0, \alpha>0 \tag{2}$$

We know that, the weighted function is

$$f_w(y) = \frac{w(y)f(y)}{E(w(y))}\,; x>0 \tag{3}$$

Where, $w(x)$ be a non-negative weight function. $E\big(w(y)\big) = \int_0^\infty w(y)f(y)dy < \infty$, $w(y) = y^c$

$$f_w(y) = \frac{y^c f(y)}{E(y^c)}\,; y>0 \tag{4}$$

where, $E(y^c) = \int_0^\infty y^c\, f(y)\, dy$

For area-biased function, put c=2,

$$f_{AB}(y) = \frac{y^2 f(y)}{E(y^2)}\,; y>0 \tag{5}$$

where,

$$E(y) = \int_0^\infty y^2\, f(y)dy \tag{6}$$

substitute (1) in (6), we get

$$E(y) = \frac{2(\alpha^4+360)}{\alpha^2(\alpha^4+24)} \tag{7}$$

Thus, the p.d.f. and c.d.f. of the area-biased Suja distribution (ABSD) can be obtained as

$$f_{ABSD}(y;\alpha) = \frac{\alpha^7}{2(\alpha^4+24)}y^2(1+y^4)e^{-\alpha y}\,;\; y>0; \alpha>0 \tag{8}$$

$$F_{ABSD}(y;\alpha) = \frac{1}{2\alpha^4 + 720}[\alpha^4\gamma(3,\alpha y) + \gamma(7,\alpha y)] \quad (9)$$

The graphical representation of the p.d.f. and c.d.f. of area-biased Suja distribution (ABSD) are also shown below:

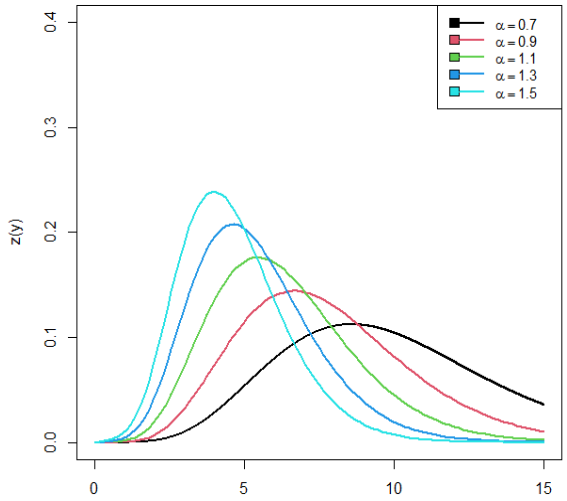


Fig. 1: *Pdf plot of ABSD*

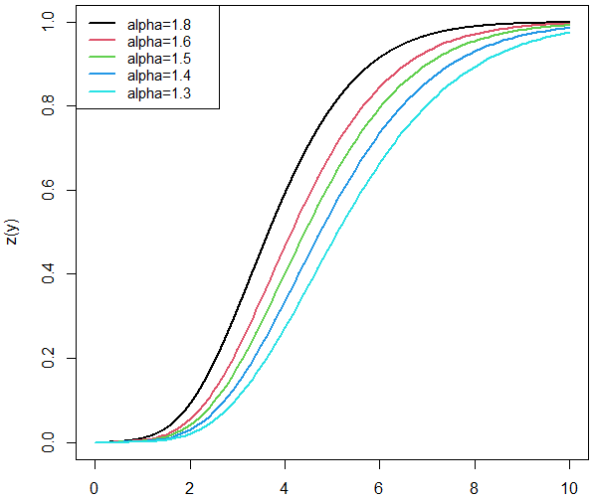


Fig. 2: *Cdf plot of ABSD*

3. Reliability Analysis

3.1 Reliability Function

The survival function of the area-biased Suja distribution is given by

$$S(y) = 1 - F(y)$$

$$S(y) = 1 - \frac{1}{2\alpha^4 + 720}[\alpha^4\gamma(3,\alpha y) + \gamma(7,\alpha y)]$$

3.2 Hazard Function

The hazard function is also known as the hazard rate, instantaneous failure rate or force of mortality and is given by

$$h(y) = \frac{f_{ABSD}(y;\alpha)}{1 - F_{ABSD}(y;\alpha)}$$

$$h(y) = \frac{\alpha^7 y^2(1+y^4)e^{-\alpha y}}{(2\alpha^4 + 720) - [\alpha^4\gamma(3,\alpha y) + \gamma(7,\alpha y)]}$$

3.3 Reverse Hazard Function

The reverse hazard function or reverse hazard rate is given by

$$h_r(y) = \frac{f_{ABSD}(y;\alpha)}{F_{ABSD}(y;\alpha)}$$

$$h_r(y) = \frac{\alpha^7 y^2(1+y^4)e^{-\alpha y}}{[\alpha^4\gamma(3,\alpha y) + \gamma(7,\alpha y)]}$$

3.4 Mill's Ratio

The Mills ratio of the area-biased Suja distribution is

$$\text{Mills Ratio} = \frac{1}{h_r(y)}$$

$$= \frac{[\alpha^4\gamma(3,\alpha y) + \gamma(7,\alpha y)]}{\alpha^7 y^2(1+y^4)e^{-\alpha y}}$$

4. Moments

The moments of ABSD have been derived to describe the characteristic of the proposed model. Then, the r^{th} order moment E(y^r) of ABSD is derived as

$$\mu_r' = E(y^r) = \int_0^\infty y^r F_{ABSD}(y;\alpha)dy \tag{10}$$

$$\mu_r' = \frac{\alpha^7}{2\alpha^4+720}\int_0^\infty y^{r+2}(1+y^4)e^{-\alpha y}\,dy \tag{11}$$

$$\mu_r' = \frac{\alpha^4\Gamma(\mathrm{r}+3) + \Gamma(\mathrm{r}+7)}{\alpha^r(2\alpha^4+720)} \tag{12}$$

In equation (12), when *r*=1, the mean of ABSD which is given by

$$\mu_1' = \frac{3(\alpha^4+840)}{\alpha(\alpha^4+360)}$$

Similarly, when *r*=2, 3, 4 in equation (4.1), we will get

$$\mu_2' = \frac{12(\alpha^4+1680)}{\alpha^2(\alpha^4+360)}$$

$$\mu_3' = \frac{60(\alpha^4+3024)}{\alpha^3(\alpha^4+360)}$$

$$\mu_4' = \frac{360(\alpha^4+5040)}{\alpha^4(\alpha^4+360)}$$

The central moments about the mean of this distribution are given as

$$\mu_0 = \mu_0' = 1$$

$$\mu_0 = 0$$

$$\mu_2 = \mu_2' - (\mu_1')^2 = \frac{12(\alpha^4+1680)}{\alpha^2(\alpha^4+360)} - \left(\frac{3(\alpha^4+840)}{\alpha(\alpha^4+360)}\right)^2$$

Therefore, the variance of ABSD is

$$\mu_2 = \frac{9(\alpha^8+2160\alpha^4+571200)}{\alpha^2(\alpha^4+360)^2}$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3 = \frac{60(\alpha^{12}+8280\alpha^8+388800\alpha^4+108864000)}{\alpha^3(\alpha^4+360)^3}$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^3 - 3(\mu_1')^4$$

$$\mu_4 = \frac{45(\alpha^4+360)(\alpha^{12}+21048\alpha^8+15871680\alpha^4+741208320)}{\alpha^4(\alpha^4+360)^4}$$

The standard deviation (σ), co-efficient of variation (C.V.), co-efficient of skewness ($\sqrt{\beta_1}$), co-efficient of kurtosis (β_2) and index of dispersion (γ) of ABSD are obtained as

$$\sigma = \frac{3\sqrt{\alpha^8+2160\alpha^4+571200}}{\alpha(\alpha^4+360)}$$

$$C.V. = \frac{\sigma}{\mu_1'} = \frac{\sqrt{\alpha^8+2160\alpha^4+571200}}{(\alpha^4+840)}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = \frac{2(\alpha^{12}+8280\alpha^8+388800\alpha^4+108864000)}{9(\alpha^8+2160\alpha^4+571200)^{\frac{3}{2}}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2} = \frac{5(\alpha^{12}+21048\alpha^8+15871680\alpha^4+7412083200)}{\alpha^2(\alpha^4+360)(\alpha^8+2160\alpha^4+571200)}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{3(\alpha^8+2160\alpha^4+571200)}{\alpha(\alpha^4+360)(\alpha^4+840)} \tag{13}$$

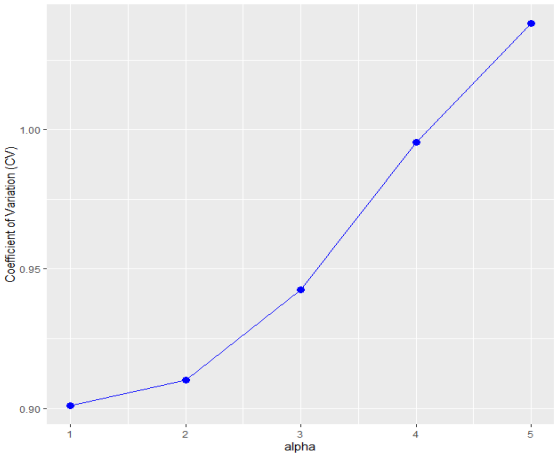


Fig. 3: *Coefficient of variation for ABSD*

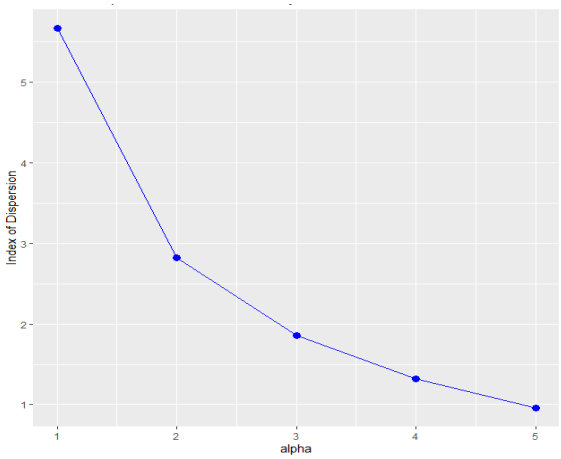


Fig. 4: *Index of dispersion for ABSD*

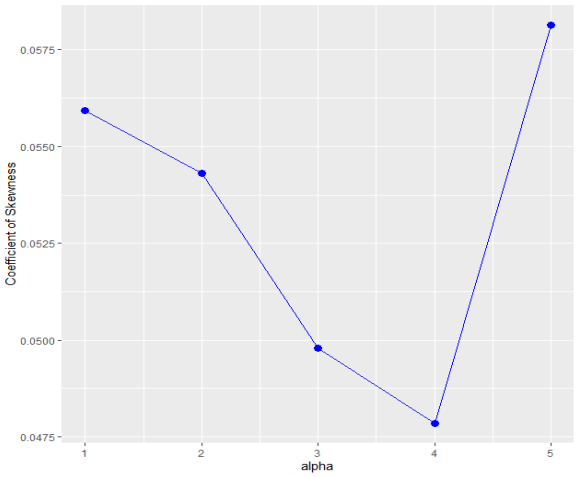


Fig. 5: *Coefficient of skewness for ABSD*

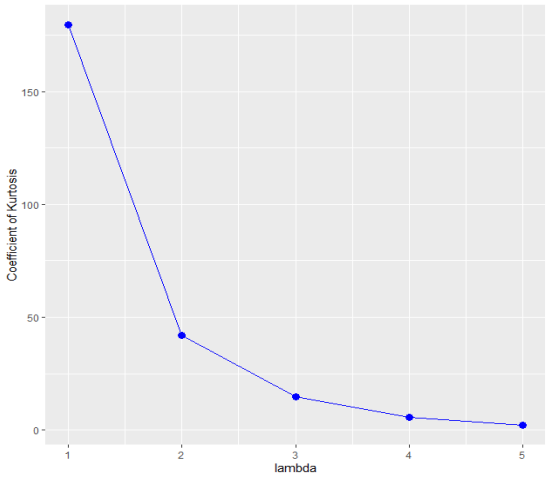


Fig. 6: *Coefficient of kurtosis for ABSD*

4.1 Harmonic Mean

The Harmonic mean of the aspired model can be derived as

$$H.M. = E\left[\frac{1}{y}\right] = \int_0^\infty \frac{1}{y} f_{ABSD}(y)dy \tag{14}$$

$$H.M. = \int_0^\infty \frac{1}{y}\frac{\alpha^7}{2(\alpha^4+360)} y^2(1+y^4)e^{-\alpha y}\,dy \tag{15}$$

Therefore,

$$H.M. = \frac{\alpha(\alpha^4+120)}{2(\alpha^4+360)}$$

4.2 Moment Generating Function and Characteristic Function

Assume Y have area-biased Suja distribution, we will get the moment generating function of Y as

$$M_y(t) = E(e^{ty}) = \int_0^\infty e^{ty} f_{ABSD}(y)dy \tag{16}$$

Using Taylor's expansion,

$$M_y(t) = \int_0^\infty \sum_{j=0}^\infty \frac{(ty)^j}{j!} f_{ABSD}(y)dy = \sum_{j=0}^\infty \frac{t^j}{j!}\mu_j' \tag{17}$$

$$M_y(t) = \frac{1}{2(\alpha^4+360)} \sum_{j=0}^\infty \frac{t^j}{j!\alpha^j}\left(\alpha^4\Gamma(\mathrm{r}+3) + \Gamma(\mathrm{r}+7)\right) \tag{18}$$

In the same way, we will get the characteristics function of SBJD can be obtained as

$$\varphi_y(t) = E[e^{ity}] = \frac{1}{2(\alpha^4+360)} \sum_{j=0}^\infty \frac{(it)^j}{j!\alpha^j}\left(\alpha^4\Gamma(\mathrm{r}+3) + \Gamma(\mathrm{r}+7)\right) \tag{19}$$

5. Order Statistics

Let $Y_1, Y_2, \ldots\ldots\ldots, Y_n$ be the random variable drawn from the continuous population. Their p.d.f. be $f_y(y)$ and cumulative density function be $F_y(y)$. Then, assume $Y_{(1)}, Y_{(2)}, \ldots\ldots\ldots, Y_{(n)}$ be the order statistics of a random sample.

Thus, the probability density function of rth order statistics $Y_{(r)}$ is given by

$$f_{y_{(r)}}(y) = \frac{n!}{(n-r)!\,(r-1)!} f(y)\left[F_y(y)\right]^{r-1}\left[1-F_y(y)\right]^{n-r} \tag{20}$$

Putting the equation (8) and (9) in equation (20), the probability density function of rth order statistics $Y_{(r)}$ of ABSD is given by

$$f_{y_{(r)}}(y) = \frac{n!}{(n-r)!(r-1)!}\left\{\frac{\alpha^7y^2(1+y^4)e^{-\alpha y}}{2(\alpha^4+360)}\right\} \mathrm{x} \left\{\frac{[\alpha^4\gamma(3,\alpha y)+\gamma(7,\alpha y)]}{2(\alpha^4+360)}\right\}^{r-1} \mathrm{x}\left\{1-\frac{[\alpha^4\gamma(3,\alpha y)+\gamma(7,\alpha y)]}{2(\alpha^4+360)}\right\}^{n-r} \tag{21}$$

Then, the probability density function of higher order statistics $Y_{(n)}$ can be derived as

$$f_{y_{(n)}}(y) = \left\{\frac{n\alpha^7y^2(1+y^4)e^{-\alpha y}}{2(\alpha^4+360)}\right\} \mathrm{x} \left\{\frac{[\alpha^4\gamma(3,\alpha y)+\gamma(7,\alpha y)]}{2(\alpha^4+360)}\right\}^{n-1} \tag{22}$$

Hence, the probability density function of 1st order statistics $Y_{(1)}$ can be obtained as

$$f_{y_{(1)}}(y) = \left\{\frac{n\alpha^7y^2(1+y^4)e^{-\alpha y}}{2(\alpha^4+360)}\right\} \mathrm{x} \left\{1-\frac{[\alpha^4\gamma(3,\alpha y)+\gamma(7,\alpha y)]}{2(\alpha^4+360)}\right\}^{n-1} \tag{23}$$

6. Maximum Likelihood Estimator and Fisher Information Matrix

The maximum likelihood estimator is the best numerical stability estimator to estimate the parameters of the distribution when compared with other estimating methods. Thus, we used this method to estimate the parameters of ABSD which is derived below:

Let $Y_{(1)}, Y_{(2)}, \ldots\ldots\ldots, Y_{(n)}$ be the random sample of size n drawn from the ABSD, then, the likelihood function of ABSD is

$$L(y;\alpha) = \prod_{i=1}^n f_{ABSD}(y;\alpha) = \frac{\alpha^{7n}}{2^n(\alpha^4+360)^n}\prod_{i=1}^n y_i^2(1+y_i^4)e^{-\alpha y_i}$$

The natural log likelihood function is

$$\log L(y;\alpha) = 7n\log\alpha - n\log(2\alpha^4+720) + 2\sum_{i=1}^n \log y_i + \sum_{i=1}^n \log(1+y_i^4) - \alpha\sum_{i=1}^n y_i \tag{24}$$

By differentiating equation (24) with respect to α, the maximum likelihood estimates of α can be attained as

$$\frac{\partial \log L}{\partial \alpha} = \frac{7n}{\alpha} - \frac{8n\alpha^3}{(2\alpha^4+720)} - \sum_{i=1}^n y_i = 0 \tag{25}$$

Because of the complicated form of likelihood equation (25), algebraically it is very difficult to solve the system of non-linear equation. Therefore, we use R and wolfram Mathematica for estimating the required parameters.

7. Likelihood ratio Test

This test is used to compare the goodness of fit of the two models based on the ratio of their likelihoods. Suppose $Y_{(1)}, Y_{(2)}, \ldots\ldots\ldots, Y_{(n)}$ be a random sample from the ABSD. To test, the random sample of size n for ABSD, the hypothesis is

$$H_0: f(y) = f_{SD}(y;\alpha) \quad against \quad H_1: f(y) = f_{ABSD}(y;\alpha)$$

To check whether the random sample of size n comes from the Suja distribution or size-biased Suja distribution, the likelihood ratio is

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{f_{ABSD}(y;\alpha)}{f_{SD}(y;\alpha)} = \left[\frac{\alpha^2(\alpha^2+24)}{2(\alpha^4+360)}\right]^n \prod_{i=1}^{n} y_i^2$$

Therefore, the null hypothesis is rejected if

$$\Delta = \left[\frac{\alpha^2(\alpha^2+24)}{2(\alpha^4+360)}\right]^n \prod_{i=1}^{n} y_i^2 > k$$

$$\Delta^* = \prod_{i=1}^{n} y_i^2 > k^* \; where \; k^* = k\left[\frac{\alpha^2(\alpha^2+24)}{2(\alpha^4+360)}\right]^n > 0$$

We can conclude, for large sample size n, 2log is distributed as chi-square distribution with one degree of freedom. Also, p-value is attained from the chi-square distribution. If p (Δ^*>k^*), where $k^* = \prod_{i=1}^{n} y_i$ is less than the specified level of significance and $\prod_{i=1}^{n} y_i$ is the observed value of the statistic Δ^*, then, reject null hypothesis.

8. Applications

It is a measure to check how a statistical model fits a data set. Here, we will discuss how the proposed model is fit to a data set which is illustrated below. Also, compare with Suja distribution to show better fit of area-biased Suja distribution. Let us represent the data set of weather temperature in Bangladesh from 2016 to 2019 by Shawkat Sujon in the website: https://www.kaggle.com/datasets/shawkatsujon/bangladesh-weather-dataset-from-1901-to2019. The data set is given as follows:

Table 1: *Data of weather temperature in Bangladesh from 2016 to 2019*

Year	2016	2016	2016	2016	2016	2016	2016
Month	1	2	3	4	5	6	7
Temperature	17.34	22.12	25.93	28.25	27.94	28.96	28.17
Year	2016	2016	2016	2016	2016	2017	2017
Month	8	9	10	11	12	1	2
Temperature	28.85	28.58	27.74	23.45	21.43	19.36	21.31
Year	2017	2017	2017	2017	2017	2017	2017
Month	3	4	5	6	7	8	9
Temperature	23.69	26.91	28.76	28.58	28.27	28.57	28.54
Year	2017	2017	2017	2018	2018	2018	2018
Month	10	11	12	1	2	3	4
Temperature	27.16	24.14	20.59	17.59	21.18	25.41	26.79
Year	2018	2018	2018	2018	2018	2018	2018
Month	5	6	7	8	9	10	11
Temperature	27.32	28.68	28.59	28.88	28.66	26.17	22.76

Year	2018	2019	2019	2019	2019	2019	2019
Month	12	1	2	3	4	5	6
Temperature	19.13	19.38	20.71	24.4	27.41	28.92	29.18
Year	2019	2019	2019	2019	2019	2019	
Month	7	8	9	10	11	12	
Temperature	28.72	29	28.28	26.94	23.88	18.51	

In order to compare the distributions, we study the criteria like Akaike Information Criterion (AIC), Akaike Information Criterion Corrected (AICC), Bayesian Information Criterion (BIC) and -2logL. The better distribution is which compatible to lesser values of AIC, BIC, AICC and -2logL.

Table 2: *Comparison of distributions*

Distribution	MLE	S.E	-2logL	AIC	BIC	AICC
Suja	$\hat{\alpha} = 0.1965312$	0.0126848	327.644	329.644	331.5152	329.730
Length-biased Suja	$\hat{\alpha} = 0.2358445$	0.0138967	319.7202	321.7202	323.5914	321.8071
Area-biased Suja	$\hat{\alpha} = 0.2751537$	0.0150105	313.2269	315.2269	317.0981	315.3138

9. Conclusion

In conclusion, this article presents a comprehensive investigation of the area-biased Suja distribution (ABSD), a noteworthy extension of the weighted distribution paradigm. The obtained probability density function (p.d.f.) and cumulative distribution function (c.d.f.) enrich the theoretical foundations of the ABSD, laying a solid groundwork for its application in various domains.

One significant advantage of the ABSD is its robust parameter resiliency, which leads to enhanced performance and more accurate results compared to other distributions. The maximum likelihood estimator proves to be an effective tool for estimating the distribution's parameter, and its validity is confirmed through the likelihood ratio test, reinforcing the reliability of our findings. The in-depth analysis of various statistical properties provides valuable insights into the ABSD's behavior and characteristics, fostering a deeper understanding of this novel model.

In particular, when applied to weather temperature data, the ABSD demonstrates superior compatibility compared to both the standard Suja distribution and the length-biased Suja distribution. The results suggest that the ABSD better captures the intricate patterns and variations inherent in weather temperature datasets.

The findings presented in this article not only enhance the understanding of the ABSD but also contribute to the broader field of statistical modeling. The improved performance in weather temperature analysis highlights the practical applicability of the ABSD in real-world scenarios. This distribution's versatility and ability to handle diverse datasets make it a promising candidate for future research and application across various domains. As the field of statistical modeling continues to evolve, the ABSD offers a valuable addition to the toolkit of data analysts and researchers alike.

References

[1] Arun Kumar Rao and Himanshu Pandey, (2020), *Parameter Estimation of Area-biased Rayleigh Distribution*, International Journal of Statistics and Applied Mathematics, Vol. 5, No. 4, pp.27-33.

[2] Ayesha Fazal, (2018), *Area-biased Poisson Exponential Distribution with Applications*, Biometrics and Biostatistics International Journal, Vol. 7, No. 3, pp. 256-261.

[3] Bahir, S. and Rasul, M., (2016), *Poisson Area-biased Lindley Distribution and its Applications on Biological Data*, Biometrics and Biostatistics International Journal, Vol. 5, No.1, pp 1-10.

[4] Elangovan, R. and Mohanasundari, R., (2019), *A New Area-biased Distribution with Applications in Cancer Data*, Journal of Science, Technology and Development, Vol.8, No. 12, pp. 1-14.

[5] Fisher, R.A., (1934), *The Effects of Methods of Ascertainment Upon the Estimation of Frequencies*, Annual Eugenics, Vol. 6, No. 1, pp. 13-25.

[6] Ibrahim Al-Omari, A. and Islam Khaled, A., (2019*), Length-biased Suja Distribution and its Applications*, Journal of Applied Probability and Statistics, Vol. 14, No.3, pp. 95-116.

[7] Ibrahim Al-Omari, A., Alhyasat, K., Ibrahim, K. and Abu Bakar, M. A., (2019), *Power Length-biased Suja Distribution: Properties and Applications*, Electronical Journal of Applied Statistical Analysis, Vol.12, no. 2, pp. 429-452.

[8] Islam Khaled, A. and Ibrahim Al-Omari, A., (2020), *Weighted Suja Distribution with Application to Bell Bearings Data*, Journal of Life Cycle Reliability and Safety Engineering, Vol. 9, pp. 195-211.

[9] Nuri Celik, (2022), *Applications in some Area-biased distribution*s, Asian Journal of Probability and Statistics, Vol. 19, No. 3, pp. 42-50.

[10] Rama Shanker, (2017), *Suja Distribution and its Application*, International Journal of Probability and Statistics, Vol. 6, No. 2, pp. 11-19.

[11] Rather, A. A. & Subramanian C. (2018). Characterization and estimation of length biased weighted generalized uniform distribution, *International Journal of Scientific Research in Mathematical and Statistical Science*, vol.5, issue. 5, pp. 72-76.

[12] Rather, A. A. & Subramanian C. (2018). Length Biased Sushila distribution, *Universal Review*, vol. 7, issue. XII, pp. 1010-1023.

[13] Rather, A. A. & Ozel G. (2021). A new length-biased power Lindley distribution with properties and its applications, *Journal of Statistics and Management Systems*, DOI: 10.1080/09720510.2021.1920665

[14] Rather, A. A. & Subramanian, C. (2019), The Length-Biased Erlang–Truncated Exponential Distribution with Life Time Data, *Journal of Information and Computational Science*, vol-9, Issue 8, pp 340-355.

[15] Rather, A. A., Subramanian, C., Shafi, S., Malik, K. A., Ahmad, P. J., Para, B. A., and Jan, T. R. (2018). A new Size Biased Distribution with applications in Engineering and Medical Science, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, Vol.5, Issue.4, pp.75-85.

[16] Rao, C. R., (1955), *On Discrete Distributions Arising Out of Methods of Ascertainment*, In Classical and Contagious Discrete Distributions, Patil, G.P., Eds., Pergamon Press.

[17] Shanker, R., Upadhyay, R. and Shukla, K.K., (2022), *A Quasi Suja Distribution*, Reliability: Theory and Applications, Vol. 17, No. 3, pp. 162-178.

[18] Zahida Praveen, Zulfiqar Ahmed and Munir Ahmad, (2016), *On Area-biased Weighted Weibull Distribution*, Science International (Lahore), Vol. 28, No. 4, pp. 3669-3679.

STATISTICAL MODELS FOR FORECASTING EMERGENCY SITUATIONS OF A BIOLOGICAL AND SOCIAL CHARACTER

Valery Akimov, Ekaterina Ivanova, Irina Oltyan

•

Federal State Budgetary Institution "All-Russian Research Institute for Civil Defense and Emergencies of the Ministry of Emergency Situations of Russia" (Federal Center for Science and High Technologies)
akimov@vniigochs.ru

Abstract

The article considers a statistical model for predicting emergency situations of a biological and social nature. Particular attention is paid to the calculation of indicators of resource provision of the medical care system and mortality rates during the spread of the epidemic.

Keywords: forecasting model; Bayesian classifiers; emergency situations of a biological and social nature; indicators of resource provision of the medical care system; mortality rates during the spread of the epidemic.

I. Introduction

The spread of the new coronavirus infection COVID-19 in the world in three years has led to more than five hundred million infections and more than six million deaths. The high dynamics of growth in morbidity and mortality has led not only to a serious burden on the healthcare system in almost all countries of the world, but also to the development of scientific methods for modeling epidemics and pandemics. In particular, in [1], a statistical model for predicting epidemics using Bayesian classifiers is proposed.

The main input data for the formation of the basic training set of this model are the following groups of parameters: input data characterizing the settlement (hereinafter - NP); input data characterizing the socio-economic indicators of the NP; input data characterizing the composition and size of the population of the NP; input data characterizing the demographic indicators of the NP; input data characterizing the symptoms and course of a respiratory viral disease (hereinafter referred to as RVD); input data characterizing the level of provision of the population with medical care resources in the settlement; input data characterizing the epidemiological situation in the NP (with the development of an epidemic caused by RVD); results of forecasting and modeling of the spread of the epidemic caused by the PR in the NP.

II. Methods

Calculation of indicators of resource provision of the medical care system.

The indicators of the availability of resources for the system of providing medical care in the territory during the spread of the epidemic include [2]: the indicator of the availability of hospitalized infectious patients with beds; indicator of availability of oxygen concentrators; indicator of availability of intensive care beds; the indicator of availability of artificial lung ventilation devices (hereinafter - ALV); indicator of provision of medical institutions with senior medical personnel; indicator of provision of medical institutions with paramedical personnel.

The indicator of bed capacity of hospitalized infectious patients (I_{ib}) is determined by the formula [3]:

$$I_{ib} = \frac{K_{\text{к}.il}}{K_{gi}},$$

Where:

$K_{\text{к}.il}$ - the number of beds for hospitalized infectious patients, units;

K_{gi} - the total number of hospitalized infectious patients, people.

The index of provision with oxygen concentrators ((I_{kk}) is determined by the formula:

$$I_{\text{кк}} = \frac{K_{\text{к.кк}}}{K_{\text{кк}}},$$

Where:

$K_{\text{к.кк}}$ — number of oxygen concentrators, units;

$K_{\text{кк}}$ — the total number of hospitalized, with the need for oxygen supply, pers.

The indicator of provision with an intensive care bed fund (I_{it}) is determined by the formula:

$$I_{it} = \frac{K_{\text{к}.it}}{K_{tb}},$$

Where:

$K_{\text{к}.it}$ — number of beds in the intensive care unit (hereinafter referred to as RIIT), units.;

K_{tb} — total number of hospitalized infectious patients, pers.

The indicator of provision with ventilators (or similar devices) (I_{ivl}) is determined by the formula:

$$I_{ivl} = \frac{K_{\text{а}.ivl}}{K_{ivl}},$$

Where:

$K_{\text{а}.ivl}$ — number of ventilators (or similar devices), units;

K_{ivl} — total number of hospitalized infectious patients with the need to connect to ventilators (or similar devices), people

The indicator of provision of medical institutions with senior medical personnel (I_{stmp}) is determined by the formula:

$$I_{stmp} = \frac{K_{stmp}}{K_{gi}}$$

Where:

K_{stmp} — indicator of senior medical personnel, people;

K_{gi} — total number of hospitalized infectious patients, pers.

The indicator of provision of medical institutions with paramedical personnel (I_{srmp}, units per 1000 people) is determined by the formula:

$$I_{srmp} = \frac{K_{srmp}}{K_{gi}}$$

Where:

K_{srmp} — indicator of the average medical personnel, people;

K_{gi} — total number of hospitalized infectious patients, pers.

Calculation of mortality rates during the spread of the epidemic

The indicators of mortality during the spread of the epidemic and in the conditions of its absence include [2]: the mortality rate due to respiratory diseases during the spread of the epidemic; mortality rate due to diseases of the circulatory system during the spread of the epidemic; mortality rate due to neoplasms during the spread of the epidemic; overall mortality rate due to respiratory diseases, diseases of the circulatory system and neoplasms during the spread of the epidemic (in the absence of it).

The mortality rate due to respiratory diseases ((I_{sod}, units per 1000 people) is determined by the formula [4]:

$$I_{sod} = \frac{K_{sod}}{K_{np}} \cdot 1000,$$

Where:

K_{sod} — the number of deaths due to respiratory diseases per month, people;

K_{np} — population of the settlement, pers.

Mortality rate due to diseases of the circulatory system (I_{sbk}, unit per 1000 people.) is determined by the formula:

$$I_{sbk} = \frac{K_{sbk}}{K_{np}} \cdot 1000,$$

Where:

K_{sbk} — the number of deaths due to diseases of the circulatory system per month, people;

K_{np} — population of the settlement, pers.

The mortality rate due to neoplasms (I_{sn}) is determined by the formula:

$$I_{sn} = \frac{K_{sn}}{K_{np}} \cdot 1000,$$

Where:

K_{sn} — the number of deaths due to neoplasms per month, people;

K_{sp} — population of the settlement, pers.

Then the overall mortality rate due to respiratory diseases, diseases of the circulatory system and neoplasms during the spread of the epidemic is determined by the following relationship:

$$I_{sm} = I_{sod} + I_{sbk} + I_{sn},$$

Where:

I_{sod} — mortality rate due to respiratory diseases, units. per 1000 people;

I_{sbk} — mortality rate due to diseases of the circulatory system, units. per 1000 people;

I_{sn} — mortality rate due to neoplasms, units per 1000 people.

III. Results

Thus, this article considers a statistical model for predicting emergency situations of a biological and social nature. Particular attention is paid to the calculation of indicators of resource provision of the medical care system and mortality rates during the spread of the epidemic.

IV. Discussion

The discussion of the verbal and mathematical foundations of predictive modeling of emergency situations of a biological and social nature during the development of an epidemic caused by EIA is quite active in the scientific literature [5–10].

At the previous conference RISK - 2022, the authors presented reports on the issues of predictive modeling of natural and man-made emergencies [11, 12].

References

[1] Predictive and analytical solutions for natural, man-made and biological and social threats of a unified system of information and analytical support for the safety of the environment and public order "Safe City" / V. A. Akimov, A. V. Mishurny, O. V. Yakimyuk [and etc.]. - Moscow: All-Russian Research Institute for Civil Defense and Emergency Situations of the Ministry of Emergency Situations of Russia, 2022. - 315 p. – ISBN 978-5-93970-278-2. – EDN MGXNYI.

[2] Akimov, V. A. Determining the indicators of resource provision of the medical care system and mortality rates during the spread of the epidemic / V. A. Akimov, O. A. Derendyaeva, E. O. Ivanova // Application of mathematical methods to problem solving EMERCOM of Russia: Proceedings of Section No. 14 of the XXIII International Scientific and Practical Conference, Khimki, March 01, 2023. - Khimki: Academy of Civil Protection of the Ministry of the Russian Federation for Civil Defense, Emergency Situations and Elimination of Consequences of Natural Disasters named after Lieutenant General D.I. Mikhailika, 2023. - S. 18-22. – EDNETMRTH.

[3] Akimov, V. A. Mathematical models for predicting the consequences of mass diseases of people / V. A. Akimov, E. O. Ivanova, O. A. Derendyaeva // Civil defense on guard of peace and security: Proceedings of the VII International Scientific and Practical conference dedicated to World Civil Defense Day. In the year of the 90th anniversary of the formation of the Academy of the State Fire Service of the Ministry of Emergency Situations of Russia: in 5 parts, Moscow, March 01, 2023. Volume Part V. - Moscow: Academy of the State Fire Service of the Ministry of the Russian Federation for Civil Defense, Emergencies and Disaster Relief, 2023. - P. 227-233. – EDN TLQAQX.

[4] Akimov, V. A. Mathematical models of epidemics and pandemics as sources of emergency situations of a biological and social nature / V. A. Akimov, M. V. Bedilo, E. O. Ivanova // Civil Security Technologies. - 2022. - T. 19, No. 3 (73). - P. 10-14. – EDN IPFEND.

[5] Akimov, V. A. Modern methods of studying emergency situations of natural, technogenic and biological and social nature / V. A. Akimov, M. V. Bedilo // Civil defense on guard of peace and security: Materials of the VI International scientific and practical conference dedicated to World Civil Defense Day. In 4 parts, Moscow, March 01, 2022 / Comp. V.S. Butko, M.V. Aleshkov, S.V. Podkosov, A.G. Zavorotny [i dr.]. Volume Part I. - Moscow: Academy of the State Fire Service of the Ministry of the Russian Federation for Civil Defense, Emergencies and Disaster Relief, 2022. - P. 16-24. – EDN HSSYVB.

[6] Akimov, V. A., Bedilo M. V., Sushchev S. P. Study of emergency situations of natural, technogenic and biological and social nature by modern scientific methods. - Moscow: All-Russian Research Institute for Civil Defense and Emergencies of the Ministry of Emergency Situations of Russia, 2021. - 179 p. – ISBN 978-5-93970-249-2. – EDN WUKXKC.

[7] Security of Russia. Legal, socio-economic and scientific and technical aspects. Analysis and provision of security from emergency situations / V. A. Akimov, A. A. Antyukhov, E. V. Arefieva [and others]; Security Council of the Russian Federation, Russian Academy of Sciences, EMERCOM of Russia, Rostekhnadzor, Russian Science Foundation, Rostec State Corporation, Rosatom State Corporation, Rosneft Oil Company PJSC, Russian Railways OJSC, Transneft PJSC, Gazprom PJSC. - Moscow: MGOF "Knowledge", 2021. - 500 p. – ISBN 978-5-87633-199-1. – EDN FXIJPZ.

[8] Akimov, V. A. Technique of ranking emergency situations of natural, technogenic and biological-social nature according to the degree of their catastrophicity / V. A. Akimov, I. Yu. Oltyan, E. O. Ivanova // Civil Security Technologies. - 2021. - T. 18, No. 1 (67). - P. 4-7. – DOI 10.54234/CST.19968493.2021.18.1.67.1.4. – EDN IOGGXC.

[9] Akimov, V. A. Applications of the General Theory of Safety to the Study of Emergency Situations of a Natural, Technogenic and Biological and Social Character / V. A. Akimov // Civil Security Technologies. - 2021. - T. 18, No. S. - P. 13-28. – DOI 10.54234/CST.19968493.2021.18.S.2.13. – EDN LRYKFU.

[10] Akimov, V. A. Investigation of emergency situations of natural, technogenic and biological and social nature by methods of post-non-classical science: problem statement / V. A. Akimov, A. I. Ovsyanik, E. O. Ivanova // Fires and emergency situations : prevention, elimination. - 2021. - No. 2. - P. 95-100. – DOI 10.25257/FE.2021.2.95-100. – EDN YBDTJO.

[11] Akimov, V. Forecast of natural emergency situations with modern methods / V. Akimov, M. Bedilo, O. Derendiaeva, E. Ivanova, I. Oltyan // RT&A, Special Issue No. 4 (70), Volume 17, November 2022. - S. 71 - 77. - DOI: https://doi.org/10.24412/1932-2321-2022-470-71-77.

[12] Akimov, V. Forecast modeling of man-made emergencies with modern methods/ V. Akimov, M. Bedilo, O. Derendiaeva, E. Ivanova, I. Oltyan // RT&A, Special Issue No. 4 (70), Volume 17, November 2022. - S. 318 - 323. - DOI: https://doi .org/10.24412/1932-2321-2022-470-318-323

A NEW EXTENDED EXPONENTIATED DISTRIBUTION WITH PROPERTIES AND APPLICATIONS

K. Selvakumar[1], Jameel A. Ansari[2], K. Kavitha[3], C. Subramanian[4], D. Vedavathi Saraja[5], Arshad Ahmad Khan[6], Rashid A. Ganaie[7], Aafaq A. Rather[8,*], Maryam Mohiuddin[9]

•

[1]Department of Mathematics, University College of Engineering-Nagercoil-629004, Tamil Nadu, India

[2]Department of Engineering Sciences, Faculty of Science and Technology, Vishwakarma University, Pune, India

[3]Department of Mathematics, S.A. Engineering College, Chennai, Tamil Nadu, India

[4,5,7]Department of Statistics, Faculty of Science, Annamalai University, Tamil Nadu, India

[6]Department of Mathematics, School of Chemical Engineering and Physical Sciences, Lovely Professional University, Phagwara, Punjab, India

[8,*]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India

[9]Department of Mathematics, National Institute of Technology, Calicut, Kerala, India

[1]selvakumaruce@gmail.com, [2]jameelahmad.ansari@vupune.ac.in, [3]kavidass.math@gmail.com, [4]manistat@yahoo.co.in, [5]sarajayoganand@gmail.com, [6]karshid.maths@gmail.com, [7]rashidau7745@gmail.com, [8]aafaq7741@gmail.com, [9]masmariam7@gmail.com

Abstract

This manuscript focuses on the statistical properties and estimation methods of the exponentiated Suja distribution, which is characterized by two parameters: scale and shape. From a frequentist perspective, our primary emphasis is on estimation techniques. Additionally, we derive statistical and reliability characteristics for the model. We explore various estimation procedures, including order statistics, entropies, reliability analysis, and the maximum likelihood method. To assess the model's superiority and practical utility, we analyze real lifetime data sets. Overall, this study provides a comprehensive analysis of the exponentiated Suja distribution, offering insights into its statistical properties, estimation techniques, and real-life applications.

Keywords: Exponentiated technique, Suja distribution, Order statistics, Entropies, Reliability analysis, Maximum likelihood Estimation.

1. Introduction

A new theory of distributions was introduced by Gupta et al. [1], who discussed a new family of distributions namely the exponentiated exponential distribution. The family has two parameters scale and shape, which are similar to the weibull and gamma family. Later Gupta and Kundu [2] studied some properties of the distribution. They observed that many properties of the new family are similar to those of the weibull or gamma family. Hence the distribution can be used an alternative to a weibull

or gamma distribution. The two parametric gamma and weibull are the most popular distributions for analyzing any lifetime data. The gamma distribution has a lot of applications in different fields other than lifetime distributions. the two parameters of gamma distribution represent the scale and the shape parameter and because of the scale and shape parameter, it has quite a bit of flexibility to analyze any positive real data. But one major disadvantage of the gamma distribution is that, if the shape parameter is not an integer, the distribution function or survival function cannot be expressed in a closed form. This makes gamma distribution little bit unpopular as compared to the weibull distribution, whose survival function and hazard function are simple and easy to study. Nowadays exponential distributions and their mathematical properties are widely studied for applied science experimental data sets. Rodrigues *et al.* [3] studied the exponentiated generalized Lindley distribution. Hassan et al. [4] discussed the exponentiated Lomax geometric distribution with its properties and applications. Nasiru *et al*. [5] obtained exponentiated generalized power series family of distributions. Rather and Subramanian [6] discussed the exponentiated Mukherjee-Islam distribution. Rather and Subramanian [7] obtained the exponentiated Ishita distribution with properties and applications. Maradesa Adeleke [10] discussed exponentiated exponential Lomax distribution and its Properties. Nasir *et al* [11] obtained the exponentiated Burr XII power series distribution with properties and its applications. Recently, Rather and Subramanian [8] discussed the exponentiated Garima distribution with applications in engineering sciences.

Suja distribution is a newly introduced one parametric lifetime model proposed by Shanker [9] for engineering sciences. The potentiality and usefulness of the proposed distribution in modeling lifetime data was greater as compared to other one parametric distributions namely Lindley, exponential, sujatha, Shanker and Aradhana. The different statistical properties of the proposed model have been derived and discussed such as order statistics, moments and associated measures, hazard and mean residual life function, stochastic ordering, Bonferroni and Lorenz curves and stress strength reliability. The parameters of the proposed distribution are estimated by employing the maximum likelihood estimation method. Finally, the goodness of fit of the proposed Suja distribution has been described by analyzing the real life data set and the fit has been found quite satisfactory over Lindley, exponential, Shanker, Aradhana, Sujatha and Amarendra distributions.

2. Exponentiated Suja Distribution (ESD)

The probability density function (pdf) of Suja distribution is given by

$$g(x)=\frac{\theta^5}{\theta^4+24}\left(1+x^4\right)e^{-\theta x};\ x>0,\theta>0 \tag{1}$$

and the cumulative distribution function (cdf) of the Suja distribution is given by

$$G(x)=1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x};\ x>0,\theta>0 \tag{2}$$

A random variable X is said to have an exponentiated distribution, if its cumulative distribution function is given by

$$F_\alpha(x)\ =\ \left(G(x)\right)^\alpha\ ;\ x\in R',\ \alpha>0 \tag{3}$$

Then X is said to have an exponentiated distribution.

The probability density function of X is given by

$$f_\alpha(x)=\alpha\left(G(x)\right)^{\alpha-1}g(x) \tag{4}$$

on Substituting (2) in (3), we will get the cumulative distribution function of Exponentiated Suja distribution

$$F_{\alpha}(x)=\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha};\ x>0,\theta>0,\alpha>0 \tag{5}$$

and the probability density function of Exponentiated Suja distribution can be obtained as

$$f_{\alpha}(x)=\frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1} \tag{6}$$

The graphical representation of Pdf and Cdf are shown in Fig. 1 and Fig. 2.

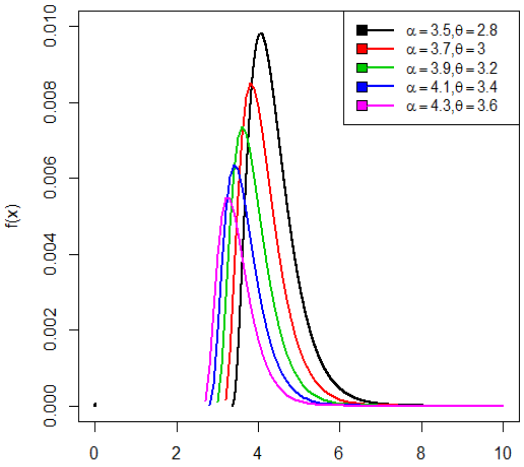


Fig. 1: *Pdf plot of exponentiated Suja distribution*

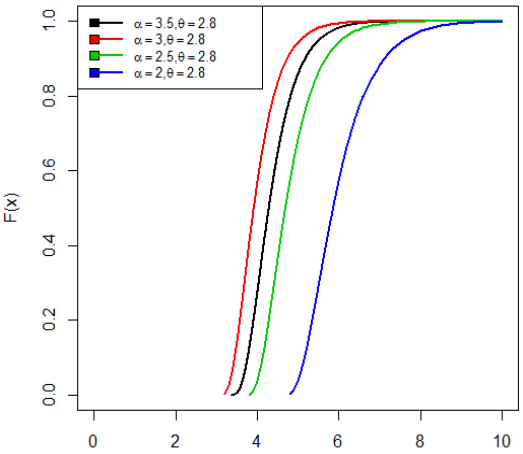


Fig. 2: *Cdf plot of exponentiated Suja distribution*

3. Reliability Analysis

In this section, we will discuss the survival function, hazard function and Reverse hazard rate function of the Exponentiated Suja distribution. Many researchers have discussed system reliability by using different techniques like Tillman et al [14], Wei et al. [15] and Wang [16].

The survival function of Exponentiated Suja distribution is given below and its graphical representation is in Fig. 3.

$$S(x)=1-\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha} \tag{7}$$

The hazard function is also known as hazard rate, instantaneous failure rate or force of mortality and is given by

$$h(x)=\left(\frac{\dfrac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\dfrac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}}{1-\left(1-\left(1+\dfrac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha}}\right) \tag{8}$$

The reverse hazard rate is given by

$$h_r(x) = \frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{(\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x)e^{-\theta x}}$$

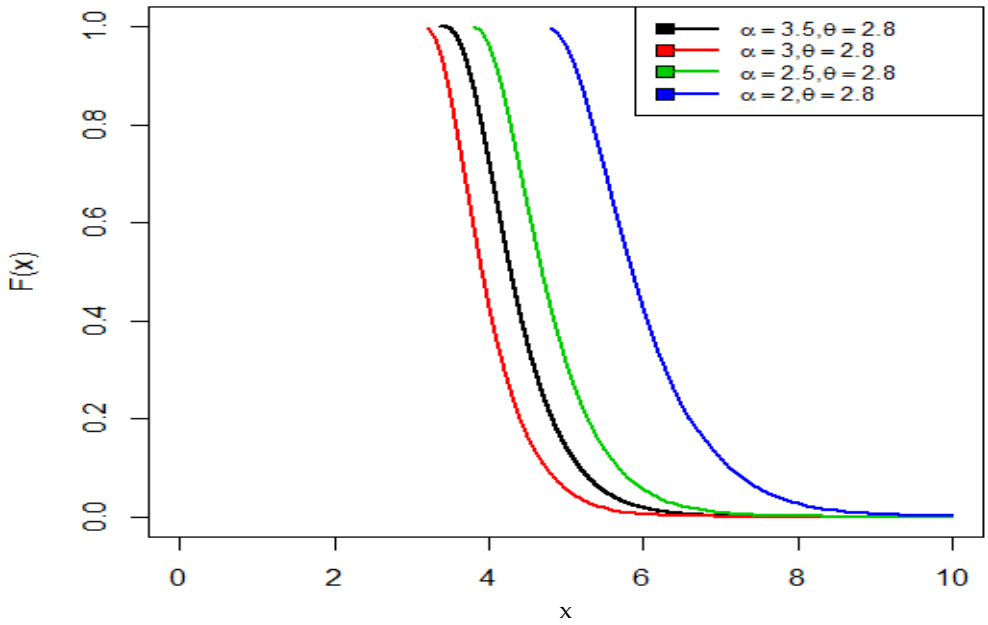


Fig. 3: *Survival function of exponentiated Suja distribution*

4. Statistical Properties

4.1 Moments

Suppose X is a random variable following exponentiated Suja distribution with parameters α and θ, then the rth order moment $E(X^r)$ for a given probability distribution is given by

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_\alpha(x)dx$$

$$E(X^r) = \int_0^\infty x^r \frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}dx \tag{9}$$

$$E(X^r) = \frac{\alpha\theta^5}{\theta^4+24}\int_0^\infty x^r\left(1+x^4\right)e^{-\theta x}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}dx \tag{10}$$

Using Binomial expansion of

$$\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1} = \sum_{i=0}^{\infty}\binom{\alpha-1}{i}\left\{\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right\}^i(-1)^i \tag{11}$$

Equation (10) will become

$$E(X^r) = \frac{\alpha\theta^5}{\theta^4+24}\sum_{i=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\int_0^\infty x^r\left(1+x^4\right)e^{-\theta x(1+i)}\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^i dx \tag{12}$$

Again using Binomial expansion of

$$\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^i=\sum_{k=0}^{\infty}\binom{i}{k}\left(\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^k \tag{13}$$

Equation (12) becomes

$$E(X^r)=\frac{\alpha\theta^5}{\theta^4+24}\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\binom{i}{k}\left(\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^k\int_o^{\infty}x^r\left(1+x^4\right)e^{-\theta x(1+i)}dx \tag{14}$$

After simplification, we obtain

$$E(X^r)=\alpha\theta^5\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\binom{i}{k}\frac{\left(\theta^4+4\theta^3+12\theta^2+24\theta\right)^k}{(\theta^4+24)^{k+1}}\left(\frac{\theta(1+i)^4\Gamma(r+10k+1)+\Gamma(r+10k+5)}{\theta(1+i)^{r+10k+5}}\right) \tag{15}$$

Since equation (15) is a convergent series for all $r \geq 0$, therefore all the moments exist.
Therefore

$$E(X)=\alpha\theta^5\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\binom{i}{k}\frac{\left(\theta^4+4\theta^3+12\theta^2+24\theta\right)^k}{(\theta^4+24)^{k+1}}\left(\frac{\theta(1+i)^4\Gamma(10k+2)+\Gamma(10k+6)}{\theta(1+i)^{10k+6}}\right) \tag{16}$$

$$E(X^2)=\alpha\theta^5\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\binom{i}{k}\frac{\left(\theta^4+4\theta^3+12\theta^2+24\theta\right)^k}{(\theta^4+24)^{k+1}}\left(\frac{\theta(1+i)^4\Gamma(10k+3)+\Gamma(10k+7)}{\theta(1+i)^{10k+7}}\right) \tag{17}$$

Therefore, the Variance of X can be obtained as

$$V(X)=E(X^2)-(E(X))^2$$

4.2 Harmonic mean

The Harmonic mean for the proposed Exponentiated Suja distribution can be obtained as

$$H.M=E\left(\frac{1}{x}\right)=\int_0^{\infty}\frac{1}{x}f_{\alpha}(x)dx$$

$$=\int_0^{\infty}\frac{1}{x}\frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}dx \tag{18}$$

$$=\frac{\alpha\theta^5}{\theta^4+24}\int_0^{\infty}\frac{1}{x}\left(1+x^4\right)e^{-\theta x}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}dx \tag{19}$$

Using Binomial expansion in equation (19), we get

$$=\frac{\alpha\theta^5}{\theta^4+24}\sum_{i=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\int_0^{\infty}\frac{1}{x}\left(1+x^4\right)e^{-\theta x(1+i)}\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^i dx \tag{20}$$

On using Binomial expansion in equation (20), we obtain

$$H.M=\frac{\alpha\theta^5}{\theta^4+24}\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\binom{i}{k}\left(\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^k\int_0^{\infty}\frac{1}{x}\left(1+x^4\right)e^{-\theta x(1+i)}dx \tag{21}$$

After the simplification of equation (21), we obtain

$$H.M=\alpha\theta^5\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\alpha-1}{i}\binom{i}{k}\frac{\left(\theta^4+4\theta^3+12\theta^2+24\theta\right)^k}{\left(\theta^4+24\right)^{k+1}}\left(\frac{\theta(1+i)^3\Gamma(10k+1)+\Gamma(10k+4)}{\theta(1+i)^{10k+4}}\right) \tag{22}$$

4.3 Moment Generating Function and Characteristics Function

Let X have an ESD, then the moment generating function of X is obtained as

$$M_X(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx} f_\alpha(x)dx$$

Using Taylor's series, we get

$$M_X(t) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + ...\right) f_\alpha(x)dx$$

$$M_X(t) = \alpha\theta^5 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i \binom{\alpha-1}{i}\binom{i}{k}\frac{t^j}{j!}\frac{\left(\theta^4+4\theta^3+12\theta^2+24\theta\right)^k}{(\theta^4+24)^{k+1}}\left(\frac{\theta(1+i)^4\Gamma(j+10k+1)+\Gamma(j+10k+5)}{\theta(1+i)^{j+10k+5}}\right) \quad (23)$$

Similarly, the Characteristics function of Exponentiated Suja distribution is given by

$$\varphi_X(t) = \alpha\theta^5 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i \binom{\alpha-1}{i}\binom{i}{k}\frac{mt^j}{j!}\frac{\left(\theta^4+4\theta^3+12\theta^2+24\theta\right)^k}{(\theta^4+24)^{k+1}}\left(\frac{\theta(1+i)^4\Gamma(j+10k+1)+\Gamma(j+10k+5)}{\theta(1+i)^{j+10k+5}}\right) \quad (24)$$

5. Order Statistics

Order statistics has wide field in reliability and life testing. There is also an extensive role of order statistics in several aspects of statistical inference. Let $X_{(1)}$, $X_{(2)}$, ….., $X_{(n)}$ be the order statistics of a random sample X_1, X_2, ….X_n drawn from the continuous population with probability density function *fx(x)* and cumulative distribution function Fx(x), then the pdf of rth order statistics $X_{(r)}$ can be written as

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x)\left(F_X(x)\right)^{r-1}\left(1-F_X(x)\right)^{n-r} \quad (25)$$

Substitute the values of equation (5) and (6) in equation (25), we will obtain the pdf of rth order statistics X(r) for exponentiated Suja distribution and is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\frac{\alpha\theta^5(1+x^4)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}$$
$$\times\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha(r-1)} \quad (26)$$
$$\times\left(1-\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha}\right)^{n-r}$$

The probability density function of higher order statistics $X_{(n)}$ can be obtained as

$$f_{x(n)}(x) = n\frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}$$
$$\times\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha(n-1)} \quad (27)$$

Similarly, the pdf of first order statistics $X_{(1)}$ can be obtained as

$$f_{x(1)}(x) = n\frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha}\right)^{n-1} \quad (28)$$

6. Maximum Likelihood Estimation

In this section, we will discuss the maximum likelihood estimators of the parameters of exponentiated Suja distribution. Let X1, X2,….Xn be the random sample of size n from the Exponentiated Suja distribution, then the likelihood function can be written as

$$L(\alpha,\theta) = \frac{\left(\alpha\theta^5\right)^n}{(\theta^4+24)^n}\prod_{i=1}^{n}\left((1+x^4)e^{-\theta x}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}\right) \quad (29)$$

The log likelihood function is given by

$$\log L(\alpha,\theta) = n\log\alpha + 5n\log\theta - n\log(\theta^4+24) + \sum_{i=1}^{n}\log\left(1+x_i^4\right) - \theta\sum_{i=1}^{n}x_i$$
$$+(\alpha-1)\sum_{i=1}^{n}\log\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right) \quad (30)$$

The maximum likelihood estimates of α, θ which maximizes (30), must satisfy the normal equations given by

$$\frac{\partial\log L}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right) = 0 \quad (31)$$

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\log\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)} \quad (32)$$

$$\frac{\partial\log L}{\partial\theta} = \frac{5n}{\theta} - n\left(\frac{4\theta^3}{\theta^4+24}\right) - \sum_{i=1}^{n}x_i + (\alpha-1)\psi\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right) = 0 \quad (33)$$

Where ψ (.) is the digamma function.

At this point it is important to mention that the analytical solution of the above system of non-linear equation is unknown. Algebraically it is very difficult to solve the complicated form of likelihood system of nonlinear equations. Therefore, we use R and wolfram mathematics for estimating the required parameters.

7. Information Measures

7.1 Renyi Entropy

Entropies quantify the diversity, uncertainty, or randomness of a system. The Renyi entropy is named after Alfred Renyi in the context of fractual dimension estimation, the Renyi entropy forms the basis of the concept of generalized dimensions. The Renyi entropy is important in ecology and statistics as index of diversity. The Renyi entropy is also important in quantum information, where it can be used as a measure of entanglement. For a given probability distribution, Renyi entropy is given by

$$e(\beta)=\frac{1}{1-\beta}\log\left(\int_0^\infty f^{\beta}(x)dx\right)$$

Where, $\beta>0$ and $\beta\neq 1$

$$=\frac{1}{1-\beta}\log\left(\int_0^\infty\left\{\frac{\alpha\theta^5\left(1+x^4\right)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}\right\}^{\beta}\right)dx \qquad (34)$$

$$=\frac{1}{1-\beta}\log\left(\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\beta}\int_0^\infty\left(1+x^4\right)^{\beta}e^{-\theta\beta x}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\beta(\alpha-1)}\right)dx \qquad (35)$$

Using binomial expansion in (35), we get

$$=\frac{1}{1-\beta}\log\left(\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\beta}\sum_{i=0}^{\infty}(-1)^i\binom{\beta(\alpha-1)}{i}\int_0^\infty\left(1+x^4\right)^{\beta}e^{-\theta x(\beta+i)}\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^{i}dx\right) \qquad (36)$$

Again using binomial expansion in (16), we get

$$=\frac{1}{1-\beta}\log\left(\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\beta}\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\beta(\alpha-1)}{i}\binom{i}{k}\left(\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^{k}\int_0^\infty\left(1+x^4\right)^{\beta}e^{-\theta x(\beta+i)}dx\right) \qquad (37)$$

After the simplification of (37) we obtain

$$e(\beta)=\frac{1}{1-\beta}\log\left(\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\beta}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\beta(\alpha-1)}{i}\binom{i}{k}\binom{\beta}{j}\left(\frac{\theta^4+4\theta^3+12\theta^2+24\theta}{\theta^4+24}\right)^{k}\frac{\Gamma(4j+10k+1)}{\theta(\beta+i)^{4j+10k+1}}\right) \qquad (38)$$

7.2 Tsallis Entropy

A generalization of Boltzmann-Gibbs (B-G) statistical mechanics initiated by Tsallis has gained a great deal to attention. This generalization of B-G statistics was proposed firstly by introducing the mathematical expression of Tsallis entropy for a continuous random variable it is defined as

$$S_\lambda=\frac{1}{\lambda-1}\left(1-\int_0^\infty f^{\lambda}(x)dx\right)$$

$$= \frac{1}{\lambda-1}\left(1-\int_0^{\infty}\left\{\frac{\alpha\theta^5(1+x^4)e^{-\theta x}}{\theta^4+24}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\alpha-1}\right\}^{\lambda}\right)dx \quad (39)$$

$$= \frac{1}{\lambda-1}\left(1-\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\lambda}\int_0^{\infty}(1+x^4)^{\lambda}e^{-\lambda\theta x}\left(1-\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)e^{-\theta x}\right)^{\lambda(\alpha-1)}dx\right) \quad (40)$$

Using binomial expansion in (40), we get

$$= \frac{1}{\lambda-1}\left(1-\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\lambda}\sum_{i=0}^{\infty}(-1)^i\binom{\lambda(\alpha-1)}{i}\int_0^{\infty}(1+x^4)^{\lambda}e^{-\theta x(\lambda+i)}\left(1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^{i}dx\right) \quad (41)$$

Again using binomial expansion in (41), we obtain

$$= \frac{1}{\lambda-1}\left(1-\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\lambda}\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\lambda(\alpha-1)}{i}\binom{i}{k}\left(\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right)^{k}\int_0^{\infty}(1+x^4)^{\lambda}e^{-\theta x(\lambda+i)}dx\right) \quad (42)$$

After the simplification of (42), we get

$$S_{\lambda} = \frac{1}{\lambda-1}\left(1-\left(\frac{\alpha\theta^5}{\theta^4+24}\right)^{\lambda}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^i\binom{\lambda(\alpha-1)}{i}\binom{i}{k}\binom{\lambda}{j}\left(\frac{\theta^4+4\theta^3+12\theta^2+24\theta}{\theta^4+24}\right)^{k}\frac{\Gamma(4j+10k+1)}{\theta(\lambda+i)^{4j+10k+1}}\right) \quad (43)$$

8. Data Analysis

In this section, we use two real-life data sets in exponentiated Suja distribution and the model has been compared with Suja and exponential distributions

Data Set 1: The following data set of 40 patients suffering from blood cancer (leukemia) is reported by one of ministry of health hospitals in Saudi Arabia by Abouammah et al. [13]. The ordered lifetimes (in years) is provided below in table 1.

Table 1: *Data represents the blood cancer patients (leukemia)*

0.315	0.496	0.616	1.145	1.208	1.263	1.414	2.025	2.036	2.162
2.211	2.37	2.532	2.693	2.805	2.91	2.912	2.192	3.263	3.348
3.348	3.427	3.499	3.534	3.767	3.751	3.858	3.986	4.049	4.244
4.323	4.381	4.392	4.397	4.647	4.753	4.929	4.973	5.074	5.381

Data set 2: The second data set represents the tensile strength, measured in GPa, of 69 carbon fibers tested under tension at gauge lengths of 20mm which were originally reported by M. G. Bader and A. M. Priest [12]. The data set is provided below in table 2.

Table 2: *Data regarding the tensile strength (GPa) of 69 carbon fibers*

1.312	1.314	1.479	1.552	1.700	1.803	1.861	1.865	1.944	1.958
1.966	1.997	2.006	2.021	2.027	2.055	2.063	2.098	2.14	2.179
2.224	2.240	2.253	2.270	2.272	2.274	2.301	2.301	2.359	2.382
2.382	2.426	2.434	2.435	2.478	2.490	2.511	2.514	2.535	2.554

2.566	2.57	2.586	2.629	2.633	2.642	2.648	2.684	2.697	2.726
2.770	2.773	2.800	2.809	2.818	2.821	2.848	2.88	2.954	3.012
3.067	3.084	3.090	3.096	3.128	3.233	3.433	3.585	3.585	

In order to compare the exponentiated Suja distribution with Suja and exponential distribution, we consider the Criteria like BIC (Bayesian information criterion), AIC (Akaike information criterion), AICC (Corrected Akaike information criterion) and -2logL. The better distribution is which corresponds to lesser values of AIC, BIC, AICC and -2logL. For calculating AIC, BIC, AICC and -2logL can be evaluated by using the formulas as follows.

$$AIC = 2k - 2\log L \qquad AICC = AIC + \frac{2k(k+1)}{n-k-1} \quad \text{and} \quad BIC = k\log n - 2\log L$$

Where k is the number of parameters in the statistical model, n is the sample size and -2log$L$ is the maximized value of the log-likelihood function under the considered model.

Table 3: *Fitted distributions of the two data sets and criteria for comparison*

Data sets	Distribution	MLE	S.E	-2logL	AIC	BIC	AICC
1	Exponentiated Suja	$\hat{\alpha} = 3.9557520$ $\hat{\theta} = 1.9201488$	$\hat{\alpha} = 1.2186387$ $\hat{\theta} = 0.1531425$	120.792	124.792	128.169	125.117
	Suja	$\hat{\theta} = 1.41132109$	$\hat{\theta} = 0.08834353$	142.669	144.669	146.358	144.775
	Exponential	$\hat{\theta} = 0.31839887$	$\hat{\theta} = 0.05034278$	171.557	173.557	175.246	173.67
2	Exponentiated Suja	$\hat{\alpha} = 91.3523215$ $\hat{\theta} = 3.6448593$	$\hat{\alpha} = 45.077905$ $\hat{\theta} = 0.2307272$	77.5453	81.5453	85.6314	81.728
	Suja	$\hat{\theta} = 1.6558639$	$\hat{\theta} = 0.0831825$	164.252	166.252	168.295	166.312
	Exponential	$\hat{\theta} = 0.40927188$	$\hat{\theta} = 0.0542091$	215.846	217.846	219.889	217.905

From table 3, it can be easily seen that the exponentiated Suja distribution have the lesser AIC, BIC, AICC and -2logL values as compared to Suja and exponential distributions. Hence we can conclude that the exponentiated Suja distribution leads to a better fit than the Suja and exponential distributions.

9. Conclusion

In conclusion, this study has introduced a new generalization of the Suja distribution, known as the exponentiated Suja distribution, which incorporates two parameters: scale and shape. The distribution was generated using the exponentiated technique, and the parameters were estimated using the maximum likelihood estimator. Various statistical properties and reliability measures of the exponentiated Suja distribution were discussed.

Furthermore, the study demonstrated the practical applications of the new distribution in real-life time data. The results of two real lifetime data sets were compared with the Suja and exponential distributions, revealing that the exponentiated Suja distribution provides a better fit than both alternatives.

These findings highlight the potential of the exponentiated Suja distribution to enhance the modeling and analysis of data in various fields, such as clinical trials, epidemiological studies, and public health research. The improved fit observed in real-life applications suggests that the exponentiated Suja distribution can offer more accurate predictions and better capture the underlying characteristics of the data.

Overall, this study contributes to the field of biostatistics by introducing a novel distribution and demonstrating its advantages over existing models. Further research and applications of the exponentiated Suja distribution are encouraged to explore its full potential in various domains of statistical analysis and decision-making.

References

[1] Gupta, R. C., Gupta, P. L. and Gupta, R. D (1998). Modeling failure time data by Lehman alternatives, *Communications in Statistics, Theory and Methods*, 27, 887–904.

[2] Gupta, R.D. and Kundu (2001). Exponentiated exponential family: an alternative to gamma and Weibull, *Biometrical Journal*., 43(1), 117- 130.

[3] Rodrigues, J., Percontini A. and Hamedani, G., (2017). The Exponentiated Generalized Lindley Distribution, *Asian Research Journal of Mathematics*, 5(3), 1-14.

[4] Hassan, A. S. and Abdelghafar, M. A., (2017). Exponentiated Lomax Geometric Distribution: Properties and Applications, *Pakistan Journal of Statistics and Operation Research*, 13(3).

[5] Nasiru, S., Mwita, P. N. and Ngesa, O., (2018). Exponentiated Generalized Half Logistic Burr X Distribution, *Advances and Applications in Statistics*, 52(3), 145-169.

[6] Rather, A. A. and Subramanian, C., (2018). Exponentiated Mukherjee-Islam Distribution, *Journal of Statistics Applications & Probability*, 7(2), 357-361.

[7] Rather, A. A. and Subramanian, C., (2019). Exponentiated Ishita Distribution with Properties and Application, *International Journal of Management, Technology and Engineering*, 9(5), 2473-2484.

[8] Rather, A. A. and Subramanian, C., (2020). A New Exponentiated Distribution with Engineering Science Applications, Journal of Statistics, Applications and Probability, 9(1), 127-137.

[9] Shanker, R. (2017). Suja distribution and its Application, *International Journal of probability and Statistics*, 6(2), 11-19.

[10] Maradesa Adeleke (2019). Exponentiated Exponential Lomax Distribution and its Properties, *Mathematical Theory and Modeling*. doi:10.7176/mtm/9-1-01.

[11] Nasir, A., Yousof, H., Jamal, F. and Korkmaz, M., (2018). The Exponentiated Burr XII Power Series Distribution: Properties and Applications. *Stats*, 2(1), 15-31.

[12] Bader, M. G. and Priest, A. M., (1982). Statistical aspects of fiber and bundle strength in hybrid composites, In; hayashi T, Kawata K. Umekawa S (Eds.), *Progress in Science in Engineering Composites,* ICCM-IV Tokyo., 1129 -1136.

[13] Abouammoh, A. M., Ahmed, R. and Khalique, A., (2000). On new renewal better than used classes of life distribution. *Statistics and Probability Letters*, 48, 189-194.

[14] Tillman, F.A., Hwag, C.L. and Kuo, W. (1977). Determining component reliability and redundancy for optimum system reliability, *IEEE Trans*, R- 26 (3), 162- 165.

[15] Wei, V.K., Hwang, F.K. and Sos, V.T. (1983). Optimal sequencing of items in a consecutive 2-out-of-n system. *IEEE Transaction on Reliability*, R-32(1), pp. 30–34.

[16] Wang, Z.H., (1992). Reliability engineering theory and practice (5th ed.). Taipei: *Quality Control Society of Republic of China.*

LEVERAGING AUXILIARY VARIABLES: ADVANCING MEAN ESTIMATION THROUGH CONDITIONAL AND UNCONDITIONAL POST-STRATIFICATION

G.R.V. Triveni[1] and Faizan Danish[2]

•

[1,2]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati AP-522237, India

[1]trivenigullinkala@gmail.com ,[2]danishstat@gmail.com

Abstract

This article presents a novel class of estimators designed for post-stratification to estimate the mean of a study variable using information from auxiliary variables. Through a rigorous examination of bias and Mean Square Error (MSE), we demonstrate the potential to improve estimation accuracy up to the first order of approximation. We also thoroughly explore both Conditional and Unconditional post-stratification properties, enhancing our understanding of the estimator's performance. To assess the effectiveness of our proposed estimator, we conduct a comprehensive numerical illustration. The results affirm its superiority over existing estimators in both Conditional and Unconditional Post-stratification scenarios, exhibiting the highest Percentage Relative Efficiency. Additionally, graphical analysis reveals that Conditional post- stratification outperforms Unconditional post-stratification. These findings underscore the significant practical value of our proposed estimator in enhancing the accuracy of mean estimation in post-stratification studies. By accurately estimating population parameters, our novel class of estimators contributes to more informed decision-making in various fields of study. The utilization of auxiliary variables allows for better utilization of available information and leads to more reliable and robust conclusions. Overall, the novel class of estimators introduced in this article represents a valuable contribution to the field of post-stratification. As researchers continue to explore and apply these estimators, they have the potential to revolutionize data analysis methods, becoming indispensable tools for survey and research design. The improvements in estimation accuracy brought about by these estimators are particularly crucial in situations where reliable data is scarce or challenging to obtain, making them invaluable for decision-makers and researchers alike. With the increased accuracy and efficiency of our proposed estimators, they provide a pathway for better resource allocation, cost-effective decision-making, and improved policy formulation. Policymakers and researchers can confidently rely on these estimators to produce more accurate results and achieve better outcomes in various domains. In conclusion, the novel class of estimators for post-stratification presented in this article opens up new avenues for advancing statistical estimation methods. The fusion of auxiliary variables with traditional post-stratification techniques represents a powerful approach to enhance estimation accuracy. Embracing and incorporating these estimators into research practices will undoubtedly bring us closer to making data-driven decisions that have a meaningful impact on society.

Keywords: Conditional post stratification, Unconditional post stratification, Mean square error.

I. Introduction

A common statistical method used in research studies to increase the precision and representativeness of survey data is post-stratification. It entails breaking down the sample population into discrete subgroups according to certain traits or factors, such age, gender, income level, or geography. Researchers can reduce potential biases and improve the generalizability of the results by stratifying the population to make sure that each subgroup is appropriately represented in the sample. After stratifying the sample, researchers can determine the population parameters by giving each subgroup the proper weights depending on its relative size. This method enables researchers to account for differences and generate more accurate and trustworthy estimates when the sample does not properly reflect the makeup of the target population. By

taking into account population changes and enhancing the precision of inferential statistics, post-stratification is a useful tool that improves the validity and robustness of study findings.

Auxiliary variable information is used in various types of literature to estimate population mean or variance. The significance of post-stratification and the proper framework for statistical inference were covered in [1]. By utilizing auxiliary data and empirical research, [2] proposed estimators for population mean and shown that the proposed estimator outperformed the alternative. For judgement post-stratification, [3] offered an alternative estimate that regularly beats the usual non-parametric mean estimator, and they noted a decrease in Mean Square Error (MSE) in their proposed estimator compared to the standard estimator. [4] proposed a class of estimators and compared them with a few already in use. They came to the conclusion that their suggested class of estimators performed well based on a numerical investigation. In post-stratified sampling, [5] created a new family of combined estimators of the population mean, and the outcomes are empirically demonstrated. Exponential estimators are later proposed in post-stratification by [6], and its bias and MSE equations are obtained. The theoretical findings are further supported by a numerical analysis. [7] suggested an estimator of the population mean utilizing information from an auxiliary variable and demonstrated the superiority of the proposed estimator over others through comparison analysis. Additionally, [8] constructed a generalized class of estimators for population variance and demonstrated the effectiveness of the suggested estimator through a numerical investigation. Through empirical research, [9] demonstrated the superiority of the suggested estimator and developed a new family of exponential estimators. The ratio and product type exponential estimators were improved in the case of post-stratification by [10]. They demonstrated that the suggested estimator performed more effectively after stratification than unbiased, ratio, and product estimators. Additionally, [11] raises the issue of estimating a population proportion in a decision following stratification. They conducted Monte Carlo simulation research to evaluate the performance of proportion estimators. According to [12], a family of Ratio estimators and the formulations for bias and an MSE are constructed in the case of the non-response issue. It is demonstrated by numerical analysis that the suggested estimator has reduced MSE values. A novel class of estimators was recently developed [13], and by numerical analysis, under ideal circumstances, the proposed class of estimators outperformed the previously taken into consideration existing estimators.[14] suggested some post-stratification enhanced estimators. They demonstrated the effectiveness of the proposed estimator using two real data sets. In this paper, a class of estimators for estimating population mean under post-stratification has been developed.

II. Terminology

Consider a finite population $\chi = \{1,2, \dots N\}$ of size N is stratified into K strata with h^{th} stratum each of which has $\mathrm{N_h}$units such that $\sum_{h=1}^{K} \mathrm{N_h} = N$. With the use of Simple random sampling without replacement, a sample with the dimension $\mathrm{n_h}$ is taken from h^{th} stratum. Let d serve as study(dependent) variable and i serve as auxiliary (independent) variable. We have used the following notations:

- $\overline{\mathrm{D}} = \sum_{\mathrm{h=1}}^{\mathrm{K}} \mathrm{W_h}\overline{\mathrm{D}}_\mathrm{h}$ is the population mean of study variable
- $\overline{\mathrm{I}} = \sum_{\mathrm{h=1}}^{\mathrm{K}} \mathrm{W_h}\overline{\mathrm{I}}_\mathrm{h}$ is the population mean of auxiliary variable
- $\overline{\mathrm{D}}_\mathrm{h} = \frac{1}{\mathrm{N_h}}\sum_{\mathrm{i=1}}^{\mathrm{N_h}} \mathrm{d_{hi}}$ and $\overline{\mathrm{I}}_\mathrm{h} = \frac{1}{\mathrm{N_h}}\sum_{\mathrm{i=1}}^{\mathrm{N_h}} \mathrm{i_{hi}}$ are the stratum means of study and auxiliary variables
- $\mathrm{S_{dh}^2} = \frac{1}{(\mathrm{N_h}-1)}\sum_{\mathrm{i=1}}^{\mathrm{N_h}}(\mathrm{d_{hi}} - \overline{\mathrm{D}}_\mathrm{h})^2$ is the variance of study variable at h^{th} stratum
- $\mathrm{S_{ih}^2} = \frac{1}{(\mathrm{N_h}-1)}\sum_{\mathrm{i=1}}^{\mathrm{N_h}}(\mathrm{i_{hi}} - \overline{\mathrm{I}}_\mathrm{h})^2$ is the variance of auxiliary variable at h^{th} stratum
- $\mathrm{S_{dih}} = \frac{1}{(\mathrm{N_h}-1)}\sum_{\mathrm{i=1}}^{\mathrm{N_h}}(\mathrm{d_{hi}} - \overline{\mathrm{D}}_\mathrm{h})((\mathrm{i_{hi}} - \overline{\mathrm{I}}_\mathrm{h})$ is the covariance at h^{th} stratum
- $\mathrm{C_{dh}^2} = \frac{1}{\overline{\mathrm{D}}^2(\mathrm{N_h}-1)}\sum_{\mathrm{i=1}}^{\mathrm{N_h}}(\mathrm{d_{hi}} - \overline{\mathrm{D}}_\mathrm{h})^2$ be the square of coefficient of variation of d

- $C_{ih}^2 = \frac{1}{\bar{I}^2(N_h-1)}\sum_{i=1}^{N_h}(i_{hi}-\bar{I}_h)^2$ be the square of coefficient of variation of i
- $\rho_{dih} = \frac{\frac{1}{N_h-1}\sum_{i=1}^{N_h}(d_{hi}-\bar{D}_h)(i_{hi}-\bar{I}_h)}{(\bar{D}_h * C_{dh})(\bar{I}_h * I_{dh})}$ be the correlation coefficient of d and i.
- $W_h = \frac{N_h}{N}$ represents stratum weight.

I. Properties of Estimators in Unconditional Post-Stratification

To derive bias and mean squared error (MSE), we write
$\bar{d}_h = \bar{D}_h(1+e_{0h})$, $\bar{i}_h = \bar{I}_h(1+e_{1h})$,
$e_0 = \frac{\sum_{h=1}^{K} W_h \bar{D}_h e_{0h}}{\bar{D}}$ and $e_1 = \frac{\sum_{h=1}^{K} W_h \bar{I}_h e_{1h}}{\bar{I}}$

Where

$e_{0h} = \frac{\bar{d}_h-\bar{D}_h}{\bar{D}_h}$ and $e_{1h} = \frac{\bar{i}_h-\bar{I}_h}{\bar{I}_h}$

$E(e_{0h}) = E(e_{1h}) = 0$

$E(e_{0h}^2) = \left[\frac{1}{nW_h}-\frac{1}{N_h}\right]C_{dh}^2$

$E(e_{1h}^2) = \left[\frac{1}{nW_h}-\frac{1}{N_h}\right]C_{ih}^2$

$E(e_{0h}e_{1h}) = \left[\frac{1}{nW_h}-\frac{1}{N_h}\right]\rho_{dih}C_{ih}C_{dh}$.

we will find the expected values of error terms as

$$E(e_0) = E\left(\frac{\sum_{h=1}^{K} W_h F_h(y) e_{0h}}{F(y)}\right) = \frac{1}{F(y)}\left(\sum_{h=1}^{K} W_h F_h(y) E(e_{0h})\right) = 0$$

Similarly,

$E(e_0) = E(e_1) = 0$

$$E(e_0^2) = E\left(\frac{\sum_{h=1}^{K} W_h \bar{D}_h^2 e_{0h}}{\bar{D}}\right)^2 = \frac{1}{\bar{D}^2}\sum_{h=1}^{K} W_h^2 \bar{D}_h^2 E(e^2{}_{0h})$$

$$= \frac{1}{\bar{D}^2}\sum_{h=1}^{K} W_h^2 \bar{D}_h^2 \left[\frac{1}{nW_h}-\frac{1}{N_h}\right]C_{dh}^2$$

$$= \frac{1}{\bar{D}^2}\sum_{h=1}^{K} W_h^2 \bar{D}_h^2 \left[\frac{1}{nW_h}-\frac{1}{N_h}\right]C_{dh}^2$$

$$= \frac{1}{\bar{D}^2}\left[\frac{1}{n}-\frac{1}{N}\right]\sum W_h S_{dh}^2 = V_D(say)$$

Similarly,

$$E(e_1^2) = \frac{1}{\bar{I}^2}\left[\frac{1}{n}-\frac{1}{N}\right]\sum_{h=1}^{K}\sum W_h S_{ih}^2 = V_I$$

$$E(e_0e_1) = \frac{1}{\bar{D}\bar{I}}\left[\frac{1}{n}-\frac{1}{N}\right]\sum_{h=1}^{K} W_h S_{dih} = V_{DI} \quad (1)$$

II. Properties of Estimators in Conditional Post-Stratification

$$E_1(e_0^2) = \frac{1}{\bar{D}^2}\sum_{h=1}^{K} W_h^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)S_{dh}^2 = V_{1D}(say)$$

$$E_1(e_1^2) = \frac{1}{\bar{I}^2}\sum_{h=1}^{K} W_h^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)S_{ih}^2 = V_{1I}$$

$$E_1(e_0e_1) = \frac{1}{\bar{D}\bar{I}}\sum_{h=1}^{K} W_h\left(\frac{1}{n_h} - \frac{1}{N_h}\right) S_{dih} = V_{1DI} \quad (2)$$

III. Estimators in Literature

We write the following estimators in terms of Unconditional case in post-stratification as

a. The usual unbiased estimator of population mean $\bar{D} = \sum_{h=1}^{K} W_h \bar{D}_h$ is given by

$$u_1 = \bar{d}_{ps} = \sum_{h=1}^{K} W_h \bar{d}_h \quad (3)$$

Using the results from Stephen (1945), the variances of $\bar{d}_{ps}$ to the first degree of approximation is given by

For Unconditional post-stratification,

$$Var(u_{1a}) = \left[\frac{1}{n} - \frac{1}{N}\right]\sum_{h=1}^{K}\sum W_h S_{dh}^2 + \frac{1}{n^2}\sum_{h=1}^{K}(1 - W_h) S_{dh}^2 \quad (4)$$

Where $Var(u_{1a})$ is the Unconditional variance of post stratified estimator $\bar{d}_{ps}$.

and $S_{dh}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(d_{hi} - \bar{D}_h)^2$

For Conditional post- stratification,

$$Var(u_{1b}) = \sum_{h=1}^{K} W_h^2\left[\frac{1}{n_h} - \frac{1}{N_h}\right] S_{dh}^2 \quad (5)$$

$Var(u_{1b})$ is the Conditional variance of post stratified estimator $\bar{d}_{ps}$.

b. The Ratio estimator for population mean according to Naik and Gupta [2] is given by

$$u_2 = \bar{d}_{ps}\left(\frac{\bar{I}}{\bar{i}_{ps}}\right)$$

Where $\bar{i}_{ps} = \sum_{h=1}^{K} W_h \bar{i}_h$

Up to the first degree of approximation, the MSE of estimator u_2 is given by

$$MSE(u_2) = \bar{D}^2\left[V_D + V_I\left[1 - 2\left(\frac{V_{DI}}{V_I}\right)\right]\right] \quad (6)$$

c. The usual product estimator is given by

$$u_3 = \bar{d}_{ps}\left(\frac{\bar{i}_{ps}}{\bar{I}}\right)$$

Up to the first degree of approximation, the MSE of estimator u_3 is given by

$$MSE(u_3) = \bar{D}^2\left[V_D + V_I\left[1 + 2\left(\frac{V_{DI}}{V_I}\right)\right]\right] \quad (7)$$

d. The Usual regression estimator for $\bar{D}$ is given by

$$u_4 = \bar{d}_{ps} + b_{ps}(\bar{I} - \bar{i}_{ps})$$

The MSE of estimator u_4 is given by

$$MSE(u_4) = \bar{D}^2 V_D(1 - \xi_{DI}^2) \quad (8)$$

Where $\xi_{DI}^2 = \frac{V_{DI}^2}{V_D V_I}$

e. Koyuncu [7] proposed a class of estimators as

$$u_5 = \left[r_1\bar{d}_{ps} + r_2(\bar{I} - \bar{i}_{ps})\right]\left(\frac{a_{ps}\bar{I}+b_{ps}}{a_{ps}\bar{i}_{ps}+b_{ps}}\right)$$

Its MSE is given by

$$\text{MSE}(u_5) = \bar{D}^2[1 + r_1^2A_1 + r_2^2A_2 + 2r_1r_2A_3 - 2r_1A_4 - 2r_2A_5] \quad (9)$$

Where $A_1 = 1 + V_D + \varphi_{ps}V_I\left[3\varphi_{ps} - 4\left(\frac{V_{DI}}{V_I}\right)\right]$

$$A_2 = \frac{V_I}{R^2}$$

$$A_3 = \frac{V_I}{R}\left(2\varphi_{ps} - \frac{V_{DI}}{V_I}\right)$$

$$A_4 = 1 + \varphi_{ps}V_I\left(\varphi_{ps} - \frac{V_{DI}}{V_I}\right)$$

$A_5 = \left(\frac{V_I}{R}\right)\varphi_{ps}$, $R=\frac{\bar{D}}{\bar{I}}$, $\varphi_{ps} = \frac{a_{ps}\bar{I}}{a_{ps}\bar{I}+b_{ps}}$,

$r_1 = \frac{A_2A_4-A_3A_5}{A_1A_2-A_3^2}$ and $r_2 = \frac{A_1A_5-A_3A_4}{A_1A_2-A_3^2}$.

f. Sharma and Singh [9] proposed exponential type estimators as

Ratio type exponential estimator

$$u_{6a} = \bar{d}_{ps}\exp\left(\frac{\bar{I} - \bar{i}_{ps}}{\bar{I} + \bar{i}_{ps}}\right)$$

Product type exponential estimator as

$$u_{6b} = \bar{d}_{ps}\exp\left(\frac{\bar{i}_{ps} - \bar{I}}{\bar{i}_{ps} + \bar{I}}\right)$$

$$u_{6c} = \bar{d}_{ps}\exp\left(\frac{\alpha(\bar{I} - \bar{i}_{ps})}{\bar{I} + \bar{i}_{ps}}\right)$$

Where α being a suitable constant

The MSEs of the above estimators as

$$\text{MSE}(u_{6a}) = \bar{D}^2\left\{V_D + \frac{V_I}{4}\left[1 - 4\left(\frac{V_{DI}}{V_I}\right)\right]\right\}$$

$$\text{MSE}(u_{6b}) = \bar{D}^2\left\{V_D + \frac{V_I}{4}\left[1 + 4\left(\frac{V_{DI}}{V_I}\right)\right]\right\}$$

$$\text{MSE}(u_{6c}) = \bar{D}^2\left\{V_D + \frac{\alpha V_I}{4}\left[\alpha - 4\left(\frac{V_{DI}}{V_I}\right)\right]\right\} \quad (10)$$

Where $\alpha = 2 * \left(\frac{V_{DI}}{V_I}\right)$

g. Sharma and Singh [9] suggested a class of estimators as

$$u_7 = \left[r_1\bar{d}_{ps} + r_2(\bar{I} - \bar{i}_{ps})\right]\exp\left(\frac{a_{ps}(\bar{I} - \bar{i}_{ps})}{a_{ps}(\bar{I} + \bar{i}_{ps)} + 2b_{ps}}\right)$$

Where a_{ps}, b_{ps} are either real numbers or the functions of the auxiliary variable.

Its MSE is given by

$$\text{MSE}(u_7) = \overline{D}^2\left[1 + r_1^2 B_1 + r_2^2 B_2 + 2r_1 r_2 B_3 - 2r_1 B_4 - 2r_2 B_5\right] \quad (11)$$

Where $B_1 = 1 + V_D + \varphi_{ps} V_I \left[\varphi_{ps} - 2\left(\frac{V_{DI}}{V_I}\right)\right]$

$B_2 = \frac{V_I}{R^2}$

$B_3 = \frac{V_I}{R}\left(\varphi_{ps} - \frac{V_{DI}}{V_I}\right)$

$B_4 = 1 + \varphi_{ps}\frac{V_I}{8}\left(3\varphi_{ps} - 4\frac{V_{DI}}{V_I}\right)$

$B_5 = \left(\frac{V_I}{2R}\right)\varphi_{ps}$

$r_1 = \frac{B_2 B_4 - B_3 B_5}{B_1 B_2 - B_3^2}$ and $r_2 = \frac{B_1 B_5 - B_3 B_4}{B_1 B_2 - B_3^2}$.

h. Singh et al. [13] suggested another class of estimators for population mean as

$$u_8 = \left[r_1 \bar{d}_{ps} + r_2 \exp\left(\frac{\delta a_{ps}(\bar{I} - \bar{i}_{ps})}{a_{ps}(\bar{I} + \bar{i}_{ps}) + 2b_{ps}}\right)\right]\left(\frac{a_{ps}\bar{I} + b_{ps}}{a_{ps}\bar{i}_{ps} + b_{ps}}\right)^{\eta}$$

Where $(\delta, \eta)$are constants belongs to real numbers like (-1,0,1).

$$\text{MSE}(u_8) = \overline{D}^2\left[1 + r_1^2 C_1 + r_2^2 C_2 + 2r_1 r_2 C_3 - 2r_1 C_4 - 2r_2 C_5\right] \quad (12)$$

Where, $C_1 = 1 + V_D - 4\eta\varphi_{ps} V_{DI} + \eta(2\eta + 1)\varphi_{ps}^2 V_I$

$C_2 = \frac{1}{\bar{I}^2 R^2}[\,1 + \theta(2\theta + 1)\varphi_{ps}^2 V_I]$

$C_3 = \frac{1}{\bar{I}R}\left[1 + \frac{(\eta + \theta)(\eta + \theta + 1)}{2}\varphi_{ps}^2 V_I - (\eta + \theta)\varphi_{ps} V_{DI}\right]$

$C_4 = 1 + \varphi_{ps}\frac{\eta}{2}\left(\frac{(\eta + 1)}{2}\varphi_{ps} V_I - 2V_{DI}\right)$

$C_5 = \frac{1}{\bar{I}R}\left[1 + \frac{\theta(\theta+1)}{2}\varphi_{ps}^2 V_I\right]$

$\theta = (2S_{dih} - 1)/2$, $r_1 = \frac{C_2 C_4 - C_3 C_5}{C_1 C_2 - C_3^2}$ and $r_2 = \frac{C_1 C_5 - C_3 C_4}{C_1 C_2 - C_3^2}$.

We have written the above considered pre-existing estimators in Unconditional case. If we change the expectations of error terms like in equation (2), we get the estimators in Conditional case.

IV. Suggested class of estimators in post stratification

We propose a class of estimators for population mean $\overline{D}$ as

$$u_{prop} = \left[r_1 \bar{d}_{ps} + r_2\left(\bar{I} - \bar{i}_{ps}\right) - r_3 \bar{d}_{ps}\left(\frac{a_{ps}\bar{I} + b_{ps}}{a_{ps}\bar{i}_{ps} + b_{ps}}\right)\right]\exp\left(\frac{\bar{I} - \bar{i}_{ps}}{\bar{I} + \bar{i}_{ps}}\right)^{\eta} \quad (13)$$

Where (r_1, r_2, η) are suitable constants and (a_{ps}, b_{ps}) are either constants or functions of auxiliary variable.

Expressing the equation (13) in terms of e_{0h} and e_{1h}, we have

$$u_{prop} = \left[r_1 \sum_{h=1}^{K} W_h \bar{D}_h \,(1 + e_{0h}) + r_2 \left(\sum_{h=1}^{K} W_h \bar{I}_h - \sum_{h=1}^{K} W_h \bar{I}_h \,(1 + e_{1h}) \right) - r_3 \sum_{h=1}^{K} W_h \bar{D}_h \,(1 + e_{0h})(1 + \varphi_{ps} e_1)^{-1} \right] \exp\left(\frac{\sum_{h=1}^{K} W_h \bar{I}_h - \sum_{h=1}^{K} W_h \bar{I}_h (1 + e_{1h})}{\sum_{h=1}^{K} W_h \bar{I}_h + \sum_{h=1}^{K} W_h \bar{I}_h (1 + e_{1h})} \right)^{\eta}$$

Where $\varphi_{ps} = \frac{a_{ps}\bar{I}}{a_{ps}\bar{I}_{ps} + b_{ps}}$

$$= \left[r_1 \bar{D}(1 + e_0) + r_2 \left(\bar{I} - \bar{I}(1 + e_1) \right) - r_3 \bar{D}(1 + e_0)(1 + \varphi_{ps} e_1)^{-1} \right] \exp\left(\frac{\bar{I} - \bar{I}(1+e_1)}{\bar{I} + \bar{I}(1+e_1)} \right)^{\eta}$$

$$= \left[r_1 \bar{D}(1 + e_0) - r_2 \bar{I}\, e_1 - r_3 \bar{D}(1 + e_0)(1 - \varphi_{ps} e_1 + \varphi_{ps}^2 e_1^2 \right] \exp\left(\frac{-e_1}{2 + e_1} \right)^{\eta}$$

$$= \left[r_1 \bar{D}(1 + e_0) - r_2 \bar{I}\, e_1 - r_3 \bar{D}(1 + e_0 - \varphi_{ps} e_1 - \varphi_{ps} e_0 e_1 + \varphi_{ps}^2 e_1^2) \right] \left[1 - \frac{\eta}{2} e_1 + \frac{3}{8} \eta e_1^2 \right]$$

$$u_{prop} - \bar{D} = \bar{D} \left\{ \left[(r_1 - 1) + r_1 e_0 - r_2 m\, e_1 - r_3 (1 + e_0 - \varphi_{ps} e_1 - \varphi_{ps} e_0 e_1 + \varphi_{ps}^2 e_1^2) \right] \left[1 - \frac{\eta}{2} e_1 + \frac{3}{8} \eta e_1^2 \right] \right\} \tag{14}$$

Where $m = \frac{\bar{I}}{\bar{D}}$

$$u_{prop} - \bar{D} = \bar{D} \left\{ (r_1 - 1) + e_0 (r_1 - r_3) - r_3 + e_1 \left(-r_2 m + r_3 \varphi_{ps} - \frac{\eta (r_1 - 1)}{2} - \frac{\eta r_3}{2} \right) + e_1^2 \left(\frac{3\eta (r_1 - 1)}{8} + r_2 \left(\frac{m\eta}{2} \right) - r_3 \varphi_{ps}^2 - r_3 \frac{\eta \varphi_{ps}}{2} - r_3 \frac{3\eta}{8} \right) + e_0 e_1 \left(-r_1 \frac{\eta}{2} + r_3 \varphi_{ps} + r_3 \frac{\eta}{3} \right) \right\}$$

By taking expectation on both sides of equation (14), we get bias as

$$\text{Bias}\,(u_{prop}) = \bar{D} \left\{ (r_1 - 1) - r_3 + V_I \left[\frac{3\eta (r_1 - 1)}{8} + r_2 \left(\frac{m\eta}{2} \right) - r_3 \left(\varphi_{ps}^2 + \frac{\eta \varphi_{ps}}{2} + \frac{3\eta}{8} \right) \right] + V_{DI} \left(-r_1 \frac{\eta}{2} + r_3 \varphi_{ps} + r_3 \frac{\eta}{3} \right) \right\}$$

By taking square on both sides of equation (14), we have

$$\left(u_{prop} - \bar{D} \right)^2 = \bar{D}^2 \left\{ \left((r_1 - 1) - r_3 \right)^2 + e_0^2 (r_1 - r_3)^2 + e_1^2 \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} + \frac{\eta r_3}{2} \right)^2 - 2 e_0 e_1 \left[(r_1 - r_3) \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} + \frac{\eta r_3}{2} \right) \right] + 2 \left((r_1 - 1) + r_3 \right) \left[e_0 (r_1 - r_3) - e_1 \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} - \frac{\eta r_3}{2} \right) \right] + e_1^2 \left[\left(\frac{3\eta (r_1 - 1)}{8} + r_2 \left(\frac{m\eta}{2} \right) - r_3 \varphi_{ps}^2 - r_3 \frac{\eta \varphi_{ps}}{2} - r_3 \frac{3\eta}{8} \right) \right] + e_0 e_1 \left(-r_1 \frac{\eta}{2} + r_3 \varphi_{ps} + r_3 \frac{\eta}{3} \right) \right\} \tag{15}$$

By considering expectation on both sides of equation (15), we get MSE as

$$\text{MSE}\,(u_{prop}) = \bar{D}^2 \left\{ \left((r_1 - 1) - r_3 \right)^2 + V_D (r_1 - r_3)^2 + V_I \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} - \frac{\eta r_3}{2} \right)^2 - 2 V_{DI} \left[(r_1 - r_3) \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} - \frac{\eta r_3}{2} \right) \right] + V_I \left[\left(\frac{3\eta (r_1 - 1)}{8} + r_2 \left(\frac{m\eta}{2} \right) - r_3 \varphi_{ps}^2 - r_3 \frac{\eta \varphi_{ps}}{2} - r_3 \frac{3\eta}{8} \right) \right] + V_{DI} \left(-r_1 \frac{\eta}{2} + r_3 \varphi_{ps} + r_3 \frac{\eta}{3} \right) \right\}$$

$$= \bar{D}^2 \left\{ (r_1 - 1)^2 + (r_3)^2 - 2 r_3 (r_1 - 1) + V_D (r_1 - r_3)^2 + V_I \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} - \frac{\eta r_3}{2} \right)^2 - 2 V_{DI} \left[(r_1 - r_3) \left(r_2 m - r_3 \varphi_{ps} + \frac{\eta (r_1 - 1)}{2} - \frac{\eta r_3}{2} \right) \right] + V_I \left[\left(\frac{3\eta (r_1 - 1)}{8} + r_2 \left(\frac{m\eta}{2} \right) - r_3 \varphi_{ps}^2 - r_3 \frac{\eta \varphi_{ps}}{2} - r_3 \frac{3\eta}{8} \right) \right] +$$

$V_{DI}\left(-r_1\frac{\eta}{2}+r_3\varphi_{ps}+r_3\frac{\eta}{3}\right)\Big\}$

We rewrite the above equation as

$$MSE\,(u_{prop})=\overline{D}^2\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right] \quad (16)$$

Where $\vartheta_1=-2+\frac{3\eta}{8}V_I+\eta V_{DI}-\frac{\eta^2}{2}V_I-\frac{\eta}{2}V_{DI}$

$\vartheta_2=1+V_D+\frac{\eta^2}{4}V_I-\eta V_{DI}$

$\vartheta_3=-m\,\eta V_I+\frac{m\eta}{2}V_I$

$\vartheta_4=m^2V_I$

$\vartheta_5=2+\frac{\eta^2}{2}V_I+\varphi_{ps}\,\eta V_I-\varphi_{ps}^2V_I-\frac{\eta}{2}\varphi_{ps}V_I-\frac{3\eta}{8}V_I+\eta V_{DI}+\varphi_{ps}V_{DI}+\frac{\eta}{3}V_{DI}$

$\vartheta_6=1+V_D+\varphi_{ps}^2V_I+\frac{\eta^2}{4}V_I+\varphi_{ps}V_I-2\varphi_{ps}V_{DI}-\eta V_{DI}$

$\vartheta_7=m\,\eta V_I-2mV_{DI}$

$\vartheta_8=-2-V_D-\frac{\eta^2}{2}V_I-\varphi_{ps}\,\eta V_I+2\varphi_{ps}V_{DI}+2\eta V_{DI}$

$\vartheta_9=-2m\varphi_{ps}V_I-m\eta V_I+2mV_{DI}$

To get the values of r_1, r_2 and r_3, differentiate equation (16) with respect to r_1, r_2 and r_3 and equate them to zero. We get,

$r_3=\frac{(2\vartheta_2\vartheta_3-\vartheta_1\vartheta_7)(2\vartheta_4\vartheta_8-\vartheta_7\vartheta_9)-(\vartheta_5\vartheta_7-\vartheta_3\vartheta_8)(\vartheta_7{}^2-4\vartheta_2\vartheta_4)}{(\vartheta_7\vartheta_8-2\vartheta_2\vartheta_9)(2\vartheta_4\vartheta_8-\vartheta_7\vartheta_9)-(\vartheta_7{}^2-4\vartheta_2\vartheta_4)(\vartheta_8\vartheta_9-2\vartheta_6\vartheta_7)}=\theta_1(say)$

$r_2=\frac{(2\vartheta_2\vartheta_3-\vartheta_1\vartheta_7)-\theta_1(\vartheta_7\vartheta_8-2\vartheta_2\vartheta_9)}{(\vartheta_7{}^2-4\vartheta_2\vartheta_4)}=\theta_2(say)$

$r_1=\frac{-(\vartheta_7\theta_2+\vartheta_8\theta_1+\vartheta_1)}{2\vartheta_2}$

V. Efficiency comparison

Theoretically, we establish the following criteria to assess the effectiveness of suggested estimator and those estimators taken into consideration in the literature.

By comparing equations (4) and (16), we have

$MSE\,(u_{prop})-MSE(u_{1a})<0$

$\overline{D}^2\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right] < \left[\frac{1}{n}-\frac{1}{N}\right]\sum_{h=1}^{K}\sum W_h S_{dh}^2+\frac{1}{n^2}\sum_{h=1}^{K}(1-W_h)\,S_{dh}^2$

By comparing equations (5) and (16),

$MSE\,(u_{prop})-MSE(u_{1b})<0$

$\overline{D}^2\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right] < \sum_{h=1}^{K}W_h^2\left[\frac{1}{n_h}-\frac{1}{N_h}\right]S_{dh}^2$

By comparing equations (6) and (16),

$MSE\,(u_{prop})-MSE\,(u_2)<0$

$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right] < \left[V_D+V_I\left[1-2\left(\frac{V_{DI}}{V_I}\right)\right]\right]$

By comparing equations (7) and (16),

MSE $(u_{prop}) - \text{MSE}\,(u_3) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<\left[V_D+V_I\left[1+2\left(\frac{V_{DI}}{V_I}\right)\right]\right]$$

By comparing equations (8) and (16),

MSE $(u_{prop}) - \text{MSE}\,(u_4) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<V_D(1-\xi_{DI}^2)$$

By comparing equations (9) and (16),

MSE $(u_{prop}) - \text{MSE}\,(u_5) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<[1+r_1^2A_1+r_2^2A_2+2r_1r_2A_3-2r_1A_4-2r_2A_5]$$

By comparing equations (10) and (16), we have

MSE $(u_{prop}) - \text{MSE}\,(u_{6a}) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<\left\{V_D+\frac{V_I}{4}\left[1-4\left(\frac{V_{DI}}{V_I}\right)\right]\right\}$$

MSE $(u_{prop}) <$ MSE (u_{6b})

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<\left\{V_D+\frac{V_I}{4}\left[1+4\left(\frac{V_{DI}}{V_I}\right)\right]\right\}$$

MSE $(u_{prop}) - \text{MSE}\,(u_{6c}) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<\left\{V_D+\frac{Av_I}{4}\left[\alpha-4\left(\frac{V_{DI}}{V_I}\right)\right]\right\}$$

By comparing equations (11) and (16),

MSE $(u_{prop}) - \text{MSE}\,(u_7) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<[1+r_1^2B_1+r_2^2B_2+2r_1r_2B_3-2r_1B_4-2r_2B_5]$$

By comparing equations (12) and (16),

MSE $(u_{prop}) - \text{MSE}\,(u_8) < 0$

$$\left[1-\frac{3\eta}{8}V_I+\vartheta_1 r_1+\vartheta_2 r_1^2+\vartheta_3 r_2+\vartheta_4 r_2^2+\vartheta_5 r_3+\vartheta_6 r_3^2+\vartheta_7 r_1 r_2+\vartheta_8 r_1 r_3+\vartheta_9 r_2 r_3\right]<[1+r_1^2C_1+r_2^2C_2+2r_1r_2C_3-2r_1C_4-2r_2C_5]$$

VI. Empirical study

We use information from the Ministry of Education of the Turkish Republic from 2007 on the number of teachers as the study variable (d) and the number of students classifying more or less than 750 in primary and secondary schools as the auxiliary attribute (i) for 923 districts across 6 regions (as 1: Marmara) 2, Atlantic, 3, Mediterranean, and 4, Central Anatolia Black Sea 5 and 6: East and Southeast Anatolia). Table 1 provides the data's summary statistics. We used Neyman allocation to place the samples in different strata.

The functions of auxiliary variable which we used in numerical calculation are:

$\sum W_h C_{ih} = 0.266448$, $\sum W_h S_{ih} = 0.246447$ and $\sum W_h \rho_{dih} = 0.145833$

In the case of unconditional post stratification, Table.2 shows the MSE values for our suggested estimator and the other estimators that were taken into consideration, together with the PRE values. It has been noted that the proposed estimator exhibits the maximum relative efficiency. It is also same in case of Conditional post stratification by observing Table.3.

Table 1. *Data Descriptive Statistics*

Stratum no.	$\mathbf{N_h}$	$\mathbf{n_h}$	$\mathbf{\bar{D}_h}$	$\mathbf{\bar{I}_h}$	$\mathbf{S_{dh}}$	$\mathbf{S_{ih}}$	$\mathbf{S_{dih}}$	$\mathbf{\beta_{2(ih)}}$
1	127	31	703.74	0.952	883.835	0.213	25.267	16.922
2	117	21	413	0.974	644.922	0.159	9.982	35.579
3	103	29	573.17	0.932	1033.467	0.253	37.453	10.34
4	170	38	424.66	0.888	810.585	0.316	44.625	4.231
5	205	22	267.03	0.912	403.654	0.284	21.04	6.675
6	201	39	393.84	0.95	711.723	0.218	18.66	15.56

Table 2. *Unconditional case: MSE and PRE values of existing estimators and proposed estimator.*

S. No.	**Estimator**	**MSE value**	**Percentage Relative Efficiency**
1.	u_{1a}	2539.82	100
2.	u_2	2400.58	105.80
3.	u_3	2629.67	96.58
4.	u_4	2398.50	105.89
5.	u_5	2397.75	105.92
6.	u_{6a}	2405.89	105.57
7.	u_{6b}	2520.44	100.77
8.	u_{6c}	2398.50	105.89
9.	u_7	2397.75	105.93
10.	u_8	1931.21	131.51
11.	u_{prop}	1714.78	148.11

Table 3. *Conditional case: MSE and PRE values of existing estimators and proposed estimator.*

S. No.	**Estimator**	**MSE value**	**Percentage Relative Efficiency**
1.	u_{1b}	2229.27	100
2.	u_2	1638.65	136.10
3.	u_3	2983.40	77.32
4.	u_4	846.88	263.23
5.	u_5	846.78	263.26
6.	u_{6a}	1913.52	116.50
7.	u_{6b}	2585.89	86.21
8.	u_{6c}	2204.72	101.11
9.	u_7	846.78	263.26
10.	u_8	519.11	429.44
11.	u_{prop}	68.54	3252.51

We may conclude that conditional post stratification outperformed unconditional post stratification by comparing MSE and PRE values in Tables.2 and 3.

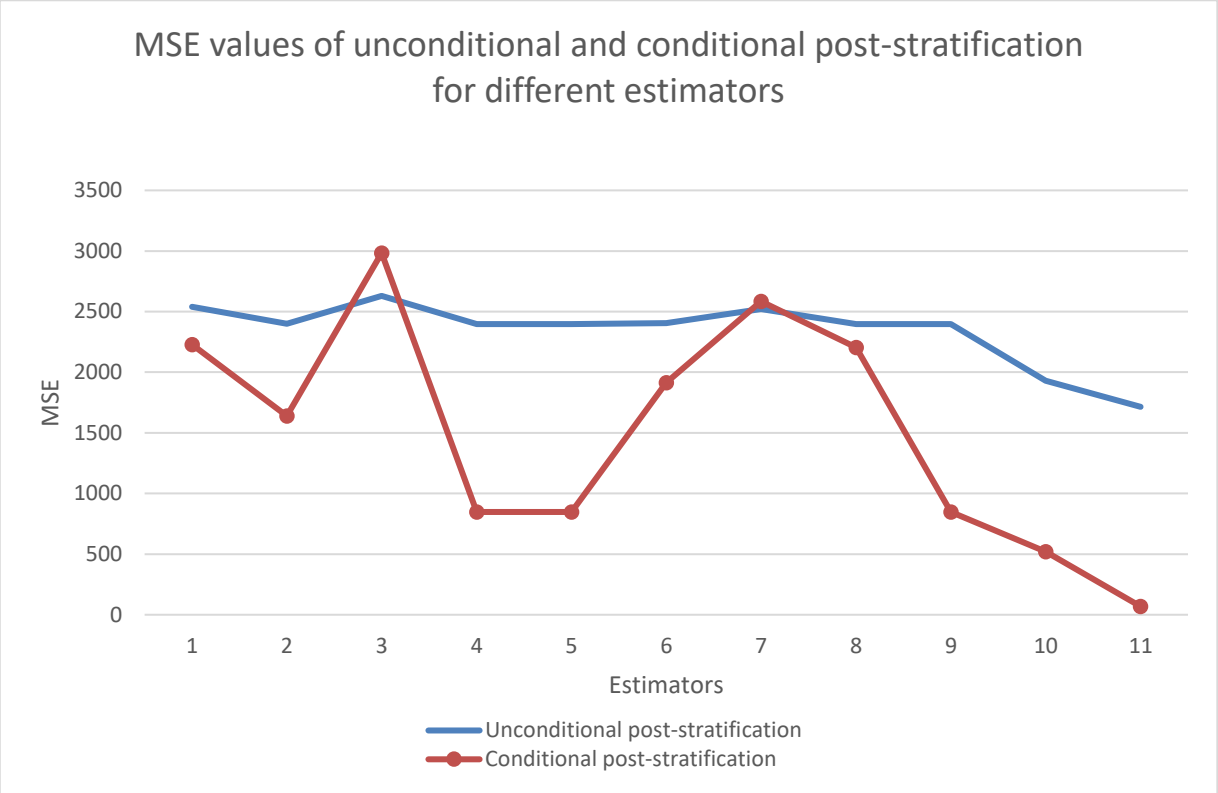


Figure 1. *MSE values for both Unconditional and Conditional Post-stratified estimators*.

We have presented the estimators in Table 2 and 3 in the above figure graphically. A line that represents the MSE values of Unconditional post stratified estimators can be seen in the picture together with a line with markers that represents the MSE values of conditional post stratified estimators. According to the graphic, conditional post stratified estimators have lower MSE values than unconditional post stratified estimators.

VII. Conclusion

In this research paper, we introduced a novel class of estimators and derived their Mean Square Error (MSE). We also investigated existing estimators and considered two cases in post-stratification: conditional and unconditional. Through a real data analysis, we computed the MSE and Percentage Relative Efficiency (PRE) values for all estimators presented in this study. The results, as shown in Table 2 and 3, clearly demonstrate that our proposed estimator exhibits the highest relative efficiency compared to the other estimators considered. Furthermore, we observed from the figure 1 that conditional post-stratification outperformed unconditional post-stratification in our analysis. These findings highlight the potential of our proposed estimators for enhancing mean estimation accuracy in post-stratification studies.

References

[1] Holt, D. and Smith, T. F. (1979). Post stratification. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 142(1):33-46.

[2] Naik, V. D. and Gupta, P. C. (1996). A note on estimation of mean with known population proportion of an auxiliary character. *Journal of Indian Society of Agricultural Stat*istics, *48*(2):151-158.

[3] Frey, J. and Feeman, T. G. (2012). An improved mean estimator for judgment post-stratification. *Computational Statistics and Data Analysis*, 56(2):418-426.

[4] Singh, H. P. and Solanki, R. S. (2012). Improved estimation of population mean in simple random sampling using information on auxiliary attribute. *Applied Mathematics and Computation*, 218(15):7798-7812.

[5] Onyeka, A. C. (2012). Estimation of population mean in post-stratified sampling using known value of some population parameter (s). *Statistics in Transition. New Series*, 13(1): 65-78.

[6] Onyeka, A. C., Nlebedim, V. U. and Izunobi, C. H. (2013). Exponential Estimators of Population Mean in Post-Stratified Sampling using Known Value of Some Population Parameters. *Global Journal of Science Frontier Research Mathematics and Decision Sciences*, *13*(9).

[7] Koyuncu, N. (2013). Improved estimation of population mean in stratified random sampling using information on auxiliary attribute. In *Proceeding of 59-th ISI world Statistics Congress, Hong Kong, China, August* (pp. 25-30).

[8] Singh, H. P. and Solanki, R. S. (2013). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical papers*, 54:479-497.

[9] Sharma, P. and Singh, R. (2013). Efficient estimator of population mean in stratified random sampling using auxiliary attribute. *World Applied Sciences Journal*, *27*(12):1786-1791.

[10] Tailor, R., Tailor, R. and Chouhan, S. (2017). Improved ratio-and product-type exponential estimators for population mean in case of post-stratification. *Communications in Statistics-Theory and Methods*, 46(21):10387-10393.

[11] Zamanzade, E. and Wang, X. (2017). Estimation of population proportion for judgment post-stratification. *Computational Statistics and Data Analysis*, *112*:257-269.

[12] Shahzad, U., Hanif, M., Koyuncu, N. and Luengo, A. G. (2019). A family of ratio estimators in stratified random sampling utilizing auxiliary attribute alongside the nonresponse issue. *Journal of Statistical Theory and Applications*, *18*(1):12-25.

[13] Singh, H. P., Gupta, A., and Tailor, R. (2023). Efficient class of estimators for finite population mean using auxiliary attribute in stratified random sampling. *Scientific Reports*, *13*(1):10253.

[14] Kumari, A., Kumar, U. and Singh, R. (2023). Improved estimators of population mean using auxiliary variables in post-stratification. *Journal of Scientific Research*, *67*(1):94-98.

RT&A, No 4 (76)
Volume 18, December 2023

ASSESSMENT OF GENERALIZED LIFETIME PERFORMANCE INDEX FOR LINDLEY DISTRIBUTION USING PROGRESSIVE TYPE-II SAMPLES

ABHIMANYU S YADAV[1], MAHENDRA SAHA[2], AMARTYA BHATTACHARYA[3], ARINDAM GUPTA[4]

•

[1] Department of Statistics, Banaras Hindu University, Varanasi, India. asybhu10@gmail.com
[2] Department of Statistics, Central University of Rajasthan, Ajmer, India. mahendrasaha@curaj.ac.in
[3] National Atlas and Thematic Mapping Organisation, GoI, Kolkata, India. amartya1991@gmail.com
[4] Department of Statistics, The University of Burdwan, Burdwan, India. guptaarin@gmail.com

Abstract

A meaningful subject of discourse in manufacturing industries is the assessment of the lifetime performance index. In manufacturing industries, the lifetime performance index is used to measure the performance of the product. A generalized lifetime performance index (GLPI) is defined by taking into consideration the median of the process measurement when the lifetime of products follow a parametric distribution may serve better the need of quality engineers and scientists in industry. The present study constructs various point estimators of the GLPI based on progressive type II right censored data for the Lindley distributed lifetime in both classical and Bayesian setup. We perform Monte Carlo simulations to compare the performances of the maximum likelihood and Bayes estimates with a gamma prior of $C_Y(L)$ under progressive type-II right censoring scheme. Finally, the validity of the model is adjudged through analysis of a data set.

Keywords: Bayesian estimation, Metropolis-Hastings method, Process Capability Index, Maximum likelihood estimator.

1. Introduction

Process Capability Indices (PCIs) have wide use in industries for evaluating a manufacturing process and whether or not it can produce articles within the specified limits. PCIs aim to quantify the capability of a process (X) to meet some specifications related to a measurable characteristic of its produced items. These specifications are determined through the lower specification limit (L), the upper specification limit(U) and the target value (T). PCI is an effective means to measure a process's performance and potential capabilities. In the manufacturing industry, PCIs are utilized to assess whether product quality meets customer expectations. Since capability is typically defined in dictionaries as the ability to carry out a task or achieve a goal, a better process capability implies better product quality. If the process capability is evaluated with product survival lifetime, it is clear that a larger lifetime means better product quality, higher reliability, and the process is capable. Hence, the lifetime of products exhibit the larger- the better quality characteristic of time

orientation. It should be noted that the lifetime of products does not follow a normal distribution. For instance, [14], [34], [6], [24] and [17] pointed out that the product lifetime possesses an exponential distribution. The description of the lifetime by the Weibull distribution was noted by [34], [40], [41] and [16]. In addition, [34] also mentioned that lifetime follows a gamma distribution. Since the lifetime of products exhibits the larger-the-better quality characteristic of time orientation, [35] and [23] recommended the use of the capability index (lifetime performance index) for evaluating the lifetime performance of electronic components, where L is the lower specification limit. Also, there are many different PCIs available in the literature. The hypothesis testing procedures are developed by [20] and [45] using the maximum likelihood estimator of PCI for Pareto distribution under type-II censored sampling and progressive type-I interval censoring, respectively. A hypothesis testing procedure was developed by [42] for the PCI of the Gompertz distribution based on the progressive type-I interval censored sample. The MLE was used by [46] to estimate the PCI of Rayleigh distribution based on the progressive type-I interval censored sample and developed a new hypothesis testing procedure utilizing an asymptotic distribution of this estimator. Some classical estimations and bootstrap confidence interval methods for the PCIs are derived by [36, 37] when the process follows exponentiated exponential and normal distributions, respectively. The classical and Bayes estimates of PCIs are obtained by [12, 13] for generalized exponential and normal distributions, respectively. For an expository review, the reader may follow the following articles of bibliography of the literature on PCIs, viz., [25], [38], [48] and [1].

The lifetime performance index (LPI) is defined by [23], denoted by C_L, which mainly originated from the concept of symmetry of lifetime distributions. The uniformly minimum variance unbiased estimator (UMVUE) for C_L was obtained by [39] and considered the problem of the hypothesis testing procedure for the exponential distribution. The UMVUE of C_L to develop the confidence interval under exponential distribution was obtained by [9]. The maximum likelihood estimator of the lifetime performance index based on first-failure progressive right type-II censored sample for Lindley distribution was obtained by [18]. The maximum likelihood estimates of the lifetime performance index based on progressive first failure censoring scheme Weibull, exponential and two-parameter exponential distributions, were obtained by [2, 3, 4] respectively. The lifetime performance index of products based on progressively Type-II censored for the Pareto samples was evaluated by [5]. The MLE, some Bayesian estimators, and credible intervals were given by [49] for the lifetime performance index of the Pareto distribution based on the general progressive type-II censored data. Approximate and exact parametric bootstrap confidence intervals are proposed by [50] for the process performance index of power-normal distribution. Most of the time, lifetime distributions are not necessarily symmetric. In this case, the median of the process distribution plays an important role than the process mean (μ). Therefore, it should be better if the index deals with the distribution's median (μ_e). If μ is replaced by μ_e, the inferential aspects and their property studies will be somewhat complicated. Using the median, [32] proposed a generalized process capability index (GPCI) that is the ratio of the proportion of specification conformance (or process yield) to the proportion of desired (or natural) conformance. In the same tune, [33] defined a GLPI given as

$$\begin{aligned} C_Y(L) &= \frac{0.5 - F(L)}{0.5 - \alpha} \\ &= \frac{1 - 2F(L)}{1 - 2\alpha}. \end{aligned} \tag{1}$$

Here $F(\cdot)$ denotes the cumulative distribution function (CDF) of the process distribution and $\alpha = P(X < LDL)$ with LDL being the lower desirable limit (practitioners sometimes take it as a lower tolerance limit). Here $(1 - \alpha)$ is the confidence level close to unity. Statistical inference for $C_Y(L)$, viz., properties of GLPI order, testing procedure for GLPI and parametric bootstrap confidence intervals based on a complete sample for the Lindley and in particular for exponential distribution have been obtained by [33].

In the case of a complete sample, it is necessary to continue the experiment until the last item (or product) fails. Sometimes, many articles have very long lifetimes, and the experiment continues for a very long period, so that the results may be of little interest or use. Then, it may be desirable to terminate the test before all the items under test fail, and the resulting observations will be called the censored sample. Various types of censored samples exist, including type II, progressive type II, and progressive first-failure censored samples. The testing of the hypothesis problem was proposed by [43] based on the maximum likelihood estimator (MLE) of C_L for two-parameter exponential distribution under type-II right censored sample. Based on a type-II right censored sample, the confidence interval using Pareto distribution was obtained by [20, 21]. A hypothesis testing procedure was proposed by [27, 28] based on MLE and UMVUE with the exponential distribution under progressive type-II right censored samples, respectively. The MLE of C_L was obtained by [19] under progressive first-failure censored samples from two-parameter exponential distributions. The C_L for the exponential lifetime products are evaluated by [29, 30] based on type-II censored data. They obtained Bayes's estimate of C_L for the Rayleigh lifetime products based on upper record values, respectively. Recently, [11] assessed the lifetime performance index for Weibull distributed products based on progressive type-II right censored samples.

In this article, we consider a progressive type-II right censoring scheme, which is helpful in a specific fraction of individuals at risk that may be removed from the experiment at each of several ordered failure times. Therefore, a progressive censoring scheme allows us to incorporate the removals before the experiment's termination into analysis, which is a very common situation in life-testing experiments. To the best of our knowledge thus far, an attempt has yet to be made to study the GLPI $C_Y(L)$ based on a progressive type-II right censoring scheme. Filling up this gap is the aim of the present study. In this article, we consider the GLPI $C_Y(L)$, introduced by [33] that could be used for either normal or non-normal and either continuous or discrete characteristics and is very simple and could be used comfortably by the practitioners.
The paper is arranged as follows. In section 2, The MLE and the Bayes estimate of $C_Y(L)$ are suggested based on progressive type II right censored sample for the Lindley distributed lifetime. In section 3, the testing procedure due to the GLPI is done. In section 4, an extensive Monte Carlo study is carried out to compare the performances of $C_Y(L)$ based on considered methods of estimation (MLE and Bayes) in terms of their corresponding mean squared errors (MSEs). A real-world application has been discussed to illustrate the proposed index under progressive type-II right censored samples in section 5. A brief concluding remark is made in section 6.

2. Estimation of $C_Y(L)$ under progressively type-II censored sample for Lindley products

Suppose that the lifetime of products may be modeled by Lindley distribution and let X denote the lifetime of such product. Hence, the probability density function (PDF) and cumulative distribution function (CDF), specified by Lindley distribution (see, [31]) are given as

$$f(x) = \frac{\theta^2}{\theta+1}(1+x)e^{-x\theta}, \quad x>0,\ \theta>0 \tag{2}$$

and

$$F(x) = 1-\frac{1+\theta+x\theta}{\theta+1}e^{-x\theta}, \quad x>0,\ \theta>0 \tag{3}$$

respectively, where, θ is the parameter. Now, for a process whose distribution can be regarded as Lindley, the GLPI is given as

$$\begin{aligned} C_Y(L) &= \left\{\frac{1-2F(L)}{1-2\alpha}\right\} \\ &= \left\{\frac{1-2\left(1-\frac{1+\theta+\theta L}{1+\theta}e^{-\theta L}\right)}{1-2\alpha}\right\}. \end{aligned} \tag{4}$$

Here, in the following subsections, we derived the maximum likelihood estimate (MLE) and the Bayes estimate of $C_Y(L)$ under progressively type-II right censoring scheme for Lindley distributed products, respectively.

2.1. Maximum likelihood estimate of $C_Y(L)$

The experimenter may not always observe the lifetimes of all the products (or items) put on tests for conducting life testing experiments. The reason may be time limitation and/or other restrictions (such as money, mechanical or experimental difficulties, material resources, etc.) on data collection. Therefore, censored samples may arise in practice. In an industrial experiment, products (or items) may break accidentally. These lead us into the area of progressive type-II censoring. Under this scheme, n units are placed on test at time zero, and m failures are observed. When the first failure is observed, r_1 of the surviving units are randomly selected and removed. At the second observed failure, r_2 of the surviving units are randomly selected and removed. Termination of the experiment occurs when the m-th failure is observed, and the remaining $r_m = n - \sum_{j=1}^{m-1} r_j - m$ surviving units are all removed. Inferences for the data obtained by progressive censoring have been investigated, among others, by [10], [7], [15], and [44]. So, in this paper, we consider the case of the progressive type-II right censoring.

Let $x_{1:m:n}$, $x_{2:m:n}$, ..., $x_{m:m:n}$ be a progressive type-II right censored sample where $x_{1:m:n}$, $x_{2:m:n}$, ..., $x_{m:m:n}$ denote the observed failure times and r_1, r_2, ..., r_m denote the corresponding numbers of items removed (withdrawn) from the test. If m be the number of failures observed before termination, then $x_{1:m:n} \leq x_{2:m:n} \leq$ ..., $\leq x_{m:m:n}$ be the observed ordered lifetimes. For convenience, we will write $x_{i:m:n}$ as $x_{(i)}$. Let r_i denote the number of items removed at the time of the ith failure, $0 \leq r_i \leq n - \sum_{j=1}^{i-1} r_j - i$, $i = 2, 3, ..., m-1$ with $0 \leq r_1 \leq n-1$ and $r_m = n - \sum_{j=1}^{m-1} r_j - m$, where r_i's and m are pres-specified integers [see, Viveros and Balakrishnan (1994)]. The complete sample ($r_1 = r_2 = ... = r_m = 0$) and type-II right censored samples ($r_1 = r_2 = ... = r_{m-1} = 0$, $r_m = n - m$) are special cases of this scheme. For further details and for relevant references the reader may follow the article of [7]. The likelihood function of θ under progressive type-II right censoring scheme is given by

$$\begin{aligned} l(\theta) &= A\prod_{i=1}^{m} f(x_{(i)};\theta)[1-F(x_{(i)};\theta)]^{r_i}, \text{ where } A = n(n-r_1-1)...(n-\sum_{j=1}^{m-1} r_j - m + 1) \\ &= A.\frac{\theta^{2m}}{(1+\theta)^{m+\sum_{i=1}^{m} r_i}} e^{-\theta\sum_{i=1}^{m}(1+r_i)x_{(i)}} \prod_{i=1}^{m}(1+x_{(i)}) \prod_{i=1}^{m}(1+\theta+\theta x_{(i)})^{r_i} \end{aligned} \tag{5}$$

Therefore, the log-likelihood function is given by

$$\begin{aligned} L(\theta) = \ln l(\theta) &= k + 2m\ln\theta - (m+\sum_{i=1}^{m} r_i)\ln(1+\theta) - \theta\sum_{i=1}^{m}(1+r_i)x_{(i)} \\ &\quad + \sum_{i=1}^{m} r_i \ln(1+\theta+\theta x_{(i)}), \end{aligned} \tag{6}$$

where k is constant, independent of θ. Now, for MLE of the parameter θ, $\frac{\partial L(\theta)}{\partial \theta} = 0$

$$\implies \frac{2m}{\theta} - \frac{m + \sum_{i=1}^{m} r_i}{1+\theta} - \sum_{i=1}^{m}(1+r_i)x_{(i)} + \sum_{i=1}^{m} \frac{r_i(1+x_{(i)})}{1+\theta+\theta x_{(i)}} = 0. \tag{7}$$

The explicit solution for the parameter θ through the above non-linear equation is not possible. Hence, to solve this equation for θ, we have to proceed by some numerical method, from the previous equation,

$$\frac{\partial^2 L(\theta)}{\partial \theta^2} = -\frac{2m}{\theta^2} + \frac{m + \sum_{i=1}^{m} r_i}{(1+\theta)^2} - \sum_{i=1}^{m} \frac{r_i(1+x_{(i)})^2}{(1+\theta+\theta x_{(i)})^2}.$$

Hence, the Fisher's information $I(\theta)$ is obtained as;

$$I(\theta) = -E(\frac{\partial^2 L(\theta)}{\partial \theta^2}) = \frac{2m}{\theta^2} - \frac{m + \sum_{i=1}^{m} r_i}{(1+\theta)^2} + \sum_{i=1}^{m} r_i E\left[\frac{(1+X_{(i)})^2}{(1+\theta+\theta X_{(i)})^2}\right] \tag{8}$$

Further, after obtaining the solution above mentioned non-linear equation, the MLE of the index $C_Y(L)$ can be directly computed using invariance property of MLE. Let $\hat{\theta}$ be the solution of the non-linear equation, then the MLE $\hat{C}_Y(L)$ of $C_Y(L)$ is given by

$$\hat{C}_Y(L) = \left\{\frac{1 - 2\left(1 - \frac{1+\hat{\theta}+\hat{\theta}L}{1+\hat{\theta}} e^{-\hat{\theta}L}\right)}{1-2\alpha}\right\}. \tag{9}$$

The asymptotic distribution of MLE $\hat{\theta}$ is normal distribution $N(\theta, I^{-1}(\theta))$. Hence, the two-sided tail asymptotic $100(1-\alpha)\%$ confidence interval for the parameter θ is given by

$$\left\{\hat{\theta} \mp \tau_{\alpha/2}\sqrt{I^{-1}(\hat{\theta})}\right\},$$

where $I(\hat{\theta}) = \frac{2m}{\hat{\theta}^2} - \frac{m+\sum_{i=1}^{m} r_i}{(1+\hat{\theta})^2} + \sum_{i=1}^{m} \frac{r_i(1+x_{(i)})^2}{(1+\hat{\theta}+\hat{\theta}x_{(i)})^2}$ and $\tau_{\alpha/2}$ is the upper $\alpha/2$-point of standard normal deviate. The asymptotic distribution of MLE $\hat{C}_Y(L)$ is also normal distribution $N(C_Y(L),$ $Var(\hat{C}_Y(L))$ with

$$Var(\hat{C}_Y(L)) = \left[\frac{\partial C_Y(L)}{\partial \theta}\right]^2 \times Var(\hat{\theta}) = \left[\frac{-2\theta L\{1+(1+\theta)(1+L)\}e^{-\theta L}}{(1-2\alpha)(1+\theta)^2}\right]^2 \times I^{-1}(\theta).$$

The asymptotic variance is approximated as

$$\hat{Var}(\hat{C}_Y(L)) \approx \left[\frac{-2\hat{\theta} L\{1+(1+\hat{\theta})(1+L)\}e^{-\hat{\theta}L}}{(1-2\alpha)(1+\hat{\theta})^2}\right]^2 \times I^{-1}(\hat{\theta}).$$

2.2. Bayes estimate of $C_Y(L)$

In this section, we obtain the Bayes estimator of $C_Y(L)$ under the assumption that the parameter θ is random variable and follows some prior distribution. Let the prior distribution of θ is assumed to be Gamma with parameter (k, a). Then the distribution of θ is given as

$$g(\theta) = \frac{a^k}{\Gamma(k)} e^{-a\theta}\theta^{k-1}, \quad \theta > 0 \tag{10}$$

Now, the posterior distribution of θ by using Equations (2.5) and (2.9) is given as

$$g(\theta \mid \underline{x}) \propto \frac{\theta^{2m+k-1}}{(1+\theta)^{m+\sum_{i=1}^{m} r_i}} e^{-\theta(a+\sum_{i=1}^{m}(1+r_i)x_{(i)})} \prod_{i=1}^{m}(1+\theta+\theta x_{(i)})^{r_i} \quad ; \ \theta > 0. \tag{11}$$

Hence, the Bayes estimate of the parameter θ under squared error loss function is obtained by the following Equation:

$$E_\theta(\theta) = \zeta \int_\theta \frac{\theta^{2m+k}}{(1+\theta)^{m+\sum_{i=1}^m r_i}} e^{-\theta(a+\sum_{i=1}^m (1+r_i)x_{(i)})} \prod_{i=1}^m (1+\theta+\theta x_{(i)})^{r_i} \, d\theta \tag{12}$$

where ζ is a proportionality constant. The computation of the Bayes estimate of the index $C_Y(L)$ under the same assumption of prior and loss function is not possible directly from the above posterior distribution. Since, the explicit form of the posterior PDF is not available but the associated plot exhibit a more or less assume the shape of normal probability distribution. Thus, the Metropolis-Hastings method with normal proposal distribution is to be used to generate random numbers from respective posterior distribution using the Gibbs algorithm. The following steps are taken to generate the posterior random deviates from the above posterior is as follows:

- Start with an initial guess $\theta^{(0)}$.
- Set $t = 1$.
- Using the Metropolis-Hastings, generate $\theta^{(t)}$ from $g(\theta \mid \underline{x})$ with the $N\left(\theta^{(t-1)}, 1\right)$ proposal distribution.
- Compute $C_Y(L)^{(t)}$ from Equation (1)
- Set $t = t + 1$.
- Repeat steps 3-5, T times.

Note that in step 3, we use the Metropolis-Hastings algorithm with $q(\theta^{(t-1)}, \sigma^2)$ proposal distribution as follows:

1. Let $x = \theta^{(t-1)}$.
2. Generate y from the proposal distribution q.
3. Let $p(x, y) = \min\left(1, \frac{g_\theta(y)q(x)}{g_\theta(x)q(y)}\right)$.
4. Accept y with the probability $p(x, y)$ or accept x with the probability $1 - p(x, y)$.

The posterior deviates for $C_Y(L)$ using the above mentioned steps is simulated using the random deviates of θ by plug-in principal. Let $C_Y(L)^1, C_Y(L)^2, \cdots, C_Y(L)^T$ be the T simulated posterior deviates, then the approximate posterior mean, and posterior variance of $C_Y(L)$ are given by

$$\hat{E}(C_Y(L)|\underline{x}) = \frac{1}{T}\sum_{t=1}^T C_Y(L)^t$$

and

$$MSE(C_Y(L)|\underline{x}) = \frac{1}{T}\sum_{t=1}^T \left(C_Y(L)^t - C_Y(L)\right)^2$$

respectively.

3. Testing procedure for the generalized lifetime performance index using progressive type-II samples

In this section, following statistical hypothesis testing will be performed to access whether the $C_Y(L)$ adheres the required level. The proposed hypothesis testing procedure using progressive type-II samples can be performed for $C_Y(L)$, summarized as follows:

1. Determine the lower specification limit L and the GLPI, C_Y^0.
2. Construct the null hypothesis $H_0 : C_Y(L) \leq C_Y^0$ against the alternative $H_1 : C_Y(L) > C_Y^0$.
3. Specify the level of significant α.
4. Compute $\hat{C}_Y(L)$ and $\hat{V}ar(\hat{C}_Y(L)) \approx [\frac{-2\hat{\theta}L\{1+(1+\hat{\theta})(1+L)\}e^{-\hat{\theta}L}}{(1-2\alpha)(1+\hat{\theta})^2}]^2.I^{-1}(\hat{\theta})$.
5. Set critical region $\omega : \hat{C}_Y(L) > C_Y^0 + \tau_\alpha \sqrt{\hat{V}ar(\hat{C}_Y(L))}$.

4. Simulation and discussion based on progressively type-II censored sample.

In this section, a comparison study has been carried out through simulation study under progressively type II censoring scheme between maximum likelihood and the Bayes estimate of $C_Y(L)$ in terms of their corresponding mean squared errors (MSEs) for the considered set up. All the calculations have been made by using R software (see, [22]). The progressive type-II right censored sample from the considered lifetime distribution for the different variation of the parameter (θ), censoring parameters (n, m), censoring schemes $(r_i's)$ and lower desired limit (L) is generated by following the algorithm suggested by [8] algorithm, as stated below:

1. Generate m independent Uniform$(0, 1)$ observations $W_1,\ W_2,\ ...,\ W_m$.
2. Set $V_i = W_i^{1/(i+r_m+r_{m-1}+...+r_{m-i+1})}$ for $i = 1, 2, ..., m$.
3. Set $U_i = 1 - V_m.V_{m-1}...V_{m-i+1}$ for $i = 1, 2, ..., m$. Then $U_1, U_2, ..., U_m$ is the required progressive type-II censored sample from the Uniform $(0, 1)$ distribution.
4. Finally, we set $U_i = 1 - \frac{1+\theta+X_i\theta}{\theta+1}e^{-X_i\theta}$, and solve this equation by Newton-Raphson method to get X_i for $i = 1, 2, ..., m$. Then $X_1, X_2, ..., X_m$ is the required progressive type-II censored sample from the distribution (2).

The simulated MSE of the MLE and the Bayes estimate of $C_Y(L)$ have been presented in Tables 1-6 for different censoring schemes with some particular choices of θ and L. From the Tables 1-6, it is found that the MSE of each estimator is decreasing with n, the sample size. This verifies the consistency property of all the estimators. It is also observed that the performance of the Bayes estimation is relatively better than the MLE under all the considered choices.

5. Applications

A real data set is cited to illustrate the MLE, Bayes estimate of LPI C_L (see, [23]) and the proposed GLPI $C_Y(L)$ for progressive type-II right censoring scheme for the Lindley distributed lifetime. The considered data set is described in detail by [33] (also available in [26]), and the goodness of fit test to Lindley distribution is discussed therein. The data set is primarily fitted to the exponential model in Lawless. [33] have checked the data set with the Lindley distribution and found it to be a better fit. Thus, from the same data set, the progressive type-II censored data are generated for the different values of $m = 10, 15$, L=100 and the censoring schemes, and the corresponding MLE and Bayes estimates of $C_Y(L)$ is reported in the Table 7. The Bayes estimates for the real data set is computed under non-informative prior.
Now, the proposed testing procedure of the GLPI $C_Y(L)$ is performed for the above chosen schemes as follows. For the considered data set, the progressive type-II censored data are generated for the same censoring schemes, mentioned in Table 8, respectively. The MLE of the parameter θ for the Lindley distribution is obtained from the Eqn. (2.7) for all the considered schemes and the same are reported in Table 8. The statistical test for testing the null hypothesis $H_0 : C_Y(L) \leq 1$ against the alternative hypothesis $H_1 : C_Y(L) > 1$ has been performed for the

Table 1: *MLEs of $C_Y(L)$ and their MSEs with $L = 0.1$ and $L = 0.3$, samples generated from the Lindley distribution for $\theta = 0.50$ under progressively type-II censoring scheme [True $C_Y(L) = 1.073193$ when $L = 0.1$ and $C_Y(L) = 0.992842$ when $L = 0.3$].*

2[0]*n,m	2[0]*Schemes	L=0.1		L=0.3	
		$\hat{C}_Y(L)_{MLE}$	$MSE[\hat{C}_Y(L)_{MLE}]$	$\hat{C}_Y(L)_{MLE}$	$MSE[\hat{C}_Y(L)_{MLE}]$
7[0]*10, 8	0*10	1.07058	0.00149	0.98230	0.01431
	2, 0*7	1.06040	0.00316	0.99986	0.01086
	1,1, 0*6	1.06686	0.00326	0.99877	0.01036
	0*7, 2	1.06560	0.00272	0.98734	0.01451
	0*6, 1*2	1.07203	0.00191	0.99791	0.00939
	1, 0*6, 1	1.06470	0.00424	0.97192	0.01438
	0*3, 1*2, 0*3	1.06315	0.00349	0.99770	0.00973
7[0]*20, 16	0*20	1.07081	0.00072	0.99550	0.00439
	4, 0*15	1.06834	0.00083	1.00409	0.00546
	1*4, 0*12	1.06991	0.00108	0.99523	0.00572
	0*15, 4	1.07087	0.00106	0.98304	0.00653
	0*12, 1*4	1.07045	0.00138	1.00387	0.00463
	2,0*14,2	1.06791	0.00096	1.06689	0.00177
	0*7, 2*2, 0*7	1.06902	0.00092	0.99288	0.00449
7[0]*30, 24	0*30	1.07523	0.00048	0.99840	0.00301
	6,0*23	1.06878	0.00068	0.97993	0.00392
	1*6, 0*18	1.06788	0.00077	0.99505	0.00416
	0*23, 6	1.06747	0.00088	0.99042	0.00439
	0*18, 1*6	1.06858	0.00057	1.00475	0.00349
	3, 0*22, 3	1.06891	0.00063	0.99489	0.00358
	0*9, 1*6, 0*9	1.07182	0.00050	0.98809	0.00379
7[0]*50, 40	0*50	1.07370	0.00028	0.99244	0.00218
	10, 0*39	1.06526	0.00050	0.98831	0.00264
	2*5, 0*35	1.07372	0.00028	0.98873	0.00256
	0*39, 10	1.07123	0.00045	0.99057	0.00281
	0*30, 1*10	1.06765	0.00052	0.99992	0.00220
	5, 0*38, 5	1.07045	0.00032	0.98705	0.00375
	0*18, 2*5, 0*17	1.07152	0.00038	0.98859	0.00329

Note: In the table, Scheme $(0 * 3, r)$ indicates that at 1st, 2nd and 3rd failure, no active unit is withdrawn or removed but at 4th failure, r active units are drawn or removed.

Table 2: *Bayes estimates of $C_Y(L)$ and their MSEs with $L = 0.1$ and $L = 0.3$, samples generated from the Lindley distribution for $\theta = 0.50$ under progressively type-II censoring scheme [True $C_Y(L) = 1.073193$ when $L = 0.1$ and $C_Y(L) = 0.992842$ when $L = 0.3$].*

2[0]*n,m	2[0]*Schemes	L=0.1		L=0.3	
		$\hat{C}_Y(L)_{Bayes}$	$MSE[\hat{C}_Y(L)_{Bayes}]$	$\hat{C}_Y(L)_{Bayes}$	$MSE[\hat{C}_Y(L)_{Bayes}]$
7[0]*10, 8	0*10	1.07597	0.00135	0.97546	0.01372
	2, 0*7	1.06649	0.00283	0.99215	0.01049
	1,1, 0*6	1.07202	0.00310	0.99148	0.00999
	0*7, 2	1.07260	0.00245	0.97842	0.01403
	0*6, 1*2	1.07847	0.00168	0.98854	0.00911
	1, 0*6, 1	1.07092	0.00416	0.96363	0.01415
	0*3, 1*2, 0*3	1.06895	0.00315	0.99046	0.00958
7[0]*20, 16	0*20	1.07378	0.00066	0.99127	0.00437
	4, 0*15	1.07176	0.00077	0.99953	0.00533
	1*4, 0*12	1.07317	0.00101	0.99098	0.00568
	0*15, 4	1.07475	0.00098	0.97784	0.00658
	0*12, 1*4	1.07413	0.00128	0.99871	0.00449
	2,0*14,2	1.07175	0.00088	1.06302	0.00169
	0*7, 2*2, 0*7	1.07239	0.00085	0.98837	0.00439
7[0]*30, 24	0*30	1.07714	0.00047	0.99551	0.00297
	6,0*23	1.07127	0.00063	0.97659	0.00390
	1*6, 0*18	1.07023	0.00072	0.99216	0.00409
	0*23, 6	1.07036	0.00081	0.98673	0.00420
	0*18, 1*6	1.07131	0.00052	1.00128	0.00341
	3, 0*22, 3	1.07160	0.00059	0.99127	0.00345
	0*9, 1*6, 0*9	1.07408	0.00048	0.98499	0.00360
7[0]*50, 40	0*50	1.07486	0.00028	0.99075	0.00209
	10, 0*39	1.06679	0.00047	0.98640	0.00240
	2*5, 0*35	1.07516	0.00027	0.98682	0.00247
	0*39, 10	1.07291	0.00044	0.98830	0.00273
	0*30, 1*10	1.06939	0.00049	0.99764	0.00193
	5, 0*38, 5	1.07212	0.00030	0.98480	0.00357
	0*18, 2*5, 0*17	1.07286	0.00037	0.98670	0.00292

Table 3: *MLEs of $C_Y(L)$ and their MSEs with $L = 0.1$ and $L = 0.3$, samples generated from the Lindley distribution for $\theta = 0.75$ under progressively type-II censoring scheme [True $C_Y(L) = 1.038898$ when $L = 0.1$ and $C_Y(L) = 0.891517$ when $L = 0.3$].*

2[0]*n,m	2[0]*Schemes	L=0.1		L=0.3	
		$\hat{C}_Y(L)_{MLE}$	$MSE[\hat{C}_Y(L)_{MLE}]$	$\hat{C}_Y(L)_{MLE}$	$MSE[\hat{C}_Y(L)_{MLE}]$
10, 8	0*10	1.02408	0.00490	0.90298	0.01973
	2, 0*7	1.03005	0.00589	0.88015	0.02835
	1,1, 0*6	1.02637	0.00580	0.90158	0.02385
	0*7, 2	1.02057	0.00629	0.91253	0.02451
	0*6, 1*2	1.02306	0.00557	0.92649	0.02055
	1, 0*6, 1	1.03016	0.00484	0.88536	0.03157
	0*3, 1*2, 0*3	1.01554	0.00776	0.90067	0.01950
20, 16	0*20	1.03593	0.00230	0.89003	0.01112
	4, 0*15	1.03237	0.00296	0.88569	0.01545
	1*4, 0*12	1.03649	0.00246	0.90247	0.01130
	0*15, 4	1.02421	0.00318	0.89690	0.01192
	0*12, 1*4	1.02827	0.00291	0.88864	0.01237
	2,0*14,2	1.03430	0.00229	1.03440	0.00274
	0*7, 2*2, 0*7	1.03789	0.00200	0.89309	0.01031
30, 24	0*30	1.03208	0.00159	0.88711	0.00637
	6,0*23	1.03609	0.00180	0.89093	0.00982
	1*6, 0*18	1.03655	0.00173	0.88934	0.00759
	0*23, 6	1.03475	0.00164	0.89663	0.00958
	0*18, 1*6	1.03212	0.00182	0.89877	0.00774
	3, 0*22, 3	1.03816	0.00202	0.89079	0.01046
	0*9, 1*6, 0*9	1.03469	0.00183	0.89383	0.00753
50, 40	0*50	1.03908	0.00077	0.89749	0.00487
	10, 0*39	1.03609	0.00114	0.89273	0.00453
	2*5, 0*35	1.03278	0.00128	0.89426	0.00483
	0*39, 10	1.03953	0.00110	0.90031	0.00463
	0*30, 1*10	1.03555	0.00109	0.89815	0.00530
	5, 0*38, 5	1.03578	0.00128	0.89571	0.00574
	0*18, 2*5, 0*17	1.03867	0.00089	0.89162	0.00504

Table 4: *Bayes estimates of $C_Y(L)$ and their MSEs with $L = 0.1$ and $L = 0.3$, samples generated from the Lindley distribution for $\theta = 0.75$ under progressively Type-II censoring scheme [True $C_Y(L) = 1.038898$ when $L = 0.1$ and $C_Y(L) = 0.891517$ when $L = 0.3$].*

2[0]*n,m	2[0]*Schemes	L=0.1		L=0.3	
		$\hat{C}_Y(L)_{Bayes}$	$MSE[\hat{C}_Y(L)_{Bayes}]$	$\hat{C}_Y(L)_{Bayes}$	$MSE[\hat{C}_Y(L)_{Bayes}]$
10, 8	0*10	1.03297	0.00455	0.89338	0.01833
	2, 0*7	1.03916	0.00579	0.87174	0.02609
	1,1, 0*6	1.03618	0.00560	0.89210	0.02173
	0*7, 2	1.03254	0.00582	0.90036	0.02258
	0*6, 1*2	1.03506	0.00509	0.91270	0.01865
	1, 0*6, 1	1.04124	0.00444	0.87482	0.02897
	0*3, 1*2, 0*3	1.02560	0.00733	0.88980	0.01799
20, 16	0*20	1.04061	0.00223	0.88503	0.01078
	4, 0*15	1.03777	0.00285	0.88032	0.01488
	1*4, 0*12	1.04167	0.00237	0.89705	0.01079
	0*15, 4	1.03095	0.00294	0.88995	0.01139
	0*12, 1*4	1.03464	0.00272	0.88194	0.01196
	2,0*14,2	1.04039	0.00220	1.02836	0.00293
	0*7, 2*2, 0*7	1.04343	0.00193	0.88711	0.00992
30, 24	0*30	1.03536	0.00152	0.88355	0.00628
	6,0*23	1.03974	0.00175	0.88705	0.00954
	1*6, 0*18	1.04011	0.00168	0.88558	0.00743
	0*23, 6	1.03937	0.00157	0.89173	0.00930
	0*18, 1*6	1.03660	0.00174	0.89408	0.00747
	3, 0*22, 3	1.04224	0.00197	0.88630	0.01015
	0*9, 1*6, 0*9	1.03840	0.00177	0.88977	0.00734
50, 40	0*50	1.04110	0.00076	0.89530	0.00479
	10, 0*39	1.03837	0.00111	0.89024	0.00445
	2*5, 0*35	1.03504	0.00124	0.89184	0.00473
	0*39, 10	1.04223	0.00109	0.89728	0.00452
	0*30, 1*10	1.03823	0.00106	0.89516	0.00518
	5, 0*38, 5	1.03836	0.00125	0.89290	0.00565
	0*18, 2*5, 0*17	1.04098	0.00088	0.88919	0.00497

Table 5: *MLEs of $C_Y(L)$ and their MSEs with $L = 0.1$ and $L = 0.3$, samples generated from the Lindley distribution for $\theta = 1.50$ under progressively type-II censoring scheme [True $C_Y(L) = 0.916334$ when $L = 0.1$ and $C_Y(L) = 0.560892$ when $L = 0.3$].*

2[0]*n,m	2[0]*Schemes	L=0.1		L=0.3	
		$\hat{C}_Y(L)_{MLE}$	$MSE[\hat{C}_Y(L)_{MLE}]$	$\hat{C}_Y(L)_{MLE}$	$MSE[\hat{C}_Y(L)_{MLE}]$
7[0]*10, 8	0*10	0.92111	0.01876	0.59670	0.06439
	2, 0*7	0.92292	0.02211	0.58234	0.07653
	1,1, 0*6	0.92693	0.01922	0.55731	0.06999
	0*7, 2	0.92967	0.01968	0.57833	0.07527
	0*6, 1*2	0.93835	0.01868	0.60658	0.06007
	1, 0*6, 1	0.92819	0.02075	0.59201	0.06345
	0*3, 1*2, 0*3	0.91772	0.02080	0.57585	0.06962
7[0]*20, 16	0*20	0.92597	0.00852	0.57983	0.02738
	4, 0*15	0.91773	0.01063	0.56602	0.03939
	1*4, 0*12	0.92573	0.01097	0.57893	0.03167
	0*15, 4	0.91871	0.01025	0.58104	0.03375
	0*12, 1*4	0.91498	0.01218	0.59792	0.03509
	2,0*14,2	0.91435	0.01054	0.58679	0.03709
	0*7, 2*2, 0*7	0.91880	0.00977	0.57989	0.03197
7[0]*30, 24	0*30	0.91462	0.00617	0.56864	0.01894
	6,0*23	0.91960	0.00780	0.57341	0.02488
	1*6, 0*18	0.92418	0.00739	0.57963	0.02501
	0*23, 6	0.92321	0.00829	0.57843	0.02103
	0*18, 1*6	0.92471	0.00667	0.56706	0.02135
	3, 0*22, 3	0.91766	0.00781	0.57184	0.01984
	0*9, 1*6, 0*9	0.91798	0.00590	0.55905	0.02281
7[0]*50, 40	0*50	0.91497	0.00387	0.57058	0.01225
	10, 0*39	0.91995	0.00436	0.56369	0.01427
	2*5, 0*35	0.92503	0.00390	0.56944	0.01577
	0*39, 10	0.92130	0.00401	0.56377	0.01520
	0*30, 1*10	0.91975	0.00443	0.57915	0.01139
	5, 0*38, 5	0.91687	0.00477	0.56287	0.01436
	0*18, 2*5, 0*17	0.91364	0.00370	0.55129	0.01093

Table 6: *Bayes estimates of $C_Y(L)$ and their MSEs with $L = 0.1$ and $L = 0.3$, samples generated from the Lindley distribution for $\theta = 1.5$ under progressively Type-II censoring scheme [True $C_Y(L) = 0.916334$ when $L = 0.1$ and $C_Y(L) = 0.560892$ when $L = 0.3$].*

2[0]*n,m	2[0]*Schemes	L=0.1		L=0.3	
		$\hat{C}_Y(L)_{Bayes}$	$MSE[\hat{C}_Y(L)_{Bayes}]$	$\hat{C}_Y(L)_{Bayes}$	$MSE[\hat{C}_Y(L)_{Bayes}]$
7[0]*10, 8	0*10	0.90365	0.01723	0.58724	0.05271
	2, 0*7	0.90402	0.01989	0.57706	0.05956
	1,1, 0*6	0.90828	0.01743	0.55465	0.05433
	0*7, 2	0.90508	0.01810	0.57052	0.05563
	0*6, 1*2	0.91331	0.01689	0.59403	0.04676
	1, 0*6, 1	0.90561	0.01861	0.58232	0.05013
	0*3, 1*2, 0*3	0.89825	0.01857	0.57098	0.05346
7[0]*20, 16	0*20	0.91626	0.00821	0.57459	0.02488
	4, 0*15	0.90710	0.01022	0.56201	0.03504
	1*4, 0*12	0.91540	0.01042	0.57481	0.02815
	0*15, 4	0.90527	0.01001	0.57409	0.02974
	0*12, 1*4	0.90220	0.01193	0.59034	0.03104
	2,0*14,2	0.90213	0.01034	0.57034	0.03124
	0*7, 2*2, 0*7	0.90799	0.00946	0.57521	0.02814
7[0]*30, 24	0*30	0.90805	0.00612	0.56529	0.01775
	6,0*23	0.91221	0.00761	0.57023	0.02298
	1*6, 0*18	0.91709	0.00714	0.57655	0.02317
	0*23, 6	0.91407	0.00807	0.57355	0.01931
	0*18, 1*6	0.91574	0.00648	0.56287	0.01980
	3, 0*22, 3	0.90917	0.00764	0.56766	0.01829
	0*9, 1*6, 0*9	0.91040	0.00581	0.55630	0.02116
7[0]*50, 40	0*50	0.91096	0.00385	0.56855	0.01179
	10, 0*39	0.91526	0.00428	0.56163	0.01363
	2*5, 0*35	0.92053	0.00379	0.56757	0.01504
	0*39, 10	0.91552	0.00393	0.56076	0.01451
	0*30, 1*10	0.91423	0.00437	0.57613	0.01084
	5, 0*38, 5	0.91159	0.00473	0.56055	0.01368
	0*18, 2*5, 0*17	0.90905	0.00370	0.54963	0.01048

pre-chosen level of significance $\alpha = 5\%$ and $L = 100$. Also, the estimate of GLPI $C_Y(L)$ and the one-sided 95% confidence interval for $C_Y(L)$, i.e., $[L_B, \infty)$ are computed and are reported in Table 8. From the obtained result it can be verified that the value of $C_Y^0 = 1$ does not belong to the one-sided confidence interval, thus the null hypothesis H_0 is rejected. Hence, the rejection of the null hypothesis indicates that the GLPI for the considered censored observations meets the required level.

Table 7: *Real data estimates of C_L and $C_Y(L)$ for different censoring schemes where $L = 100$.*

n	m	Schemes	$\hat{C}_{LMLE}$	$\hat{C}_{LBayes}$	$\hat{C}_Y(L)_{MLE}$	$\hat{C}_Y(L)_{Bayes}$
6[0]*19	3[0]*10	0,1*9	1.29976	1.31230	1.06033	1.06053
		1*9,0	1.32879	1.30256	1.07422	1.07421
		0*4,3*3,0*3	1.33256	1.32036	1.08983	1.08686
	3[0]*15	0*11,1*4	1.26546	1.27359	1.05995	1.05736
		1*4,0*11	1.24328	1.25370	1.06700	1.05697
		0*6,1*4,0*5	1.25042	1.23734	1.08076	1.07896

Table 8: *Testing of hypothesis of C_L, $C_Y(L)$ for different censoring schemes based on considered data set.*

n	m	Schemes	$\hat{\theta}$	$\hat{C}(L)_{MLE}$	L'_B	$\hat{V}ar(\hat{C}(L))$	$\hat{C}_Y(L)_{MLE}$	L_B	$\hat{V}ar(\hat{C}_Y(L))$
6[0]*19	3[0]*10	0,1*9	0.001602602	1.29976	1.145873	0.00032479	1.06033	1.06053	0.00012319
		1*9,0	0.001193257	1.32879	1.267304	0.00025823	1.07422	1.07421	0.00004212
		0*4,3*3,0*3	0.001438218	1.33256	1.192645	0.00025178	1.08983	1.08686	0.00008373
	3[0]*15	0*11,1*4	0.002093421	1.26546	1.238082	0.0000432	1.05995	1.05736	0.00030325
		1*4,0*11	0.002393421	1.24328	1.253402	0.00089531	1.06700	1.05697	0.00024422
		0*6,1*4,0*5	0.002293421	1.25042	1.210789	0.000045286	1.08076	1.07896	0.00030897

6. Conclusions

The present article considers the problem of estimating the GLPI, introduced by [32], under progressive type-II right censored sample where the lower specification limit is given for the Lindley distributed products. The model parameter and the GLPI are obtained by the MLE and Bayes estimation methods, respectively. A comparison study has been carried out through the Monte Carlo simulation study under a progressive type-II censoring scheme between MLE and the Bayes estimate of GLPI in their corresponding MSEs. A real data set is analyzed to study the performance of the proposed index. Though the approach of classical estimation and Bayes estimation are different in direction, assuming the gamma prior and using MCMC method, we discussed the Bayes estimation of $C_Y(L)$. The Bayes estimate of the performance index is relatively better than MLE in terms of corresponding MSEs. The proposed procedure can be extended to obtain the confidence interval of $C_Y(L)$ based on MLE and Bayes estimate to evaluate whether the product quality meets the required level. In our upcoming course of work, the problem will be attempted. We may use the MLE and the Bayes estimate of $C_Y(L)$ based on the progressive type-II right censored sample to draw conclusions about additional lifetime distributions in future.

Declarations
Disclosure of Conflicts of interest/competing interests: The authors declare that they have no conflict of interest.
Authors contributions: Each author has equal contribution. All authors jointly write, review and edit the manuscript.
Funding: The authors received no specific funding for this study.
Data Availability Statements: All cited data analysed in the article are included in References. Data sets are also provided in the article.
Ethical Approval: This article does not contain any studies with human participants performed by any of the authors.
Code availability: Codes are available on request.

References

[1] Ahmed, S., Alatefi, M., Alkahtani, M. and Anwar, S. (2018): Bibliometric analysis for process capability research. *Quality Technology and Quantitative Management*, 459-477.
[2] Ahmadi, M. V., Doostparast, M. and Ahmadi, J. (2013): Estimating the lifetime performance index with Weibull distribution based on progressive first-failure censoring scheme. *Journal of Computational and Applied Mathematics*, 239, 93-102.
[3] Ahmadi, M. V., Doostparast, M. and Ahmadi, J. (2015): Statistical inference for the lifetime performance index based on generalised order statistics from exponential distribution. *International Journal of Systems Science*, 46 (6), 1094-107.
[4] Ahmadi, M. V., Ahmadi, J. and Abdi, M. (2019): Evaluating the lifetime performance index of products based on generalized order statistics from two-parameter exponential model. *International Journal of System Assurance Engineering and Management*, 10 (2), 251“75.
[5] Ahmadi, M. V., and Doostparast, M. (2021): Evaluating the lifetime performance index of products based on progressively Type-II censored Pareto samples: A new bayesian approach. *Quality and Reliability Engineering International*, doi:10.1002/qre.3040.
[6] Anderson, D. R., Sweeney, D. J. and Williams, T. A. (1990): Statistics for Business and Economics. *Saint Paul*, Minnesota, West Publishing Company.
[7] Balakrishnan, N. and Agarwala, R. (2000): Progressive censoring theory, Methods and applications, Birkhauser, Boston.
[8] Balakrishnan, N. and Sandhu, R. A. (1996): Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive type II censored samples. *Sankhya*, 58, 1-9.

[9] Chen, H. T., Tong, L. I. and Chen, K. S. (2002): Assessing the lifetime performance of electronic components by confidence interval. *Journal of the Chinese Institute of Industrial Engineers*, 19, 53-60.

[10] Cohen, A. C. (1963): Progressively censored samples in life testing. *Technometrics*, 5, 327-329.

[11] Dey, S., Sharma, V. K., Anis, A. Z. and Yadav, B. (2017): Assessing lifetime performance index of Weibull distributed products using progressive type II right censored samples. *International Journal of System Assurance Engineering and Management*, 8(2), 318-333.

[12] Dey, S., M. Saha, and S. Kumar. (2021a): Parametric confidence intervals of S_{pmk} for generalized exponential distribution. *American Journal of Mathematical and Management Sciences*, 1-22.

[13] Dey, S., C. Zhang, and M. Saha. (2021b): Classical and Bayesian estimation of the index C_{pmk} and its confidence intervals for normally distributed quality characteristic. *Journal of Statistical Computation and Simulation*, 91(10), 1911“34.

[14] Epstein, B. and M. Sobel, M. (1953): Life- Testing. *Journal of American Statistical Association*, 48, 486-502.

[15] Fernandez, A. Z. (2004): On estimating exponential parameters with general type II progressive censoring. *Journal of Statistical Planning and Inference*, 121, 135-147.

[16] Giri, B. C., Jalan A. K. and Chaudhuri, K. S. (2000): Economic order quantity model with Weibull deterioration distribution, shortage and ramp-type demand. *International Journal of Systems Science*, 34, 237-243.

[17] Gupta, R., Kishan, R. and Kumar, P. A. (1999): Two non-identical unit parallel system with correlated lifetimes. *International Journal of Systems Science*, 30, 1123-1129.

[18] Hassanein, W. A. (2018): Statistical inference of the lifetime performance index for Lindley distribution under progressive first failure censoring scheme applied to HPLC data. *International Journal of Biomathematics*, 11(5), doi: 10.1142/S1793524518500730.

[19] Hong, C. W., Lee, W. C. and Wu, J.W. (2012): Computational procedure of performance assessment of lifetime index of products for the Weibull distribution with the progressive first-failure-censored sampling plan. *Journal of Applied Mathematics*, doi:10.1155/2012/717184.

[20] Hong, C. W., Wu, J. W., Cheng, C. H. (2007): Computational procedure of performance assessment of lifetime index of businesses for the pareto lifetime model with the right type II censored sample. *Applied Mathematics and Computation* , 184, 336-350.

[21] Hong, C. W., Wu, J.W. and Cheng, C. H. (2008): Computational procedure of performance assessment of lifetime index of pareto lifetime businesses based on confidence interval. *Applied Soft Computing*, 8, 698-705.

[22] Ihaka, R. and Gentleman, R (1996): R: A Language for Data analysis and Graphics”, *Journal of Computational and Graphical Statistics*, **5**, 299-314, 1996.

[23] Kane, V. E. (1986): Process capability indices. *Journal of Quality Technology*, 18, 41-52.

[24] Keller, G., Warrack, B. and Bartel H. (1994): Statistics for Management and Economics. *Belmont*, CA, Duxbury Press.

[25] Kotz, S. and Johnson, N. L. (2002): Process Capability Indices - A Review, *Journal of Quality Technology*, 34(1), 2-19.

[26] Lawless, J. F. (2003): Statistical models and methods for lifetime data. Wiley, New York.

[27] Lee, W. C., Wu, J. W. and Hong, C. W. (2009): Assessing the lifetime performance index of products with the exponential distribution under progressively type-ii right censored samples. *Journal of Computational and Applied Mathematics*, 231, 648-656.

[28] Lee, W. C., Wu, J. W. and Lei, C. L. (2010): Evaluating the lifetime performance index for the exponential lifetime products. *Applied Mathematical Modelling*, 34, 1217-1224.

[29] Lee, H. M., Wu, J. and Lei, C. (2013a): Assessing the lifetime performance index of exponential products with step-stress accelerated life testing data. *IEEE Transaction on Reliability*, 62, 296-304.

[30] Lee, W. C., Wu, J. W., Hong, C. W. and Hong, S. F. (2013b): Evaluating the lifetime performance index based on the Bayesian estimation for the Rayleigh lifetime products with the upper record values. *Journal of Applied Mathematics*, doi:10.1155/2013/547209.

[31] Lindley, D. V. (1958): Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society*, 20, 102-107.

[32] Maiti, S. S., Saha, M. and Nanda, A. K. (2010): On generalizing process capability indices. *Journal of Quality Technology and Quantitative Management*, 7(3), 279-300.

[33] Maiti, S. S., Bhattarchya, A. and Saha, M. (2021): On generalizing lifetime performance index. *Life Cycle Reliability and Safety Engineering*, 10, 31-38.

[34] Meyer, P. L. (1965): Introductory Probability and Statistical Applications. *Reading*, MA, Addison-Wesley.

[35] Montgomery, D. C. (1985): Introduction to statistical quality control. *New York*, John Wiley and Sons.

[36] Saha, M., Dey, S. and S. Nadarajah, S. (2021a): Parametric inference of the process capability index Cpc for exponentiated exponential distribution. *Journal of Applied Statistics*, 1"25.

[37] Saha, M., S. Dey, S., Yadav, A. S. and S. Ali. (2021b): Confidence intervals of the index Cpk for normally distributed quality characteristics using classical and Bayesian methods of estimation. *Brazilian Journal of Probability and Statistics*, 35 (1), 138-57.

[38] Spiring, F., Leung, B., Cheng, S. and Yeung, A. (2003): A Bibliography of Process Capability Papers, *Quality and Reliability Engineering International*, 19(5), 445-460.

[39] Tong, L. I., Chen, K. S. and Chen, H. T. (2002): Statistical testing for assessing the performance of lifetime index of electronic components with exponential distribution. *International Journal of Quality and Reliability Management*, 19, 812-824.

[40] Wu, J. W., Lin, C., Tan, B. and Lee, W. C. (2000): An EOQ inventory model with time-varying demand and Weibull deterioration with shortages. *International Journal of Systems Science*, 31, 677-683.

[41] Wu, J. W. (2002): EOQ inventory model for items with Weibull distribution deterioration, time-varying demand and partial backlogging. *International Journal of Systems Science*, 33, 323-329.

[42] Wu, S. F. and Hsieh, Y. T. (2019): The assessment on the lifetime performance index of products with Gompertz distribution based on the progressive type I interval censored sample. *Journal of Computational and Applied Mathematics*, 351, 66-76.

[43] Wu, J. W., Lee, H. M. and Lei, C. L. (2007): Computational testing algorithmic procedure of assessment for lifetime performance index of products with two-parameter exponential distribution. *Applied Mathematics and Computation*, 190, 116-125

[44] Wu, S. J., Chen, D. H. and Chen, S. T. (2006): Bayesian inference for Rayleigh distribution under progressive type censored sample. *Applied Stochastic Business and Industry*, 22, 269-278.

[45] Wu, S. F., and Lu, J. Y. (2017): Computational testing algorithmic procedure of assessment for lifetime performance index of Pareto products under progressive type I interval censoring. *Computational Statistics*, 32(2), 647-66.

[46] Wu, S. F. and Chang, W. T. (2021): Power comparison of the testing on the lifetime performance index for Rayleigh lifetime products under progressive type I interval censoring. *Communications in Statistics - Simulation and Computation*, 1"14.

[47] Viveros, R. and Balakrishnan, N. (1994): Interval estimation of parameters of life from progressively censored data. *Technometrics*, 36, 84-91.

[48] Yum, B. J. and Kim, K. W. (2011). A bibliography of the literature on process capability indices: 2000-2009. *Quality and Reliability Engineering International*, 27, 251-268.

[49] Zhang, Y. and Gui, W. (2021). Statistical inference for the lifetime performance index of products with Pareto distribution on basis of general progressive type II censored sample. *Communications in Statistics - Theory and Methods*, 50(16), 3790-808.

[50] Zhu, J., Xin, H., Zheng, C. and Tsai, T. R. (2022). Inference for the process performance index of products on the basis of power-normal distribution. *Mathematics* 10(1), 35.

EXPLORING THE LENGTH BIASED TORNUMONKPE DISTRIBUTION: PROPERTIES, ESTIMATIONS AND PRACTICAL APPLICATIONS

D. Vedavathi Saraja[1], B. Jayakumar[2], Madhulika Mishra[3], Priya Deshpande[4], C. Subramanian[5], Rashid A. Ganaie[6], Aafaq A. Rather[7,*], Bilal Ahmad Bhat[8]

•

[1,5,6]Department of Statistics, Annamalai University, Tamil nadu, India
[2]Department of Statistics, Arignar Anna Govt Arts College, Villupuram, Tamil nadu, India
[3,4,7]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[8]Department of Statistics, University of Kashmir, Srinagar-190006, J&K, India
[1]sarajayoganand@gmail.com, [2]jaiphdau@gmail.com, [3]madhulika1707@gmail.com,
[4]priyadeshpande06@gmail.com, [5]manistat@yahoo.co.in, [6]rashidau7745@gmail.com,
[7]aafaq7741@gmail.com, [8]bilal3819md@gmail.com

Abstract

In this study, we introduce a novel extension of the Tornumonkpe distribution, known as the length biased Tornumonkpe distribution. This distribution holds particular significance as it belongs to the family of weighted distributions, specifically the length biased variant. Through an in-depth analysis, we explore the mathematical and statistical properties of this distribution, shedding light on its unique characteristics. To estimate the model parameters of this new distribution, we employ the well-established technique of maximum likelihood estimation. This allows us to accurately determine the parameters and enhance our understanding of the distribution's behavior. To demonstrate the practical applicability and advantages of the length biased Tornumonkpe distribution, we showcase its performance using a real-life time data set. Through this empirical examination, we investigate the distribution's superiority and flexibility, providing valuable insights into its potential use in various domains.

Keywords: Length biased distribution, Tornumonkpe distribution, order statistics, maximum likelihood estimation

1. Introduction

Weighted distributions have emerged as a unifying and powerful tool to address biases in unequally weighted sample data, providing a comprehensive approach for modeling and representing statistical information. This concept was initially suggested by Fisher [2], exploring the influence of ascertainment methods on the distribution of recorded observations. Subsequently, Rao [5] further developed and unified the theory, particularly in situations where standard distributions were inadequate for capturing observations with equal probabilities. The theory of weighted distributions also provides an integrative conceptualization for model stipulation and data representation problems. The weighted distributions are used as a tool in selection of appropriate models for observed data especially when samples are drawn without a proper frame. The weighted distribution reduces to length biased distribution when the weight function considers only the length of units of

interest. The concept of length biased sampling was introduced by Cox [1] and Zelen [12]. The application of length biased distributions has found widespread use in various biomedical areas, including family history analysis, survival analysis, clinical trials, intermediate events, reliability theory, and population studies. In situations where a proper sampling frame is absent, length biased distributions offer an elegant solution by sampling items at a rate proportional to their lengths, thereby granting larger values a higher probability of being sampled. Many studies on length biased distribution has been published, for example; Rather and Subramanian [6], Rather and Subramanian [7], Rather and Ozel [8], Rather and Subramanian [9], Rather et al. [10].

Tornumonkpe distribution is a recently executed one parametric continuous probability distribution proposed by Nwikpe [4]. Its various mathematical and statistical properties such as order statistics, crude and raw moments, moment generating function, hazard rate function, graphs of pdf, cdf and hazard function and Renyi entropy have been discussed. Its parameter has also been estimated by using the maximum likelihood estimation.

2. Length Biased Tornumonkpe (LBT) Distribution

The probability density function of Tornumonkpe distribution is given by

$$f(x;\theta)=\frac{\theta^3}{(\theta^2+2)}\left(x^2+x\theta\right)e^{-\theta x};\ x>0,\ \theta>0 \tag{1}$$

and the cumulative distribution function of Tornumonkpe distribution is given by

$$F(x;\theta)=\left(1-\left(1+\frac{\theta^2x^2+\theta x(\theta^2+2)}{(\theta^2+2)}\right)e^{-\theta x}\right);\quad x>0,\ \theta>0 \tag{2}$$

Suppose the random variable X following non-negative condition with probability density function $f(x)$ .Let $w(x)$ be its non-negative weight function, then the probability density function of weighted random variable X_w is given by

$$f_w(x)=\frac{w(x)f(x)}{E(w(x))},\quad x>0.$$

Where $w(x)$ be the non-negative weight function and $E(w(x))=\int w(x)f(x)dx<\infty$.

Depending upon the various choices of weighted function $w(x)$ obviously when $w(x) = x^c$, resulting distribution is known as weighted distribution. In this paper, we have to study the length biased version of Tornumonkpe distribution called as length biased Tornumonkpe distribution. So, the weight function considered at $w(x) = x$, resulting distribution is called length biased distribution with its probability density function given by

$$f_l(x)=\frac{x\,f(x)}{E(x)} \tag{3}$$

Where $E(x)=\int_0^\infty x\,f(x,\theta)dx$

$$E(x)=\frac{(6+2\theta^2)}{\theta(\theta^2+2)} \tag{4}$$

By substituting equations (1) and (4) in equation (3), we will obtain the probability density function of length biased Tornumonkpe distribution as

$$f_l(x)=\frac{x\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x} \tag{5}$$

and the cumulative distribution function of length biased Tornumonkpe distribution can be obtained as

$$F_l(x) = \int_0^x f_l(x)dx$$

$$F_l(x) = \frac{1}{\left(6+2\theta^2\right)}\left(\theta^4 \int_0^x x^3 e^{-\theta x}dx \; + \; \theta^5 \int_0^x x^2 e^{-\theta x}dx\right) \tag{6}$$

Put $\theta x = t \;\Rightarrow \theta dx = dt \;\Rightarrow dx = \frac{dt}{\theta}$, Also $x = \frac{t}{\theta}$

When $x \to x,\ t \to \theta x$ and When $x \to 0,\ t \to 0$

After the simplification of equation (6), we will obtain the cumulative distribution function of length biased Tornumonkpe distribution as

$$F_l(x) = \frac{1}{\left(6+2\theta^2\right)}\left(\gamma(4,\ \theta x) + \theta^2\gamma(3,\ \theta x)\right) \tag{7}$$

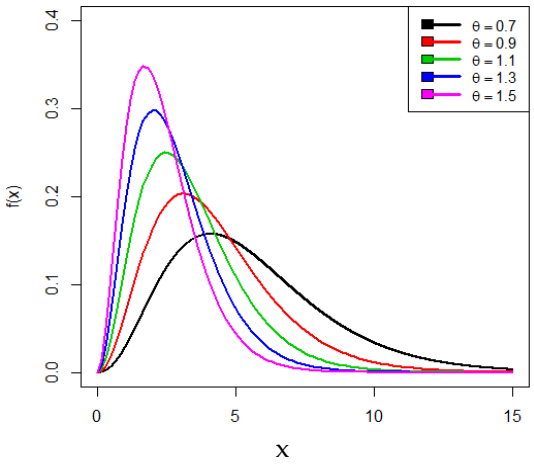


Figure 1: *Pdf plot of LBT distribution*

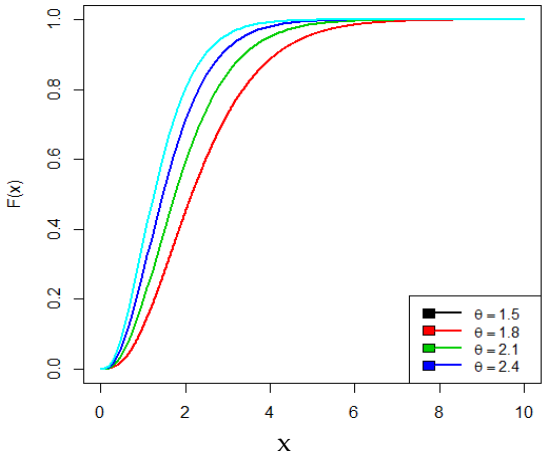


Figure 2: *Cdf plot of LBT distribution*

3. Survival Analysis

In this section, we will discuss about the survival function, hazard rate function, reverse hazard rate function and Mills ratio of proposed length biased Tornumonkpe distribution.

3.1 Survival function

The survival or reliability function of the length biased Tornumonkpe distribution can be obtained as

$$S(x) = 1 - F_l(x)$$

$$S(x) = 1 - \frac{1}{\left(6+2\theta^2\right)}\left(\gamma(4,\ \theta x) + \ \theta^2\gamma(3,\ \theta x)\right) \tag{8}$$

3.2 Hazard function

The hazard function is also known as hazard rate or failure rate or force of mortality and is given by

$$h(x) = \frac{f_l(x)}{1 - F_l(x)}$$

$$h(x) = \frac{x\theta^4 (x^2 + x\theta)\, e^{-\theta x}}{(6 + 2\theta^2) - (\gamma(4,\, \theta x) + \theta^2 \gamma(3,\, \theta x))} \tag{9}$$

3.3 Reverse hazard function

The reverse hazard rate function is given by

$$h_r(x) = \frac{f_l(x)}{F_l(x)}$$

$$h_r(x) = \frac{x\theta^4 (x^2 + x\theta)\, e^{-\theta x}}{(\gamma(4,\, \theta x) + \theta^2 \gamma(3,\, \theta x))} \tag{10}$$

3.4 Mills Ratio

$$M.R = \frac{1}{h_r(x)} = \frac{(\gamma(4,\, \theta x) + \theta^2 \gamma(3,\, \theta x))}{x\theta^4 (x^2 + x\theta)\, e^{-\theta x}} \tag{11}$$

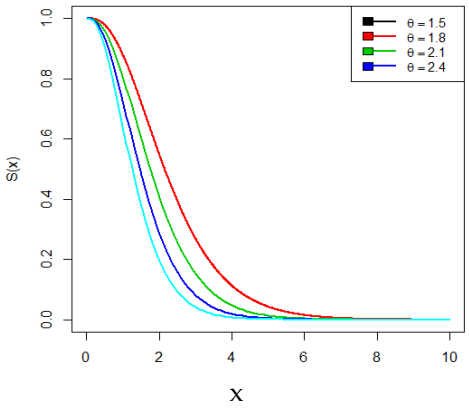


Fig. 3: Survival plot of LBT distribution

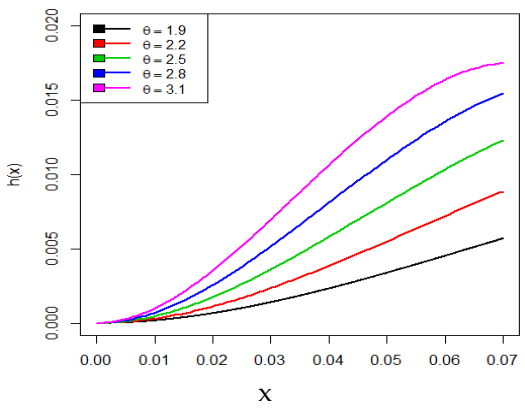


Fig. 4: Hazard plot of LBT distribution

4. Statistical Properties

In this section, we will discuss various statistical properties of length biased Tornumonkpe distribution those include moments, harmonic mean, moment generating function and characteristic function.

4.1 Moments

Let X be the random variable following length biased Tornumonkpe distribution with parameter θ, then the r^{th} order moment $E(X^r)$ of proposed distribution can be obtained as

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_l(x)dx$$

$$E(X^r) = \mu_r' = \int_0^\infty x^r \frac{x\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}dx \tag{12}$$

$$E(X^r) = \mu_r' = \int_0^\infty \frac{x^{r+1}\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}dx \tag{13}$$

$$E(X^r) = \mu_r' = \frac{\theta^4}{\left(6+2\theta^2\right)}\int_0^\infty x^{r+1}\left(x^2+x\theta\right)e^{-\theta x}dx \tag{14}$$

$$E(X^r) = \mu_r' = \frac{\theta^4}{\left(6+2\theta^2\right)}\left(\int_0^\infty x^{(r+4)-1}e^{-\theta x}dx + \theta\int_0^\infty x^{(r+3)-1}e^{-\theta x}dx\right) \tag{15}$$

After the simplification of equation (15), we obtain

$$E(X^r) = \mu_r' = \frac{\Gamma(r+4)+\theta^2\Gamma(r+3)}{\theta^r(6+2\theta^2)} \tag{16}$$

Now putting r = 1, 2, 3 and 4 in equation (16), we will obtain the first four moments of length biased Tornumonkpe distribution as

$$E(X) = \mu_1' = \frac{24+6\theta^2}{\theta(6+2\theta^2)} \tag{17}$$

$$E(X^2) = \mu_2' = \frac{120+24\theta^2}{\theta^2(6+2\theta^2)} \tag{18}$$

$$E(X^3) = \mu_3' = \frac{720+120\theta^2}{\theta^3(6+2\theta^2)} \tag{19}$$

$$E(X^4) = \mu_4' = \frac{5040+720\theta^2}{\theta^4(6+2\theta^2)} \tag{20}$$

$$\text{Variance} = \frac{120+24\theta^2}{\theta^2(6+2\theta^2)} - \left(\frac{24+6\theta^2}{\theta(6+2\theta^2)}\right)^2 \tag{21}$$

$$S.D(\sigma) = \sqrt{\left(\frac{120+24\theta^2}{\theta^2(6+2\theta^2)} - \left(\frac{24+6\theta^2}{\theta(6+2\theta^2)}\right)^2\right)} \tag{22}$$

4.2 Harmonic mean

The harmonic mean for proposed length biased Tornumonkpe distribution can be obtained as

$$H.M = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_l(x)dx \tag{23}$$

$$H.M = \int_0^\infty \frac{1}{x} \frac{x\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}dx \tag{24}$$

$$H.M = \int_0^\infty \frac{\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}dx \tag{25}$$

$$H.M = \frac{\theta^4}{\left(6+2\theta^2\right)}\int_0^\infty \left(x^2+x\theta\right)e^{-\theta x}dx \tag{26}$$

$$H.M = \frac{\theta^4}{\left(6+2\theta^2\right)}\left(\int_0^\infty x^{(3)-1}e^{-\theta x}dx \; + \; \theta\int_0^\infty x^{(2)-1}e^{-\theta x}dx\right) \tag{27}$$

After the simplification of equation (27), we obtain

$$H.M = \frac{\theta(2+\theta^2)}{\left(6+2\theta^2\right)} \tag{28}$$

4.3 Moment generating function and characteristic function

Let X be the random variable following length biased Tornumonkpe distribution with parameter θ, then the moment generating function of proposed distribution can be obtained as

$$M_X(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx} f_l(x)dx \tag{29}$$

Using Taylor's series, we obtain

$$M_X(t) = \int_0^\infty \left(1+tx+\frac{(tx)^2}{2!}+\ldots\right) f_l(x)dx \tag{30}$$

$$M_X(t) = \int_0^\infty \sum_{j=0}^\infty \frac{t^j}{j!}x^j f_l(x)dx \tag{31}$$

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!}\mu_j' \tag{32}$$

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!}\left(\frac{\Gamma(j+4)+\theta^2\Gamma(j+3)}{\theta^j(6+2\theta^2)}\right) \tag{33}$$

$$M_X(t) = \frac{1}{(6+2\theta^2)}\sum_{j=0}^\infty \frac{t^j}{j!\theta^j}\left(\Gamma(j+4)+\theta^2\Gamma(j+3)\right) \tag{34}$$

Similarly, the characteristic function of length biased Tornumonkpe distribution can be obtained as

$$\varphi_x(t) = M_X(it)$$

$$M_X(it) = \frac{1}{(6+2\theta^2)} \sum_{j=0}^{\infty} \frac{it^j}{j!\theta^j}\left(\Gamma(j+4) + \theta^2\Gamma(j+3)\right) \tag{35}$$

5. Order Statistics

Consider X(1), X(2),…, X(n) be the order statistics of a random sample X1, X2,…, Xn from a continuous population with probability density function fx (x) and cumulative distribution function FX(x), then the probability density function of rth order statistics X(r) is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x)\left(F_X(x)\right)^{r-1}\left(1-F_X(x)\right)^{n-r} \tag{36}$$

By using the equations (5) and (7) in equation (36), we will obtain the probability density function of rth order statistics of length biased Tornumonkpe distribution as

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\left(\frac{x\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}\right) \times \left(\frac{1}{\left(6+2\theta^2\right)}\left(\gamma(4,\theta x)+\theta^2\gamma(3,\theta x)\right)\right)^{r-1} \times \left(1-\frac{1}{\left(6+2\theta^2\right)}\left(\gamma(4,\theta x)+\theta^2\gamma(3,\theta x)\right)\right)^{n-r} \tag{37}$$

Therefore, the probability density function of higher order statistic X(n) of length biased Tornumonkpe distribution can be obtained as

$$f_{x(n)}(x) = \frac{nx\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}\left(\frac{1}{\left(6+2\theta^2\right)}\left(\gamma(4,\theta x)+\theta^2\gamma(3,\theta x)\right)\right)^{n-1} \tag{38}$$

and the probability density function of first order statistic X(1) of length biased Tornumonkpe distribution can be obtained as

$$f_{x(1)}(x) = \frac{nx\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x} \times \left(1-\frac{1}{\left(6+2\theta^2\right)}\left(\gamma(4,\theta x)+\theta^2\gamma(3,\theta x)\right)\right)^{n-1} \tag{39}$$

6. Likelihood Ratio Test

Let X1, X2,…., Xn be the random sample of size n from length biased Tornumonkpe distribution. To examine its significance, we use the hypothesis.

$$H_o : f(x) = f(x;\theta) \quad \text{against} \quad H_1 : f(x) = f_1(x;\theta)$$

In order to investigate, whether the random sample of size n comes from the Tornumonkpe distribution or length biased Tornumonkpe distribution, the following test statistic is used.

$$\Delta = \frac{L_1}{L_o} = \prod_{i=1}^{n} \frac{f_l(x;\theta)}{f(x;\theta)}$$

$$\Delta = \frac{L_1}{L_o} = \prod_{i=1}^{n} \left(\frac{x_i\, \theta(\theta^2+2)}{\left(6+2\theta^2\right)} \right) \tag{40}$$

$$\Delta = \frac{L_1}{L_o} = \left(\frac{\theta(\theta^2+2)}{\left(6+2\theta^2\right)} \right)^n \prod_{i=1}^{n} x_i \tag{41}$$

We should refuse to accept the null hypothesis, if

$$\Delta = \left(\frac{\theta(\theta^2+2)}{\left(6+2\theta^2\right)} \right)^n \prod_{i=1}^{n} x_i > k \tag{42}$$

Equivalently we should also refuse to retain the null hypothesis, where

$$\Delta^* = \prod_{i=1}^{n} x_i > k \left(\frac{\left(6+2\theta^2\right)}{\theta(\theta^2+2)} \right)^n \tag{43}$$

$$\Delta^* = \prod_{i=1}^{n} x_i > k^*, \text{Where } k^* = k \left(\frac{\left(6+2\theta^2\right)}{\theta(\theta^2+2)} \right)^n \tag{44}$$

For large sample of size *n*, *2log Δ* is distributed as chi-square distribution with one degree of freedom and also *p*-value is obtained by using the chi square distribution. Thus, we should refuse to accept the null hypothesis, when the probability value is given by

$p\left(\Delta^* > \alpha^*\right)$, Where $\alpha^* = \prod_{i=1}^{n} x_i$ is less than a specified level of significance and $\prod_{i=1}^{n} x_i$ is the observed value of the statistic Δ^*.

7. Bonferroni and Lorenz Curves

The bonferroni and Lorenz curves are known as income distribution or classical curves are mostly being used to measure the distribution of inequality in income or poverty. The bonferroni and Lorenz curves are defined as

$$B(p) = \frac{1}{p\mu_1'} \int_0^q x\, f_l(x)dx$$

$$L(p) = pB(p) = \frac{1}{\mu_1'}\int_0^q x\, f_l(x)dx \qquad \text{and} \quad q = F^{-1}(p)$$

Where $\mu_1' = \dfrac{\left(24+6\theta^2\right)}{\theta(6+2\theta^2)}$

$$B(p) = \frac{\theta(6+2\theta^2)}{p\,(24+6\theta^2)}\int_0^q \frac{x^2\,\theta^4}{\left(6+2\theta^2\right)}\left(x^2+x\theta\right)e^{-\theta x}dx \tag{45}$$

$$B(p) = \frac{\theta^5}{p(24+6\theta^2)}\int_0^q x^2\left(x^2+x\theta\right)e^{-\theta x}dx \tag{46}$$

$$B(p) = \frac{\theta^5}{p(24+6\theta^2)}\left(\int_0^q x^{(5)\,-1}e^{-\theta x}dx \;+\; \theta\int_0^q x^{(4)-1}e^{-\theta x}dx\right) \tag{47}$$

After simplification, we obtain

$$B(p) = \frac{\theta^5}{p(24+6\theta^2)}\left(\gamma(5,\,\theta q) + \theta\,\gamma(4,\,\theta q)\right) \tag{48}$$

$$L(p) = \frac{\theta^5}{(24+6\theta^2)}\left(\gamma(5,\,\theta q) + \theta\,\gamma(4,\,\theta q)\right) \tag{49}$$

8. Parameter Estimation

In this section, we will discuss the method of maximum likelihood estimation to estimate the parameters of length biased Tornumonkpe distribution. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from length biased Tornumonkpe distribution, then the likelihood function can be defined as

$$L(x) = \prod_{i=1}^{n} f_l(x)$$

$$L(x) = \prod_{i=1}^{n}\left(\frac{x_i\,\theta^4}{\left(6+2\theta^2\right)}\left(x_i^{\,2}+x_i\,\theta\right)e^{-\theta x_i}\right) \tag{50}$$

$$L(x) = \frac{\theta^{4n}}{(6+2\theta^2)^n}\prod_{i=1}^{n}\left(x_i\left(x_i^{\,2}+x_i\,\theta\right)e^{-\theta x_i}\right) \tag{51}$$

The log likelihood function is given by

$$\log L = 4n\log\theta - n\log(6+2\theta^2) + \sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log\left(x_i^{\,2}+x_i\,\theta\right) - \theta\sum_{i=1}^{n}x_i \tag{52}$$

Now differentiating the log likelihood equation (52) with respect to parameter θ. We must satisfy the following normal equation

$$\frac{\partial \log L}{\partial\theta} = \frac{4n}{\theta} - n\left(\frac{4\theta}{6+2\theta^2}\right) + \sum_{i=1}^{n}\left(\frac{x_i}{\left(x_i^{\,2}+x_i\,\theta\right)}\right) - \sum_{i\,=\,1}^{n}x_i = 0 \tag{53}$$

The above likelihood equation is too complicated to solve it algebraically. Therefore, we use numerical technique like Newton-Raphson method for estimating the required parameter of proposed distribution.

In order to use the asymptotic normality results for obtaining the confidence interval. We have that if $(\hat{\beta} = \hat{\theta})$ denotes the MLE *of* $(\beta = \theta)$. We can state the results as

$$\sqrt{n}(\hat{\beta} - \beta) \to N(0,\ I^{-1}(\beta))$$

Where $I^{-1}(\beta)$ is Fisher' s information matrix i.e.,

$$I(\beta) = -\frac{1}{n}\left(E\left(\frac{\partial^2 \log L}{\partial \theta^2} \right) \right) \tag{54}$$

Here, we define

$$E\left(\frac{\partial^2 \log L}{\partial \theta^2} \right) = -\frac{4n}{\theta^2} - n\left(\frac{(4)(6+2\theta^2) - (16\theta^2)}{(6+2\theta^2)^2} \right) - \sum_{i=1}^{n} \left(\frac{E\left(x_i\right)^2}{(x_i^{\,2} + x_i\,\theta)^2} \right) \tag{55}$$

Since β being unknown, we estimate $I^{-1}(\beta)$ by $I^{-1}(\hat{\beta})$ and this can be used to obtain asymptotic confidenc interval for θ.

9. Applications

In this section, we have fitted a real life data set in length biased Tornumonkpe distribution in order to show that length biased Tornumonkpe distribution fits better over Tornumonkpe, exponential and Lindley distributions. The real life data set is given below as

The following data set represents the failure times in hours of an accelerated life test of 59 conductors without any censored observation is obtained by Lawless [3] and the data set is given below in table 1.

Table 1: *Data regarding the failure times in hours*

2.997	4.137	4.288	4.531	4.700	4.706	5.009	5.381	5.434	5.459	5.589	5.640
5.807	5.923	6.033	6.071	6.087	6.129	6.352	6.369	6.476	6.492	6.515	6.522
6.538	6.545	6.573	6.725	6.869	6.923	6.948	6.956	6.958	7.024	7.224	7.365
7.398	7.459	7.489	7.495	7.496	7.543	7.683	7.937	7.945	7.974	8.120	8.336
8.532	8.591	8.687	8.799	9.218	9.254	9.289	9.663	10.092	10.491	11.038	

To compute the model comparison criterion values along with the estimation of unknown parameters, the technique of R software is used. In order to compare the performance of length biased Tornumonkpe distribution over Tornumonkpe, exponential and Lindley distributions, we use the criterion values such as *AIC* (Akaike Information Criterion), *BIC* (Bayesian Information Criterion), *AICC* (Akaike Information Criterion Corrected) and *-2logL*. The better distribution is which corresponds to the lesser values of *AIC*, *BIC*, *AICC* and *-2logL*. For calculating the criterion values like *AIC*, *BIC*, *AICC* and *-2logL* following formulas are used.

$$AIC = 2k - 2\log L, \qquad BIC = k\log n - 2\log L \qquad \text{and} \qquad AICC = AIC + \frac{2k(k+1)}{n-k-1}$$

Where n is the sample size, k is the number of parameters in statistical model and $-2\log L$ is the maximized value of log-likelihood function under the considered model.

Table 2: *Shows Comparison and Performance of fitted distributions*

Distributions	**MLE**	**S.E**	**-2logL**	**AIC**	**BIC**	**AICC**
Length Biased Tornumonkpe	$\hat{\theta} = 0.55730745$	$\hat{\theta} = 0.03557811$	270.4884	272.4884	274.5659	272.5585
Tornumonkpe	$\hat{\theta} = 0.41550194$	$\hat{\theta} = 0.03053944$	285.505	287.505	289.5825	287.5751
Exponential	$\hat{\theta} = 0.14326745$	$\hat{\theta} = 0.01865093$	347.2809	349.2809	351.3585	349.3510
Lindley	$\hat{\theta} = 0.25722036$	$\hat{\theta} = 0.02393044$	316.7054	318.7054	320.783	318.7755

From results given above in table 2, it has been clearly realized and observed that the length biased Tornumonkpe distribution has lesser *AIC*, *BIC*, *AICC* and *-2logL* values as compared to the Tornumonkpe, exponential and Lindley distributions. Hence, it can be revealed that length biased Tornumonkpe distribution leads to a better fit over Tornumonkpe, exponential and Lindley distributions.

10. Conclusion

In this article, we have introduced the length biased Tornumonkpe distribution as a novel approach within the field. By applying the length biased technique to its baseline distribution, we have generated a new distribution that exhibits unique properties and characteristics. Throughout our analysis, we have explored and derived various structural properties of the length biased Tornumonkpe distribution. These properties include moments, the shape of the probability density function (PDF) and cumulative distribution function (CDF), mean and variance, harmonic mean, survival function, hazard rate function, moment generating function, reverse hazard rate function, order statistics, Bonferroni and Lorenz curves. Through a comprehensive study of these properties, we have gained valuable insights into the behavior and statistical properties of the length biased Tornumonkpe distribution. To estimate the parameters of the length biased Tornumonkpe distribution, we employed the widely used technique of maximum likelihood estimation. This allowed us to obtain reliable estimates and enhance our understanding of the distribution's characteristics.

Furthermore, we have validated the practical applicability and superiority of the length biased Tornumonkpe distribution through its examination with real-life data sets. By comparing its fit with other well-known distributions such as the Tornumonkpe, exponential, and Lindley distributions, we have demonstrated that the length biased Tornumonkpe distribution outperforms these alternatives in terms of goodness-of-fit.

Finally, the length biased Tornumonkpe distribution represents a significant advancement in statistical modeling. Its distinctive properties, derived through rigorous analysis, and its superior fit with real-life data sets emphasize its potential for various applications. Researchers and practitioners can benefit from incorporating the length biased Tornumonkpe distribution into their analyses, enabling more accurate and robust modeling in diverse fields.

References

[1] Cox, D. R. (1969). Some sampling problems in technology, In New Development in Survey Sampling, Johnson, N. L. and Smith, H., Jr .(eds.) *New York Wiley- Interscience*, 506-527.

[2] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13-25.

[3] Lawless, J. F. (2003). Statistical Models and methods for Lifetime data, 2nd ed., *John Wiley and Sons*, New Jersey.

[4] Nwikpe, B. J. (2021). Tornumonkpe distribution: statistical properties and goodness of fit. *Asian Journal of Probability and Statistics*, 14(1), 12-22.

[5] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; *Pergamum Press and Statistical Publishing Society*, Calcutta. 320-332.

[6] Rather, A. A. & Subramanian C. (2018). Characterization and estimation of length biased weighted generalized uniform distribution, *International Journal of Scientific Research in Mathematical and Statistical Science*, vol.5, issue. 5, pp. 72-76.

[7] Rather, A. A. & Subramanian C. (2018). Length Biased Sushila distribution, *Universal Review*, vol. 7, issue. XII, pp. 1010-1023.

[8] Rather, A. A. & Ozel G. (2021). A new length-biased power Lindley distribution with properties and its applications, *Journal of Statistics and Management Systems*, DOI: 10.1080/09720510.2021.1920665

[9] Rather, A. A. & Subramanian, C. (2019), The Length-Biased Erlang–Truncated Exponential Distribution with Life Time Data, *Journal of Information and Computational Science*, vol-9, Issue 8, pp 340-355.

[10] Rather, A. A., Subramanian, C., Shafi, S., Malik, K. A., Ahmad, P. J., Para, B. A., and Jan, T. R. (2018). A new Size Biased Distribution with applications in Engineering and Medical Science, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, Vol.5, Issue.4, pp.75-85.

[11] Zelen, M. (1974). Problems in cell kinetic and the early detection of disease, in Reliability and Biometry, F. Proschan & R.J. Sering, eds, *SIAM, Philadelphia*, 701-706.

ENHANCING ENGINEERING SCIENCES WITH UMA DISTRIBUTION: A PERFECT FIT AND VALUABLE CONTRIBUTIONS

R. A. Ganaie[1], C. Subramanian[2], V. P. Soumya[3], R. Shenbagaraja[4], Mahfooz Alam[5], D. Vedavathi Saraja[6], Rushika Kinjawadekar[7], Aafaq A. Rather[8,*], Showkat A. Dar[9]

•

[1,2, 3 ,6]Department of Statistics, Annamalai University, Tamil nadu, India
[4] Department of Statistics, Dr. M.G.R. Govt. Arts and Science College for Women, Tamil Nadu, India
[5]Department of Statistics, Faculty of Science and Technology, Vishwakarma University, Pune, India
[7]Marathwada Mitra Mandal, Pune, India
[8,*]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[9]Department of Computer Science and Engineering, GITAM University Bangalore Campus, India
[1]rashidau7745@gmail.com, [2]manistat@yahoo.co.in, [3]soumyaponnan@gmail.com, [4]statraja@gmail.com, [5]mahfooz.alam@vupune.ac.in, [6]sarajayoganand@gmail.com, [7]rushika.kinjawadekar@vupune.ac.in, [8,*]aafaq7741@gmail.com, [9]showkatme2009@gmail.com

Abstract

In this study, we introduce a novel class of distributions called the length biased Uma distribution. This distribution is a specific instance of the broader weighted distribution family, known for its versatility in various applications. We explore the structural properties of the length biased Uma distribution and propose a robust parameter estimation technique based on maximum likelihood estimation. To assess its efficacy, we apply the newly introduced distribution to two real-world datasets, evaluating its flexibility and performance in comparison to existing models. The results obtained demonstrate the potential of the length biased Uma distribution as a valuable addition to the repertoire of statistical distributions, offering valuable insights for a wide range of practical applications.

Keywords: Length biased distribution, Uma distribution, Reliability analysis, Order statistics, Maximum likelihood estimation.

1. Introduction

In probability distribution theory, the concept of weighted probability models have attained a great importance for modeling different lifetime data sets occurring from various practical and applied fields like engineering, medical sciences, insurance, finance etc. There are several situations were classical distributions may not provide best fit to lifetime data. In such situation attempt has been made to generalize standard distribution by introducing an extra parameter to it. This extra parameter can be introduced through various techniques. One of such technique is of weighted technique. Fisher [7] introduced the concept of weighted distribution to model the ascertainment bias which was later formalized by Rao [16] in a unifying theory for problems were the observations fall in non-experimental, non-replicated and non-random categories. The weighted distributions are remarkable

for efficient modeling of statistical data and prediction obviously when classical distributions are not suitable. The weighted distributions were formulated in such a situation to record the observation according to some weight function. The weighted distribution reduces to length biased distribution when the weight function considers only the length of units of interest. The probability of selecting an individual in a population is proportional to its magnitude is called length biased sampling. However, when observations are selected with probability proportional to their length, resulting distribution is called length biased distribution. The concept of length biased distribution was introduced by Cox [5] in renewal theory. Length biased sampling situation occurs were a proper sampling frame is absent. In such cases items are sampled at a rate proportional to their lengths so that the larger values could be sampled with higher probability.

A significant and remarkable contribution has been done by several authors to develop some important length biased probability models along with their applications in various fields. Oluwafemi and Olalekan [15] discussed on length and area biased exponentiated weibull distribution based on forest inventories. Ekhosuehil et al. [6] proposed the weibull length biased exponential distribution with statistical properties and applications. Atikankul et al. [2] discussed on the length biased weighted Lindley distribution with applications. Ratnaparkhi and Nimbalkar [18] presented the length biased lognormal distribution and its application in the analysis of data from oil field exploration studies. Mathew [14] proposed the reliability test plan for the Marshall Olkin length biased Lomax distribution. Mustafa and Khan [13] obtained the length biased power hazard rate distribution with its properties and applications. Chaito et al. [3] discussed on the length biased Gamma-Rayleigh distribution with applications. Al-omari and Alanzi [1] introduced the inverse length biased Maxwell distribution with statistical inference and application. Ghorbal [11] discussed on properties of length biased exponential model with applications. Ganaie and Rajagopalan [9] presented the weighted power Garima distribution with applications in blood cancer and relief times. Recently, Chaito and Khamkong [4] discussed on length-biased weibull-Rayleigh distribution for application to hydrological data.

Uma distribution is a recently introduced one parametric continuous lifetime distribution proposed by Shanker [19]. Its various statistical properties like moments, mean residual life function, hazard rate function, reverse hazard function, stochastic ordering, coefficient of variation, skewness, kurtosis and index of dispersion have been discussed. For estimating its parameter the method of maximum likelihood estimation has been discussed.

2. Length Biased Uma (LBU) Distribution

The probability density function of Uma distribution is given by

$$f(x;\theta)=\frac{\theta^4}{\theta^3+\theta^2+6}\left(1+x+x^3\right)e^{-\theta x};\ x>0,\ \theta>0 \tag{1}$$

and the cumulative distribution function of Uma distribution is given by

$$F(x;\theta)=1-\left(1+\frac{\theta x\left(\theta^2x^2+3\theta x+\theta^2+6\right)}{\theta^3+\theta^2+6}\right)e^{-\theta x};\quad x>0,\ \theta>0 \tag{2}$$

Consider X be the non-negative random variable with probability density function $f(x)$ .Let its non-negative weight function be $w(x)$, then the probability density function of weighted random variable X_w is given by

$$f_w(x) = \frac{w(x) f(x)}{E(w(x))}, \quad x > 0.$$

Where $w(x)$ be the non-negative weight function and $E(w(x)) = \int w(x) f(x) dx < \infty$.

For various forms of weight function $w(x)$ obviously when $w(x) = x^c$, resulting distribution is called weighted distribution. In this paper, we have considered the weight function at $w(x) = x$ to obtain the length biased version of Uma distribution and its probability density function is given by

$$f_l(x) = \frac{x f(x)}{E(x)} \tag{3}$$

Where $E(x) = \int_0^\infty x f(x, \theta) dx$

$$E(x) = \frac{\left(\theta^3 + 2\theta^2 + 24\right)}{\theta(\theta^3 + \theta^2 + 6)} \tag{4}$$

By using the equations (1) and (4) in equation (3), we will obtain the probability density function of length biased Uma distribution as

$$f_l(x) = \frac{x\theta^5}{(\theta^3 + 2\theta^2 + 24)}\left(1 + x + x^3\right)e^{-\theta x} \tag{5}$$

and the cumulative distribution function of length biased Uma distribution can be obtained as

$$F_l(x) = \int_0^x f_l(x) dx$$

$$F_l(x) = \int_0^x \frac{x\theta^5}{(\theta^3 + 2\theta^2 + 24)}\left(1 + x + x^3\right)e^{-\theta x} dx$$

$$F_l(x) = \frac{1}{(\theta^3 + 2\theta^2 + 24)} \int_0^x x\theta^5 \left(1 + x + x^3\right)e^{-\theta x} dx$$

$$F_l(x) = \frac{1}{(\theta^3 + 2\theta^2 + 24)}\left(\theta^5 \int_0^x x e^{-\theta x} dx + \theta^5 \int_0^x x^2 e^{-\theta x} dx + \theta^5 \int_0^x x^4 e^{-\theta x} dx\right) \tag{6}$$

Put $\theta x = t \Rightarrow \theta dx = dt \Rightarrow dx = \frac{dt}{\theta}$, Also $x = \frac{t}{\theta}$

When $x \to x,\ t \to \theta x$ and When $x \to 0,\ t \to 0$

After the simplification of equation (6), we will obtain the cumulative distribution function of length biased Uma distribution as

$$F_l(x) = \frac{1}{(\theta^3 + 2\theta^2 + 24)}\left(\theta^3 \gamma(2, \theta x) + \theta^2 \gamma(3, \theta x) + \gamma(5, \theta x)\right) \tag{7}$$

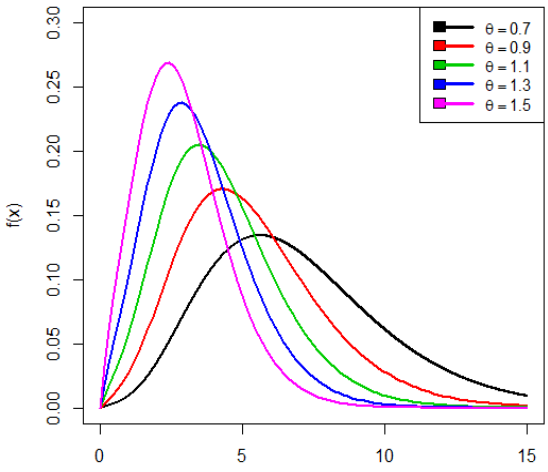


Fig. 1: *Pdf plot of LBU distribution*

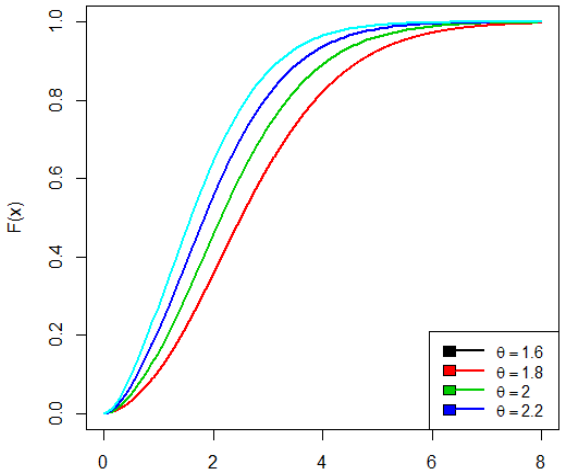


Fig. 2: *Cdf plot of LBU distribution*

3. Reliability Analysis

In this section, we will discuss about the reliability function, hazard rate function, reverse hazard rate function and Mills ratio of the proposed length biased Uma distribution.

3.1 Reliability function

The reliability function is termed as survival function and the reliability function of executed length biased Uma distribution can be determined as

$$R(x) = 1 - F_l(x)$$

$$R(x) = 1 - \frac{1}{(\theta^3 + 2\theta^2 + 24)}\left(\theta^3\gamma(2,\ \theta x) + \ \theta^2\gamma(3,\ \theta x) + \gamma(5,\ \theta x)\right) \tag{8}$$

3.2 Hazard function

The hazard function is also known as failure rate or force of mortality and is given by

$$h(x) = \frac{f_l(x)}{1 - F_l(x)}$$

$$h(x) = \frac{x\theta^5(1 + x + x^3)\ e^{-\theta x}}{(\theta^3 + 2\theta^2 + 24) - (\theta^3\gamma(2,\ \theta x) + \theta^2\gamma(3,\ \theta x) + \gamma(5,\ \theta x))} \tag{9}$$

3.3 Reverse hazard function

The reverse hazard rate function is given by

$$h_r(x) = \frac{f_l(x)}{F_l(x)}$$

$$h_r(x) = \frac{x\theta^5(1+x+x^3)\ e^{-\theta x}}{(\theta^3\gamma(2,\ \theta x)+\theta^2\gamma(3,\ \theta x)+\gamma(5,\ \theta x))} \tag{10}$$

3.4 Mills Ratio

The Mills Ratio of proposed model is given by

$$M.R = \frac{1}{h_r(x)} = \frac{(\theta^3\gamma(2,\ \theta x)+\theta^2\gamma(3,\ \theta x)+\gamma(5,\ \theta x))}{x\theta^5(1+x+x^3)\ e^{-\theta x}} \tag{11}$$

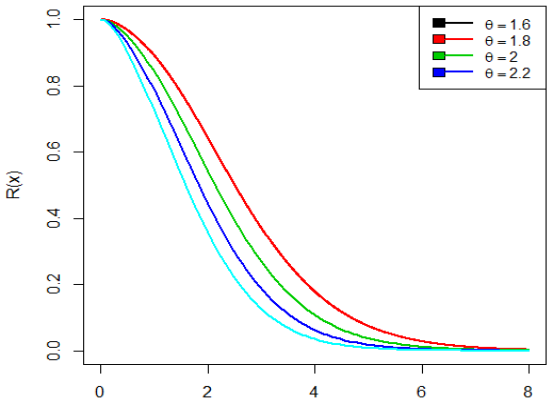


Fig. 3: *Reliability plot of LBU distribution*

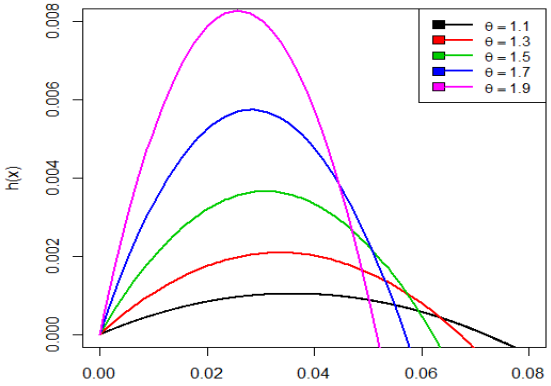


Fig. 4: *Hazard rate plot of LBU distribution*

4. Order Statistics

Suppose X(1), X(2),…, X(n) be the order statistics of a random sample X1, X2,…, Xn from a continuous population with cumulative distribution function FX(x) and probability density function fx (x) then the probability density function of rth order statistics X(r) is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\ f_X(x)\left(F_X(x)\right)^{r-1}\left(1-F_X(x)\right)^{n-r} \tag{12}$$

Now substituting the equations (5) and (7) in equation (12), we will obtain the probability density function of r^{th} order statistics $X_{(r)}$ of length biased Uma distribution as

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\left(\frac{x\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}\right)\left(\frac{1}{(\theta^3+2\theta^2+24)}\left(\theta^3\gamma(2,\ \theta x)+\theta^2\gamma(3,\ \theta x)+\gamma(5,\theta x)\right)\right)^{r-1}$$
$$\times\left(1-\frac{1}{(\theta^3+2\theta^2+24)}\left(\theta^3\gamma(2,\ \theta x)+\theta^2\gamma(3,\ \theta x)+\gamma(5,\ \theta x)\right)\right)^{n-r}$$

Therefore, the probability density function of higher order statistic $X_{(n)}$ of length biased Uma distribution can be determined as

$$f_{x(n)}(x)=\frac{n\,x\,\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}\left(\frac{1}{(\theta^3+2\theta^2+24)}\left(\theta^3\gamma(2,\theta x)+\theta^2\gamma(3,\theta x)+\gamma(5,\theta x)\right)\right)^{n-1}$$

and probability density function of first order statistic $X_{(1)}$ of length biased Uma distribution can be determined as

$$f_{x(1)}(x)=\frac{n\,x\,\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}\left(1-\frac{1}{(\theta^3+2\theta^2+24)}\left(\theta^3\gamma(2,\theta x)+\theta^2\gamma(3,\theta x)+\gamma(5,\theta x)\right)\right)^{n-1}$$

5. Test for Length biasedness of Length biased Uma distribution

Consider X1, X2,…., Xn be the random sample of size n from length biased Uma distribution. To analyse its flexibility consider the hypothesis.

$$H_o: f(x)=f(x;\theta) \quad \text{against} \quad H_1: f(x)=f_l(x;\theta)$$

In order to determine, whether the random sample of size n comes from Uma distribution or length biased Uma distribution, the following test statistic rule is employed.

$$\Delta=\frac{L_1}{L_o}=\prod_{i=1}^{n}\frac{f_l(x;\theta)}{f(x;\theta)} \tag{13}$$

$$\Delta=\frac{L_1}{L_o}=\prod_{i=1}^{n}\left(\frac{x_i\,\theta(\theta^3+\theta^2+6)}{(\theta^3+2\theta^2+24)}\right) \tag{14}$$

$$\Delta=\frac{L_1}{L_o}=\left(\frac{\theta(\theta^3+\theta^2+6)}{(\theta^3+2\theta^2+24)}\right)^n\prod_{i=1}^{n}x_i \tag{15}$$

We should refuse to accept the null hypothesis, if

$$\Delta=\left(\frac{\theta(\theta^3+\theta^2+6)}{(\theta^3+2\theta^2+24)}\right)^n\prod_{i=1}^{n}x_i>k \tag{16}$$

Equivalently, we should also refrain to retain the null hypothesis where

$$\Delta^*=\prod_{i=1}^{n}x_i>k\left(\frac{(\theta^3+2\theta^2+24)}{\theta(\theta^3+\theta^2+6)}\right)^n \tag{17}$$

$$\Delta^*=\prod_{i=1}^{n}x_i>k^*,\ \text{Where}\ k^*=k\left(\frac{(\theta^3+2\theta^2+24)}{\theta(\theta^3+\theta^2+6)}\right)^n$$

Whether, the $2log\,\Delta$ is distributed as chi-square distribution with one degree of freedom if the sample is large of size n and also p-value is determined by using chi-square distribution. Thus, we should refuse to accept the null hypothesis, if the probability value is given by

$p\left(\Delta^*>\gamma^*\right)$, Where $\gamma^*=\prod_{i=1}^{n}x_i$ is less than a specified level of significance and $\prod_{i=1}^{n}x_i$ is the experimental value of the statistic Δ^*.

6. Statistical properties

In this section, we will derive about the different structural properties of length biased Uma distribution those include moments, harmonic mean, moment generating function and characteristic function.

6.1 Moments

Let X be the random variable following length biased Uma distribution with parameter θ, then the r^{th} order moment $E(X^r)$ of proposed distribution can be obtained as

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_l(x)dx \tag{18}$$

$$E(X^r) = \mu_r' = \int_0^\infty x^r \frac{x\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{19}$$

$$E(X^r) = \mu_r' = \int_0^\infty \frac{x^{r+1}\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{20}$$

$$E(X^r) = \mu_r' = \frac{\theta^5}{(\theta^3+2\theta^2+24)}\int_0^\infty x^{r+1}\left(1+x+x^3\right)e^{-\theta x}dx \tag{21}$$

$$E(X^r) = \mu_r' = \frac{\theta^5}{(\theta^3+2\theta^2+24)}\left(\begin{array}{l}\int_0^\infty x^{(r+2)-1}e^{-\theta x}dx + \int_0^\infty x^{(r+3)-1}e^{-\theta x}dx \\ +\int_0^\infty x^{(r+5)-1}e^{-\theta x}dx\end{array}\right) \tag{22}$$

After simplification, we obtain from equation (22)

$$E(X^r) = \mu_r' = \frac{\theta^3\Gamma(r+2)+\theta^2\Gamma(r+3)+\Gamma(r+5)}{\theta^r(\theta^3+2\theta^2+24)} \tag{23}$$

Now substituting r = 1, 2, 3 and 4 in equation (23), we will obtain the first four moments of length biased Uma distribution as

$$E(X) = \mu_1' = \frac{2\theta^3+6\theta^2+120}{\theta(\theta^3+2\theta^2+24)} \tag{24}$$

$$E(X^2) = \mu_2' = \frac{6\theta^3+24\theta^2+720}{\theta^2(\theta^3+2\theta^2+24)} \tag{25}$$

$$E(X^3) = \mu_3' = \frac{24\theta^3+120\theta^2+5040}{\theta^3(\theta^3+2\theta^2+24)} \tag{26}$$

$$E(X^4) = \mu_4' = \frac{120\theta^3+720\theta^2+40320}{\theta^4(\theta^3+2\theta^2+24)} \tag{27}$$

$$\text{Variance} = \frac{6\theta^3+24\theta^2+720}{\theta^2(\theta^3+2\theta^2+24)} - \left(\frac{2\theta^3+6\theta^2+120}{\theta(\theta^3+2\theta^2+24)}\right)^2$$

$$S.D(\sigma)=\sqrt{\left(\frac{6\theta^3+24\theta^2+720}{\theta^2(\theta^3+2\theta^2+24)}-\left(\frac{2\theta^3+6\theta^2+120}{\theta(\theta^3+2\theta^2+24)}\right)^2\right)}$$

6.2 Harmonic mean

The harmonic mean for proposed length biased Uma distribution can be obtained as

$$H.M=E\left(\frac{1}{x}\right)=\int_0^\infty\frac{1}{x}f_l(x)dx \tag{28}$$

$$H.M=\int_0^\infty\frac{1}{x}\frac{x\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{29}$$

$$H.M=\int_0^\infty\frac{\theta^5}{(\theta^3+2\theta^2+24)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{30}$$

$$H.M=\frac{\theta^5}{(\theta^3+2\theta^2+24)}\int_0^\infty\left(1+x+x^3\right)e^{-\theta x}dx \tag{31}$$

$$H.M=\frac{\theta^5}{(\theta^3+2\theta^2+24)}\left(\int_0^\infty x^{(2)-2}e^{-\theta x}dx+\int_0^\infty x^{(2)-1}e^{-\theta x}dx+\int_0^\infty x^{(4)-1}e^{-\theta x}dx\right) \tag{32}$$

After simplification, we obtain

$$H.M=\frac{\theta(\theta^2+\theta^2+6)}{(\theta^3+2\theta^2+24)} \tag{33}$$

6.3 Moment generating function and characteristic function

Let X be the random variable following length biased Uma distribution with parameter θ, then the moment generating function of executed distribution can be determined as

$$M_X(t)=E\left(e^{tx}\right)=\int_0^\infty e^{tx}f_l(x)dx \tag{34}$$

Using Taylor's series, we obtain

$$M_X(t)=\int_0^\infty\left(1+tx+\frac{(tx)^2}{2!}+\ldots\right)f_l(x)dx$$

$$M_X(t)=\int_0^\infty\sum_{j=0}^\infty\frac{t^j}{j!}x^j f_l(x)dx$$

$$M_X(t)=\sum_{j=0}^\infty\frac{t^j}{j!}\mu_j{}'$$

$$M_X(t)=\sum_{j=0}^\infty\frac{t^j}{j!}\left(\frac{\theta^3\Gamma(j+2)+\theta^2\Gamma(j+3)+\Gamma(j+5)}{\theta^j(\theta^3+2\theta^2+24)}\right) \tag{35}$$

$$M_X(t)=\frac{1}{(\theta^3+2\theta^2+24)}\sum_{j=0}^{\infty}\frac{t^j}{j!\theta^j}\left(\theta^3\Gamma(j+2)+\theta^2\Gamma(j+3)+\Gamma(j+5)\right) \tag{36}$$

Similarly, the characteristic function of length biased Uma distribution can be determined as

$$\varphi_x(t)=M_X(it)$$

$$M_X(it)=\frac{1}{(\theta^3+2\theta^2+24)}\sum_{j=0}^{\infty}\frac{it^j}{j!\theta^j}\left(\theta^3\Gamma(j+2)+\theta^2\Gamma(j+3)+\Gamma(j+5)\right) \tag{37}$$

7. Bonferroni and Lorenz Curves

The bonferroni and Lorenz curves are also termed as income distribution or classical curves are mostly being used in economics to measure the distribution of inequality in income or poverty. The bonferroni and Lorenz curves can be written as

$$B(p)=\frac{1}{p\mu_1'}\int_0^q x\,f_l(x)dx \tag{38}$$

$$L(p)=pB(p)=\frac{1}{\mu_1'}\int_0^q x\,f_l(x)dx \qquad \text{and} \quad q=F^{-1}(p)$$

Where $\mu_1'=\dfrac{\left(2\theta^3+6\theta^2+120\right)}{\theta(\theta^3+2\theta^2+24)}$

$$B(p)=\frac{\theta(\theta^3+2\theta^2+24)}{p(2\theta^3+6\theta^2+120)}\int_0^q x\frac{x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{39}$$

$$B(p)=\frac{\theta(\theta^3+2\theta^2+24)}{p(2\theta^3+6\theta^2+120)}\int_0^q \frac{x^2\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{40}$$

$$B(p)=\frac{\theta^6}{p(2\theta^3+6\theta^2+120)}\int_0^q x^2\left(1+x+x^3\right)e^{-\theta x}dx \tag{41}$$

$$B(p)=\frac{\theta^6}{p(2\theta^3+6\theta^2+120)}\left(\int_0^q x^{(3)-1}e^{-\theta x}dx+\int_0^q x^{(4)-1}e^{-\theta x}dx+\int_0^q x^{(6)-1}e^{-\theta x}dx\right) \tag{42}$$

After simplification, we obtain

$$B(p)=\frac{\theta^6}{p(2\theta^3+6\theta^2+120)}\left(\gamma(3,\theta q)+\gamma(4,\theta q)+\gamma(6,\theta q)\right) \tag{43}$$

$$L(p)=\frac{\theta^6}{(2\theta^3+6\theta^2+120)}\left(\gamma(3,\theta q)+\gamma(4,\theta q)+\gamma(6,\theta q)\right) \tag{44}$$

8. Maximum Likelihood Estimation and Fisher's Information Matrix

In this section, we will discuss the method of maximum likelihood estimation to estimate the parameters of length biased Uma distribution. Consider $X_1, X_2, \ldots, X_n$ be a random sample of size n from the length biased Uma distribution, then the likelihood function can be defined as

$$L(x) = \prod_{i=1}^{n} f_l(x)$$

$$L(x) = \prod_{i=1}^{n} \left(\frac{x_i \, \theta^5}{\left(\theta^3 + 2\theta^2 + 24\right)} \left(1 + x_i + x_i{}^3\right) e^{-\theta x_i} \right) \tag{45}$$

$$L(x) = \frac{\theta^{5n}}{(\theta^3 + 2\theta^2 + 24)^n} \prod_{i=1}^{n} \left(x_i \left(1 + x_i + x_i{}^3\right) e^{-\theta x_i} \right) \tag{46}$$

The log likelihood function is given by

$$\log L = 5n \log \theta - n \log(\theta^3 + 2\theta^2 + 24) + \sum_{i=1}^{n} \log x_i$$
$$+ \sum_{i=1}^{n} \log\left(1 + x_i + x_i{}^3\right) - \theta \sum_{i=1}^{n} x_i \tag{47}$$

By differentiating log likelihood equation (47) with respect to parameter θ and must satisfy the following normal equation

$$\frac{\partial \log L}{\partial \theta} = \frac{5n}{\theta} - n \left(\frac{(3\theta^2 + 4\theta)}{\left(\theta^3 + 2\theta^2 + 24\right)} \right) - \sum_{i=1}^{n} x_i = 0 \tag{48}$$

The above likelihood equation is too complicated to solve it algebraically. Therefore, we use numerical technique like Newton-Raphson method for estimating required parameter of proposed distribution. To use the asymptotic normality results for determining the confidence interval. We have that if $(\hat{\lambda} = \hat{\theta})$ denotes the MLE $of\,(\lambda = \theta)$. We can determine the results as

$$\sqrt{n}(\hat{\lambda} - \lambda) \to N(0,\ I^{-1}(\lambda))$$

Where $I^{-1}(\lambda)$ is Fisher' s informatio n matrix.i.e .,

$$I(\lambda) = -\frac{1}{n}\left(E\left(\frac{\partial^2 \log L}{\partial \theta^2} \right) \right)$$

Here, we can define that

$$E\left(\frac{\partial^2 \log L}{\partial \theta^2} \right) = -\frac{5n}{\theta^2} - n \left(\frac{(\theta^3 + 2\theta^2 + 24)(6\theta + 4) - (3\theta^2 + 4\theta)^2}{(\theta^3 + 2\theta^2 + 24)^2} \right) \tag{49}$$

Since λ being unknown, we estimate $I^{-1}(\lambda)$ by $I^{-1}(\hat{\lambda})$ and this can be used to obtain asymptotic confidenc interval for θ.

9. Application

In this section, we have applied the two real data sets in length biased Uma distribution to determine its goodness of fit and then comparison has been developed in order to realize that the length biased Uma distribution provides quite satisfactory results over Uma, Shanker, Garima and Lindley distributions. The two real data sets are given below as

The following first real data set reported from Lawless [12] represents the failure times in hours of an accelerated life test of 59 conductors without any censored observation. and the data set is given below in table 1.

Table 1: *Data regarding the failure times in hours of an accelerated life test of 59 conductors by Lawless (2003)*

2.997	4.137	4.288	4.531	4.700	4.706	5.009	5.381	5.434	5.459	5.589	5.640
5.807	5.923	6.033	6.071	6.087	6.129	6.352	6.369	6.476	6.492	6.515	6.522
6.538	6.545	6.573	6.725	6.869	6.923	6.948	6.956	6.958	7.024	7.224	7.365
7.398	7.459	7.489	7.495	7.496	7.543	7.683	7.937	7.945	7.974	8.120	8.336
8.532	8.591	8.687	8.799	9.218	9.254	9.289	9.663	10.092	10.491	11.038	

The following data set obtained from Folks and Chhikara [8] and studied by Gadde et al. [10] represents runoff amounts at Jug Bridge, Maryland and the data set is given as under in table 2

Table 2: *Data regarding the runoff amounts at Jug Bridge, Maryland reported by Gadde et al. (2019)*

0.17	1.19	0.23	0.33	0.39	0.39	0.40	0.45	0.52	0.56
0.59	0.64	0.66	0.70	0.76	0.77	0.78	0.95	0.97	1.02
1.12	1.24	1.59	1.74	2.92					

To compute the model comparison criterions along with estimation of unknown parameters, the technique of R software is applied. In order to compare the performance of length biased Uma distribution over Uma, Shanker, Garima and Lindley distributions, we use criterion values like *AIC* (Akaike Information Criterion), *BIC* (Bayesian Information Criterion), *AICC* (Akaike Information Criterion Corrected), *CAIC* (Consistent Akaike Information Criterion), Shannon's entropy *H(X)* and *-2logL*. The distribution is better which shows corresponding criterions *AIC*, *BIC*, *AICC*, *CAIC*, *H(X)* and *-2logL* values smaller as compared with other distributions. For determining the criterion values *AIC, BIC, AICC*, *CAIC, H(X)* and *-2logL* given below following formulas are used.

$$AIC = 2k - 2\log L, \qquad BIC = k\log n - 2\log L, \qquad AICC = AIC + \frac{2k(k+1)}{n-k-1}$$

$$CAIC = -2\log L + \frac{2kn}{n-k-1} \qquad \text{and} \qquad H(X) = -\frac{2\log L}{n}$$

Where n is the sample size, k is the number of parameters in statistical model and $-2\log L$ is the maximized value of log-likelihood function under the considered model.

Table 3: *Shows MLE and S. E Estimate for Data set 1 and Data set 2*

Data set 1		
Distributions	MLE	S.E
Length Biased Uma	$\hat{\theta} = 0.69944543$	$\hat{\theta} = 0.04014727$
Uma	$\hat{\theta} = 0.54873604$	$\hat{\theta} = 0.03478584$
Shanker	$\hat{\theta} = 0.27241572$	$\hat{\theta} = 0.02435319$
Garima	$\hat{\theta} = 0.21941841$	$\hat{\theta} = 0.02422882$
Lindley	$\hat{\theta} = 0.25722036$	$\hat{\theta} = 0.02393044$
Data set 2		
Length biased Uma	$\hat{\theta} = 3.5827918$	$\hat{\theta} = 0.3376072$
Uma	$\hat{\theta} = 2.3356047$	$\hat{\theta} = 0.2501545$

Shanker	$\hat{\theta} = 1.5164881$	$\hat{\theta} = 0.2188903$
Garima	$\hat{\theta} = 1.5342883$	$\hat{\theta} = 0.2647162$
Lindley	$\hat{\theta} = 1.6358862$	$\hat{\theta} = 0.2574658$

Table 4: *Shows Comparison and Performance of fitted distributions for Data 1*

Distributions	-2logL	AIC	BIC	AICC	CAIC	H(X)
LBU	261.0041	263.0041	265.0817	263.0742	263.0742	4.4237
Uma	274.1516	276.1516	278.2291	276.2217	276.2217	4.6466
Shanker	308.9524	310.9524	313.03	311.0225	311.0225	5.2364
Garima	335.3371	337.3371	.339.4147	337.4072	337.4072	5.6836
Lindley	316.7054	318.7054	320.783	318.7755	318.7755	5.3678

Table 5: *Shows Comparison and Performance of fitted distributions for Data 2*

Distributions	-2logL	AIC	BIC	AICC	CAIC	H(X)
LBU	32.62635	34.62635	35.84523	34.8002	34.8002	1.3050
Uma	42.32838	44.32838	45.54726	44.5022	44.5022	1.6931
Shanker	40.29075	42.29075	43.50963	42.4646	42.4646	1.6116
Garima	40.39727	42.39727	43.61614	42.5711	42.5711	1.6158
Lindley	39.60251	41.60251	42.82138	41.7764	41.7764	1.5841

From results given above in table 4 and table 5, it has been clearly revealed and observed that the length biased Uma distribution has lesser *AIC*, *BIC*, *AICC, CAIC, H(X)* and *-2logL* values as compared to the Uma, Shanker, Garima and Lindley distributions. Hence, it can be concluded that the length biased Uma distribution leads to a better fit over Uma, Shanker, Garima and Lindley distributions.

9. Conclusion

This article introduces a novel distribution known as the length-biased Uma distribution, which has been carefully developed and compared to its baseline distribution. Throughout the study, various essential characteristics of the length-biased Uma distribution, including moments, reliability function, hazard rate function, reverse hazard function, shapes of PDF, CDF, hazard, and reliability function, order statistics, Bonferroni, and Lorenz curves, have been thoroughly explored and presented.

Furthermore, the article employs the maximum likelihood estimation method to estimate the parameters of the length-biased Uma distribution, enhancing the robustness and applicability of the proposed model.

To assess its practical effectiveness, the newly introduced distribution has been tested with two real-world datasets. The results demonstrate the superiority of the length-biased Uma distribution over existing distributions such as Uma, Shanker, Garima, and Lindley distributions. This superiority is evidenced by the significantly improved performance and satisfactory outcomes obtained from the proposed length-biased Uma distribution.

In conclusion, the study provides valuable insights into the characteristics and applications of the length-biased Uma distribution, making it a promising addition to the field of statistical modeling and probability theory. The findings open up new avenues for researchers and practitioners to explore the potential of this distribution in various real-world scenarios and data analyses.

References

[1] Al-Omari, A. I. and Alanzi, A. R. A. (2021). Inverse length biased Maxwell distribution: Statistical inference with an application. *Computer Systems Science & Engineering*, 39(1), 147-164.

[2] Atikankul, Y., Thongteeraparp, A., Bodhisuwan, W. and Volodin, A. (2020). The length-biased weighted Lindley distribution with applications. *Lobachevskii Journal of Mathematics*, 41(3), 308-319.

[3] Chaito, T., Nanthaprut, P., Nakharutai, N. and Khamkong, M. (2022). The length-biased Gamma-Rayleigh distribution with applications. *Thailand Statistician*, 20(2), 293-307.

[4] Chaito, T. and Khamkong, M. (2022). The length-biased Weibull-Rayleigh distribution for application to hydrological data. *Lobachevskii Journal of Mathematics*, 42, 3253-3265.

[5] Cox, D. R. (1962). Renewal theory, *Barnes and Noble, New York*.

[6] Ekhosuehil, N., Kenneth, G. E. and Kevin, U. K. (2020). The Weibull length biased exponential distribution: Statistical properties and applications. *Journal of Statistical and Econometric Methods*, 9(4), 15-30.

[7] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13-25.

[8] Folks, J. L. and Chhikara, R. S. (1978). The inverse Gaussian distribution and its statistical application a review. *Journal of the Royal Statistical Society: Series B (Methodological)*, 40(3), 263-275.

[9] Ganaie, R. A., Rajagopalan, V., Nair, S. R. and Kanan, M. (2023). A new generalization of power Garima distribution with applications in blood cancer and relief times. *Journal of Statistics Applications & Probability*, 12(2), 371-394.

[10] Gadde, S., Rosaiah, K. and Prasad, S. (2019). Bootstrap confidence intervals of cnpk for type-ii generalized log-logistic distribution. *Journal of industrial Engineering International*, 15, 87-94.

[11] Ghorbal, A. B. (2022). On properties of length-biased exponential model. *Mathematical Problems in Engineering*. 2022 https://doi.org/10.1155/2022/9779767.

[12] Lawless, J. F. (2003). Statistical Models and methods for Lifetime data, 2nd ed., John Wiley and Sons, New Jersey.

[13] Mustafa, A. and Khan, M. I. (2022). The length-biased power hazard rate distribution: Some properties and applications. *STATISTICS IN TRANSITION new series*, 23(2), 1-16.

[14] Mathew, J. (2020). Reliability test plan for the Marshall Olkin length biased Lomax distribution. *Reliability: Theory & Applications*, 15, No. 2(57), 36-49.

[15] Oluwafemi, O. S. and Olalekan, D. M. (2017). Length and area biased exponentiated weibull distribution based on forest inventories. *Biometrics & Biostatistics International Journal*, 6(2), 311-320.

[16] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; Pergamum Press and Statistical publishing Society, Calcutta. 320-332.

[17] R core team (2019). R version 3.5.3: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria.URL https:// www.R-project .org/.

[18] Ratnaparkhi, M. V. and Naik-Nimbalkar, U. V. (2012). The length-biased lognormal distribution and its application in the analysis of data from oil field exploration studies. *Journal of Modern Applied Statistical Methods*, 11(1), 255-260.

[19] Shanker, R. (2022). Uma distribution with Properties and Applications. *Biometrics & Biostatistics International Journal*, 11(5), 165-169.

RT&A, No 4 (76)
Volume 18, December 2023

A NOVEL EXTENSION OF INVERSE EXPONENTIAL DISTRIBUTIONS: A HEAVY-TAILED MODEL WITH UPSIDE DOWN BATHTUB SHAPED HAZARD RATE

Jabir Bengalath

•

Department of Statistics, Govt. Arts and Science College, Calicut, University of Calicut, Kerala, India.
jabirbengu@gmail.com

Bindu Punathumparambath

•

Department of Statistics, Govt. Arts and Science College, Calicut, University of Calicut, Kerala, India.
ppbindukannan@gmail.com

Abstract

Heavy-tailed distributions have garnered interest due to their advantageous statistical and reliability characteristics, rendering them valuable in applied fields such as economics, finance, and risk management. Such distributions offer robust properties, making them pertinent to studies in various areas like econometrics, statistics, and insurance. Thus, the primary objective of this paper is to propose a Two parameter right skewed- upside down bathtub type, heavy tailed distribution, which is a generalisation of Inverse Exponential distribution and is referred to as Modi Inverse Exponential distribution. We derive several mathematical and statistical features, including quantile function, mode, median, skewness, kurtosis, and mean deviation. Additionally, the reliability and hazard rate functions are also derived. Stochastic ordering and order statistics of the proposed distribution were derived. We also investigate the tail behaviour of the proposed model. Furthermore, estimation methods such as maximum likelihood estimation and its asymptotic confidence bound, percentile method, and Cramer-von-Mises method were examined. To demonstrate the appropriateness of the suggested model, we have considered two distinct real datasets along with three distinct models and concluded that the proposed model is more adaptable.

Keywords: Inverse Exponential distribution, Modi Inverse Exponential distribution, Moments, Tail Behaviour, Order Statistics, Parameter estimation.

1. Introduction

During the recent years, heavy-tailed distributions have gained attention as an attractive subject for various research and studies. References to some works on these distributions can be found in [1],[2],[3],[4],[5]. These distributions possess excellent statistical and reliability properties, making them practical for many applied sciences such as economics, finance, econometrics, statistics, risk management, and insurance. Several authors have developed inferential results under financial modeling, as seen in [6],[7],[8],[9],[10]. There exist various heavy-tailed distributions in many practical situations, such as financial sciences, reliability engineering and bio-medical science,data are usually positive, and their distribution is uni-modal hump shaped and extreme values yielding heavier tails than the classical models. For example, in health science research (1). The medical

expenditure that exceed a given threshold and (2). Length of stay in hospital, presents highly skewed, heavy tailed data for which standard classical distributions and simple variable transform are insufficient to provide an adequate fit to such data. The Exponential distribution has been widely used for analyzing life-time data. However, its usefulness is restricted to scenarios with a constant hazard rate, which can be difficult to justify in practical situations. To address this problem, so many alternative models for life-time data have been developed. Among them, distributions like Weibull and gamma have been extensively employed when dealing with life-time data exhibiting a monotonically increasing or decreasing hazard rate. The UBT failure rate distributions commonly appear in medical and biological fields like in lung cancer patient data (see [11], in bladder cancer patient data (see [28]) and in breast carcinoma patient data (see [12]. The inverse transform method is a widely used approach to derive the inverse form of different lifetime distributions. The distribution family obtained by this method, known as the generalized inverted family, often exhibits the characteristic"upside-doun bathtub" hazard rate pattern. These distributions have the advantage of being the number of parameters required and are straightforward to apply. For instance, Notable examples of such a distribution include the inverted gamma distribution (IED) proposed by Lin et al[13] and the Inverse Lindley Distribution (ILD), introduced by Sharma et al[14]. In addition,The transmuted inverse exponential distribution was presented by Oguntunde and Adejumo [15]. In the same year, Khan et al [16] propose the transmuted inverse Weibull distribution. In addition, in 2014, Sharma et al [17]introduce the transmuted inverse Rayleigh distribution. Then, in 2016, Sharma et al [18] further introduced the generalized inverse Lindley distribution. The purpose of this study is to propose a new inverted probability model with UBT type of failure rate. For this purpose , we consider A one parameter Inverse exponential distribution. Let X be the random variable having the probability density function (pdf) is given by

$$g(x) = \frac{\theta}{x^2} e^{-\frac{\theta}{x}}, \quad x > 0, \quad \theta > 0. \tag{1}$$

and the cumulative distribution function (cdf) is

$$G(x) = e^{-\frac{\theta}{x}}, \quad x > 0, \quad \theta > 0. \tag{2}$$

for all $x > 0$, where $\theta > 0$. Furthermore, K. Modi et al [19] introduced a new family of distributions name Modi family of distribution in his paper titled a new family of distributions with applications on two real data sets on the survival problem. where he proposed a new probability distribution by taking base line distribution by exponential distribution with one parameter. This paper aims to substitute α^{β} with γ within the established family, resulting in a modified probability density function(pdf) and cumulative distribution function(cdf) is given by.

$$f_Y(x) = \frac{(1+\gamma)(\gamma g(x))}{(\gamma + G(x))^2}, \quad \gamma > 0. \tag{3}$$

and

$$F_Y(x) = \frac{(1+\gamma)G(x)}{\gamma + G(x)}, \quad \gamma > 0. \tag{4}$$

where $g(x)$ and $G(x)$ are the pdf and cdf of the baseline distribution. Moreover, We can easily verified that the given Modi family satisfies the identifiable properties and other properties which is required for a probability distribution. Hence, this familiy of distribution can be used to generate more flexible probability distributions. The hazard function of equation (1) is given by,

$$h_Y(x) = \left[\frac{1+\gamma}{\gamma + G(x)}\right] h_X(x), \quad \gamma > 0.$$

Where $h_X(x)$ is the hazard function of the baseline distribution.

Acknowledging the need for more flexible lifetime distributions, we introduce a new family of probability distribution known as the Modi Inverse Exponential distribution with two parameters. which can be extensively used to fit and analyze data in a variety of field. The paper is

organized as follows: Section 2 specifies the Modi Inverse Exponential distribution, whereas Section 3 provides the Modi Inverse Exponential distribution Properties, which include the Hazard function, Survival function, Quantile function,Mode, Median, Skewness, Kurtosis, Mean Deviation, Stochastic Ordering, Order statistics. In Section 4, we look at the Modi inverted exponential distribution's tail behaviour. Section 5 We investigated different method of estimation which includes Maximum likelihood estimation and its asymptotic confidence bound, Percentile method and Cramer-von Mises method is discussed Section 5, we conduct a simulation study to validate the proposed model's estimations, and two real data sets are analysed to demonstrate the efficacy of the proposed model. The conclusion is provided in Section 6.

2. Modi Inverse Exponential Distribution

A random variable X is said to have Modi Inverse Exponential distribution (MIE) if its cumulative distribution function (cdf) is given by

$$F(x) = \frac{(\gamma+1)}{1+\gamma e^{\frac{\theta}{x}}} \quad , \quad x > 0, \quad \gamma > 0, \quad \theta > 0. \tag{5}$$

and its pdf is,

$$f(x) = \frac{\theta\gamma(\gamma+1)e^{\frac{\theta}{x}}}{(x\gamma e^{\frac{\theta}{x}}+x)^2} \quad , \quad x > 0, \quad \gamma > 0, \quad \theta > 0. \tag{6}$$

The shape of the distribution might provide important insights into its characteristics, such as whether it is symmetrical or skewed. In this context, the MIE(γ, θ) distribution is represented by its cumulative distribution function (cdf) in Figure 2 and its probability density function (pdf) in Figure 1 for different values parameter.

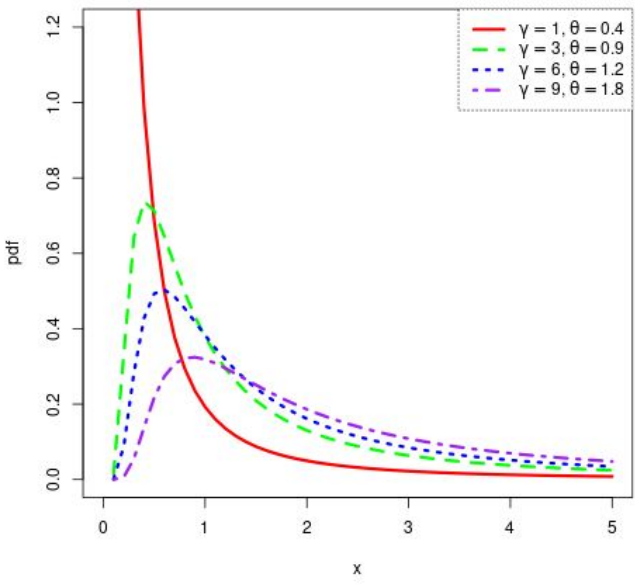


Figure 1: *Pdf plot of MIE distribution for different parameter values*

Theorem 1. Given that X follows the MIE (γ, θ) distribution with $f(x)$ and $F(x)$ as given in (6) and (5) respectively, then:

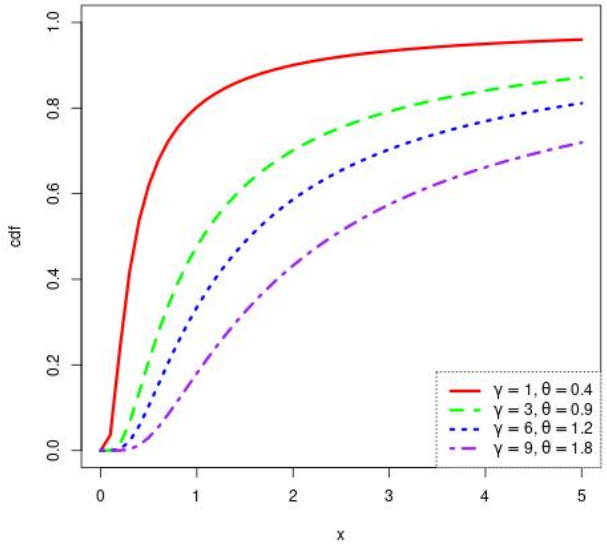


Figure 2: *cdf plot of MIE distribution for different parameter values*

1. $\lim_{x\to\infty} f(x) = 0$
2. $\lim_{x\to\infty} F(x) = 1$

proof. Trivial and hence omitted.

3. Some Statistical Properties of MIE Distribution

This section contains various statistical features associated with the new distribution that have been derived.

3.1. Hazard Rate Function

The hazard function characterizing a specific phenomenon elucidates the inherent nature of the failure rate that is associated with lifetime of the specific equipment. For the cdf and pdf provided in equations (5) and (6), respectively, the expression for $h(x)$ is as follows:

$$h(x) = \frac{(\gamma+1)\theta e^{\frac{\theta}{x}}}{x^2(e^{\frac{\theta}{x}}-1)(\gamma e^{\frac{\theta}{x}}+1)} \tag{7}$$

The hazard rate function plot in Figure 3 shows various curves indicating different values of the parameters γ and θ. We can gain useful insights about the nature of the model's failure rate by using this visual depiction, which reveals a distinct right-skewed pattern and UBT type failure model.The mathematical verification of this assertion may also be established through the utilization of the outcome presented by Glaser [20]. Glaser demonstrated that the Condition for UBT can be established if and only if the following conditions are met: $\varphi'(t) > 0$ for all $t \in (0, t_0)$, $\varphi'(t_0) = 0$, $\varphi'(t) < 0$ for all $t > t_0$, and satisfying $\lim_{t\to 0} f(t) = 0$ where φ is equal to $-\frac{f'(t)}{f(t)}$ and $f(t)$ is the first derivative of the density function with respect to t. For our proposed model, it is evident that.

$$\varphi(t) = -\frac{2t(1+e^{\frac{\theta}{t}}\gamma)+\theta-e^{\frac{\theta}{t}}\gamma\theta}{t^2(1+e^{\frac{\theta}{t}}\gamma)}$$

and

$$\varphi'(t) = \frac{2((t+e^{\frac{\theta}{t}}t\gamma)^2-e^{\frac{\theta}{t}}\gamma\theta^2+t(\theta-e^{\frac{2\theta}{t}}\gamma^2\theta)}{t^4(1+e^{\frac{\theta}{t}}\gamma)^2}$$

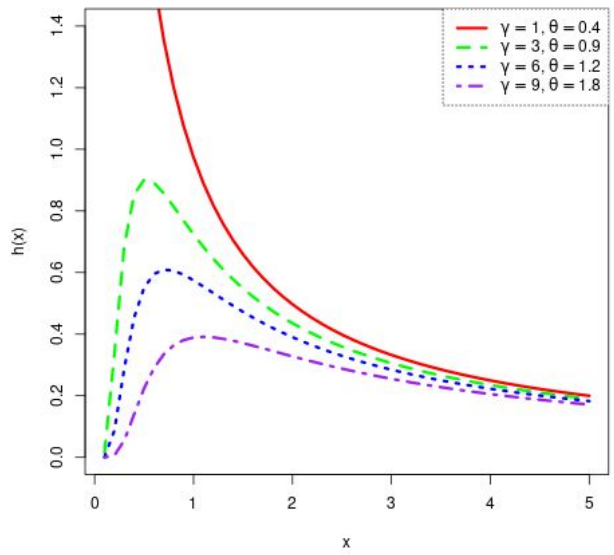


Figure 3: *Hazard plot of MIE distribution for different parameter values*

Since, the above equation is not provided in an explicit form to derive the solution, a simulation study was performed.It was observed that for $t_0 \approx 0.4488$, $\varphi'(t) > 0$ for all $t \in (0, t_0)$, $\varphi'(t_0) = 0$, $\varphi'(t) < 0$ for all $t > t_0$. Also, from Equation (6), we verified that $\lim_{x\to 0} f(t) = 0$. Therefore, it can be deduced that the MIE(γ,θ) distribution proposed exhibits a right-skewed distribution, which is characterized by an UBT shape of hazard rate. This distribution is particularly useful when analyzing medical and reliability data.

3.2. Survival Function

The survival function describes the probability that a unit, component, or individual will not fail at a given time. The expression for survival function $S(x)$ is stated as follows, and its corresponding survival plot is presented in Fiqure 4:

$$S(x) = \frac{\gamma(e^{\frac{\theta}{x}} - 1)}{\gamma e^{\frac{\theta}{x}} + 1}, \tag{8}$$

Theorem 2. The limit of the hazard rate function of MIE(γ, θ) distribution as $x \to \infty$ is zero.

$$i.e, \lim_{x\to\infty} h(x) = 0.$$

Proof. Trivial and hence omitted. ■

3.3. The Odd Function

The Odd Function is obtained using the relation $Q(x) = \frac{F(x)}{s(x)}$ and is given by

$$Q(x) = \frac{\gamma + 1}{\gamma(e^{\frac{\theta}{x}} - 1)}, \tag{9}$$

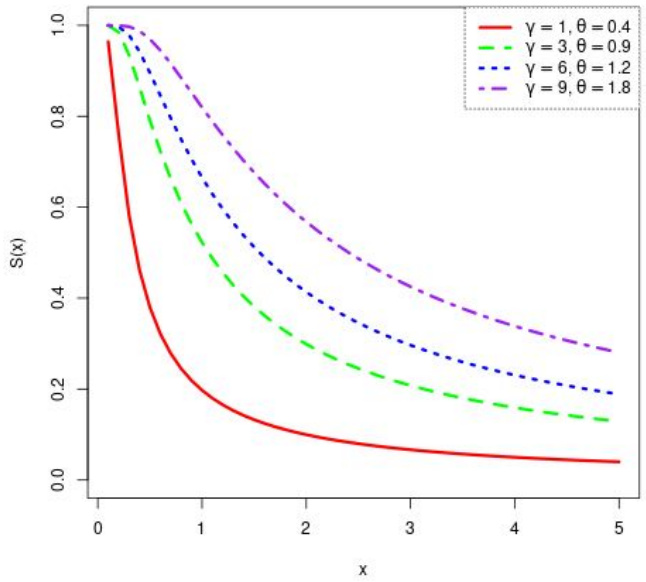


Figure 4: *Survival plot of MIE distribution for different parameter values*

3.4. Reverse Hazard Rate Function

The Revised Hazard Rate Function is obtained by using the relation,

$$\begin{aligned} \phi(x) &= \frac{f(x)}{F(x)} \\ &= \frac{\theta\gamma e^{\frac{\theta}{x}}}{x^2(\gamma e^{\frac{\theta}{x}} + 1)}, \end{aligned} \tag{10}$$

3.5. Cumulative Hazard Function

The Cumulative Hazard Function is obtained using the relation,

$$\begin{aligned} C(x) &= -\log(S(x)) \\ &= -\log\left[\frac{\gamma(e^{\frac{\theta}{x}} - 1)}{\gamma e^{\frac{\theta}{x}} + 1}\right], \end{aligned} \tag{11}$$

3.6. Quantile Function, Skewness and Kurtosis

The MIE distribution can be simulated using the inverse cdf method,

$$X = \left[\frac{\theta}{\ln\left(\frac{\gamma+1-u}{\gamma u}\right)}\right], \tag{12}$$

where, u is a uniform random variable, $0 < u < 1$. The q^{th} quantile of the MIE distribution is obtained as:

$$x_q = \left[\frac{\theta}{\ln\left(\frac{\gamma+1-q}{\gamma q}\right)}\right], \tag{13}$$

By making use of equation (13), we are able to calculate the first and third quartiles by substituting $q = 0.25$ and $q = 0.75$, correspondingly. Once we have obtained these values, we can subsequently

calculate Galton's [[21]] skewness (S_k) and Moor's [[22]] Kurtosis (K_r) by means of the given formulae.
The measure of skewness S_k,

$$S_k = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)}, \tag{14}$$

and the measure of Kurtosis, K_r

$$K_r = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)}, \tag{15}$$

3.7. Median

Since the distribution proposed is a heavily tailed, right-skewed distribution, the most appropriate measure of central tendency is the median.The median of the proposed distribution can be obtained by utilizing $q = 0.5$ in the quantile function, as delineated in equation (13). So median M_d

$$M_d = \left[\frac{\theta}{\ln\left(\frac{\gamma+1-0.5}{\gamma(0.5)}\right)} \right], \tag{16}$$

3.8. Mode

if a random variable X has the PDF given by equation (6), then the corresponding mode is given by $f'(x) = 0$, thus we obtain

$$f'(x) = \frac{e^{\frac{\theta}{x}}\theta\gamma(1+\gamma)\left(e^{\frac{\theta}{x}}(\theta-2x) - \gamma(\theta+2x)\right)}{\left(e^{\frac{\theta}{x}}+\gamma\right)^3 x^4} = 0$$

$$\implies \left(e^{\frac{\theta}{x}}(\theta-2x) - \gamma(\theta+2x)\right) = 0$$

For various values of γ and θ, we can estimate the value of x by using an optimization technique in R. If $\gamma = 3$ and $\theta = 4$, we obtain the mode as 1.93755.

3.9. Mean Deviation

The mean deviation from the median is a statistical measure, serves as an indicator of population dispersion. Let "M" stand in for the median of the MIE Distributions specified in equation (16). The mean deviation from the median may be computed as follows:

$$\rho(x) = E|x - M| = \int_0^\infty |x - M| f(x)\, \mathrm{dx},$$

it can be obtained the following equation $\rho = \mu - 2W(M)$ where W(M) is

$$W(M) = \theta\gamma(\gamma+1)\int_0^M \frac{e^{\frac{\theta}{x}}}{x^2(\gamma e^{\frac{\theta}{x}}+1)^2}\mathrm{dx}, \tag{17}$$

This integral may be readily computed numerically using tools such as R, MATLAB, Mathcad, and others. Thus, obtaining the mean deviation from the median is desired.

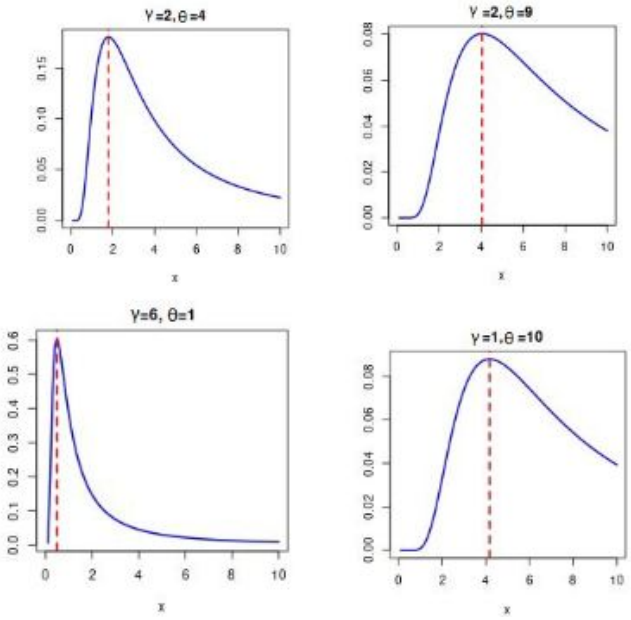


Figure 5: *Mode plot of MIE distribution for different parameter value*

3.10. Stochastic Ordering

Let X_1 and X_2 be random variables with cumulative distribution functions (cdf's) $F_1(x)$ and $F_2(x)$, respectively. X_1 is said to be stochastically greater than or equal to X_2 if $F_1(x) \leq F_2(x)$ for all x. (see Gupta et al [23] for more detail).

Theorem 3. Let $X_1 \sim \text{MIE}(\gamma_1, \theta_1)$ and $X_2 \sim \text{MIE}(\gamma_2, \theta_2)$. X_1 is said to be stochastically greater than X_2 if $\gamma_1 = \gamma_2 = \gamma$ and $\theta_1 > \theta_2$.
Proof. Let's consider $\theta_1 > \theta_2$ and $\gamma_1 = \gamma_2 = \gamma$, the ratio simplifies to:

$$\frac{F_1(x)}{F_2(x)} = \frac{\frac{\gamma+1}{1+\gamma e^{\frac{\theta_1}{x}}}}{\frac{\gamma+1}{1+\gamma e^{\frac{\theta_2}{x}}}}$$

Since $\theta_1 > \theta_2$, we have:

$$\frac{e^{\frac{\theta_1}{x}}}{e^{\frac{\theta_2}{x}}} > 1$$

Therefore, when $\theta_1 > \theta_2$ and $\gamma_1 = \gamma_2 = \gamma$, $F_1(x)$ is stochastically smaller than $F_2(x)$ for all $x > 0$. ■

3.11. Order Statistics

In this section, we derive a compact expression for the pdf of the i^{th} order statistic of the Modi inverse exponential distribution. Let $X_1, X_2, X_3, ..., X_n$ be a simple random sample from the Modi inverse exponential distribution with cdf and pdf given by equations (5) and (6), respectively. Let $X_{1:n} \leq X_{2:n} \leq X_{3:n} \leq ... \leq X_{n:n}$ denote the order statistics obtained from this sample. We now give the pdf of $X_{r:n}$, denoted as $f_{r:n}(x)$, and the r^{th} moments of $X_{r:n}$, for $i = 1, 2, ..., n$, which are given by:

$$f_{r:n}(x) = C_{r:n} \left[F(x;\gamma,\theta)\right]^{r-1} \left[1 - F(x;\gamma,\theta)\right]^{n-r} f(x;\gamma,\theta), \tag{18}$$

for all $x > 0$, where F and f are given by equations (5) and (6), respectively, and $C_{r:n} = \frac{n!}{(n-1)!(n-r)!}$. Thus, using the binomial series expansion:

$$(1-x)^{\alpha-1} = \sum_{j=0}^{\infty}(-1)^j \binom{\alpha-1}{j} x^j.$$

We obtain:

$$\begin{aligned} f_{r:n}(x) &= C_{r:n}\sum_{s=0}^{\infty}(-1)^s \binom{n-r}{s} [F(x;\gamma,\theta)]^{r+s-1} f(x;\gamma,\theta), \\ &= C_{r:n}\sum_{s=0}^{\infty}(-1)^s \binom{n-r}{s} \left[\frac{\gamma+1}{1+\gamma e^{\frac{\theta}{x}}}\right]^{r+s-1} \frac{\theta\gamma(\gamma+1)e^{\frac{\theta}{x}}}{(x\gamma e^{\frac{\theta}{x}}+x)^2}, \\ &= C_{r:n}\sum_{s=0}^{\infty}(-1)^s \binom{n-r}{s} \theta\gamma(\gamma+1)^{r+s}(1+\gamma e^{\frac{\theta}{x}})^{r+s-1} e^{\frac{\theta}{x}}(x\gamma e^{\frac{\theta}{x}}+x)^{-2}. \end{aligned} \tag{19}$$

4. Tail Area Property

According to Klugman et al [24]and Nair et al [25] a distribution is classified as a heavy-tailed distribution when it displays the heavy tail property. These types of distributions are characterized by the lack of one or more orders of moments. Specifically, the absence of the first moment, which represents the distribution's arithmetic mean, indicates the presence of the distribution's heavy tail property. The proposed distribution's arithmetic mean can be derived by solving:

$$\theta\gamma(\gamma+1)\int_0^{\infty} \frac{e^{\frac{\theta}{x}}}{x(\gamma e^{\frac{\theta}{x}}+1)^2} \mathrm{dx},$$

which is a divergent integral,then the arithmetic mean of the corresponding distribution cannot be determined. Consequently, based on this criterion, the proposed distribution can be classified as a heavy-tailed distribution. Another method for evaluating the heavy tail attribute of the distribution is to examine whether the ratio of the hazard rate to x approaches zero as x approaches infinity; if it does, then the distribution displays the characteristic of a heavy-tailed distribution. For the proposed distribution:

$$\frac{(\gamma+1)\theta e^{\frac{\theta}{x}}}{x^3(e^{\frac{\theta}{x}}-1)(\gamma e^{\frac{\theta}{x}}+1)} \to 0$$

This fact can be proven by implementing L'Hpital's rule. As a result, the distribution put forward exhibits a heavy-tailed distribution.

In our analysis, we additionally consider the examination of the heavy tail characteristic of the distribution via an alternative methodology. This methodology involves the observation of the ratio of two survival functions. If the ratio of survivals approaches infinity as x approaches infinity, then one survival function is considered to be heavier than the other. Moreover, the limiting case of the ratio of two survival functions provides the limiting case of two probability density functions. Therefore, this ratio also indicates the ratio of two density functions.

$$\lim_{x\to\infty}\frac{S_1(x)}{S_2(x)} = \lim_{x\to\infty}\frac{S_1'(x)}{S_2'(x)} = \lim_{x\to\infty}\frac{f_1(x)}{f_2(x)}$$

here, we conduct a comparison between the proposed distribution and the Pareto Type II distribution. The survival function of the Pareto Type II distribution is expressed as $S(x) = P(X > x) = \left(1+\frac{x}{\lambda}\right)^{-\alpha}$. For $\alpha > 1$, the ratio between the two distributions tends to infinity as x approaches infinity. This suggests that the tail of the proposed distribution. Figure 6 provides the plot of the tail density for the proposed distribution in comparison with two other distributions, namely the normal distribution and the Parato Type II distribution.

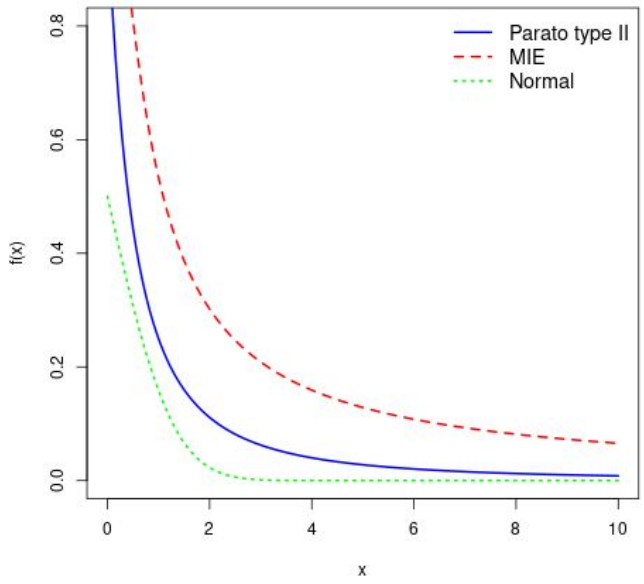


Figure 6: *Tail behaviour of Normal, Parato type II and MIE densities*

Remark 1: A distribution is said to be heavy-tailed distribution if and only if

$$\int_0^\infty e^{\lambda x} f(x)\, dx = \infty \quad \text{for all } \lambda > 0.$$

Hence MIE(γ,θ) is heavy tailed becuase it satisfies condition.

Remark 2: if a distribution is said to be heavy-tailed, if the right probabilities are heavier than the exponential distribution if

$$\lim_{x\to\infty} \frac{1 - F(x)}{e^{-\lambda x}} = \infty \quad \text{for all} \quad \lambda > 0.$$

Hence MIE(γ,θ) distribution also satisfies this condition.

Definition: An ultimately positive function f is called regularly varying at infinity with index $\gamma \in \mathbb{R}$ if for any fixed $c > 0$:

$$\lim_{x\to\infty} \frac{f(cx)}{x} = c^\gamma.$$

The aforementioned theorem demonstrates that the density function derived from the MIE(γ,θ) in equation (6) exhibits regularly varying tails.

Theorem 4. The density function of the MIE(γ,θ) distribution is a function with regularly varying tails.

Proof. Using the density function (6), we have:

$$\lim_{x\to\infty} \frac{f(cx)}{f(x)} = \lim_{x\to\infty} \frac{e^{\frac{\theta}{cx}} x^2 (\gamma e^{\frac{\theta}{x}} + 1)^2}{e^{\frac{\theta}{x}} (x + cx)^2 (\gamma e^{\frac{\theta}{cx}} + 1)^2} = 1,$$

applying limits, the above simplifies to:

$$\lim_{x\to\infty} \frac{f(cx)}{f(x)} = 1,$$

Hence the proof. ■

Definition: An ultimately positive function f is long-tailed and is said to belong to class L if and only if:

$$\lim_{x\to\infty}\frac{f(x+y)}{f(x)}=1,\quad \text{for all } y>0.$$

Theorem 5. The MIE(γ,θ)Distribution belongs to the class L.

Proof.

$$\lim_{x\to\infty}\frac{f(x+y)}{f(x)}=\frac{e^{\frac{\theta}{x+y}}x^2(\gamma e^{\frac{\theta}{x}}+1)^2}{e^{\frac{\theta}{x}}(x+y)^2(\gamma e^{\frac{\theta}{x+y}}+1)^2}=1,$$

Hence, f belongs to the class L. ■

Definition: An ultimately positive function f belongs to the class D of dominated variation distributions if there exists $c>0$ such that:

$$\lim_{x\to\infty}\frac{f(x)}{f(2x)}=c,\quad \text{for all } x>0.$$

Theorem 6. The MIE(γ,θ) Distribution belongs to the class D of dominated variation distributions.

Proof.

$$\lim_{x\to\infty}\frac{f(x)}{f(2x)}=\frac{2^2e^{\frac{\theta}{x}}\left(\gamma e^{\frac{\theta}{2x}}+1\right)^2}{e^{\frac{\theta}{2x}}\left(\gamma e^{\frac{\theta}{x}}+1\right)^2}=4,$$

hence f belongs to the class of dominated variation distributions. ■

4.1. Different method of Estimation

In this section, we are looking at three estimation methods for estimating the unknown model parameters of the proposed model. The procedures are detailed below

4.2. Maximum Likelihood Estimation

Let X be a random variable with the pdf of the Modi inverse exponential distribution defined in equation (6), then its log-likelihood function is defined by:

$$\begin{aligned}\log L(x;\gamma,\theta)&=n\log\theta+n\log\gamma+n\log(\gamma+1)\\&\quad+\theta\sum_{i=0}^{\infty}\frac{1}{x_i}-2\sum_{i=0}^{\infty}\log(x_i\gamma e^{\frac{\theta}{x_i}}+x_i)\end{aligned}\tag{20}$$

Thus, the non-linear normal equations are given as follows:

$$\frac{\partial\log L(x;\gamma,\theta)}{\partial\theta}=\frac{n}{\theta}+\sum_{i=0}^{n}\frac{1}{x_i}-2\sum_{i=0}^{n}\frac{e^{\frac{\theta}{x_i}}\gamma}{(x_i+e^{\frac{\theta}{x_i}}\gamma x_i)}\tag{21}$$

$$\frac{\partial\log L(x;\gamma,\theta)}{\partial\gamma}=\frac{n}{\gamma}+\frac{n}{1+\gamma}-2\sum_{i=0}^{n}\frac{e^{\frac{\theta}{x_i}}x_i}{(x_i+e^{\frac{\theta}{x_i}}\gamma x_i)}\tag{22}$$

The equations from (21) to (22) above are not in closed form. We refer to using some iterative procedure such as Newton Raphson, Bisection methods, or some other method to obtain the approximate maximum likelihood estimates (MLE) of these parameters for the solution of these explicit equations.

4.3. The Asymptotic Confidence Bounds:

The maximum likelihood estimators (MLE) of the unknown parameters γ, θ in the MIE (γ, θ) do not have closed-form solutions. As a result, the exact distribution of the MLE cannot be derived. However, asymptotic confidence bounds can be obtained based on the asymptotic distribution of the MLE. The information matrix is calculated by taking the second partial derivatives of equations (21) to (22) and is given as:

$$\frac{\partial^2 \log L(x;\gamma,\theta)}{\partial \theta^2} = -\frac{n}{\theta^2} - 2\sum_{i=0}^{n}\left(\frac{e^{\frac{2\theta}{x_i}}\gamma^2}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)^2} + \frac{e^{\frac{\theta}{x_i}}\gamma}{x_i(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)}\right) \tag{23}$$

$$\frac{\partial^2 \log L(x;\gamma,\theta)}{\partial \gamma^2} = -\frac{n}{\gamma^2} - \frac{n}{(1+\gamma)^2} - 2\sum_{i=0}^{n}\left(-\frac{e^{\frac{2\theta}{x_i}}x_i^2}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)^2}\right) \tag{24}$$

$$\frac{\partial^2 \log L(x;\gamma,\theta)}{\partial \gamma,\theta} = -2\sum_{i=0}^{n}\left(-\frac{e^{\frac{2\theta}{x_i}}\gamma x_i}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)^2} + \frac{e^{\frac{\theta}{x_i}}}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)}\right) \tag{25}$$

So that the observed information matrix is given by:

$$I = -\begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}$$

Hence, the variance-covariance matrix is approximated as:

$$V \approx I^{-1} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

To obtain the estimate of V, we replace the parameters by their corresponding maximum likelihood estimators (MLEs) to get:

$$\hat{V} \approx \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix}$$

Using the above variance-covariance matrix, one can derive the $(1-\beta)100\%$ confidence intervals for the parameters θ and γ as follows:

$$\hat{\gamma} \pm Z_{\frac{\gamma}{2}}\sqrt{\text{var}(\hat{\gamma})}, \quad \hat{\theta} \pm Z_{\frac{\gamma}{2}}\sqrt{\text{var}(\hat{\theta})}.$$

4.4. The Percentile Method

Let $X_{(j)}$ be the jth order statistic, i.e., $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$. if p_j denote some estimate of $F(x_j;\theta,\gamma)$, then the estimate of θ and γ can be obtained by minimizing

$$\sum_{j=1}^{n}\left(x_j - \left[\frac{\theta}{\ln\left(\frac{\gamma+1-p_j}{\gamma p_j}\right)}\right]\right)^2;$$

with respect to θ and γ. Several types of estimators for p_j can be used [26], and this paper considers $p_j = \frac{j}{n+1}$.

4.5. Method of Cramer-von Mises

Cramer-von-Mises type minimum distance estimators aim to minimize the disparity between the theoretical and empirical cumulative distribution functions. It has been demonstrated empirically

that these estimators have a lesser bias than other minimum distance estimators. The Cramer-von-Mises estimators $\hat{\gamma}_{\text{CME}}$, and $\hat{\theta}_{\text{CME}}$, are the values of γ, and θ that minimizing

$$W(\gamma,\theta) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F\left(t_i\,;\,\gamma,\theta\right) - \frac{2i-1}{2n}\right]^2 \tag{26}$$

where t_i is the i-th ordered observation, and $F(t_i;\gamma,\theta)$ is the cumulative distribution function of the MIED with parameters γ and θ. To estimate the parameters, we differentiate the above equation partially with respect to the parameters γ and θ, respectively, and equate them to zero to get the normal equations. Since the normal equations are nonlinear, we can use iterative methods to obtain the solutions.

5. Simulation Study and Data Analysis

5.1. Simulation Study

In this section, we conduct a comprehensive Monte Carlo simulation study that is repeated 1000 times in order to compare the performance of the previously discussed estimators. We evaluate these estimators using Mean Squared Error (MSE) and make comparisons across sample sizes $n = 50, 100, 150, 200$ for two distinct parameter settings: $\gamma = 1$ and $\theta = 0.05$, and $\gamma = 0.9$ and $\theta = 2.5$. The simulation-based outcomes provide estimates denoted as $\gamma_{\hat{PM}}, \theta_{\hat{PM}}, \gamma_{\hat{ML}}, \theta_{\hat{ML}}, \gamma_{\hat{CM}}$, and $\theta_{\hat{CM}}$ for the Percentile Method (PM), Maximum Likelihood Estimation (MLE), and Cramer-von Mises (CVM) method. The corresponding MSE values are displayed in parentheses. Notably, as sample size n increases, the Mean Squared Error (MSE) tends to decrease, indicating improved estimation accuracy. Table 1 and Table 2 present the simulation results.

n	PM		MLE		CVM	
	$\gamma_{\hat{PM}}$	$\theta_{\hat{PM}}$	$\gamma_{\hat{ML}}$	$\theta_{\hat{ML}}$	$\gamma_{\hat{CM}}$	$\theta_{\hat{CM}}$
50	0.0836	0.0203	0.0013	0.0206	0.0900	0.0102
	(0.3486)	(6.918×10^{-04})	(8.709×10^{-04})	(5.277×10^{-04})	(0.4250)	(0.0050)
100	0.0406	0.0097	0.0005	0.0102	0.0425	0.0049
	(0.1623)	(1.303×10^{-04})	(6.057×10^{-04})	(1.144×10^{-04})	(0.1819)	(0.0026)
150	0.0269	0.0065	0.0004	0.0067	0.0282	0.0033
	(0.1069)	(6.499×10^{-05})	(1.357×10^{-04})	(5.123×10^{-05})	(0.1195)	(0.0017)
200	0.0203	0.0049	0.0004	0.0050	0.0213	0.0024
	(0.0804)	(3.696×10^{-05})	(4.004×10^{-05})	(3.308×10^{-05})	(0.0896)	(0.0013)

Table 1: *Simulation outcomes obtained for the parameter value of $\gamma = 1$ and $\theta = 0.05$ are presented herein. The values enclosed within the parentheses denote the Mean Squared Error (MSE) values.*

5.2. Data Analysis

In this section, we demonstrate the usefulness of the proposed Modi Inverse Exponential distribution with parmeter γ and θ . We fit this distribution to a real-life data set and compare the results with some recent efficient models, namely the Inverse Generalized Weibull distribution, Generalized Inverse Generalized Weibull distribution. The corresponding PDFs are presented below:

- Inverse Generalized Weibull Distributions:

$$f(x,\alpha,\beta,\lambda) = \alpha\beta\lambda^{\beta}e^{-\left(\frac{\lambda}{x}\right)^{\beta}}x^{-(\beta+1)}\left(1-e^{-\left(\frac{\lambda}{x}\right)^{\beta}}\right)^{\alpha-1},$$

- Generalized Inverse Generalized Weibull Distribution:

n	PM		MLE		CVM	
	$\hat{\gamma_{PM}}$	$\hat{\theta_{PM}}$	$\hat{\gamma_{ML}}$	$\hat{\theta_{ML}}$	$\hat{\gamma_{CM}}$	$\hat{\theta_{CM}}$
50	0.0564	0.0725	1.9823	0.05313	0.0065	0.0362
	(3.926×10^{-03})	(3.08×10^{-02})	(0.2345)	(3.09×10^{-02})	(4.036×10^{-03})	(3.103×10^{-02})
100	0.02762	0.0354	1.2356	0.0256	0.0029	0.0177
	(2.591×10^{-05})	(2.044×10^{-04})	(0.0083)	(2.055×10^{-04})	(2.688×10^{-05})	(2.064×10^{-04})
150	0.01835	0.0232	0.8017	0.0168	0.0019	0.0116
	(2.536×10^{-07})	(2.024×10^{-06})	(0.0072)	(2.048×10^{-06})	(2.682×10^{-07})	(2.053×10^{-06})
200	0.0138	0.0174	0.2211	0.0127	0.0014	0.0087
	(4.744×10^{-09})	(3.928×10^{-08})	(0.0037)	(3.991×10^{-08})	(5.322×10^{-09})	(4.046×10^{-08})

Table 2: *Simulation outcomes obtained for the parameter value of $\gamma = 0.9$ and $\theta = 2.5$ are presented herein. The values enclosed within the parentheses denote the Mean Squared Error (MSE) values.*

$$f(x, \alpha, \beta, \lambda, \gamma) = \alpha\beta\gamma\lambda^{\beta} e^{-\gamma\left(\frac{\lambda}{x}\right)^{\beta}} x^{-(\beta+1)} \left(1 - e^{-\gamma\left(\frac{\lambda}{x}\right)^{\beta}}\right)^{\alpha-1},$$

Data Set 1: This data set has been taken from [27]. The data on survival of 40 patients suffering from leukemia, from the Ministry of Health Hospitals in Saudi Arabia, was taken from Abouammoh et al. (1994):

115	181	255	418	441	461	516	739	743	789	807	865	924	983
1024	1062	1063	1165	1191	1222	1222	1251	1277	1290	1357	1369	1408	1455
1478	1549	1578	1578	1599	1603	1605	1696	1735	1799	1815	1852		

Table 3: *Estimates and Goodness-of-fit measures based on AIC, BIC, AICC, and CAIC for Data Set 1*

Distribution	**Estimates**	**Log-Likelihood**	**AIC**	**BIC**	**AICC**	**CAIC**
MIE	$\gamma = 0.1350$	-326.978	656.822	660.199	657.146	660.199
	$\theta = 1.7856$					
IGWD	$\alpha = 0.0426$	-346.170	698.340	703.406	699.006	703.406
	$\beta = 1.3546$					
	$\theta = 2.7048$					
GIGWD	$\alpha = 0.0323$	-367.455	742.909	749.664	744.052	749.664
	$\beta = 0.9067$					
	$\theta = 1.2614$					
	$c = 4.2107$					

From Table 3, it shows that the proposed Modi Inverse Exponential distribution model has the lowest AIC, BIC, AICC, and CAIC values among the other distributions, suggesting that it provides the best fit to the dataset.

Data Set 2: This data set represents survival times in Days, from a Two-Arm Clinical Trial considered by [28] and [29]. The survival time in days for the 31 patients from Arm B are:

37	84	92	94	110	112	119	127	130	133	140	146	155	159	173	179
194	195	209	249	281	319	339	432	469	519	633	725	817	1557	1776	

From Table 4, we can see that our proposed model MIE has minimum AIC, BIC, AICC, and CAIC values compared to IGWD and GIGWD distributions. Thus, we can infer that the newly proposed model is a better fit for the given data compared to the other models.

Table 4: *Estimates and Goodness-of-fit measures based on AIC, BIC, AICC, and CAIC for Data Set 2*

Distribution	Estimates	Log-Likelihood	AIC	BIC	AICC	CAIC
MIE	$\gamma = 0.0618$	-206.657	417.315	420.183	417.744	420.183
	$\theta = 1.1340$					
IGWD	$\alpha = 0.0596$	-217.366	440.733	445.035	441.622	445.035
	$\beta = 1.8165$					
	$\theta = 4.5081$					
GIGWD	$\alpha = 0.0665$	-217.015	442.032	447.767	443.570	447.767
	$\beta = 0.8546$					
	$\theta = 4.9719$					
	$c = 1.0867$					

6. Conclusion

In this article, We establishes the Modi Inverse Exponential distribution which a right skewed heavy tailed UBT shaped probability model. The related structural properties are derived and represented in the respective sections. Furthermore, we explore the tail behavior of the suggested model and conclude that it is heavy-tailed. To estimate the distribution's parameters, different estimation methods such as method of maximum likelihood, method percentile and Method of Cramer-von Mises are used. For the simulated data set, the results are shown in Table 1 and 2. We can see that the estimated values obtained are near to the predefined parameters, and as n increases, MSE decreases, confirming the law of large numbers. However, the application to two real-life data sets shows that the MIE distribution has a better fit than other competing models, such as the Inverse Generalized Weibull distribution (IGWD) and Generalized Inverse Generalized Weibull distribution, based on goodness-of-fit measures AIC, BIC, AICC, and CAIC.

References

[1] Venturini, S., Dominici, F. and Parmigiani, G. (2012). Gamma shape mixtures for heavy-tailed distributions. *Annals of Applied Statistics*, 2(2):756–776.

[2] Ickowicz, A. and Sparks, R. (2017). Modelling hospital length of stay using convolutive mixtures distributions. *Statistics in Medicine*, 36(1):122–135.

[3] Harini, S., Subbiah, M. and Srinivasan, M. R. (2017). Fitting length of stay in hospitals using transformed distributions. *Communications in Statistics - Case Studies, Data Analysis and Applications*, 4(1):1–8.

[4] Karagrigoriou, A. and Vonta, I. (2014). Statistical Inference for Heavy-Tailed Distributions in Technical Systems. In *2014 Ninth International Conference on Availability, Reliability and Security* :412–419.

[5] Jayakumar, K. and Fasna, K. (2022). On a New Generalization of Cauchy Distribution. *Asian Journal of Statistical Science*, 2(1): 61–81.

[6] Bhati, D. and Ravi, S. (2018). Evolution by gene duplication. On generalized log-Moyal distribution: A new heavy-tailed size distribution. *Insurance: Mathematics and Economics*, 79:247–259.

[7] Beirlant, J., Matthys, G. and Dierckx, G. (2001). Heavy-tailed distributions and rating. *Astin Bulletin*, 31(1):37–58.

[8] Resnick, S. I. (1997). Discussion of the Danish data on large fire insurance losses. *Astin Bulletin*, 27(1):139–151.

[9] Dutta, K. and Perry, J. (2006). Evolution by gene duplication. A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital.

[10] Adcock, C., Eling, M. and Loperfido, N. (2015). Skewed distributions in finance and actuarial science: a review. *European Journal of Finance*, 21:1253–1281.

[11] Bennett, S. (1983). Log-logistic regression models for survival data.*Applied Statistics*, 32(2):165.

[12] Lamberson, L. R. (1974). An evaluation and comparison of some tests for the validity of the assumption that the underlying distribution of life is exponential. *AIIE Transactions*, 6(4):327–337.

[13] LIN, B. S. DURAN, T. O. LEWIS (1989). Inverted gamma as a life distribution. *Microelectronics Reliability*, 29(4):619–626.

[14] V. K. SHARMA, S. K. SINGH, U. SINGH, V. AGIWAL (2015). The inverse Lindley distribution: A stress-strength reliability model with application to head and neck cancer data. *Journal of Industrial and Production Engineering*, 32(3): 162–173.

[15] Oguntunde, P. and Adejumo, O. (2014). The transmuted inverse exponential distribution. *International Journal of Advanced Statistics and Probability*, 3(1):1–7.

[16] Khan, M. S., King, R., and Hudson, I. (2014). Characterizations of the transmuted inverse Weibull distribution. *ANZIAM Journal*, 55:197–217.

[17] Sharma, V. K., Singh, S. K., and Singh, U. (2014). A new upside-down bathtub shaped hazard rate model for survival data analysis. *Applied Mathematics and Computation*, 239:242–253.

[18] Sharma, V. K., Singh, S. K., Singh, U., and Merovci, F. (2016). The generalized inverse Lindley distribution: A new inverse statistical model for the study of upside-down bathtub data. *Communications in Statistics - Theory and Methods*, 45(19):5709–5729.

[19] Modi, K., Kumar, D., and Singh, Y. (2020). A new Family of Distribution with Application on Two Real Datasets on Survival Problem. *Science and Technology Asia*, 25: 1–10.

[20] Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 75(371): 667-672.

[21] Galton, F. (1883). Inquiries into Human Faculty and its Development,*Macmillan and Company*, London. (25).

[22] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society, Series D* , 37(1):25–32.

[23] Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by Lehmann alternatives.*Communications in Statistics - Theory and Methods*, 27(4):887–904.

[24] Klugman, S.A, Panjer, H.H, Willmot, G.E.(2012) Loss Models: From Data to Decisions, vol. 715,*John Wiley and Sons*, New York.

[25] Nair,J., Wierman, A.and Zwart,B.(2013). The fundamentals of heavy-tails: Properties, emergence, and identification. In *ACM SIGMETRICS Performance Evaluation Review*, 41:387–388.

[26] Mann, N., Schafer, R., and Singpurwalla, N. (1974). Methods for Statistical Analysis of Reliability and Life Data. *Wiley*.

[27] Abouammoh, A., Abdulghani, S., Qamber, I. (1994). On partial orderings and testing of new better than renewal used classes.*Reliability Engineering and System Safety*, 43:37–41.

[28] Efron, B. (1988). Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of the American Statistical Association*, 83:414–425.

[29] Mudholkar, G. S., Srivastava, D. K., and Kollia, G. D. (1996). A Generalization of the Weibull Distribution with Application to the Analysis of Survival Data. *Journal of the American Statistical Association*, 91.

A NOVEL METHODOLOGY IN DEVELOPING STRESS-STRENGTH RELIABILITY MODEL FOR WEIBULL DISTRIBUTION: A COMPARISON OF ARTIFICIAL NEURAL NETWORK (ANN) AND RESPONSE SURFACE ANALYSIS (RSA)

Dr. Saurabh L. Raikar[1], Prof. Rajesh S. Prabhu Gaonkar[2]

[1]Mechanical Engineering Department, Goa College of Engineering (affiliated to Goa University), Goa
[2]Indian Institute of Technology Goa (IIT Goa), Farmagudi, Goa
[1]saurabhr3003@gmail.com , [2]rpg@iitgoa.ac.in

Abstract

Stress strength interference theory is widely used in evaluating the reliability of mechanical components. Various interference models have been developed when stress and strength follow a wide range of distributions. But when stress and strength follow Weibull distribution, a closed form of interference model is not available. This paper deals with developing a methodology for obtaining a closed form interference model for a given application when the stress and strength follow Weibull distribution. The method of artificial neural network (ANN) and response surface analysis (RSA) are used in modelling and analysis. The validation experiment has been conducted and the error obtained shows that the proposed methodology performs reasonably well.

Keywords: Weibull distribution, stress strength interference, reliability, ANN, RSM

I. Introduction

Reliability is gaining increasing importance in recent years as it takes into account the uncertainty present in the properties. One of the most common theories used in estimating reliability is the stress strength interference theory. The theory says that if X and Y follow a particular distribution, then their interference area gives the probability of failure [1]. Many interference models have been developed when stress and strength follow various distributions like normal, lognormal, exponential, etc. But, when stress and strength follow Weibull distribution, the closed form of interference model is not available. Weibull distribution is widely used statistical and reliability studies because of its ability to fit a wide range of data.

II. Stress Strength interference models

Reliability models have been developed when strength and stress are seen to be following normal, lognormal, exponential distribution, etc. [2]. This section presents reliability models for some of the most widely used distributions.

I. Reliability when stress and strength follow normal distribution

Consider the strength (random variable X) and stress (random variable Y) follow normal

distribution with pdf

$$f(x) = \frac{1}{\delta_x\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\mathrm{x}-\mu_{nx}}{\delta_x}\right)^2\right) \quad (1)$$

and

$$f(y) = \frac{1}{\delta_y\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\mathrm{y}-\mu_{ny}}{\delta_y}\right)^2\right) \quad (2)$$

respectively, where μ_{nx} is the mean of strength, δ_x is the standard deviation of strength, μ_{ny} is the mean of stress and δ_y is the standard deviation of strength. As per the interference theory, the reliability of the system will be equal to [3]

$$R = \int_0^\infty \frac{1}{\delta_x\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\mathrm{x}-\mu_{nx}}{\delta_x}\right)^2\right)\left(\int_0^x \frac{1}{\delta_y\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\mathrm{y}-\mu_{ny}}{\delta_y}\right)^2\right)dy\right)dx \quad (3)$$

On simplifying the above equation, the reliability can be obtained as

$$R = \phi\left(\frac{\mu_{nx}-\mu_{ny}}{\sqrt{{\delta_x}^2+{\delta_y}^2}}\right) \quad (4)$$

II. Reliability when stress and strength follow lognormal distribution

Consider that the strength (random variable X) and stress (random variable Y) follow lognormal distribution with pdf

$$f(x) = \frac{1}{\mathrm{x}.\,\delta_x\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(\mathrm{x})-\mu_{nx}}{\delta_x}\right)^2\right) \quad (5)$$

and

$$f(y) = \frac{1}{\mathrm{y}.\,\delta_y\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(\mathrm{y})-\mu_{ny}}{\delta_y}\right)^2\right) \quad (6)$$

respectively, where μ_{nx} is the mean and δ is the standard deviation of ln(X), μ_{ny} is the mean and δ_y is the standard deviation of ln(Y). As per the interference theory, the reliability of the system will be equal to [4]

$$R = \int_0^\infty \frac{1}{\mathrm{x}.\,\delta_x\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(\mathrm{x})-\mu_{nx}}{\delta_x}\right)^2\right)\left(\int_0^x \frac{1}{\mathrm{y}.\,\delta_y\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(\mathrm{y})-\mu_{ny}}{\delta_y}\right)^2\right)dy\right)dx \quad (7)$$

On simplifying the above equation, the reliability can be obtained as

$$R = \phi\left(\frac{\mu_{nx}-\mu_{ny}}{\sqrt{{\delta_x}^2+{\delta_y}^2}}\right) \quad (8)$$

III. Reliability when stress and strength follow exponential distribution

Consider that the strength (random variable X) and stress (random variable Y) follow lognormal distribution with pdf

$$f(x) = \lambda_x \, e^{-\lambda_x x} \tag{9}$$

and

$$f(y) = \lambda_y \, e^{-\lambda_y y} \tag{10}$$

where λ is the rate parameter.
As per the interference theory, the reliability of the system will be equal to

$$R = \int_0^\infty \lambda_x \, e^{-\lambda_x x} \left(\int_0^x \lambda_y \, e^{-\lambda_y y} dy \right) dx \tag{11}$$

On simplifying the above equation, the reliability can be obtained as [5]

$$R = \frac{\lambda_y}{\lambda_x + \lambda_y} \tag{12}$$

IV. Reliability when stress and strength follow Weibull distribution
Consider that the strength (random variable X) and stress (random variable Y) follow Weibull distribution with pdf

$$f(x) = \frac{p_1}{{\sigma_1}^{p_1}} (x - \mu_1)^{p_1 - 1} \exp\left\{ -\left(\frac{x - \mu_1}{\sigma_1} \right)^{p_1} \right\}, x > \mu_1, \sigma_1 > 0, p_1 > 0 \tag{13}$$

and

$$f(y) = \frac{p_2}{{\sigma_2}^{p_2}} (y - \mu_2)^{p_2 - 1} \exp\left\{ -\left(\frac{y - \mu_2}{\sigma_2} \right)^{p_2} \right\}, y > \mu_2, \sigma_2 > 0, p_2 > 0 \tag{14}$$

where $\mu_1, \sigma_1 \; and \; p_1$ are the location, scale and shape parameter respectively for strength and $\mu_2, \sigma_2 \; and \; p_2$ are location, scale and shape parameter respectively for stress. As per the interference theory, the reliability of the system will be equal to

$$R = \int_0^\infty \frac{p_1}{{\sigma_1}^{p_1}} (x - \mu_1)^{p_1 - 1} \exp\left\{ -\left(\frac{x - \mu_1}{\sigma_1} \right)^{p_1} \right\} \left(\int_0^x \frac{p_2}{{\sigma_2}^{p_2}} (y - \mu_2)^{p_2 - 1} \exp\left\{ -\left(\frac{y - \mu_2}{\sigma_2} \right)^{p_2} \right\} dy \right) dx \tag{15}$$

$$R = \int_0^\infty \frac{p_1}{{\sigma_1}^{p_1}} (x - \mu_1)^{p_1 - 1} \exp\left\{ -\left(\frac{x - \mu_1}{\sigma_1} \right)^{p_1} \right\} \left(e^{-\frac{e^{p_2 \ln(-\mu_2)}}{{\sigma_2}^{p_2}}} - e^{-\frac{e^{p_2 \ln(x - \mu_2)}}{{\sigma_2}^{p_2}}} \right) dx \tag{16}$$

Equation 16 cannot be solved further and hence, the calculation of stress-strength reliability for Weibull distribution has to be solved using numerical or graphical methods [6]. This can sometimes lead to complications and is time consuming.
Similarly, the reliability models have been developed for other distributions of stress and strength. S. Nadarajah, 2003 [7] developed stress-strength interference for stress and strength following lifetime distributions i.e. exponential and gamma distribution. He has also developed a reliability model for stress and strength following bivariate gamma distribution [8]. Patowary et al., 2013 [9] studied and proposed a mathematical model for stress-strength reliability for stress and strength following mixture of distributions. An inference on reliability was also drawn, stating standby redundancy aids in achieving high reliability. An et al., 2008 [10] developed a discrete stress-strength interference model based on universal generating function. K. Shen, 1992 [11] proposed a new empirical approach based on the subinterval probabilities of stress and strength in the

interference region to compute the unreliability bounds. Kotz et al., 2003 [12] reviewed the stress-strength interference models and showed practical results in application of stress-strength interference concepts in industrial systems. Many studies have been carried out in developing stress-strength reliability models for various distributions. However, it has been identified that the reliability model for stress and strength following Weibull distributions has not been developed yet.

III. Design of experiments

Design of experiments (DOE) is a systematic tool to find the relations between the input variables and the response. DOE gives a significant experimental setup sufficient to find the relation between input variables and output response which helps in saving time, cost and resources. Taguchi method and response surface analysis (RSA) are some of the widely used techniques in DOE. Taguchi method is used to obtain a set of significant experiments and to find the most influential parameter towards the response. RSA is also used to obtain a set of significant parameters and analyze for the influencing parameters. Additionally, RSA gives a prediction model of input variables and the output response. Khare et al, 2018 [13] used DOE in optimizing the surface roughness of AA 6061 material in turning operation. The cutting speed, feed rate, depth of cut and rake angle were taken as the input parameters while the surface roughness was taken as the response. Taguchi's method was used for carrying out the DOE and analysis. It was found that the rake angle was the most influential parameter towards the surface roughness followed by cutting speed. The set of optimum parameters was also found. Similarly Taguchi analysis has been used by many researchers in their studies [14]–[17]. Laghari et al., 2018 [18] developed prediction models for tool wear and surface roughness in turning of Al/SiCp workpiece using response surface analysis (RSA). Cutting speed, feed rate and depth of cut were taken as the independent variables towards the response. The response surface analysis was proved to be effective in modeling the prediction equation. Ammeri et al., 2015 [19] combined Taguchi method and RSA in determining the optimal lot size for the manufactured product in supply chain. RSA has been used to develop models and carrying out analysis of parameters and response [20]–[23]. Nair et al., 2004 [24] used design of experiments for design of accelerated test experiments for reliability improvement. Rigdon et al., 2022 [25] studied on the use of design of experiments to understand and improve product reliability. A detailed description of the statistical distributions, methods to model reliability, various DOE methods that can be used and the analysis that can be carried out has been made in this research.

IV. Weibull stress strength model

Weibull distribution is most commonly used to describe mechanical systems. The interference model of reliability when the stress and strength follow Weibull distribution is given as:

$$R = \int_0^\infty \frac{p_2}{\sigma_2}\left(\frac{\mathrm{s}-\mu_2}{\sigma_2}\right)^{p_2-1} e^{-\frac{(s-\mu_2)}{\sigma_2}} \left[\int_0^s \left(\frac{l-\mu_1}{\sigma_1}\right)^{p_1-1} e^{-\frac{(l-\mu_1)}{\sigma_1}} dl\right] ds \tag{17}$$

The integration of the above model is complicated and does not have a closed form. This study attempts to obtain the closed form of stress-strength reliability using design of experiments (DOE) when the stress and strength follow Weibull distribution. Minitab 16 was the software used to conduct design of experiments and analysis. The design chosen was L27 for 6 factors of 3 level each. The parameters chosen for DOE are shown in Table 1. The results of design of experiments with response are displayed in Table 2. The stress-strength equation was partly solved manually

and partly using Wolfram Mathematica software.

Table 1 *Factors and levels for stress and strength following Weibull distribution*

Distribution	Factors	Levels
Stress	Shape Parameter: p_1	0.5, 1.5, 2.5
	Scale Parameter: σ_1	1, 1.5, 2
	Location Parameter: μ_1	0, 0.5, 1
Strength	Shape Parameter: p_2	1.5, 2.5, 3.5
	Scale Parameter: σ_2	1, 1.5, 2
	Location Parameter: μ_2	1.5, 2, 2.5

Table 2 *Design of experiments for stress and strength following Weibull distribution*

Stress			Strength			Reliability
p1	σ1	μ1	p2	σ2	μ2	R
0.5	1	0	1.5	1	1.5	0.78155
0.5	1	0	1.5	1.5	2	0.83273
0.5	1	0	1.5	2	2.5	0.86709
0.5	1.5	0.5	2.5	1	1.5	0.67035
0.5	1.5	0.5	2.5	1.5	2	0.74267
0.5	1.5	0.5	2.5	2	2.5	0.79106
0.5	2	1	3.5	1	1.5	0.56323
0.5	2	1	3.5	1.5	2	0.65846
0.5	2	1	3.5	2	2.5	0.72007
1.5	1	0.5	3.5	1	2	0.97066
1.5	1	0.5	3.5	1.5	2.5	0.996097
1.5	1	0.5	3.5	2	1.5	0.979634
1.5	1.5	1	1.5	1	2	0.722195
1.5	1.5	1	1.5	1.5	2.5	0.8851964
1.5	1.5	1	1.5	2	1.5	0.743035
1.5	2	0	2.5	1	2	0.815387
1.5	2	0	2.5	1.5	2.5	0.91893406
1.5	2	0	2.5	2	1.5	0.85073635
2.5	1	1	2.5	1	2.5	0.99773487
2.5	1	1	2.5	1.5	1.5	0.91938117
2.5	1	1	2.5	2	2	0.993511113
2.5	1.5	0	3.5	1	2.5	0.99883877
2.5	1.5	0	3.5	1.5	1.5	0.97996489
2.5	1.5	0	3.5	2	2	0.998682329
2.5	2	0.5	1.5	1	2.5	0.87792976
2.5	2	0.5	1.5	1.5	1.5	0.6762277
2.5	2	0.5	1.5	2	2	0.87038164

I. Response surface analysis

Response surface analysis has been successfully implemented in modelling studies. In this article, the response surface analysis was carried out to model an equation for stress and strength following Weibull distribution and to study the two-parameter interaction towards the reliability. The equation for reliability model developed using response surface analysis is

$$\begin{aligned} R = {} & 0.879 + 0.1346\ p1 - 0.6489\ \sigma1 - 0.3621\ \mu1 + 0.2496\ p2 + 0.5987\ \sigma2 - 0.4066\ \mu2 \\ & - 0.04577\ p1{*}p1 + 0.0502\ \sigma1{*}\ \sigma1 + 0.0215\ \mu1{*}\ \mu1 - 0.01542\ p2{*}\ p2 - 0.0504\ \sigma2{*}\ \sigma2 \\ & + 0.0009\ \mu2{*}\ \mu2 - 0.00085\ p1{*}\ \sigma2 + 0.04883\ p1{*}\mu2 - 0.0652\ \sigma1{*}\ \sigma2 + 0.2211\ \sigma1{*}\ \mu2 \\ & - 0.0930\ \mu1{*}\ \sigma2 + 0.1933\ \mu1{*}\mu2 - 0.0925\ p2{*}\ \sigma2 \end{aligned} \tag{18}$$

The R-sq value for the above equation is 99.82% which shows that the equation can predict reliability with significantly less variability. Figure 1 shows the main effects plot for reliability. It can be seen that the reliability increases with increase in location and scale parameter of strength, and decrease with increase in location and scale parameter of stress. A notable observation that can be made is that reliability increases with increase in shape parameter of both the stress and strength distribution. Figure 2 shows the interaction plot for reliability. Figure 3 shows contour plot of two parameter interaction towards reliability for stress and strength following Weibull distribution. When the parameters are held at middle values from the levels considered, a high reliability greater than 0.9 is obtained when parameter set lies in a region inscribed by the origin and the following as shown in the figure: p1 greater than 1.4 and σ1 on the minimum side preferably lesser than 1.4 in p1 x σ1 interaction, μ1 lesser than 0.25 and p1 greater than 1.7 in μ1 x p1 interaction, μ2 greater than 2.2 and p1 greater than 1.7 in μ2 x p1 interaction, μ1 lesser than 0.5 and σ1 lesser than 1.4 in μ1 x σ1 interaction, p2 greater than 2.4 and σ1 on the minium side in p2 x σ1 interaction, σ2 greater than 1.5 and σ1 lesser than 1.25 in σ2 x σ1 interaction, σ1 close to 1 in μ2 x σ1 interaction, and σ2 greater than 1.75 and μ1 lesser than 0.2 σ2 x μ1 interaction.

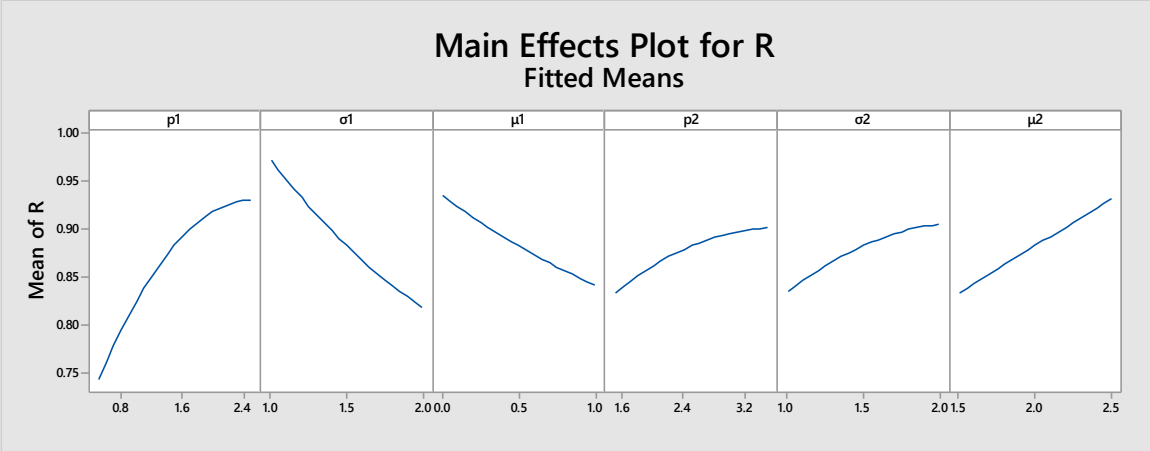


Figure 1 *Main effects plot for reliability for Weibull stress and strength*

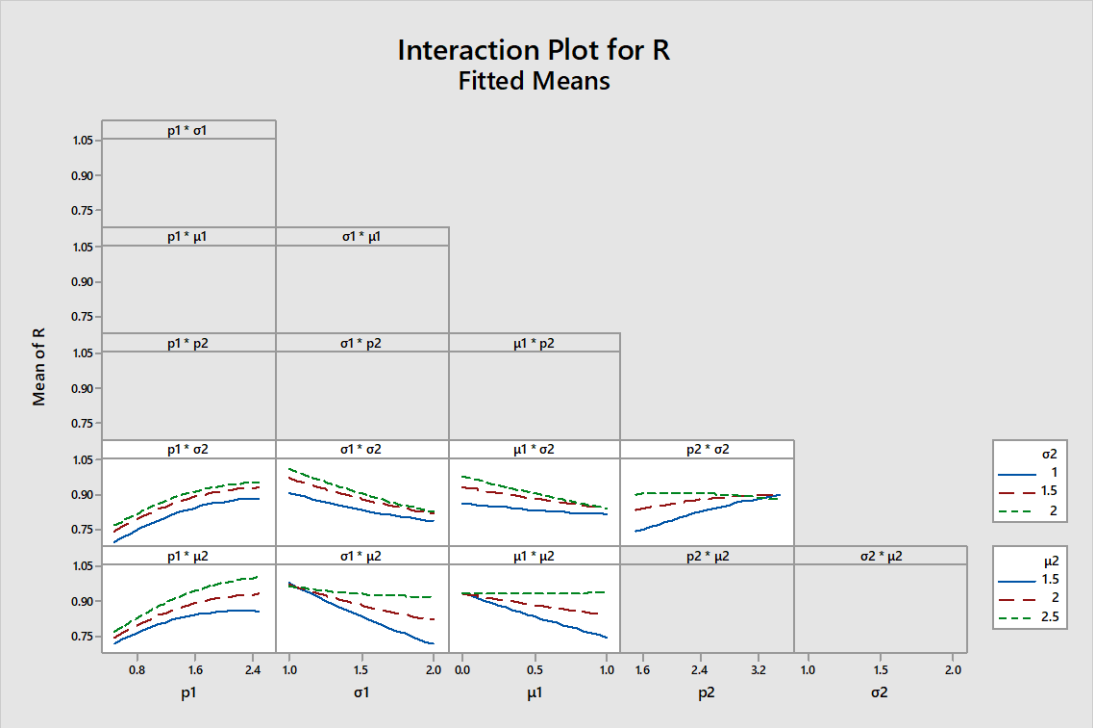


Figure 2 *Interaction plot for reliability for Weibull stress and strength*

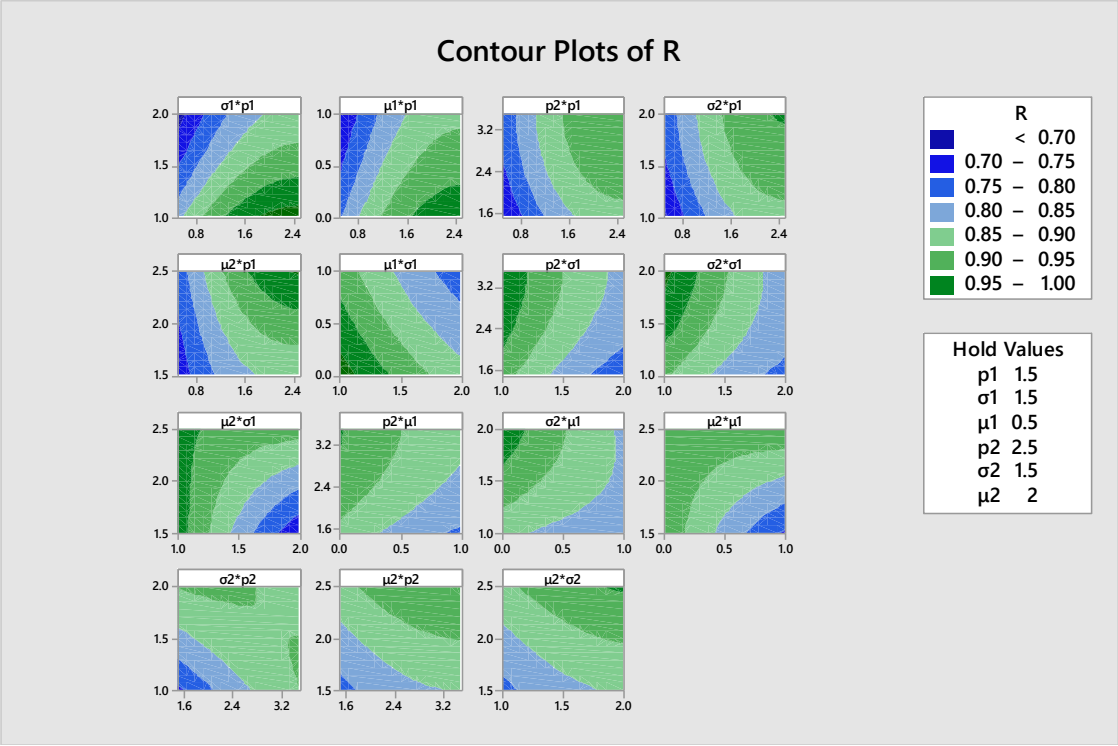


Figure 3 *Contour plot for reliability in case of stress and strength following Weibull distribution*

II. Artificial Neural Network (ANN)

Artificial Neural Network is a modelling technique based on artificial neurons which are a set of interconnected units or nodes that loosely resemble the neurons in a biological brain. Each link has

the ability to communicate with other neurons, much like the synapses in a human brain. An artificial neuron receives a signal, processes it and sends a signal to the neurons connected to it. Each neuron computes its output using a non-linear function of the sum of its inputs, where the "signal" at a connection is a real value. The method of ANN has been used in many recent applications and has shown great performance in modelling studies [26]. Abbasi et al., 2008 [27] deployed ANN method in estimation of the three parameters of Weibull distribution and obtained satisfactory results. Also, the authors compared the method with other techniques used in the application and concluded that the application of ANN is easier compared to the other methods.
In this study, the inputs to ANN model are the parameters of the distribution and the output is stress-strength reliability. ANN model consisting of 6 nodes in the input layer, a single hidden layer with 5 nodes and a single node output layer has been selected and is depicted in Figure 4.

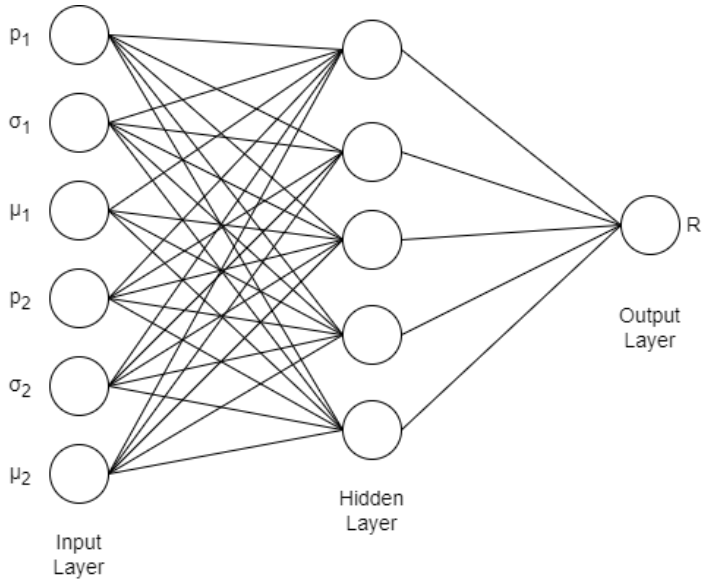


Figure 4 *ANN model with an input layer, hidden layer and output layer*

The model equation is given by

$$y(output) = b_2 + \sum_{k=1}^{h}\left[w_k f_{sig}\left(b_{1k} + \sum_{i=1}^{m} w_{ik}x_i\right)\right] \tag{19}$$

where,
y is the output, xi is the input of ith node, k is the number of nodes in hidden layer,
k is the number of nodes in the hidden layer.
wik is the weight connecting ith node in the input layer and kth node in the hidden layer.
wk is the weight connecting the kth node in the hidden layer and the output node.
b1k is the bias of the kth node in the hidden layer, b2 is the bias of the output node.
The function f_{sig} is given by

$$f_{sig}(a) = \frac{2}{1+e^{-2a}} - 1 \tag{20}$$

Around 70% of the data points have been used in training, 15% for validation and 15% for testing. Levenberg-Marquardt algorithm has been used and the training is conducted till the regression coefficients for training, validation and testing are more than 0.98 and the MSE is less than 10-3. The regression plots for training, validation and testing are shown on Figure 5. Weights and biases of the trained neural network are as follows:

$$W_{ik} = \begin{bmatrix} 0.643333 & 0.821043 & 0.609942 & -0.82412 & -0.56263 & -1.03838 \\ -1.41695 & 0.404932 & -1.0738 & -0.43627 & 0.54569 & 1.152087 \\ 2.417915 & -0.47916 & 0.179956 & 1.587887 & 0.295953 & 0.896151 \\ 0.212642 & -0.16882 & -2.00698 & -0.05127 & 0.636396 & 0.046588 \\ 1.813323 & -0.9575 & -0.69993 & -0.90945 & 0.643434 & -0.25748 \end{bmatrix}$$

$$b_{1k} = \begin{bmatrix} -1.79789 \\ 0.793825 \\ -0.1424 \\ -1.06508 \\ 1.602905 \end{bmatrix} \qquad b_2 = [-0.31431] \qquad W_k = \begin{bmatrix} -0.50302 \\ 0.031225 \\ 0.308233 \\ -0.02231 \\ 0.502355 \end{bmatrix}$$

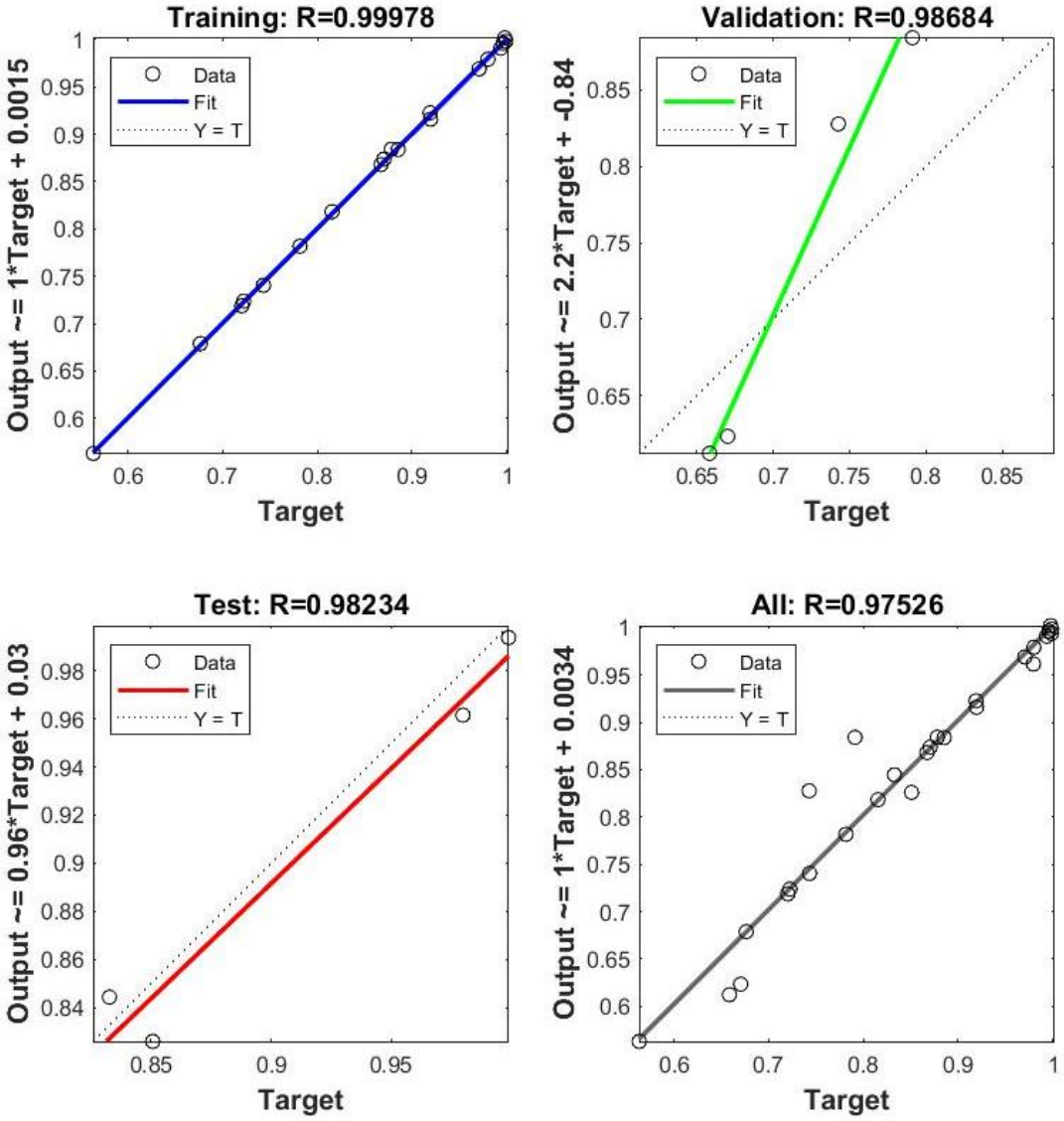


Figure 5 *Regression plot for ANN model*

III. Validation experiment

Two parameter sets are considered for the validation experiment. The parameter values are as shown below:

$p1 = 2.5 \quad \sigma1 = 1 \quad \mu1 = 0 \quad p2 = 3.5 \quad \sigma2 = 2 \quad \mu2 = 2$

$p1 = 0.5 \quad \sigma1 = 1 \quad \mu1 = 0 \quad p2 = 2.5 \quad \sigma2 = 2 \quad \mu2 = 2.5$

The results of the validation experiment using RSA and ANN have been depicted in Table 3 & Table 4 respectively. A set of 100, 1000 and 10000 numbers were generated for both the stress and strength distributions. Simulation was carried out using Matlab software. The distribution plot of corresponding stress-strength interference is shown in Figure 6. The error in probability obtained from both the models and that obtained from simulation is below 1% and is depicted in Figure 7 for RSA and Figure 8 for ANN. The optimization of equation 18 was carried out and the reliability obtained was 0.99999 with the parameter set $p1 = 2.5$, $\sigma1 = 1$, $\mu1 = 0$, $p2 = 1.5$, $\sigma2 = 1$, $\mu2 = 2.5$. The distribution plot with the optimum set of parameters is shown in Figure 6.

Table 3 Results of validation experiment for stress and strength following Weibull distribution using RSM

Sr No.	Parameters	Sample Size	Reliability using Simulation	R estimated as per proposed model	Bias	Error (%)
1	$p1 = 2.5$, $\sigma1 = 1$ $\mu1 = 0$, $p2 = 3.5$ $\sigma2 = 2$, $\mu2 = 2$	100	0.9876	0.99633	0.00873	0.88396
		1000	0.99143	0.99633	0.0049	0.49424
		10000	0.99612	0.99633	0.00021	0.02108
2	$p1 = 0.5$, $\sigma1 = 1$ $\mu1 = 0$, $p2 = 2.5$ $\sigma2 = 2$, $\mu2 = 2.5$	100	0.8732	0.8687	-0.0045	0.5153
		1000	0.87078	0.8687	-0.0021	0.24116
		10000	0.86973	0.8687	-0.0010	0.11843

Table 4 Results of validation experiment for stress and strength following Weibull distribution using ANN

Sr No.	Parameters	Sample Size	Reliability using Simulation	R estimated as per proposed model	Bias	Error (%)
1	$p1 = 2.5$, $\sigma1 = 1$ $\mu1 = 0$, $p2 = 3.5$ $\sigma2 = 2$, $\mu2 = 2$	100	0.9876	0.9961	0.0085	0.86067
		1000	0.99143	0.9961	0.00467	0.471037
		10000	0.99612	0.9961	-0.00002	0.002008
2	$p1 = 0.5$, $\sigma1 = 1$ $\mu1 = 0$, $p2 = 2.5$ $\sigma2 = 2$, $\mu2 = 2.5$	100	0.8732	0.8663	-0.0069	0.790197
		1000	0.87078	0.8663	-0.00448	0.514481
		10000	0.86973	0.8663	-0.00343	0.394375

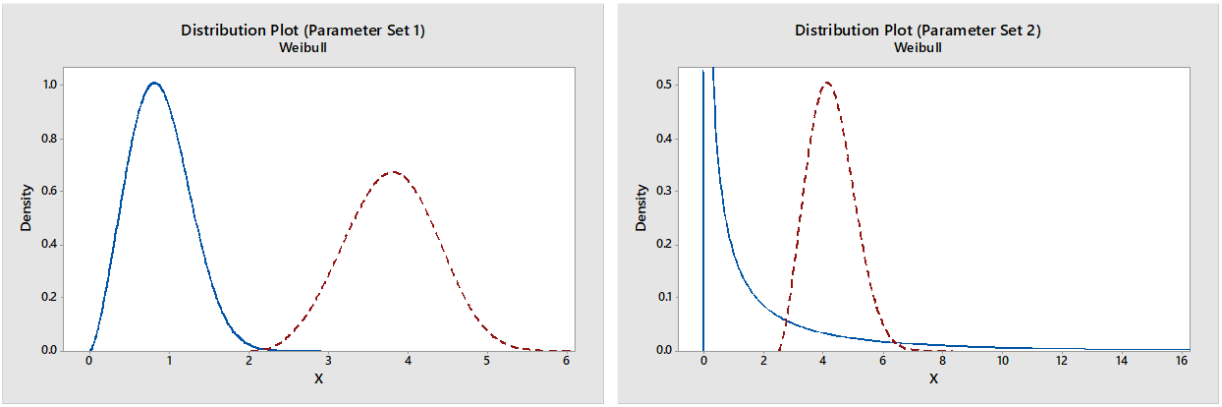


Figure 6 *Distribution plots for stress and strength following Weibull distribution*

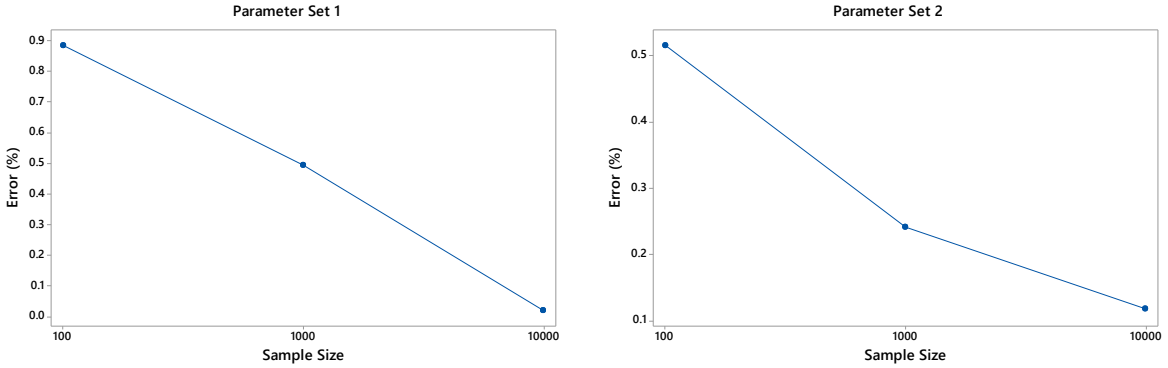


Figure 7 *Plot of error in estimation of stress-strength reliability for Weibull distribution using RSA*

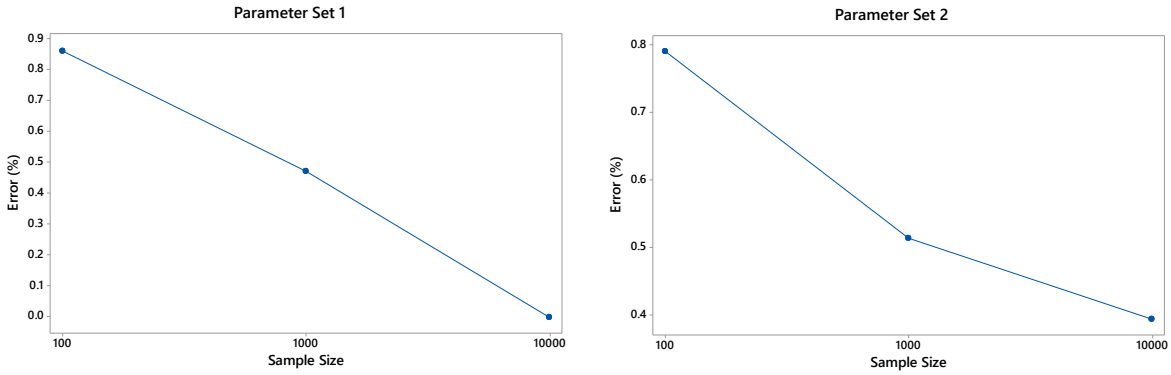


Figure 8 *Plot of error in estimation of stress-strength reliability for Weibull distribution using ANN*

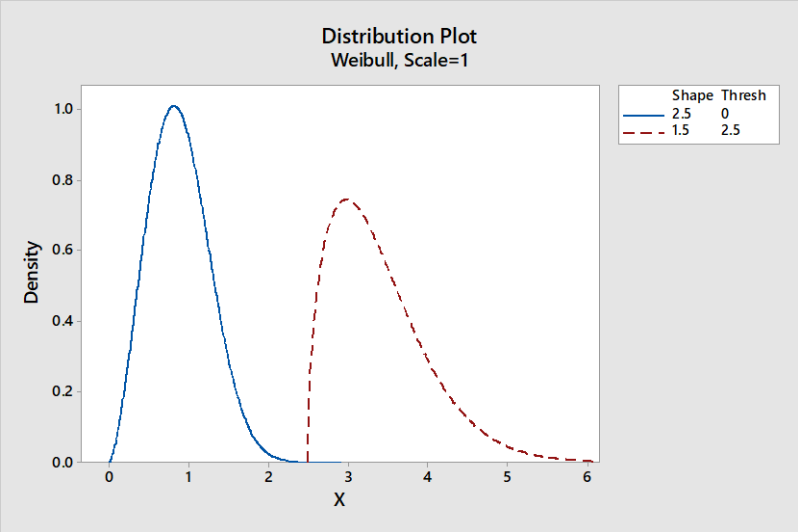


Figure 8 *Stress-strength interference with optimum set of parameters for Weibull distribution*

V. Conclusion

In this study an attempt has been made to derive a methodology in obtaining the closed form stress strength interference model for stress and strength following Weibull distribution. Design of experiments approach has been used and analysis has been carried out using Taguchi and RSA method. The reliability model has been developed for the considered set of parameters. The validation experiments with errors obtained less than 1% shows that the proposed method performs well. The optimum set of parameters was determined for the given range having reliability of 0.9999. Further studies can be carried out in this area by considering the different range of parameters and its effect on the reliability model. Also, considering dynamic nature of strength as seen in many cases, the studies can be accordingly modified.

References

[1] Raikar, S. L., & Gaonkar, R. S. P. (2022). Jaya algorithm in estimation of P[X>Y] for two parameter weibull distribution. *AIMS Mathematics*, *7*(2), 2820–2839.

[2] B. S. Dhillon. (1988). Mechanical reliability : theory, models, and applications. *American Institute of Aeronautics and Astronautics*, Wiley.

[3] Patrick D.T. O'Connor. (2013). Practical Reliability Engineering. *Wiley India*, 4th edition.

[4] Bridget, E., & Abiodun, M. (2017). Inference on Stress-Strength Reliability for Log-Normal Distribution based on Lower Record Values, *AMSE IIETA*, 22(July), 77–97.

[5] Dhillon, B. S. (1980). Stress-Strength Reliability Models. *Miroelectronics Reliability*, 20(2), 513–516.

[6] Rao, S. S. (2016). Reliability Engineering. Pearson.

[7] Nadarajah, S. (2003). Reliability for lifetime distributions. *Mathematical and Computer Modelling*, 37(7–8), 683–688.

[8] Nadarajah, Saralees. (2004). Reliability for laplace distributions. 2(October 2003), 169–183.

[9] Patowary, A. N., Hazarika, J., & Sriwastav, G. L. (2013). Interference theory of reliability: A review. *International Journal of Systems Assurance Engineering and Management*, 4(2), 146–158.

[10] An, Z. W., Huang, H. Z., & Liu, Y. (2008). A discrete stress-strength interference model based on universal generating function. *Reliability Engineering and System Safety*, 93(10), 1485–1490.

[11] Shen, K. (1992). An empirical approach to obtaining bounds on the failure probability

through stress/strength interference. *Reliability Engineering and System Safety*, 36(1), 79–84.

[12] Samuel Kotz, Lumelskii, Y., & Marianna Pensky. (2003). The Stress–Strength Model and Its Generalizations. *World Scientific*.

[13] Khare, S. K., Agarwal, S., & Srivastava, S. (2018). Analysis of Surface Roughness during Turning Operation by Taguchi Method. *Materials Today: Proceedings*, *5*(14), 28089–28097

[14] Oemar, B., & Chang, W. C. (2020). Taguchi method for optimizing process parameters in the production of activated carbon from rubber seed shell. *International Journal of Advanced Manufacturing Technology*, *107*(11–12), 4609–4620

[15] Jagatheesan, K., & Babu, K. (2020). Taguchi optimization of minimum quantity lubrication turning of AISI-4320 steel using biochar nanofluid. *Biomass Conversion and Biorefinery*.

[16] Sah, S. K., Murugesan, K., & Elangovan, R. (2021). Optimization of energy consumption for indoor climate control using Taguchi technique and utility concept. *Science and Technology for the Built Environment*, *27*(10), 1473–1491.

[17] Bement, T. R. (1989). Taguchi Techniques for Quality Engineering. *Technometrics*, *31*(2), 253–255.

[18] Laghari, R. A., Li, J., Xie, Z., & Wang, S. qi. (2018). Modeling and Optimization of Tool Wear and Surface Roughness in Turning of Al/SiCp Using Response Surface Methodology. *3D Research*, *9*(4).

[19] Ammeri, A., Chabchoub, H., Hachicha, W., & Masmoudi, F. (2015). Combining Taguchi approach and response surface methodolgy for lot-sizing problem in MTO environnement. *2015 4th IEEE International Conference on Advanced Logistics and Transport, IEEE ICALT 2015*, 252–256.

[20] Bayuo, J., Abukari, M. A., & Pelig-Ba, K. B. (2020). Optimization using central composite design (CCD) of response surface methodology (RSM) for biosorption of hexavalent chromium from aqueous media. *Applied Water Science*, *10*(6).

[21] Hazir, E., & Ozcan, T. (2019). Response Surface Methodology Integrated with Desirability Function and Genetic Algorithm Approach for the Optimization of CNC Machining Parameters. *Arabian Journal for Science and Engineering*, *44*(3), 2795–2809.

[22] Patel, K. A., & Brahmbhatt, P. K. (2018). Response Surface Methodology based Desirability Approach for Optimization of Roller Burnishing Process Parameter. *Journal of The Institution of Engineers (India): Series C*, *99*(6), 729–736.

[23] Myers, R. H., Montgomery, D. C., Geoffrey Vining, G., Borror, C. M., & Kowalski, S. M. (2004). Response Surface Methodology: A Retrospective and Literature Survey. *Journal of Quality Technology*, *36*(1), 53–78.

[24] Nair, V. N., Escobar, L. A., & Hamada, M. S. (2004). Design and Analysis of Experiments for Reliability Assessment and Improvement. In *Mathematical Reliability: An Expository Perspective*. Springer Science + Business Media New York.

[25] Rigdon, S. E., Pan, R., Montgomery, D. C., & Borror, C. M. (2022). *Design of Experiments for Reliability Achievement*, Wiley.

[26] Mohan, S., Dinesha, P., & Campana, P. E. (2022). ANN-PSO aided selection of hydrocarbons as working fluid for low-temperature organic Rankine cycle and thermodynamic evaluation of optimal working fluid. *Energy*, *259*(June), 124968.

[27] Abbasi, B., Rabelo, L., & Hosseinkouchack, M. (2008). Estimating parameters of the three-parameter Weibull distribution using a neural network. *European Journal of Industrial Engineering*, *2*(4), 428–445.

RT&A, No 4 (76)
Volume 18, December 2023

RELIABILITY MODELLING OF UTENSILS MANUFACTURING SYSTEM WITH TEMPERATURE DEPENDENT MAINTENANCE

Manisha Gaba

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
manishagaba887@gmail.com

Dalip Singh

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
dsmdur@gmail.com

Sheetal

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
rtksheetal@gmail.com

Kajal Sachdeva*

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
kajal.rs.maths@mdurohtak.ac.in

*Corresponding Author

Abstract

In this paper, a stochastic model for utensils manufacturing system with preventive maintenance (PM) is analysed in detail. The operation is affected by variation in the temperature dependent maintenance. The entire manufacturing process of utensils goes through four subsystems viz., Circle cutting subsystem 1, Pressing subsystem 2, Spinning subsystem 3 and Polishing & Packing 4. The system has series structure of all the subsystems. The system is put under PM on the winter time and after PM it operates as new. The PM time distributions are considered as arbitrary and the time to failure as well as repair of each subsystem follows a negative exponential distribution. All random variables are statistically independent. Several measures for evaluating the effectiveness of a system, including mean time to system failure (MTSF), system availability (in summer and winter), busy period of repairman and expected number of repairs (in summer and winter) are derived using a regenerative point technique and Markov process. The system is also analysed for particular values of the parameters.

Keywords: Utensils Manufacturing System; MTSF; Availability; Regenerative Point Technique, and Preventive Maintenance (PM).

1. Introduction

Over the years researchers have made significant contribution to the reliability field. With the advent of advanced technological system, the expectations of the people have increased extremely for the use of flawless system at least for considerably period. To cater the demand

and expectation of the people, researcher have developed various stochastic models considering the aspects of different repair and obtaining performance affecting measures which include Teng et al. [1], Yusuf and Yusuf [2], Manocha and Taneja [3], Gupta et al. [4], Fagge et al. [5], Rajesh et al. [6], Kumar and Malik [7], Rajesh and Taneja [8], Jain and Malik [9], Rahbi et al. [10], Sheetal and Taneja [11, 12], Sachdeva et al. [13], Rizwan et al. [14, 15]. Singh and Mahajan [16] studied reliability of utensils manufacturing plant-a case study. Kumar and Kumar [17] studied mathematical modelling of stainless-steel utensils manufacturing plant using fuzzy reliability. This research investigated several failure modes in order to increase system reliability. There may be instances where the system's operation is impacted by temperature-dependent maintenance. Thus, the objective of this paper is to make a contribution in this regard because none of the research described above examined the impact of temperature-dependent maintenance on the operation of the system.

Despite playing a significant part in our daily lives, the manufacturing facility for utensils with preventative maintenance has not yet been covered. With this in mind, the present study took into account four subsystems of the manufacturing facility for utensils, with constant rates of subsystem failure and repair, and it covered stochastic modelling of the facility for utensils with preventive maintenance using the regenerative point technique and Markov process. As a result, PM of the unit is required after a certain amount of time to increase reliability and availability. Also, an effort was made to discuss the plant's availability in relation to various failure and repair rates.

The system needs to be maintained since low temperature harm the quality of utensils. By employing lubrications, replacement of a nut, screw, or other component of the system, cleaning, or other techniques to create high-quality utensils, the system is made operational as quickly as feasible. Preventive maintenance can therefore be used to increase system reliability and availability. It's also interesting to note that there hasn't been much work documented in the reliability literature so far on reliability modelling of the utensil manufacturing facility subject to preventive maintenance. Utensils plant can have a variety of parts but mainly the plant consists for four subsystems like cutting system, pressing system, spinning system and polishing and packing system. In winter there is low temperature spinning system goes in reduced capacity because of breakdown of rubber plate which is used in dye.

The mean time to system failure (MTSF), availability (in summer and winter), busy period analysis of repairman for repair (in summer and winter), expected number of visits of the repairman for repair (in summer and winter), and expected number of visits of the repairman for preventive maintenance are a few of the several ways that system efficacy can be measured that are obtained. Further the profit incurred to the system is obtained. Graphical representations of various intriguing system efficacy behaviours have been produced.

2. System Description, Notations and Assumptions

Utensils manufacturing plants are widely used to produce various kinds of utensils. Utensils plant can have a variety of parts but mainly the plant consists of four subsystems like cutting system, pressing system, spinning system and polishing and packing system. Manufacturing of utensils entails the press or spin forming of metal, which frequently involves complex geometries with straight sides and as well as curves of various radii. Below is a list of every system and notation needed for the mathematical formulation.

2.1. Description of the System

Sub-system M_C (Circle Cutting Machine)
As needed, sheets are cut into circular shapes.
Sub-system M_P (Pressing Machine)
The circle that was cut using a circular saw is now being sent to a pressing machine. Here, it is

pressed using various dies in accordance with the size and shape of various types of kitchenware. Due to their shallow depth, some products, including as plates and bojanthal are ready for polishing right away.

Sub-system M_S (Spinning Machine)

According to their dies, the product created by pressing is sent for spinning. Some goods don't require further annealing before polishing, but others require it because of their deeper shapes. To eliminate contaminants, these items must be subjected to acid cleaning (Acid is a combination of Sulphuric and nitric acid).

Sub-system M_D (Polishing & Packing)

The final process has produced a product that is polished-ready. This stage involves packing and polishing the final product.

2.2. Notations

$m_1(t), M_1(t)$	probability and cumulative density functions by which the system go for preventive maintenance
$m_2(t), M_2(t)$	probability and cumulative density functions for completion of preventive maintenance time
$w_1(t), W_1(t)$	probability and cumulative density functions for changing the summer to winter season
$w_2(t), W_2(t)$	probability and cumulative density functions for changing the winter to summer season
a_1, a_2, a_3, a_4	rate of failure for subsystem $M_C, M_P, M_S, M_{\overline{S}}$
$b_1, b_2, b_3, , b_4$	rate of repair for subsystem $M_C, M_P, M_S, M_{\overline{S}}$
M_C, M_P, M_S	subsystem M_C, M_P, M_S operative
$M_{CUR}, M_{PUR}, M_{SUR}, M_{\overline{S}UR}$	subsystem $M_C, M_P, M_S, M_{\overline{S}}$ under repair
$M_{\overline{S}}$	subsystem M_S in reduced state
M_{DPM}	subsystem M_D under preventive maintenance
$\odot$	Laplace Stieltjes Convolution
©	Laplace Convolution
q_{ij}	probability density function of the first passage time from regenerative state i to regenerative state j
p_{ij}	steady state transition probability from state i to state j
m_{ij}	the unconditional mean time taken to transit to any regenerative state i from the epoch of entry into regenerative state j
μ_i	mean sojourn time in the regenerative state i before transiting to any other state
$\phi_i(t)$	cumulative distribution function (c.d.f.) of the first passage time from a regenerative state i to a failed state
AS_0, AW_0	availability in summer, winter
BS_0, BW_0	busy period of the repairman due to repair in summer, winter
VS_0, VW_0	expected number of visits of repairman for repair during summer, winter
PM_0	expected number of visits of the repairman for preventive maintenance

2.3. Assumptions

- The failure and repair rates are independent and exponential in general.
- None of the sub-systems are experiencing simultaneous failures.

- Subsystem M_D has never failed.
- The repaired system works just like the new system.
- Subsystems are only repaired when they are in reduced or failed state.

3. Analysis of Model

In Fig. 1, the system model's possible transition diagram is shown.

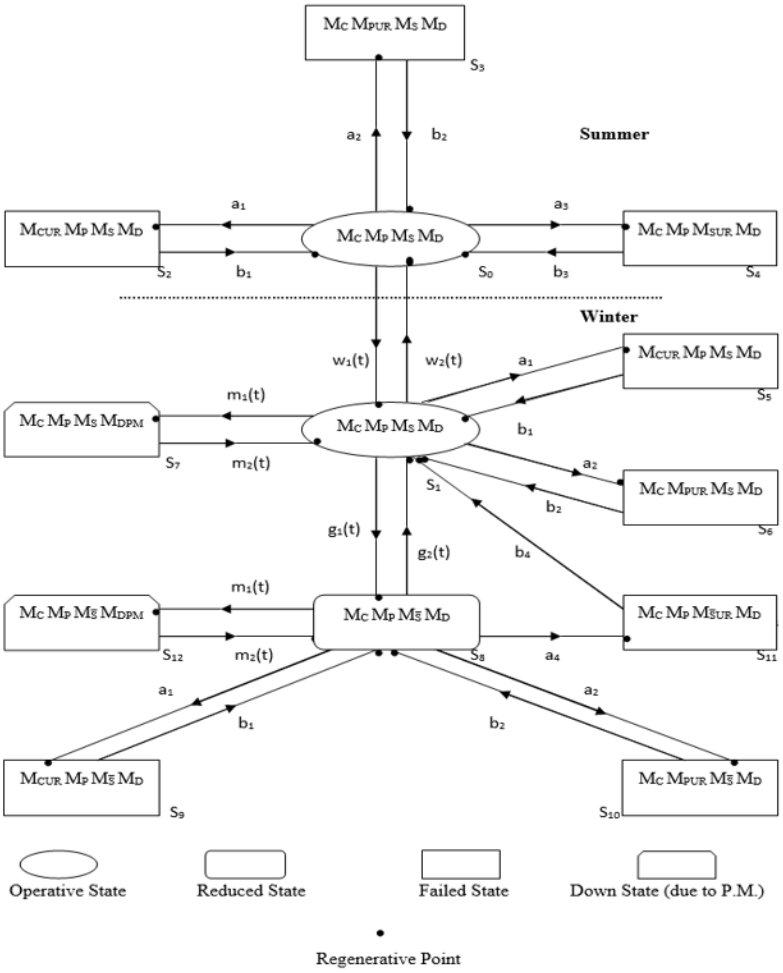


Figure 1: *State Transition Diagram*

3.1. Description of the Model and Transition Probabilities

3.1.1 Description of the model

Various states of the model for the system consisting four subsystems with season wise (summer and winter) and the state transition diagram is displayed in above Fig. 1. States 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 of the state transition diagram are regeneration points and hence these states are regenerative states for the model. States 0 and 1 are the states four subsystem work and so represents operative state during summer and winter respectively. States 2 and 5 are the states where the sub-system M_C go in failed states during summer and winter respectively so represents failed state. States 3 and 6 are the states where sub-system M_P go in failed states during summer and winter respectively so represents failed state. States 4 and 11 are the states

where sub-system M_S go in failed states during summer and winter respectively so represents failed state. State 8 is the state where sub-system M_S go in reduced capacity during winter so represents reduced state. States 9 and 10 are the states where sub-system M_C and sub-system M_P go in failed states respectively so represents failed state also in those states sub-system M_S is in reduced capacity. States 7 and 12 are the states where sub-system M_D under preventive maintenance. Also, in the state 7 the sub-system M_S work in full capacity, in the state 12 the sub-system M_S work in reduced capacity.

3.1.2 State Transition Probabilities

In Fig. 1, the system's transition diagram is depicted, indicating the various states of the system. Expressions for $q_{ij}(t)$ (for all required combinations of i and j) are found based on the state transition diagram, and same are provided below:

$q_{01}(t) = e^{-(a_1+a_2+a_3)t} w_1(t)$
$q_{02}(t) = a_1 e^{-(a_1+a_2+a_3)t} \overline{W_1}(t)$
$q_{03}(t) = a_2 e^{-(a_1+a_2+a_3)t} \overline{W_1}(t)$
$q_{04}(t) = a_3 e^{-(a_1+a_2+a_3)t} \overline{W_1}(t)$
$q_{10}(t) = e^{-(a_1+a_2)t} F_{10}(t)$
$q_{15}(t) = a_1 e^{-(a_1+a_2)t} F_1(t)$
$q_{16}(t) = a_2 e^{-(a_1+a_2+a_3)t} F_1(t)$
$q_{17}(t) = a_3 e^{-(a_1+a_2+a_3)t} F_{17}(t)$
$q_{18}(t) = e^{-(a_1+a_2+a_3)t} F_{18}(t)$
$q_{20}(t) = b_1 e^{-b_1 t}$
$q_{30}(t) = b_2 e^{-b_2 t}$
$q_{40}(t) = b_3 e^{-b_3 t}$
$q_{51}(t) = b_1 e^{-b_1 t}$
$q_{61}(t) = b_2 e^{-b_2 t}$
$q_{71}(t) = m_2(t)$
$q_{81}(t) = e^{-(a_1+a_2+a_4)t} G_{8,1}(t)$
$q_{89}(t) = a_1 e^{-(a_1+a_2+a_3)t} G_8(t)$
$q_{8,10}(t) = a_2 e^{-(a_1+a_2+a_3)t} G_8(t)$
$q_{8,11}(t) = a_4 e^{-(a_1+a_2+a_4)t} G_8(t)$
$q_{8,12}(t) = e^{-(a_1+a_2+a_4)t} G_{8,12}(t)$
$q_{98}(t) = b_1 e^{-b_1 t}$
$q_{10,8}(t) = b_2 e^{-b_2 t}$
$q_{11,1}(t) = b_4 e^{-b_4 t}$
$q_{12,8}(t) = m_2(t)$

where

$G_8(t) = \overline{M_1}(t)\overline{G_2}(t)$
$G_{8,1}(t) = \overline{M_1}(t) g_2(t)$
$G_{8,12}(t) = m_1(t)\overline{G_2}(t)$
$F_{10}(t) = \overline{M_1}(t)\overline{G_1}(t) w_2(t)$
$F_{17}(t) = m_1(t)\overline{G_1}(t)\overline{W_2}(t)$
$F_{18}(t) = \overline{M_1}(t) g_1(t)\overline{W_2}(t)$
$F_1(t) = \overline{M_1}(t)\overline{G_1}(t)\overline{W_2}(t)$

Transition probabilities $p_{ij}(t)$ from state i to state j can be calculated by taking Laplace transform of above obtained values of $q_{ij}(t)$ and then using the following mathematical relationship between p_{ij} and $q^*_{ij}(s)$
$p_{ij} = \lim_{s\to 0} q^*_{ij}(s)$
values of for all required combinations of i and j are obtained and the same are given as follows:

$p_{01} = w_1^*(a_1 + a_2 + a_3)$
$p_{02} = \frac{a_1}{(a_1+a_2+a_3)}[1 - w_1^*(a_1 + a_2 + a_3)]$
$p_{03} = \frac{a_2}{(a_1+a_2+a_3)}[1 - w_1^*(a_1 + a_2 + a_3)]$
$p_{04} = \frac{a_3}{(a_1+a_2+a_3)}[1 - w_1^*(a_1 + a_2 + a_3)]$
$p_{10} = F_{10}^*(a_1 + a_2)$
$p_{15} = a_1 F_1^*(a_1 + a_2)$
$p_{16} = a_2 F_1^*(a_1 + a_2)$
$p_{17} = F_{17}^*(a_1 + a_2)$
$p_{18} = F_{18}^*(a_1 + a_2)$
$p_{71} = m_2^*(0)$
$p_{81} = G_{81}^*(a_1 + a_2 + a_4)$
$p_{89} = a_1 G_8^*(a_1 + a_2 + a_4)$
$p_{8,10} = a_2 G_8^*(a_1 + a_2 + a_4)$
$p_{8,11} = a_4 G_8^*(a_1 + a_2 + a_4)$
$p_{8,12} = G_{8,12}^*(a_1 + a_2 + a_4)$
$p_{12,8} = m_2^*(0)$
$p_{20} = p_{30} = p_{40} = p_{51} = p_{61} = p_{98} = p_{10,8} = p_{11,1} = 1$
We may verify that
$p_{01} + p_{02} + p_{03} + p_{04} = 1$
$p_{10} + p_{15} + p_{16} + p_{17} + p_{18} = 1$
$p_{81} + p_{89} + p_{8,10} + p_{8,11} + p_{8,12} = 1$
$p_{20} = p_{30} = p_{40} = p_{51} = p_{61} = p_{71} = p_{98} = p_{10,8} = p_{11,1} = p_{12,8} = 1$

3.1.3 Mean Sojourn time (μ_i)

If T_i denotes the stay time of the system in state i, then using the following mathematical relationship between μ_i and T_i
$\mu_i = \int_0^\infty P[T_i > t]dt$
values of μ_i for all required values of i are found, and the same are provided as:
$\mu_0 = \int_0^\infty e^{-(a_1+a_2+a_3)t}\overline{W_1}(t)dt$
$\mu_1 = \int_0^\infty e^{-(a_1+a_2)t}\overline{F_1}(t)$
$\mu_2 = \mu_5 = \mu_9 = \frac{1}{b_1}$
$\mu_3 = \mu_6 = \mu_{10} = \frac{1}{b_2}$
$\mu_4 = \frac{1}{b_3}$
$\mu_7 = \mu_{12} = \int_0^\infty \overline{M_2}(t)dt$
$\mu_{11} = \frac{1}{b_4}$
The unconditional mean time (m_{ij}) which the system under consideration takes to move to state j where counting of the time starts as soon as it enters into state i can be obtained using the following mathematical relationship between m_{ij} and $q_{ij}(t)$
m_{ij}= $\int_0^\infty tq_{ij}(t)dt$,
values of for all required combinations of i and j thus obtained and given as follows:

$\mu_0 = m_{01} + m_{02} + m_{03} + m_{04}$
$\mu_1 = m_{10} + m_{15} + m_{16} + m_{17} + m_{18}$
$\mu_2 = m_{20} = \frac{1}{b_1}$
$\mu_3 = m_{30} = \frac{1}{b_2}$
$\mu_4 = m_{40} = \frac{1}{b_3}$
$\mu_5 = m_{51} = \frac{1}{b_1}$
$\mu_6 = m_{61} = \frac{1}{b_2}$
$\mu_7 = m_{71} = \int_0^\infty tm_2(t)dt$
$\mu_8 = m_{81} + m_{89} + m_{8,10} + m_{8,11} + m_{8,12}$
$\mu_9 = m_{98} = \frac{1}{b_1}$
$\mu_{10} = m_{10,8} = \frac{1}{b_2}$
$\mu_{11} = m_{11,1} = \frac{1}{b_4}$
$\mu_{12} = m_{12,8} = \int_0^\infty tm_2(t)dt$

4. System Performance Measures

4.1. Mean Time to System Failure

We retain failed states as absorbing states in order to calculate the system's MTSF. Using recursive relations for $\phi_i(t)$ can be obtained and the same are given as:
$\phi_0(t) = Q_{01}(t) \odot \phi_1(t) + Q_{02}(t) + Q_{03}(t) + Q_{04}(t)$
$\phi_1(t) = Q_{10}(t) \odot \phi_0(t) + Q_{15}(t) + Q_{16}(t) + Q_{17}(t) + Q_{18}(t) \odot \phi_8(t)$
$\phi_7(t) = Q_{71}(t) \odot \phi_1(t)$
$\phi_8(t) = Q_{81}(t) \odot \phi_1(t) + Q_{89}(t) + Q_{8,10}(t) + Q_{8,11}(t) + Q_{8,12}(t) \odot \phi_{12}(t)$
$\phi_{12}(t) = Q_{12,8}(t) \odot \phi_8(t)$
By solving these relations for $\phi_0^{**}(s)$ using the Laplace Stieltjes transformation of these relations, we get
$\phi_0^{**}(s) = \frac{N(s)}{D(s)}$,
where

$$N(s) = q^*_{8,12}(s)q^*_{12,8}(s)[(q^*_{02}(s) + q^*_{03}(s) + q^*_{04}(s))(q^*_{17}(s)q^*_{71}(s) - 1) - q^*_{01}(s)(q^*_{15}(s) + q^*_{16}(s))] + (q^*_{02}(s) + q^*_{03}(s) + q^*_{04}(s))(1 - q^*_{18}(s)q^*_{81}(s)) + q^*_{17}(s)q^*_{71}(s)) + q^*_{01}(s)(q^*_{15}(s) + q^*_{16}(s)) + q^*_{01}(s)q^*_{18}(s)(q^*_{89}(s) + q^*_{8,10}(s) + q^*_{8,11}(s))$$

$$D(s) = 1 + q^*_{01}(s)q^*_{10}(s)q^*_{8,12}(s)q^*_{12,8}(s) - q^*_{01}(s)q^*_{10}(s)q^*_{17}(s)q^*_{71}(s)q^*_{8,12}(s)q^*_{12,8}(s) - q^*_{17}(s)q^*_{71}(s) - q^*_{18}(s)q^*_{81}(s) - q^*_{8,12}(s)q^*_{12,8}(s)$$

Using above calculated value of $\phi_0^{**}(s)$, MTSF can be obtained when the system under consideration starts from the state 0 and the same is given as follows:

$T_0 = \lim_{s\to 0} \frac{1-\phi_0^{**}(s)}{s} = \frac{N}{D}$,
where
$N = \mu_0[(1-p_{17})(1-p_{8,12}) - p_{18}p_{81}] - \mu_1 p_{01}p_{8,12} + \mu_8 p_{01}p_{18} + \mu_7[p_{18}p_{8,12} + (1-p_{8,12})p_{17}]$
$D = (1-p_{8,12})(1-p_{01}p_{10}-p_{17}) - p_{18}p_{81}$

4.2. Availabilities in Summer and Winter

During Summer
To determine the availability in summer $AS_0(t)$ of the system, recursive relations thus obtained using probabilistic arguments, are given as:
$AS_0(t) = M_0(t) + q_{01}(t)\copyright AS_1(t) + q_{02}(t)\copyright AS_2(t) + q_{03}(t)\copyright AS_3(t) + q_{04}(t)\copyright AS_4(t)$
$AS_1(t) = q_{10}(t)\copyright AS_0(t) + q_{15}(t)\copyright AS_5(t) + q_{16}(t)\copyright AS_6(t) + q_{17}(t)\copyright AS_7(t) + q_{18}(t)\copyright AS_8(t)$
$AS_2(t) = q_{20}(t)\copyright AS_0(t)$
$AS_3(t) = q_{30}(t)\copyright AS_0(t)$
$AS_4(t) = q_{40}(t)\copyright AS_0(t)$
$AS_5(t) = q_{51}(t)\copyright AS_1(t)$
$AS_6(t) = q_{61}(t)\copyright AS_1(t)$
$AS_7(t) = q_{71}(t)\copyright AS_1(t)$
$AS_8(t) = q_{81}(t)\copyright AS_1(t) + q_{89}(t)\copyright AS_9(t) + q_{8,10}(t)\copyright AS_10(t) + q_{8,11}(t)\copyright AS_{11}(t) + q_{8,12}(t)\copyright AS_{12}(t)$
$AS_9(t) = q_{98}(t)\copyright AS_8(t)$
$AS_{10}(t) = q_{10,8}(t)\copyright AS_8(t)$
$AS_{11}(t) = q_{11,1}(t)\copyright AS_1(t)$
$AS_{12}(t) = q_{12,8}(t)\copyright AS_8(t)$
where,
$M_0(t) = e^{-(a_1+a_2+a_3)t}\overline{W_1}(t)$
By solving these relations for $AS_0^*(s)$ using the Laplace transform of these relations, we get
$AS_0^*(s) = \frac{N_1(s)}{D_1(s)}$
where,
$N_1(s) = M_0^*(s)[1 - q_{18}^*(s)q_{11,1}^*(s)q_{8,11}^*(s) - q_{15}^*(s)q_{51}^*(s) - q_{16}^*(s)q_{61}^*(s) - q_{17}^*(s)q_{71}^*(s) - q_{18}^*(s)q_{81}^*(s)$
$- q_{89}^*(s)q_{98}^*(s) - q_{8,10}^*(s)q_{10,8}^*(s) - q_{8,12}^*(s)q_{12,8}^*(s) + (q_{15}^*(s)q_{51}^*(s) + q_{16}^*(s)q_{61}^*(s)$
$+ q_{17}^*(s)q_{71}^*(s))(q_{89}^*(s)q_{98}^*(s) + q_{8,10}^*(s)q_{10,8}^*(s) + q_{8,12}^*(s)q_{12,8}^*(s))]$

$D_1(s) = [q_{02}^*(s)q_{20}^*(s) + q_{03}^*(s)q_{30}^*(s) + q_{04}^*(s)q_{40}^*(s)][q_{15}^*(s)q_{51}^*(s) + q_{16}^*(s)q_{61}^*(s) + q_{17}^*(s)q_{71}^*(s)$
$+ q_{18}^*(s)q_{81}^*(s) + q_{89}^*(s)q_{98}^*(s) + q_{8,10}^*(s)q_{10,8}^*(s) + q_{8,12}^*(s)q_{12,8}^*(s) - 1] + (q_{01}^*(s)q_{10}^*(s) + q_{15}^*(s)q_{51}^*(s)$
$+ q_{16}^*(s)q_{61}^*(s) + q_{17}^*(s)q_{71}^*(s))(q_{89}^*(s)q_{98}^*(s) + q_{8,10}^*(s)q_{10,8}^*(s) + q_{8,12}^*(s)q_{12,8}^*(s) - 1) + q_{18}^*(s)$
$q_{11,1}^*(s)q_{8,11}^*(s)(q_{02}^*(s)q_{20}^*(s) + q_{03}^*(s)q_{30}^*(s) + q_{04}^*(s)q_{40}^*(s) - 1) + (q_{02}^*(s)q_{20}^*(s) + q_{03}^*(s)q_{30}^*(s)$
$+ q_{04}^*(s)q_{40}^*(s))(q_{15}^*(s)q_{51}^*(s) + q_{16}^*(s)q_{61}^*(s) + q_{17}^*(s)q_{71}^*(s))(q_{89}^*(s)q_{98}^*(s) + q_{8,10}^*(s)q_{10,8}^*(s)$
$+ q_{8,12}^*(s)q_{12,8}^*(s))$

Using above calculated value of $AS_0^*(s)$ availability in summer can be obtained in steady-state and the same is given as follows:
$AS_0 = \lim_{s\to 0} sAS_0^*(s) = \frac{N_1}{D_1}$
where,
$N_1 = \mu_0 p_{10}(p_{81} + p_{8,11})$
$D_1 = (\mu_0 + \mu_2 p_{02} + \mu_3 p_{03} + \mu_4 p_{04})p_{10}(p_{81} + p_{8,11}) + (\mu_1 + \mu_5 p_{15} + \mu_6 p_{16} + \mu_7 p_{17})p_{01}(p_{81} + p_{8,11})$
$+ (\mu_8 + \mu_9 p_{89} + \mu_{10} p_{8,10} + \mu_{12} p_{8,12})p_{01}p_{18}$

During Winter
Similarly, steady-state availability during winter are given as follows:
$AW_0 = \lim_{s\to 0} sAW_0^*(s) = \frac{N_2}{D_1}$
where,
D_1 already defined and

$N_2 = \mu_1 p_{01}(p_{81} + p_{8,11})$

4.3. Busy Period Analysis

Busy period of the repairman due to repair in summer
Similarly, steady-state Busy period of the repairman due to repair in summer are given as follows:
$BS_0 = \lim_{s\to 0} sBS_0^*(s) = \frac{N_3}{D_1}$
where,
D_1 already defined and
$N_3 = (\mu_2 p_{02} + \mu_3 p_{03} + \mu_4 p_{04}) p_{10}(p_{81} + p_{8,11})$
During Winter
Similarly, steady-state Busy period of the repairman due to repair in winter are given as follows:
$BW_0 = \lim_{s\to 0} sBW_0^*(s) = \frac{N_4}{D_1}$
where,
D_1 already defined and
$N_4 = p_{01}[p_{18}(\mu_9 p_{89} + \mu_{10} p_{8,10} + \mu_{11} p_{8,11}) + (\mu_5 p_{15} + \mu_6 p_{16})(p_{81} + p_{8,11})]$

4.4. Expected Number of Visits of the Repairman for Repair

During summer
Similarly, steady-state number of visits of the repairman during summer are given as follows:
$VS_0 = \lim_{s\to 0} sVS_0^*(s) = \frac{N_5}{D_1}$
where,
D_1 already defined and
$N_5 = p_{10}(1 - p_{01})(p_{81} + p_{8,11})$
During Winter
Similarly, steady-state number of visits of the repairman during winter are given as follows:
$VW_0 = \lim_{s\to 0} sVW_0^*(s) = \frac{N_6}{D_1}$
where,
D_1 already defined and
$N_6 = p_{01} p_{18}(1 - p_{81} - p_{8,12}) + p_{01}(1 - p_{10} - p_{17} - p_{18})(p_{81} + p_{8,11})$

4.5. Expected Number of Visits of the Repairman for Preventive Maintenance

Similarly, steady-state number of visits of the repairman for preventive maintenance are given as follows:
$PM_0 = \lim_{s\to 0} sPM_0^*(s) = \frac{N_7}{D_1}$
where,
D_1 already defined and
$N_7 = p_{01} p_{17}(p_{81} + p_{8,11}) + p_{01} p_{18} p_{8,12}$

5. Cost-Benefit Analysis

Profit of the system under consideration can be obtained by subtracting the costs due to repair, per visit charges of the repairman for repair in summer and winter and per visit charges of the repairman for preventive maintenance. The same can expressed in terms of the various performance measures obtained through the model developed in this given as follows:
$Profit = CS_0 AS_0 + CW_0 AS_0 - CS_1 BS_0 - CW_1 BW_0 - CS_2 VS_0 - CW_2 VW_0 - C_3 PM_0$
where,

CS_0 : revenue during summer, per unit uptime
CW_0 : revenue during winter, per unit uptime
CS_1 : revenue during summer per unit time for repair
CW_1 : revenue during winter per unit time for repair
CS_2 : Cost per visit during summer for repair
CW_2 : Cost per visit during winter for repair
C_3 : Cost per visit for preventive maintenance

6. Numerical Interpretation

To obtain various numerical outcomes, the following specific case is used:

$w_1(t) = \alpha e^{-\alpha t}$ $\quad$ $w_2(t) = \beta e^{-\beta t}$
$m_1(t) = \gamma e^{-\gamma t}$ $\quad$ $m_2(t) = \delta e^{-\delta t}$
$g_1(t) = \lambda e^{-\lambda t}$ $\quad$ $g_2(t) = \mu e^{-\mu t}$
$\mu_0 = \frac{1}{a_1+a_2+a_3+\alpha}$ $\quad$ $\mu_1 = \frac{1}{a_1+a_2+a_3+\beta+\gamma}$
$\mu_2 = \frac{1}{b_1}$ $\quad$ $\mu_3 = \frac{1}{b_2}$
$\mu_4 = \frac{1}{b_3}$ $\quad$ $\mu_5 = \frac{1}{b_1}$
$\mu_6 = \frac{1}{b_2}$ $\quad$ $\mu_7 = \frac{1}{\delta}$
$\mu_8 = \frac{1}{a_1+a_2+a_4+\mu+\gamma}$ $\quad$ $\mu_9 = \frac{1}{b_1}$
$\mu_{10} = \frac{1}{b_2}$ $\quad$ $\mu_{11} = \frac{1}{b_4}$
$\mu_{12} = \frac{1}{\delta}$ $\quad$ $p_{01} = \frac{\alpha}{a_1+a_2+a_3+\alpha}$
$p_{02} = \frac{a_1}{a_1+a_2+a_3+\alpha}$ $\quad$ $p_{03} = \frac{a_2}{a_1+a_2+a_3+\alpha}$
$p_{04} = \frac{a_3}{a_1+a_2+a_3+\alpha}$ $\quad$ $p_{10} = \frac{\beta}{a_1+a_2+a_3+\gamma+\beta}$
$p_{15} = \frac{a_1}{a_1+a_2+a_3+\gamma+\beta}$ $\quad$ $p_{16} = \frac{a_2}{a_1+a_2+a_3+\gamma+\beta}$
$p_{17} = \frac{a_3}{a_1+a_2+a_3+\gamma+\beta}$ $\quad$ $p_{18} = \frac{\gamma}{a_1+a_2+a_3+\gamma+\beta}$
$p_{81} = \frac{\mu}{a_1+a_2+a_4+\gamma+\mu}$ $\quad$ $p_{89} = \frac{a_1}{a_1+a_2+a_4+\gamma+\mu}$
$p_{8,10} = \frac{a_2}{a_1+a_2+a_4+\gamma+\mu}$ $\quad$ $p_{8,11} = \frac{a_4}{a_1+a_2+a_4+\gamma+\mu}$
$p_{8,12} = \frac{\gamma}{a_1+a_2+a_4+\gamma+\mu}$
$p_{20} = p_{30} = p_{40} = p_{51} = p_{61} = p_{71} = p_{98} = p_{10,8} = p_{11,1} = p_{12,8} = 1$

where
$a_1 = 0.235, a_2 = 0.0381, a_3 = 0.01589, a_4 = 0.02673, b_1 = 0.887, b_2 = 0.793, b_3 = 0.821,$
$b_4 = 0.896, \alpha = 0.615, \beta = 0.83, \lambda = 0.034, \gamma = 0.00937, \delta = 0.870, \mu = 0.875, CS_0 = 15000,$
$CS_1 = 1500, CW_0 = 15000, CW_1 = 1600, CS_2 = 1450, CW_2 = 1550, C_3 = 1400.$
Various graphs have been plotted but all the graphs have not been shown here to use minimum space and to avoid repetition of similar interpretations. However, the users of such systems may plot graph of their interest as per the requirement and may take important decision regarding profitability of the system. Regarding the availability and nature of MTSF, various rates have been depicted as shown in Fig. 2 ,3, 4 and 5 which reveal that MTSF, Availability and profit decreases as failure rates increases. However, their values go in the direction δ and b_2. Some of the plotted graphs are shown as follows:
The MTSF behaviour for different values of β is shown in Fig. 2. MTSF decreases as the failure rate value (a_2) rises. Higher values of β correspond to higher values in it.

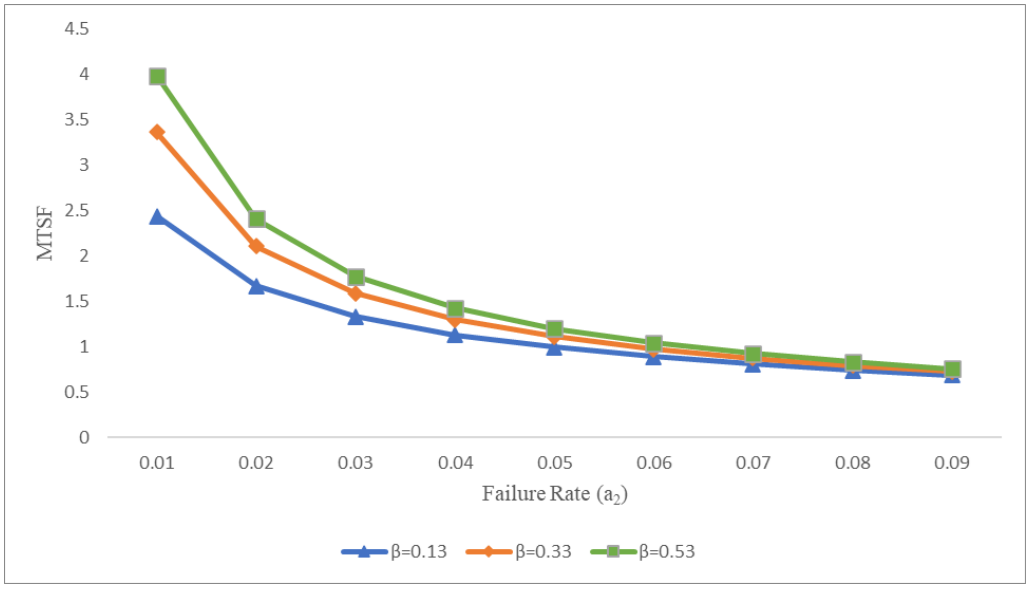


Figure 2: *MTSF versus Failure Rate* (a_2) *for different values of* (β)

The availability behaviour in the summer for various repair rate values (b_2) is displayed in Fig. 3. As the failure rate (a_2) rises, summer availability decreases. Additionally, it has been noted that when b_2 values rise, so does availability.

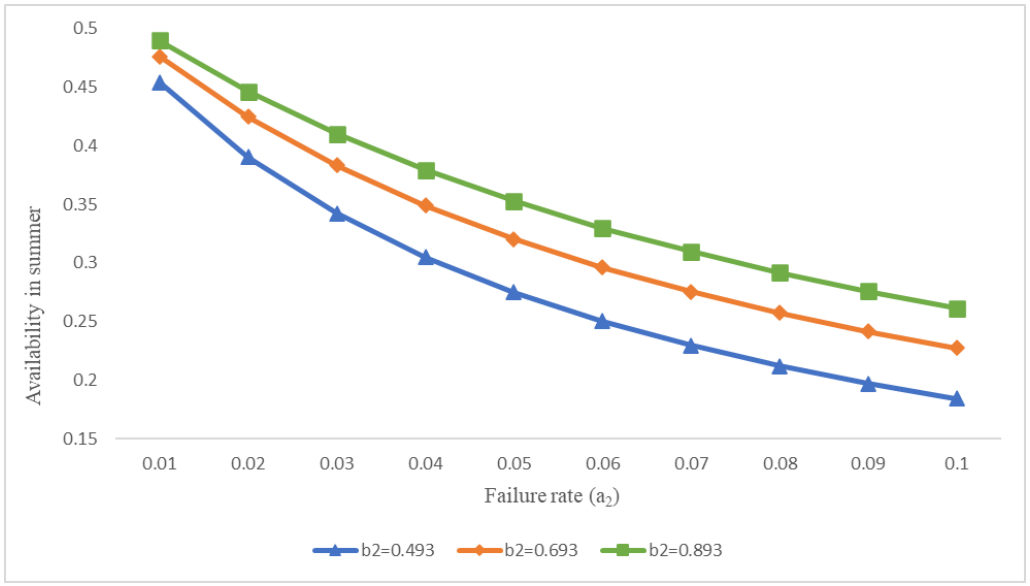


Figure 3: *Availability in Summer versus Failure rate* (a_2) *for different values of Repair rate* (b_2)

The way that profit acts in relation to revenue in the summer (CS_0) for various values of the cost paid for repair in the summer (CS_1) is shown in Fig. 4. As revenue values rise in the summer (CS_0), profit rises as well. Additionally, it has been seen that as (CS_1) values rise, the profit falls.

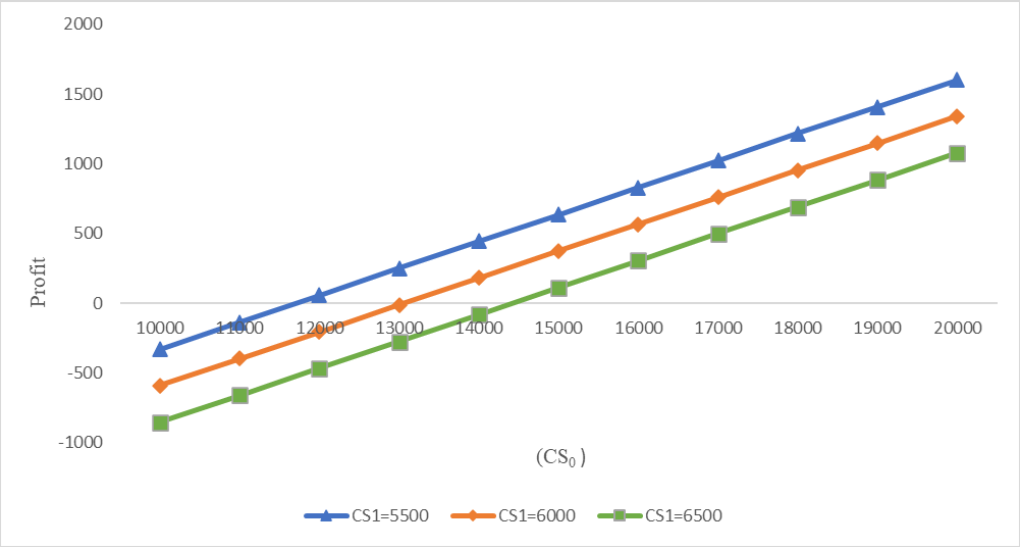


Figure 4: *Profit versus Revenue in Summer* (CS_0) *for different values of cost paid for paid in Summer* (CS_1)

Fig. 5 illustrates the behaviour of profit in relation to revenue during the winter (CW_0) for various costs (C_3) associated with preventive maintenance. With an increase in winter revenue values (CW_0), profit rises. Additionally, it has been seen that as C_3 values rise, the profit falls.

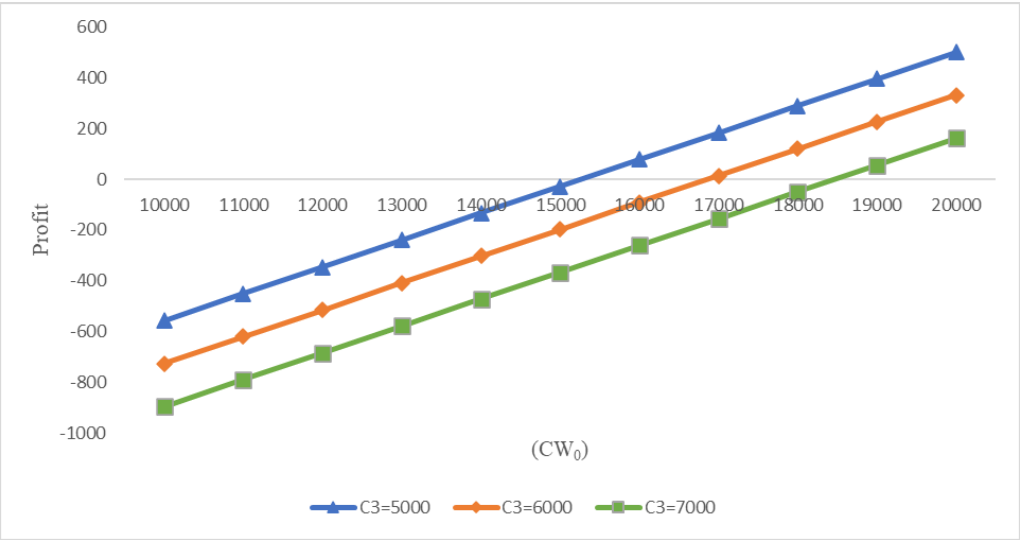


Figure 5: *Profit versus Revenue in Winter* (CW_0) *for different values of cost* (C_3) *paid for paid for Preventive Maintenance*

Values of parameters taken and cut-off points obtained from the above figures are tabulated as follows:

Fig	Varied Parameters	Condition	Interpretation
4	CS_1 = 5500	CS_0> 11700.3105	System is profitable
	CS_1 = 6000	CS_0> 13059.0062	System is profitable
	CS_1 = 6500	CS_0> 14417.1842	System is profitable
5	CW_1 =5000	CW_0>15261.2412	System is profitable
	CW_1 =6000	CW_0>16863.2744	System is profitable
	CW_1 =7000	CW_0>18465.3067	System is profitable

7. Conclusion

In the current study, a reliability model is developed using a system for producing utensils. The findings for a specific situation demonstrate the relevance of research since cut-off points may be used to set lower and upper limits for a variety of factors. For instance, setting a product's pricing so that the system is profitable depends on the cut-off point for revenue per unit uptime. The cut-off points facilitate many crucial judgments for the profits according to revenue.

Funding

The first author delightedly acknowledges the University Grants Commission (UGC), New Delhi, India for providing financial support.

Disclosure statement

The authors declare that they have no conflict of interest.

References

[1] Teng, X., Pham, H. and Jeske, D.R., (2006). Reliability modeling of hardware and software interactions, and its applications. *IEEE Transactions on Reliability*, 55(4): 571–577.

[2] Yusuf, B. and Yusuf, I., (2013). Evaluation of some reliability characteristics of a system under three types of failures with repair-replacement at failure. *American Journal of Operational Research*, 3(3): 83–91.

[3] Manocha, A. and Taneja, G.,(2015). Stochastic analysis of a two-unit cold standby system with arbitrary distribution for life, repair and waiting times. *International Journal of Performability Engineering*, 11(3): 293–299.

[4] Gupta, P., Lal, A.K., Sharma, R.K. and Singh, J., (2016). Numerical analysis of reliability and availability of the serial processes in butter oil processing plant. *International Journal of Quality and Reliability Management*, 22(3): 303–316.

[5] Fagge, N. J., Yusuf, I. and Ali, U. A., (2017). Availability evaluation of repairable system requiring two types of supporting device for operations. *International Journal of Applied Mathematical Research*, 6 (1): 14–19.

[6] Rajesh, Taneja, G. and Jagdish, P., (2018).Reliability of a gas turbine system with change in weather and optimisation of electricity price when working in single cycle. *Int. J. Agricult. Stat. Sci.*, 14(1): 119–128.

[7] Kumar, A. and Malik, S. C., (2014). Stochastic modeling of a concrete mixture plant with preventive maintenance. *An International Journal of Applications and Applied Mathematics*,09(1): 13–27.

[8] Rajesh, Sangal, P. K., Himani and Taneja, G.,(2020). Reliability and profit analysis of a gas turbine system with optimisation of electricity price of single cycle for different cut-off points. *Int. J. Agricult. Stat. Sci.*, 16(1): 2143–2150.

[9] Jain, P., Pawar, D. and Malik, S.C., (2022). Profit analysis of A 1- out Of 2 unit system with a standby unit and arrival time of server. *Int. J. Agricult. Stat. Sci.*,18(1): 29–33.

[10] Rahbi, Y., Rizwan, S. M., Alkali, B. M., Cowell, A. and Taneja, G., (2018). Reliability analysis of rodding anode plant in an aluminum industry with multiple repairmen. *Advances and Applications in Statistics*, 53(5): 569–597.

[11] Sheetal, Singh, D. and Taneja, G., (2019). A reliability model for a system with temperature-dependent working and time-dependent payment for major fault. *Journal of Advance Research in Dynamical & Control system*, 11(01): 1189–1196.

[12] Sheetal, Taneja, G. and Singh, D., (2018). Reliability and profit analysis of a system with effect of temperature on operation. *International Journal of Applied Engineering Research*, 13 (7): 4865–4870.

[13] Sachdeva, K., Taneja, G., and Manocha, A., (2022).Sensitivity and economic analysis of an insured system with extended conditional warranty. *Reliability: Theory & Applications*,315–327.

[14] Rizwan, S. M., Sachdeva, K., Alagiriswamy, S. and Al Rahbi, Y., (2023). Performability and Sensitivity Analysis of the Three Pumps of a Desalination Water Pumping Station. *International Journal of Engineering Trends and Technology* , 51(2): 283–292.

[15] Rizwan, S. M., Sachdeva, K., Al Balushi, N., Al Rashdi, S., and Taj, S. Z., (2023) Reliability and Sensitivity Analysis of Membrane Biofilm Fuel Cell. *International Journal of Engineering Trends and Technology* , 71(3): 73–80.

[16] Singh, J., and Mahajan, P., (1999). Reliability of utensils manufacturing plant-a case study. *Opsearch* , 13(24): 260–269.

[17] Kumar, K. and Kumar, P., (2010). Mathematical modelling and analysis of stainless-steel utensils manufacturing plant using fuzzy reliability. *International Journal of Engineering Science and Technology*, 2(6): 2370–2376.

OPTIMIZING MULTI-OBJECTIVE MULTI-INDEX TRANSPORTATION PROBLEMS: A SMART ALGORITHMIC SOLUTION WITH LINDO SOFTWARE

Ajjaz Maqbool Dar[1], K. Selvakumar[2], S. Ramki[3], K. M. Karuppasamy[4], Jameel A. Ansari[5], Aafaq A. Rather[6,*]

•

[1]Jammu and Kashmir Institue of Mathematical Sciences, Srinagar-190008, J&K, India
[2]Department of Mathematics, University College of Engineering-Nagercoil-629004, Tamil Nadu, India
[3,4]Department of Mathematics & Statistics, Bharath Institute of Higher Education and Research, Selaiyur, Chennai – 73
[5]Department of Engineering Sciences, Faculty of Science and Technology, Vishwakarma University, Pune, India
[6]Symbiosis statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[1]ajjazmaqbool013@gmail.com, [2]selvakumaruce@gmail.com, [3]ramki.stat24@gmai.com, [4]m.k.samystat@gmail.com, [5]jameelahmad.ansari@vupune.ac.in, [6,*]aafaq7741@gmail.com

Abstract

In the present paper, we create an algorithm to address the transportation problem with numerous objectives and indexes. The transportation problem exists when there are more supply points, more demand points, and various means of transportation are used to meet demand or when moving certain types of goods. The transportation problem may frequently be more complex than the typical form of transportation problem. We create a model that blends fuzzy multi-objective programming and the multi-index transportation problem' by using LINDO software to resolve all related problems. Additionally, the decision-maker may present a variety of data and it may be further improved. The new algorithm for addressing transport problems in fuzzy environments is demonstrated numerically.

Key Words: Fuzzy transportation problem (FTP), Linear Programming Problem(LP P), Multi-index transportation problem (MITP)

1. Introduction

The transportations prolem has a extensive range of real-world applications and can be seen as a specific example of the LPP. It is one of the best optimization techniques and has a wide range of real-world applications. a combination of various goods from any of the m origins to any of the n destination places. In order to reduce the overall cost of a transportation issue, we control the amount to be transported from all origins to all destinations. We may not have focused on a single objective function in this situation, which is multi-objective. All of the objectives of MOTP are in competition with one another, and all of the restrictions are of the equality kind. The technique for multi-objective fuzzy linear programming with uncertain goals. The multi-objective transportation problem has the best compromise solution when applied to fuzzy linear programming.

Numerous academics have evaluated the use of the MITP to optimization, mathematical modelling and industry. Wang et al. [10] established a decomposition technique for handling the standard three-index transportation problem that is entirely dependent on the successive adjustment of the optimality criterion. They looked at the transportation problem's solution using a linear and quadratic objective function. In addition to recommending the adoption of such transportation efficiency, Rautman et al. [8] find a solution to the shipment scheduling conflict by utilising a multi-index transportation problem method to optimize the integral system Bit et al. [4] used a fuzzy programming method with a hyperbolic membership function to solve the Multi-Objective Capacitated Transportation Problem, in which the targets are non-commensurable and incompatible and the deliver and demand constraints are all of the same kind. In order to identify the best and most effective compromise solution to a multi-objective capacitated transportation problem, fuzzy programming with hyperbolic membership function was applied. For the first time, all parameters are taken into consideration when using fuzzy multi index bi-criteria constant fee bottleneck transportation (FMIBCFCBTP) by Sungeeta et al. [9]. An algorithm was created to detect FMIBCFCBTP fuzzy time-value change-off pairings. A numerical example was provided to explain the said algorithm. An exponential membership function was employed in the fuzzy programming method by Kaur et al. [5] to resolve a multi-objective and multi-index transportation problem. The main emphasis is on reducing the prices, decreasing rates, and underutilized capacity of transporting raw materials via various modes of transportation from various points of origin to various destination sites. Each target function is given a unique form of non-linear membership function by employing the fuzzy programming technique to solve actual transport problems using an exponential function and creating a non-dominated compromise solution. To tackle the multi-index fixed charge bi-criterion transportation problem , Archana and Veena [2] provided a method for determining the ideal trade-off pair amongst efficient cost-time trade-off pairs.

The linear-multi-objective-solid transportation problem was approached from a fuzzy-linear programming perspective by Bit et al. [3]. The outcome is a compromise approach that is both cost-effective and ideal. The fuzzy linear programming approach was used to develop the FORTRAN programme. With profit maximisation and time minimization as the objectives, Anjana et al. [1] developed a multi-objective multi-object strong transportation problem (MMSTP) with fuzzy inequality constraints. Fuzzy chance programming and the additional specificity opportunity-necessity principle are used to create a method for changing the ambiguous MMSTP to the equivalent-deterministic shape. The fuzzy interactive satisfied method was used to develop optimal compromise solutions for the MMSTP through a generalised reduced gradient strategy. The finest non-dominating solution was created using the technique for order preference by similar to perfect solution. An algorithm was created by Porchelvi and Anitha [7] to address the multi-objective transportation problem. The source and destination parameters , along with the cost coefficients of the goal function, are expressed as interval costs. To solve MOTP, they used fuzzy programming techniques with linear membership functions for various costs.

2. Mathematical Model

The MOMITP,we assume p_{ijl} be a multidimensional array $1 \le i \le m, 1 \le j \le n$ *and* $1 \le l \le k$ and let $P = p_{ij}$, $Q = q_{jl}$ and $R = r_{il}$ be multi-matrices then the MITP is as fallows

$$\text{Min } Z = \sum_i \sum_j \sum_l p_{ijl} X_{ijl} \tag{1}$$

Such that

$$\sum_i X_{ijl} = p_{ij} \quad for\, all\, i, j$$

$$\sum_{j} X_{ijl} = r_{il} \quad for\, all\, i,l$$

$$\sum_{j} X_{ijl} = q_{jl} \quad for\, all\, j,l\,,\ \&\ X_{ijl} \geq 0,\ \ for\, all\, i,j,l \tag{2}$$

it further, implies that

$$\sum_{i} p_{ij} = \sum_{l} q_{jl} \ ;\ \sum_{j} p_{ij} = \sum_{l} r_{il} \ ;\ \sum_{j} q_{jl} = \sum_{l} r_{il} \tag{3}$$

All the 3-conditions are necessary but not sufficient. MOMITP is defined as fallows

$$\text{Minimize } Z_k = \sum_{i}^{m}\sum_{j}^{n} v_{ij}^{(1)}\, x_{ij}^{(1)} + \sum_{i}^{m}\sum_{j}^{n} v_{ij}^{(2)}\, x_{ij}^{(2)} \tag{4}$$

Such that

$$\sum_{j}^{n} x_{ij}^{(1)} = p_{1i}\,,\ \forall i \tag{5}$$

$$\sum_{j}^{n} x_{ij}^{(2)} = p_{2i}\,,\ \forall i \tag{6}$$

$$\sum_{i}^{m} x_{ij}^{(1)} = q_{1j}\,,\ \forall j \tag{7}$$

$$\sum_{i}^{m} x_{ij}^{(2)} = q_{2j}\,,\ \forall j \tag{8}$$

$$x_{ij}^{(1)} + x_{ij}^{(2)} = r_{ij}\,,\ \forall i,j \tag{9}$$

$$x_{ij}^{(1)} \geq 0,\ x_{ij}^{(2)} \geq 0 \tag{10}$$

The following set of conditions are necessary for the existance solution.

$$\sum_{j}^{n} r_{il} = p_{1i} + p_{2i},\ \forall i, \tag{11}$$

$$\sum_{j}^{n} r_{il} = q_{1j} + q_{2j},\ \forall j \tag{12}$$

$$\sum_{i}^{m} p_{1i} = \sum_{j}^{n} q_{1j}\,,\ \forall i,j \tag{13}$$

$$\sum_{i}^{m} p_{2i} = \sum_{j}^{n} q_{2j}\,,\ \forall i,j \tag{14}$$

$$\sum_{j}^{n} r_{il} \leq \min\,(p_{1i} + q_{1j}) + \min\,(p_{2i} + q_{2j}),\ \forall i,j \tag{15}$$

3. Proposed Algorithm

Step 1: Formulate a FTP.

Step 2: Solving the MOTP, k times, taking, one at a time, we first develop a matrix form inorder to get corresponding values for each objectives for each solution.

$$\begin{array}{c} \\ X_1 \\ X_2 \\ \vdots \\ X_k \end{array} \begin{array}{c} \begin{array}{cccc} Z_1 & Z_2 & \ldots & Z_k \end{array} \\ \begin{bmatrix} Z_{11} & Z_{12} & \ldots & Z_{1k} \\ Z_{21} & Z_{22} & \ldots & Z_{2k} \\ \vdots & \vdots & & \vdots \\ Z_{k1} & Z_{k2} & \ldots & Z_{kk} \end{bmatrix} \end{array}$$

Where, each $X^i, i=1,2,..k$ represent the isolated optimal solutions to the K distinct transportation problems for k distict objective functions $Z_{ij}=Z_j(X^i)\ \forall i,j,\ \ where\ i,j=1,2,3.....k$, respectavelly the ith row and jth column members of the matrix.

Step 3: Using step 2, we set upper and lower bounds for each objective and defining the range of values for the membership function that represents the degree of acceptance and rejection for a perticular solution. The values of such functions can be calculated as.

$U_k^{\mu}=Max(Z_r(X_r))$

$L_k^{\mu}=Min(Z_r(X_r)),\ 0\le r\le k,$

Where U_k^{μ} and L_k^{μ} are rtespectivelly the upper and lower bound for the (k^{th} objective function Z_k) $k=1,2,3.....K$, $d_k=U_k^{\mu}-L_k^{\mu}$,the degradation allowance for objective k.

Step 4: We define the membership function as:

$$\mu_k\{Z(X)_k\}=\begin{cases}1, & L_k^{\mu}\ge Z_k(X)\\ 1-\dfrac{(Z_1(X)-L_k^{\mu})}{d_k}, & L_k^{\mu}\le Z_k(X)\le U_k^{\mu},\ where\ d_k=U_k^{\mu}-L_k^{\mu}\\ 0, & Z_k(X)\ge U_k^{\mu}\end{cases} \tag{16}$$

Step 5: We use a LMF for the initial fuzzy model, the crisp model can be simplified as:

Minimize α

Subject to

$$Z_k(X)+\alpha d_k\le U_k^{\mu}, \tag{17}$$
$$\sum_j^n x_{ij}^{(1)}=p_{1i},\ \forall i,j$$
$$\sum_j^n x_{ij}^{(2)}=p_{2i},\ \forall j$$
$$\sum_i^m x_{ij}^{(1)}=q_{1j},\ \forall i$$
$$\sum_i^m x_{ij}^{(2)}=q_{2j},\ \forall j$$
$$x_{ij}^{(1)}+x_{ij}^{(2)}=r_{ij},\ \forall i,j$$
$$x_{ij}^{(1)}\ge 0,\ x_{ij}^{(2)}\ge 0$$

Above system of LP can be solved by using LONDO statistical software.

Step 6: Using the precise mathematical Programming approach, we are able to solve the crisp model.

Min α

$$C_{ij}^k x_{ij}+\alpha d_k\le U_k^{\mu},\ k=1,2,.....K, \tag{18}$$
$$\sum_j^n x_{ij}^{(1)}=p_{1i},\ \forall i,j$$
$$\sum_j^n x_{ij}^{(2)}=p_{2i},\ \forall i,$$
$$\sum_i^m x_{ij}^{(1)}=q_{1j},\ \forall j$$
$$\sum_i^m x_{ij}^{(2)}=q_{2j},\ \forall j$$
$$x_{ij}^{(1)}+x_{ij}^{(2)}=r_{ij},\ \forall i,j$$

$$x_{ij}^{(1)} \geq 0,\ x_{ij}^{(2)} \geq 0$$

Another membership function, like the Hyperbolic Tangent function, is one that we utlise.

Min α

$$\alpha \geq \frac{1}{2} + \frac{1}{2}\tanh\{\frac{U_k^{\mu} + L_k^{\mu}}{2} + Z_k\}\tau_k, \tag{19}$$

Where, $\tau_k = \frac{S}{U_k^{\mu} - L_k^{\mu}}$, where S is number of constraints.

$$x_{ij} \geq 0, \forall i, j \text{ and } \alpha \geq 0.$$

Step 7: An intuitionistic fuzzy optimization for MOLP is defined as

$$\mu_k^e\{Z_k(X)\} = \begin{cases} 1, & L_k \geq Z_k(X) \\ e^{-\frac{1}{2}\left(\frac{Z_k - L_k^{\mu}}{d_k}\right)}, & L_k^{\mu} \leq Z_k(X) \leq U_k^{\mu} \quad where \quad d_k = U_k^{\mu} - L_k^{\mu} \\ 0, & Z_k(X) \geq U_k^{\mu} \end{cases} \tag{20}$$

where $k = 1,2,\ldots\ldots\ldots K$.

4. Illustrations

Example 4.1: Consider the fuzzy MOMITP discussed by Lohgaonkar et al. [6]

$$\begin{aligned} Minimize\ Z_1 = {} & 4Y_{11}^{(1)} + 3Y_{12}^{(1)} + 5Y_{13}^{(1)} + 8Y_{21}^{(1)} + 6Y_{22}^{(1)} + 2Y_{23}^{(1)} + 7Y_{31}^{(1)} + 4Y_{32}^{(1)} \\ & + Y_{33}^{(1)} + 9Y_{41}^{(1)} + 10Y_{42}^{(1)} + 12Y_{43}^{(1)} + 8Y_{11}^{(2)} + 6Y_{12}^{(2)} + 3Y_{13}^{(2)} + 5Y_{21}^{(2)} \\ & + 4Y_{22}^{(2)} + Y_{23}^{(2)} + 9Y_{31}^{(2)} + 2Y_{32}^{(2)} + 6Y_{33}^{(2)} + 4Y_{41}^{(2)} + 9Y_{42}^{(2)} + 3Y_{43}^{(2)} \end{aligned}$$

Subject to

$$\begin{aligned} & Y_{11}^{(1)} + Y_{12}^{(1)} + Y_{13}^{(1)} = 9 \\ & Y_{21}^{(1)} + Y_{22}^{(1)} + Y_{23}^{(1)} = 14 \\ & Y_{31}^{(1)} + Y_{32}^{(1)} + Y_{33}^{(1)} = 6 \\ & Y_{41}^{(1)} + Y_{42}^{(1)} + Y_{43}^{(1)} = 7 \\ & Y_{11}^{(2)} + Y_{12}^{(2)} + Y_{13}^{(2)} = 6 \\ & Y_{21}^{(2)} + Y_{22}^{(2)} + Y_{23}^{(2)} = 7 \\ & Y_{31}^{(2)} + Y_{32}^{(2)} + Y_{33}^{(2)} = 5 \\ & Y_{41}^{(2)} + Y_{42}^{(2)} + Y_{43}^{(2)} = 6 \\ & Y_{11}^{(1)} + Y_{21}^{(1)} + Y_{31}^{(1)} + Y_{41}^{(1)} = 14 \\ & Y_{12}^{(1)} + Y_{22}^{(1)} + Y_{32}^{(1)} + Y_{42}^{(1)} = 12 \\ & Y_{13}^{(1)} + Y_{23}^{(1)} + Y_{33}^{(1)} + Y_{43}^{(1)} = 10 \\ & Y_{11}^{(2)} + Y_{21}^{(2)} + Y_{31}^{(2)} + Y_{41}^{(2)} = 5 \\ & Y_{12}^{(2)} + Y_{22}^{(2)} + Y_{32}^{(2)} + Y_{42}^{(2)} = 8 \\ & Y_{13}^{(2)} + Y_{23}^{(2)} + Y_{33}^{(2)} + Y_{43}^{(2)} = 11 \\ & Y_{11}^{(1)} + Y_{11}^{(2)} = 5 \\ & Y_{12}^{(1)} + Y_{12}^{(2)} = 7 \end{aligned}$$

$$
\begin{aligned}
&Y_{13}^{(1)} + Y_{13}^{(2)} = 3\\
&Y_{21}^{(1)} + Y_{21}^{(2)} = 8\\
&Y_{22}^{(1)} + Y_{22}^{(2)} = 4\\
&Y_{23}^{(1)} + Y_{23}^{(2)} = 9\\
&Y_{31}^{(1)} + Y_{31}^{(2)} = 4\\
&Y_{32}^{(1)} + Y_{32}^{(2)} = 1\\
&Y_{33}^{(1)} + Y_{33}^{(2)} = 6\\
&Y_{41}^{(1)} + Y_{41}^{(2)} = 2\\
&Y_{42}^{(1)} + Y_{42}^{(2)} = 8\\
&Y_{43}^{(1)} + Y_{43}^{(2)} = 3\\
&Y_{ij}^{(1)}, Y_{ij}^{2} \geq 0,\ i = 1,2,3,4 \quad and \quad j = 1,2,3
\end{aligned}
\tag{21}
$$

Example 4.2:

$$
\begin{aligned}
Minimize\, Z_2 = {} & 5Y_{11}^{(1)} + 6Y_{12}^{(1)} + 7Y_{13}^{(1)} + 4Y_{21}^{(1)} + 5Y_{22}^{(1)} + 2Y_{23}^{(1)} + 1Y_{31}^{(1)} + 3Y_{32}^{(1)}\\
& + 4Y_{33}^{(1)} + 4Y_{41}^{(1)} + 2Y_{42}^{(1)} + 3Y_{43}^{(1)} + 10Y_{11}^{(2)} + 9Y_{12}^{(2)} + 9Y_{13}^{(2)} + 7Y_{21}^{(2)}\\
& + 9Y_{22}^{(2)} + 2Y_{23}^{(2)} + 8Y_{31}^{(2)} + 7Y_{32}^{(2)} + 9Y_{33}^{(2)} + 8Y_{41}^{(2)} + 4Y_{41}^{(2)} + 5Y_{43}^{(2)}
\end{aligned}
$$

Subject to

$$
\begin{aligned}
&Y_{11}^{(1)} + Y_{12}^{(1)} + Y_{13}^{(1)} = 9\\
&Y_{21}^{(1)} + Y_{22}^{(1)} + Y_{23}^{(1)} = 14\\
&Y_{31}^{(1)} + Y_{32}^{(1)} + Y_{33}^{(1)} = 6\\
&Y_{41}^{(1)} + Y_{42}^{(1)} + Y_{43}^{(1)} = 7\\
&Y_{11}^{(2)} + Y_{12}^{(2)} + Y_{13}^{(2)} = 6\\
&Y_{21}^{(2)} + Y_{22}^{(2)} + Y_{23}^{(2)} = 7\\
&Y_{31}^{(2)} + Y_{32}^{(2)} + Y_{33}^{(2)} = 5\\
&Y_{41}^{(2)} + Y_{42}^{(2)} + Y_{43}^{(2)} = 6\\
&Y_{11}^{(1)} + Y_{21}^{(1)} + Y_{31}^{(1)} + Y_{41}^{(1)} = 14\\
&Y_{12}^{(1)} + Y_{22}^{(1)} + Y_{32}^{(1)} + Y_{42}^{(1)} = 12\\
&Y_{13}^{(1)} + Y_{23}^{(1)} + Y_{33}^{(1)} + Y_{43}^{(1)} = 10\\
&Y_{11}^{(2)} + Y_{21}^{(2)} + Y_{31}^{(2)} + Y_{41}^{(2)} = 5\\
&Y_{12}^{(2)} + Y_{22}^{(2)} + Y_{32}^{(2)} + Y_{42}^{(2)} = 8\\
&Y_{13}^{(2)} + Y_{23}^{(2)} + Y_{33}^{(2)} + Y_{43}^{(2)} = 11\\
&Y_{11}^{(1)} + Y_{11}^{(2)} = 5\\
&Y_{12}^{(1)} + Y_{12}^{(2)} = 7\\
&Y_{13}^{(1)} + Y_{13}^{(2)} = 3\\
&Y_{21}^{(1)} + Y_{21}^{(2)} = 8\\
&Y_{22}^{(1)} + Y_{22}^{(2)} = 4\\
&Y_{23}^{(1)} + Y_{23}^{(2)} = 9
\end{aligned}
$$

$$Y_{31}^{(1)} + Y_{31}^{(2)} = 4$$
$$Y_{32}^{(1)} + Y_{32}^{(2)} = 1$$
$$Y_{33}^{(1)} + Y_{33}^{(2)} = 6$$
$$Y_{41}^{(1)} + Y_{41}^{(2)} = 2$$
$$Y_{42}^{(1)} + Y_{42}^{(2)} = 8$$
$$Y_{43}^{(1)} + Y_{43}^{(2)} = 3$$
$$Y_{ij}^{(1)}, Y_{ij}^{2} \geq 0,\ i = 1,2,3,4 \quad and \quad j = 1,2,3$$

If no errors are found, then the LINGO Solver status window appears of the illustration 4.1 is given below (by changing variable Y to X and also, taking $X_{11}^{(1)}$ and $X_{11}^{(2)}$ respectively X11M and X11N so on. In LINGO solver)

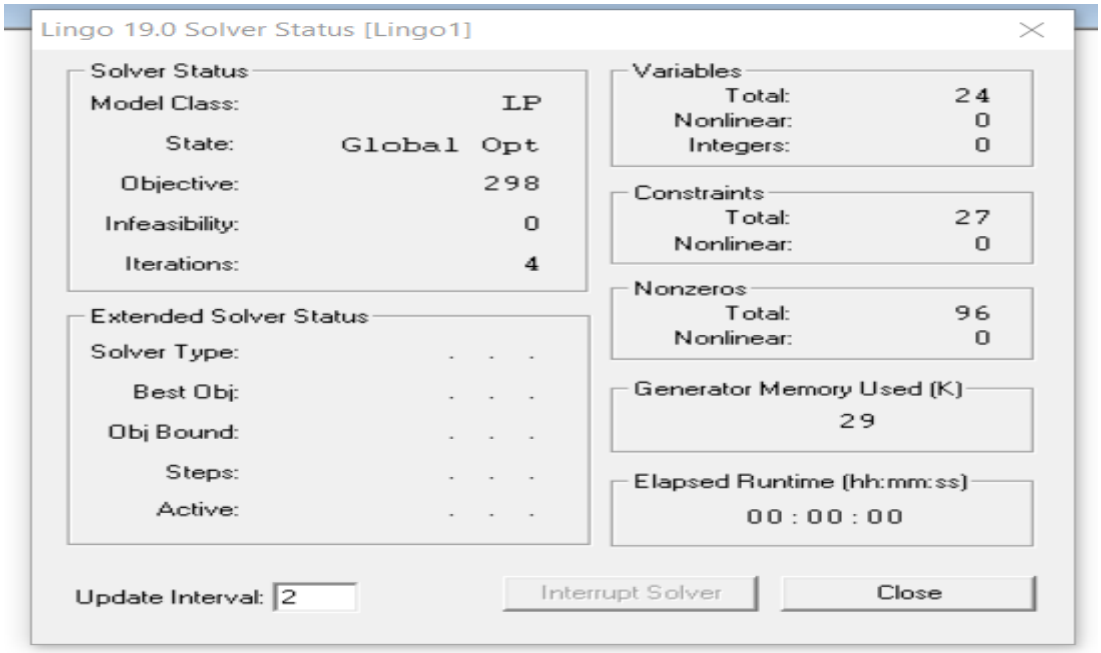


Figure 1: *Illustration of model 4.1 on LINGO window*

By closing above window, we can view the solution of the Model 4.1 is given
(Global optimal solution found).

Objective value:	298.0000
Infeasibilities:	0.000000
Total solver iterations:	4
Model Class:	LP
Total variables:	24
Nonlinear variables:	0
Integer variables:	0
Total constraints:	27
Nonlinear constraints:	0
Total nonzeros	96
Nonlinear nonzeros:	0

Table 1: *Optimal Solution of model 4.1*

Variable	Value	Reduced Cost
X11M	5.000000	0.000000
X12M	4.000000	0.000000
X13M	0.000000	6.000000
X21M	8.000000	0.000000
X22M	1.000000	0.000000
X23M	5.000000	0.000000
X31M	1.000000	0.000000
X32M	0.000000	6.000000
X33M	5.000000	0.000000
X41M	0.000000	2.000000
X42M	7.000000	0.000000
X43M	0.000000	9.000000
X11N	0.000000	3.000000
X12N	3.000000	0.000000
X13N	3.000000	0.000000
X21N	0.000000	1.000000
X22N	3.000000	0.000000
X23N	4.000000	0.000000
X31N	3.000000	0.000000
X32N	1.000000	0.000000
X33N	1.000000	0.000000
X41N	2.000000	0.000000
X42N	1.000000	0.000000
X43N	3.000000	0.000000

If no errors are found, then the LINGO Solver status window appears for the illustration 4.2 is given below

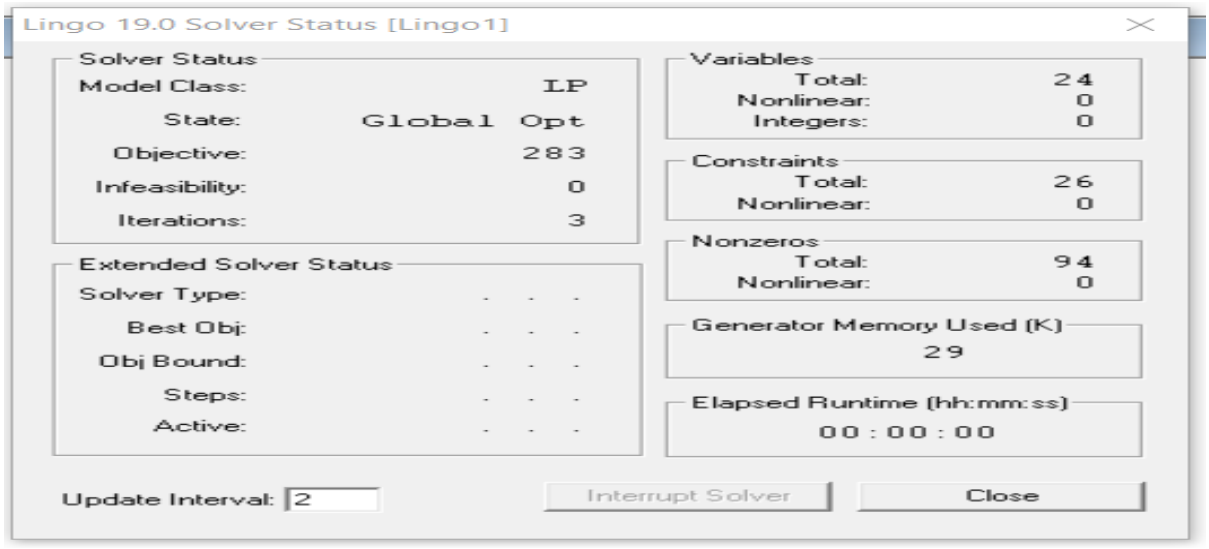


Figure 2: *Illustration of model 4.2 on LINGO window*

By closing above window, we can view the solution of the Model 4.2 is given (Global optimal solution found).

Objective value:	283.0000
Infeasibilities:	0.000000
Total solver iterations:	3
Model Class:	LP
Total variables	24
Nonlinear variables:	0
Integer variables:	0
Total constraints:	27
Nonlinear constraints:	0
Total nonzeros:	96
Nonlinear nonzeros:	0

Table 2: *Optimal Solution of model 4.2*

Variable	Value	Reduced Cost
X11M	3.000000	0.000000
X12M	6.000000	0.000000
X13M	0.000000	1.000000
X21M	8.000000	0.000000
X22M	4.000000	0.000000
X23M	2.000000	0.000000
X31M	1.000000	0.000000
X32M	0.000000	1.000000
X33M	5.000000	0.000000
X41M	2.000000	0.000000
X42M	2.000000	0.000000
X43M	3.000000	0.000000
X11N	2.000000	0.000000
X12N	1.000000	0.000000
X13N	3.000000	0.000000
X21N	0.000000	1.000000
X22N	0.000000	4.000000
X23N	7.000000	0.000000
X31N	3.000000	0.000000
X32N	1.000000	0.000000
X33N	1.000000	0.000000
X41N	0.000000	0.000000
X42N	6.000000	0.000000
X43N	0.000000	0.000000

Now, we can determine Z_1 and Z_2, for $(X^{(2)}, X^{(1)})$ respectively as given below $Z_2(X^{(1)}) = 295$ and $Z_1(X^{(2)}) = 335$, and written in the form of matrix.

$$\begin{array}{c} \\ X^{(1)} \\ X^{(2)} \end{array} \begin{array}{c} \begin{array}{cc} Z_1 & Z_2 \end{array} \\ \begin{bmatrix} 298 & 295 \\ 335 & 283 \end{bmatrix} \end{array}$$

From the above, we have $U_1^{\mu} = 335$ $U_2^{\mu} = 295$ $L_1^{\mu} = 298$ $L_2^{\mu} = 283$

We define Z_1 and Z_2 respectively

$$\mu_1(X) = \begin{cases} 1, & if\ Z_1(X) \le 298 \\ 1 - \dfrac{Z_1(X) - 298}{335 - 298}, & if\ 298 \le Z_1(X) \le 335 \\ 0, & if\ Z_1(X) \ge 335 \end{cases}$$

$$d_{k_1} = 37$$

$$\mu_2(X) = \begin{cases} 1, & if\ Z_1(X) \le 283 \\ 1 - \dfrac{Z_2(X) - 283}{295 - 283}, & if\ 283 \le Z_1(X) \le 295 \\ 0, & if\ Z_1(X) \ge 295 \end{cases}$$

$$d_{k_2} = 12$$

We find

Minimize α

$$Z_1(X) + 37a \le 335 \quad (22)$$

$$Z_2(X) + 12a \le 295 \quad (23)$$

We Solve the above crisp model and the Solver window appears given below

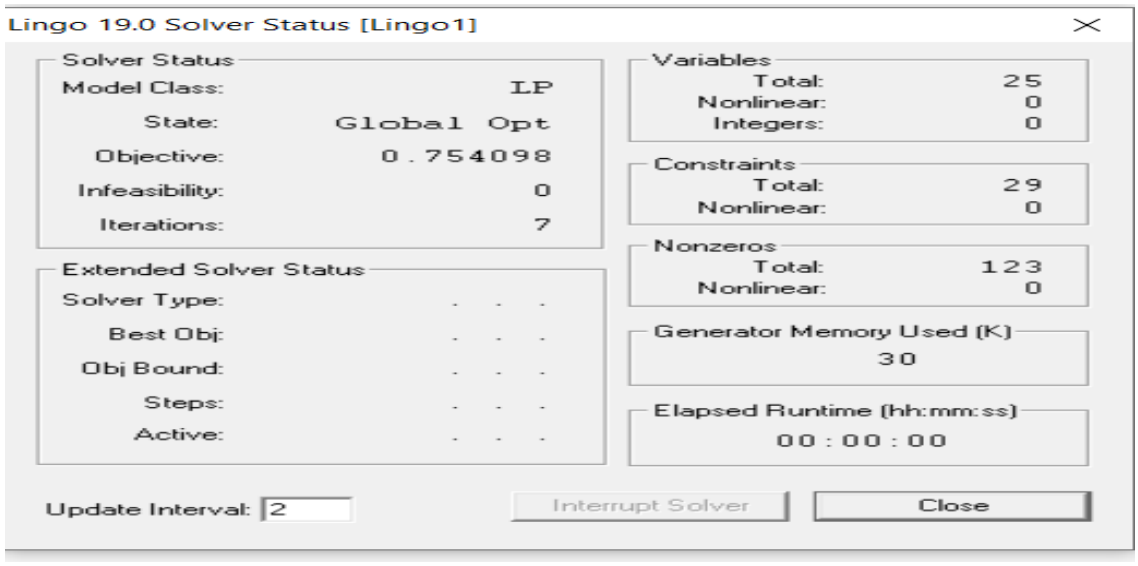


Figure 3: *Illustration of model 22 and 23 on LINGO window*

By closing above window, we can view the solution of the Model 22 and 23 is given
(Global optimal solution found).

Objective value: 0.7540984

Infeasibilities: 0.000000

Total solver iterations:	7
Elapsed runtime seconds:	0.11
Model Class:	LP
Total variables:	25
Nonlinear variables:	0
Integer variables:	0
Total constraints:	29
Nonlinear constraints:	0
Total nonzeros:	123
Nonlinear nonzeros:	0

Table 3: *Optimal Solution of crisp model 22 and 23*

Variable	Value	Reduced Cost
A	0.7540984	0.000000
X11M	5.000000	0.000000
X12M	3.316940	0.000000
X13M	0.6830601	1.000000
X21M	7.000000	0.000000
X22M	3.683060	0.000000
X23M	3.316940	0.000000
X31M	0.000000	0.4918033E-01
X32M	0.000000	0.2622951
X33M	6.000000	0.000000
X41M	2.000000	0.000000
X42M	5.000000	0.000000
X43M	0.000000	0.1639344E-01
X11N	0.000000	0.1311475
X12N	3.683060	0.000000
X13N	2.316940	0.000000
X21N	1.000000	0.000000
X22N	0.3169399	0.000000
X23N	5.683060	0.000000
X31N	4.000000	0.000000
X32N	1.000000	0.000000
X33N	0.000000	0.000000
X41N	0.000000	0.4918033E-01
X42N	3.000000	0.000000
X43N	3.000000	0.000000

$$Z_1^* = 307.04 \quad Z_2^* = 285.92 \qquad \alpha = 0.754$$

Maximize α

$$\alpha \geq \frac{1}{2} + \frac{1}{2}\tanh\{\frac{U_k^\mu + L_k^\mu}{2} - Z_k\}\tau_k$$

Further implies that

$$Z_k\tau_k + \tanh^{-1}(2\alpha - 1) \leq \frac{U_k^\mu + L_k^\mu}{2}\tau_k \tag{24}$$

Maximize w

$$6(4\cdot X_{11}^{(1)}+3\cdot X_{12}^{(1)}+5\cdot X_{13}^{(1)}+8\cdot X_{21}^{(1)}+6\cdot X_{22}^{(1)}+2\cdot X_{23}^{(1)}+7\cdot X_{31}^{(1)}+4\cdot X_{32}^{(1)}+1\cdot X_{33}^{(1)}+9\cdot X_{41}^{(1)}$$
$$+10\cdot X_{42}^{(1)}+12\cdot X_{43}^{(1)}+8\cdot X_{11}^{(2)}+6\cdot X_{12}^{(2)}+3\cdot X_{13}^{(2)}+5\cdot X_{21}^{(2)}+4\cdot X_{22}^{(2)}+1\cdot X_{23}^{(2)}+9\cdot X_{31}^{(2)}+2\cdot X_{32}^{(2)}$$
$$+6\cdot X_{33}^{(2)}+4\cdot X_{41}^{(2)}+9\cdot X_{42}^{(2)}+3\cdot X_{43}^{(2)})+37\cdot\alpha\le 1899$$
$$6(5\cdot X_{11}^{(1)}+6\cdot X_{12}^{(1)}+7\cdot X_{13}^{(1)}+4\cdot X_{21}^{(1)}+5\cdot X_{22}^{(1)}+2\cdot X_{23}^{(1)}+1\cdot X_{31}^{(1)}+3\cdot X_{32}^{(1)}+4\cdot X_{33}^{(1)}+4\cdot X_{41}^{(1)}$$
$$+2\cdot X_{42}^{(1)}+3\cdot X_{43}^{(1)}+10\cdot X_{11}^{(2)}+9\cdot X_{12}^{(2)}+9\cdot X_{13}^{(2)}+7\cdot X_{21}^{(2)}+9\cdot X_{22}^{(2)}+2\cdot X_{23}^{(2)}+8\cdot X_{31}^{(2)}+7\cdot X_{32}^{(2)}$$
$$+9\cdot X_{33}^{(2)}+8\cdot X_{41}^{(2)}+4\cdot X_{41}^{(2)}+5\cdot X_{43}^{(2)})+12\cdot\alpha\le 1734$$

Subject to the condition (21)

Solver window of model (24) appears below

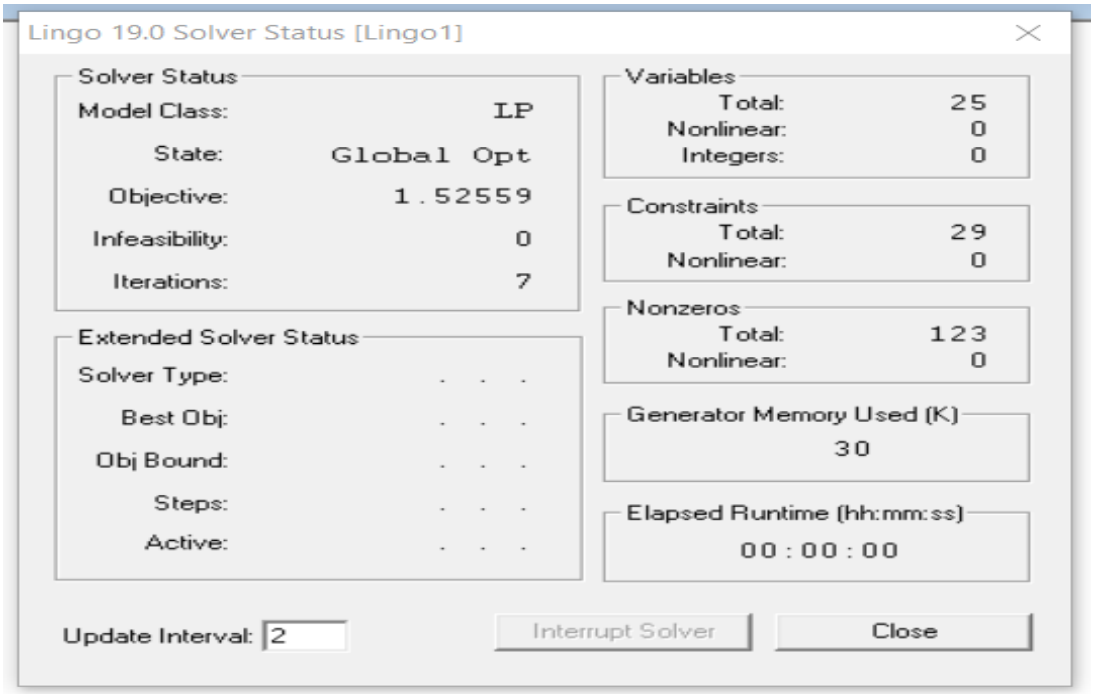


Figure 4: *Illustration of model 24 on LINGO window*

Solution of above model (24) is given by
(Global optimal solution found).

Objective value:	1.525591
Infeasibilities:	0.000000
Total solver iterations:	7
Elapsed runtime seconds:	0.17
Model Class:	LP
Total variables:	25
Nonlinear variables:	0
Integer variables:	0
Total constraints:	29
Nonlinear constraints:	0
Total nonzeros:	123
Nonlinear nonzeros:	0

Table 4 : *Optimal solution of model 24*

Variable	Value	Reduced Cost
A	1.525591	0.000000
X11M	5.000000	0.000000
X12M	3.316273	0.000000
X13M	0.6837270	0.000000
X21M	7.000000	0.000000
X22M	3.683727	0.000000
X23M	3.316273	0.000000
X31M	0.000000	0.4918033E-01
X32M	0.000000	1.574803
X33M	6.000000	0.000000
X41M	2.000000	0.000000
X42M	5.000000	0.000000
X43M	0.000000	0.9842520E-0
X11N	0.000000	0.7874016
X12N	3.683727	0.000000
X13N	2.316273	0.000000
X21N	1.000000	0.000000
X22N	0.3162730	0.000000
X23N	5.683727	0.000000
X31N	4.000000	0.000000
X32N	1.000000	0.000000
X33N	0.000000	0.000000
X41N	0.000000	0.2952756
X42N	3.000000	0.000000
X43N	3.000000	0.000000

$$Z_1^* = 307.04 \quad Z_2^* = 285.92$$

Where $w = \tanh^{-1}(2\alpha - 1)$ *and* w=1.52, $\alpha = 0.92$

$$\mu_1^e\{Z_1(X)\} = \begin{cases} 1, & 298 \le Z_1(X) \\ e^{-\frac{1}{2}\left(\frac{Z_1 - 298}{d_k}\right)} & 298 \le Z_1(X) \le 335, \quad d_{k_1} = 37 \\ 0, & Z_1(X) \ge 335 \end{cases}$$

$$\mu_2^e\{Z_2(X)\} = \begin{cases} 1, & 283 \le Z_1(X) \\ e^{-\frac{1}{2}\left(\frac{Z_2 - 283}{d_k}\right)} & 283 \le Z_2(X) \le 295, \quad d_{k_2} = 12 \\ 0, & Z_2(X) \ge 295 \end{cases}$$

Maximize α

Such that

$$\alpha \le e^{-\frac{1}{2}\left(\frac{307.04 - 298}{37}\right)} \text{ and } \alpha \le e^{-\frac{1}{2}\left(\frac{285.92 - 283}{12}\right)}$$

The solution of the problem is given by $\alpha = 0.89$

5. Conclusion

A fuzzy MOMIT algorithm is constructed in this paper, and with the help of numerical examples, a solution is demonstrated using three various kinds of membership functions, including linear, hyperbolic, and exponential membership functions. The numerous modes of transporting goods between points of origin and destination are represented by the multi-index transportation problem. The crisp model becomes a linear one when the hyperbolic membership functionis used. When compared to the linear membership function and hyperbolic membership function, the optimum compromise solution is drastically different. However, there is no significant different between the linear membership function's solution and the exponential membership function's ideal compromise solution.

There are numerous methods that future research in the field of fuzzy programming might be carried out. For problems involving many scales and several objectives in linear programming, employing fuzzy programming to design decision support systems will be particularly beneficial in real-world scenarios. Future research may take into account the use of fuzzy programming to solve MOTP's when the supply and demand factors are simply made up of fuzzy integers. There is still room for more research into duality theory and post optimality analysis in multi-objective two- and three-dimensional transportation problems. The demand parameters, agency capacity, and mode of transportation capacity can be anticipated as random variables that follow specific probability distributions in multi-index transportation problems, in addition to the two index transportation problem.

References

[1] Anjana K and Kuiri, A., Das, B. (2020). An application of FISM and TOPSIS to a multi-objective multi-item solid transportation problem. *OPSEARCH* 57**,** 1299–1318.

[2] Archana K and Veena A., (2015). On multi-index fixed charge bi-criterion transportation problem, *OPSEARCH*, 52, 733-745https://doi.org/10.1007/s12597-020-00456-7.

[3] Bit, A. K., (2004). Fuzzy programming with hyperbolic membership functions for Multi-objective capacitated transportation problem, *OPSEARCH*, 41, 106-120.

[4] Bit A.K, Biswal M.P, and Alam S.S, (1993). Fuzzy programming approach to multi-objective solid transportation problem, *Fuzzy sets and system*, volume 57, 26, pages 183-194.

[5] Kaur D., Mukherjee S and Basu. K, (2015). Solution of a multi-objective and multi-index real-life transportation probem using different fuzzy membership functions., *Journal of Optimization theory and applications*,164, 666-678.

[6] Lohgaonkar.M H, Bajaj. V.H, Jadhav V.A. and Patwari V.A, (2010). Fuzzy multi-objective multi-index transportation problem, *Advances in Information Mining*, Volume 2, Issue 1, pp-01-07.

[7] Porchelvi R.S and Anitha M., (2018). On Solving Multi Objective Interval Transportation Problem Using Fuzzy Programming Technique, *International Journal of Pure and Applied Mathematics*, Volume 118 No. 6, 483-491.

[8] Rautman C. A., Reid R. A., Ryder E. E., (1993). Scheduling the disposal of nuclear waste material in a geologic repository using the transportation model, *Operations Research*, 41, 459-469.

[9] Sungeeta S., Renu T., and Deepali S., (2018). Optimal Solutions of the Fuzzy Multi Index Bi-criteria Fixed Charge Bottleneck Transportation Problem, *Intern. J. Fuzzy Mathematical Archive*,Vol. 15, No. 1, 2018, 19-36 ISSN: 2320 –3242 (P), 2320 –3250.

[10] Wang, L.; Esenkov, A.S.; Tizik, A.P.; Torchinskaya, E.V. (2018). Decomposition Method for Solving Three-Index Transportation Problems. *J. Comput. Syst. Sci. Int*. 57, 759–765.

CHARACTERIZATION OF NEW QUASI LINDLEY DISTRIBUTION BY TRUNCATED MOMENTS AND CONDITIONAL EXPECTATION OF ORDER STATISTICS

Mohd. Amir Ansari
Department of Community Medicine
Maharishi Markandeshwar Medical College and Hospital,
Kumarhatti, Solan Himachal Pradesh-173229-India
amir_stats@live.co.uk

•

Mohammad Faizan
Department of Statistics and Operations Research
Aligarh Muslim University, Aligarh-202002 India
mdfaizan02@gmail.com

•

Rafiqullah Khan
Department of Statistics and Operations Research
Aligarh Muslim University, Aligarh-202002 India
aruke11@gmail.com

Abstract

Characterization of a probability distribution plays an important role in probability and statistics. Before a particular probability distribution model is applied to fit the real world data, it is necessary to confirm whether the given probability distribution satisfies the underlying requirements by its characterization. The aim of this paper is to find characterization results New Quasi Lindley distribution. These results are established using the relation between truncated moments and failure rate functions and conditional expectation of adjacent order statistics. The first characterization result is based on relation between left truncation moment and failure rate function while the second result is based on relation between right truncated moment and reverse failure rate function. In the third characterization result we used conditional expectation of order statistics when the conditioned one is adjacent order statistics. Further, some of its important deductions are also discussed.

Keywords: New quasi Lindley distribution, Characterization, Truncated moments, Failure rate function, Reversed failure rate function, Order statistics.

1. Introduction

Characterization is a condition involving certain properties of a random variable $X=(X_1, X_2, ..., X_n)$, which identifies the associated distribution function $F(x)$. The property that uniquely determines $F(x)$ may be based on a function of random variables whose joint distribution is related to that of $X=(X_1, X_2, ..., X_n)$. The only method of finding distribution function $F(x)$ exactly, which avoids the subjective choice, is a characterization theorem. A theorem is on a characterization of a distribution function if it concludes that a set of conditions is satisfied by $F(x)$ and only by $F(x)$. There has been a vast development in the field of

characterizing distributions through different techniques; mainly characterization through distributional properties, moments and conditional expectation.Characterization of a probability distribution plays an important role in probability and statistics. There has been a great interest, in recent years, in the characterizations of probability distributions by truncated moments (see, for example [1], [2], [3], [4], [5], [6], [7], [8], [9], [10] amongst others).

Lindley distribution was introduced by [11]. A random variable X is said to have Lindley distribution with parameter θ if its probability density function (pdf) is of the form

$$f(x)=\frac{\theta^2}{1+\theta}(1+x)e^{-\theta x}\ ;\quad x>0,\ \theta>0. \tag{1}$$

Its distribution function (df) is

$$F(x)=1-\frac{\theta+1+\theta x}{1+\theta}e^{-\theta x}\ ;\quad x>0,\ \theta>0. \tag{2}$$

Various properties of this distribution have been discussed by [12] and they showed, in many ways (1) provides a better model for some applications than the exponential distribution.

A two-parameter distribution called quasi Lindley distribution (QLD) has been introduced by [13]. A distribution with parameters α and θ is said to have quasi Lindley distribution if its pdf is of the form

$$f(x;\alpha,\theta)=\frac{\theta(\alpha+\theta x)}{\alpha+1}e^{-\theta x}\ ;\qquad x>0,\theta>0,\ \alpha>-1 \tag{3}$$

and the df is

$$F(x;\alpha,\theta)=1-\frac{1+\alpha+\theta x}{\alpha+1}e^{-\theta x}\ ;\qquad x>0,\theta>0,\ \alpha>-1. \tag{4}$$

It can easily be seen that at $\alpha=\theta$, the QLD reduces to the Lindley distribution and at $\alpha=0$, it reduces to the gamma distribution with parameters $(2,\theta)$.[13] have discussed its various properties and showed that this QLD is a better model than the Lindley distribution for modeling waiting and survival times data.

A new form of quasi Lindley distribution called new quasi Lindley distribution (NQLD) is introduced by [14]. The pdf of NQLD is given by

$$f(x;\theta,\alpha)=\frac{\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x};\quad x>0,\theta>0,\alpha<-\theta^2 \tag{5}$$

It can easily be seen that at $\alpha=\theta$, the new QLD (5) reduces to the Lindley distribution (1) and at $\alpha=0$, it reduces to the exponential distribution with parameter θ.

The df of the new QLD is obtained as

$$F(x;\theta,\alpha)=1-\frac{\theta^2+\alpha+\theta\alpha x}{\theta^2+\alpha}e^{-\theta x};\quad x>0,\theta>0,\alpha<-\theta^2\ . \tag{6}$$

The failure rate function (frf) of New Quasi Lindley distribution with parameters α and θ is given by

$$r(x)=\frac{f(x)}{1-F(x)}=\frac{\theta^2(\theta+\alpha x)}{\theta^2+\alpha+\alpha\theta x}. \tag{7}$$

The reverse failure rate function $(rfrf)$ of NQLD with parameters α and θ is given by

$$\eta(x)=\frac{f(x)}{F(x)}=\frac{\theta^2(\theta+\alpha x)}{(\theta^2+\alpha)e^{\theta x}-(\theta^2+\alpha+\alpha\theta x)}. \tag{8}$$

The k^{th} moment (about the origin) of the NQLD with parameters α and θ is given by

$$E[X^k] = \frac{k!(\alpha k + \theta^2 + \alpha)}{(\theta^2 + \alpha)\theta^k}. \quad (9)$$

Several properties of NQLD have been given by [14] and they showed that NQLD is more flexible than Lindley distribution, exponential distribution, and QLD. They also showed that NQLD is closer fit than exponential distribution, Lindley distribution and QLD in goodness of fit.

Several characterization results of Lindley distribution has been given by [6]. They characterized Lindley distribution through left and right truncated moments. Conditional expectation of order statistics is used to characterize Lindley distribution by [7]. In this paper, we have obtained characterization results for quasi Lindley distribution.

2. Characterizations through Truncated Moments:

First, we give the following two lemmas which are used to prove Theorem 1 and Theorem 2 respectively.

Lemma 1: Suppose that the random variable X has an absolutely continuous df $F(x)$ with $F(0) = 0, F(x) > 0$ for all x, pdf $f(x) = F'(x)$, frf $r(x) = \frac{f(x)}{[1 - F(x)]}$. Let $g(x)$ be a continuous function in $x > 0$ and $0 < E[g(X)] < \infty$. If

$$E[g(x) \mid X > x] = h(x)\, r(x) \qquad x > 0$$

where $h(x)$ is a differentiable function in $x > 0$, then

$$f(x) = K \exp\left[-\int_0^x \frac{g(y) + h'(y)}{h(y)} dy \right], \quad x > 0$$

where $K > 0$ is a normalizing constant [6].

Lemma 2: Suppose that the random variable X has an absolutely continuous cdf $F(x)$ with $F(0) = 0, F(x) > 0$ for all x, pdf $f(x) = F'(x)$, $rfrf$ $\eta(x) = \frac{f(x)}{F(x)}$. Let $g(x)$ be a continuous function in $x > 0$ and $0 < E[g(X)] < \infty$. If

$$E[g(x) \mid X \le x] = w(x)\, r(x) \qquad x > 0$$

where $w(x)$ is a differentiable function in $x > 0$, then

$$f(x) = K \exp\left[-\int_0^x \frac{w'(y) - g(y)}{w(y)} dy \right], \quad x > 0$$

where $K > 0$ is a normalizing constant [6].

Theorem 1: Suppose that the random variable X has absolutely continuous distribution with the pdf $f(x)$ and df $F(x)$ with $F(0) = 0$, $F(x) > 0$ for all $x > 0$ and frf $r(x) = \frac{f(x)}{1 - F(x)}$. Assume that $0 < E[X^k] < \infty$ for a given positive integer k. Then X has NQLD with parameters α and θ if and only if

$$[E(X^k \mid X \ge x)] = \frac{r(x)}{(\theta + \alpha x)} \sum_{j=0}^{k+1} \mu_j x^j \quad (10)$$

where $\mu_0 = \frac{k![\theta^2+\alpha(k+1)]}{\theta^{k+2}}$, $\mu_j = \frac{k![\theta^2+\alpha(k+1)]}{j!\theta^{k-j+2}}$, $\mu_{j+1} = \frac{\theta}{j+1}\mu_j \quad j = 0,1,2,...,k-1$ and

$\mu_{k+1} = \frac{\alpha}{\theta}$.

Proof: for necessary condition, we have

$$[E(X^k \mid X \geq x)] = \frac{1}{1-F(x)}\int_x^\infty y^k f(y)dy. \tag{11}$$

From relation among pdf, df and frf, we have

$$[E(X^k \mid X \geq x)] = \frac{r(x)}{f(x)}\int_x^\infty y^k f(y)dy. \tag{12}$$

From (5), we have

$$[E(X^k \mid X \geq x)] = \frac{r(x)}{(\theta+\alpha x)e^{-\theta x}}\left[\int_x^\infty (\theta+\alpha y)y^k e^{-\theta y}dy\right]$$

$$= \frac{r(x)}{(\theta+\alpha x)e^{-\theta x}}\left[\theta\int_x^\infty y^k e^{-\theta y}dy + \alpha\int_x^\infty y^{k+1}e^{-\theta y}dy\right]. \tag{13}$$

The following integration result given by [15] (pg-340) is used to solve the above integration

$$\int_u^\infty x^n e^{-\mu x}dy = e^{-\mu u}\sum_{k=0}^{n}\frac{n!}{k!}\frac{u^k}{\mu^{n-k+1}}. \tag{14}$$

Using (14), we can write (13) as

$$[E(X^k \mid X \geq x)] = \frac{r(x)}{(\theta+\alpha x)e^{-\theta x}}\left[\theta e^{-\theta x}\sum_{j=0}^{k}\frac{k!}{j!}\frac{x^j}{\theta^{k-j+1}} + \alpha e^{-\theta x}\sum_{j=0}^{k+1}\frac{(k+1)!}{j!}\frac{x^j}{\theta^{k+1-j+1}}\right]$$

$$= \frac{r(x)}{(\theta+\alpha x)}\left[\sum_{j=0}^{k}\frac{k![\theta^2+\alpha(k+1)]}{j!}\frac{x^k}{\theta^{k-j+2}} + x^{k+1}\right]^j$$

$$= \frac{r(x)}{(\theta+\alpha x)}\sum_{j=0}^{k+1}\mu_j x$$

where $\mu_0 = \frac{k![\theta^2+\alpha(k+1)]}{\theta^{k+2}}$, $\mu_j = \frac{k![\theta^2+\alpha(k+1)]}{j!\theta^{k-j+2}}$, $\mu_{j+1} = \frac{\theta}{j+1}\mu_j \quad j = 0,1,2,...,k-1$ and

$\mu_{k+1} = \frac{\alpha}{\theta}$.

and hence the necessary part

For sufficiency part let $g(x) = x^k$ and $h(x) = \frac{\sum_{j=0}^{k+1}\mu_j x^j}{(\theta+\alpha x)}$. (15)

Using the recurrence relations of the μ_k 's, we have

$$\theta\sum_{j=0}^{k+1}\mu_j x^j - \sum_{j=1}^{k+1} j\mu_j x^j = \sum_{j=0}^{k-1}[\theta\mu_j - (j+1)\mu_{j+1}]x^j + [\theta\mu_k - (k+1)\mu_{k+1}]x^k + \theta\mu_{k+1}x^{k+1}$$

$$= (\theta+\alpha x)x^k. \tag{16}$$

From (15) and (16), we have

$$\frac{g(x)}{h(x)}=\frac{(\theta+\alpha x)x^j}{\sum_{j=0}^{k+1}\mu_j x^j}=\frac{\theta\sum_{j=0}^{k+1}\mu_j x^j-\sum_{j=1}^{k+1}j\mu_j x^j}{\sum_{j=0}^{k+1}\mu_j x^j}==\theta-\frac{\sum_{j=1}^{k+1}j\mu_j x^j}{\sum_{j=0}^{k+1}\mu_j x^j} \tag{17}$$

Also from (15)

$$h(x)=\frac{\sum_{j=0}^{k+1}\mu_j x^j}{(\theta+\alpha x)}$$

Taking logarithm both sides of above equation, we have

$$\ln[h(x)]=-\ln[\theta+\alpha x]+\ln[\sum_{j=0}^{k+1}\mu_j x^j]. \tag{18}$$

Differentiating (18) with respect to x, we have

$$\frac{h'(x)}{h(x)}=\frac{\sum_{j=1}^{k+1}j\mu_j x^j}{\sum_{j=0}^{k+1}\mu_j x^j}-\frac{\alpha}{(\theta+\alpha x)}. \tag{19}$$

From (17) and (19), we have

$$\frac{h'(x)+g(x)}{h(x)}=-\frac{\alpha}{(\theta+\alpha x)}+\theta. \tag{20}$$

Integrating (20) over $(0,x)$, we have

$$\int_0^x\frac{h'(y)+g(y)}{h(y)}dy=\int_0^x\left(-\frac{\alpha}{(\theta+\alpha y)}+\theta\right)dy=-\ln[\theta+\alpha x]+\theta x \tag{21}$$

Using Lemma 1 given by [6] in (21), we have

$$f(x)=C\exp\left[-\int_0^x\frac{h'(y)+g(y)}{h(y)}dy\right]=C\exp[-\ln[\theta+\alpha x]+\theta x]=C(\theta+\alpha x)e^{-\theta x}$$

where C is normalizing constant i.e. $\int_0^\infty f(x)=1$ which gives $C=\frac{\theta^2}{\theta^2+\alpha}$.

Therefore,

$$f(x)=\frac{\theta^2}{\theta^2+\alpha}(\alpha+\theta x)e^{-\theta x}$$

which is the pdf of NQLD and hence sufficiency part.

Remark 1: Putting $\alpha=\theta$ in Theorem 1 we get the characterizing result for Lindley distribution obtained by [6].

Remark 2: Putting $\alpha=0$ in Theorem 1 we get the characterizing result for exponential (θ) distribution.

Theorem 2: Suppose that the random variable X has absolutely continuous distribution with the pdf $f(x)$ and df $F(x)$ with $F(0)=0$, $F(x)>0$ for all $x>0$ and the $rfrf$ $\eta(x)=\frac{f(x)}{F(x)}$.

Assume that $0<E[X^k]<\infty$ for a given positive integer k. Then X has NQLD if and only if

$$[E(X^k \mid X \le x)] = \frac{\eta(x)}{(\theta+\alpha x)e^{-\theta x}}\left(\mu_0 e^{\theta x} - \sum_{j=0}^{k+1} \mu_j x^j \right) \tag{22}$$

where $\mu_0 = \frac{k![\theta^2+\alpha(k+1)]}{\theta^{k+2}}$, $\mu_j = \frac{k![\theta^2+\alpha(k+1)]}{j!\theta^{k-j+2}}$, $\mu_{j+1} = \frac{\theta}{j+1}\mu_j \quad j = 0,1,2,\ldots,k-1$ and

$\mu_{k+1} = \frac{\alpha}{\theta}$.

Proof: for necessary part let X has NQLD with parameter α and θ.
Then we have

$$[E(X^k \mid X \le x)] = \frac{1}{F(x)}\int_x^\infty y^k f(y)dy$$

from relation among pdf df , and $rfrf$, we have

$$[E(X^k \mid X \le x)] = \frac{\eta(x)}{f(x)}\int_x^\infty y^k f(y)dy . \tag{23}$$

From (5), we have

$$[E(X^k \mid X \le x)] = \frac{\eta(x)}{(\theta+\alpha x)e^{-\theta x}}\left[\int_0^x (\theta+\alpha y)y^k e^{-\theta y}dy\right]$$

or

$$[E(X^k \mid X \le x)] = \frac{\eta(x)}{(\theta+\alpha x)e^{-\theta x}}\left[\theta\int_0^x y^k e^{-\theta y}dy + \alpha\int_0^x y^{k+1}e^{-\theta y}dy\right]. \tag{24}$$

The following integration result given by [15] (pg-340) is used to solve the above integration

$$\int_0^u x^n e^{-\mu x}dy = \frac{n!}{\mu^{n+1}} - e^{-\mu u}\sum_{k=0}^{n}\frac{n!}{k!}\frac{u^k}{\mu^{n-k+1}} . \tag{25}$$

From (24) and (25), we have

$$[E(X^k \mid X \le x)] = \frac{\eta(x)}{(\theta+\alpha x)e^{-\theta x}}$$

$$\times\left(\frac{k![\theta^2+\alpha(k+1)]}{\theta^{k+2}}e^{\theta x} - \sum_{j=0}^{k}\frac{k![\theta^2+\alpha(k+1)]}{j!}\frac{x^j}{\theta^{k-j+2}} - \frac{\alpha}{\theta}x^{k+1}\right)$$

$$= \frac{\eta(x)}{(\theta+\alpha x)e^{-\theta x}}\left(\mu_0 e^{\theta x} - \sum_{j=0}^{k+1}\mu_j x^j \right)$$

where $\mu_0 = \frac{k![\theta^2+\alpha(k+1)]}{\theta^{k+2}}$, $\mu_j = \frac{k![\theta^2+\alpha(k+1)]}{j!\theta^{k-j+2}}$, $\mu_{j+1} = \frac{\theta}{j+1}\mu_j \quad j = 0,1,2,\ldots,k-1$ and

$\mu_{k+1} = \frac{\alpha}{\theta}$.

and hence the necessary part.

For sufficient part Let $g(x) = x^k$ and $w(x) = \dfrac{\mu_0 e^{\theta x} - \sum_{j=0}^{k+1}\mu_j x^j}{(\theta+\alpha x)}$. (26)

From equations (16) and (26), we have

$$\frac{g(x)}{w(x)}=\frac{(\theta+\alpha x)x^{k}}{\mu_0 e^{\theta x}-\sum_{j=0}^{k+1}\mu_j x^j}=\frac{\theta\sum_{j=0}^{k+1}\mu_j x^j-\sum_{j=1}^{k+1}j\mu_j x^j}{\mu_0 e^{\theta x}-\sum_{j=0}^{k+1}\mu_j x^j}=\frac{\theta\mu_0 e^{\theta x}-\sum_{j=1}^{k+1}j\mu_j}{\mu_0 e^{\theta x}-\sum_{j=0}^{k+1}\mu_j x^j}-\theta\ . \tag{27}$$

Also from equation (26)

$$\ln[w(x)]=-\ln[(\theta+\alpha x)]+\ln\left[\mu_0 e^{\theta x}-\sum_{j=0}^{k+1}\mu_j x^j\right]. \tag{28}$$

Differentiating (28) with respect to x, we have

$$\frac{w'(x)}{w(x)}=-\frac{\alpha}{(\theta+\alpha x)}+\frac{\theta\mu_0 e^{\theta x}-\sum_{j=1}^{k+1}j\mu_j}{\mu_0 e^{\theta x}-\sum_{j=0}^{k+1}\mu_j x^j}.$$

Using (27) in above equation, we have

$$\frac{w'(x)}{w(x)}=-\frac{\alpha}{(\theta+\alpha x)}+\frac{g(x)}{w(x)}+\theta$$

which can be written as

$$\frac{w'(x)-g(x)}{w(x)}=-\frac{\alpha}{(\theta+\alpha x)}+\theta. \tag{29}$$

Integrating (29) over $(0,x)$, we have

$$\int_0^x\frac{w'(y)-g(y)}{w(y)}dy=\int_0^x\left(-\frac{\alpha}{(\theta+\alpha y)}+\theta\right)dy=-\ln[(\theta+\alpha x)]+\theta x \tag{30}$$

Using Lemma 2 given by [6], we have

$$f(x)=C\exp\left[-\int_0^x\frac{w'(y)-g(y)}{w(y)}dy\right]=C\exp\left[\ln[(\theta+\alpha x)]-\theta x\right]=C(\theta+\alpha x)e^{-\theta x}.$$

where C is normalizing constant i.e. $\int_0^\infty f(x)dx=1$ which gives $C=\frac{\theta^2}{\theta^2+\alpha}$.

Therefore,

$$f(x)=\frac{\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x}$$

which is the pdf of NQLD and hence the theorem.

Remark 3: Putting $\alpha=\theta$ in Theorem 2 we get the characterizing result for Lindley distribution obtained by [6].

Remark 4: Putting $\alpha=0$ in Theorem 2 we get the characterizing result for exponential (θ) distribution.

Characterization through conditional expectation of order statistics

Theorem 3: Let X be a continuous random variable with $df\ F(x)$ and $pdf\ f(x)$. Then for $r<n$

$$E[X_{r+1:n}\mid X_{r:n}=x]=x+\frac{\Gamma[n-r+1,(n-r)(\theta^2+\alpha+\theta\alpha x)]}{\theta\alpha(n-r)^{n-r+1}(\theta^2+\alpha+\theta\alpha x)^{n-r}e^{-(n-r)(\theta^2+\alpha+\theta\alpha x)}}, \tag{31}$$

$$r=1,2,...,n-1$$

if and only if X has the df given in (6).

Proof: First we will prove (6) implies (31).

We have

$$E[X_{r+1:n} \mid X_{r:n} = x] = \frac{(n-r)}{[1-F(x)]^{n-r}} \int_x^{\beta} y[1-F(y)]^{n-r-1} f(y)\, dy .$$

From (5) and (6), we have

$$E[X_{r+1:n} \mid X_{r:n} = x] = \frac{(n-r)\theta^2}{\left[(\theta^2+\alpha+\theta\alpha x)e^{-\theta x}\right]^{n-r}} \times \int_x^{\infty} y[(\theta^2+\alpha+\theta\alpha y)e^{-\theta y}]^{n-r-1}(\theta+\alpha y)e^{-\theta y} dy. \tag{32}$$

Integrating (32) by parts and then re arranging, we have

$$E[X_{r+1:n} \mid X_{r:n} = x] = x + \frac{\int_x^{\infty} (\theta^2+\alpha+\theta\alpha y)^{n-r} e^{-\theta(n-r)y} dy}{\left[(\theta^2+\alpha+\theta\alpha x)e^{-\theta x}\right]^{n-r}}$$

which reduces to

$$E[X_{r+1:n} \mid X_{r:n} = x] = x + \frac{\Gamma[n-r+1,(n-r)(\theta^2+\alpha+\theta\alpha x)]}{\theta\alpha(n-r)^{n-r+1}(\theta^2+\alpha+\theta\alpha x)^{n-r} e^{-\frac{(n-r)}{\alpha}(\theta^2+\alpha+\theta\alpha x)}}$$

and hence necessary part.

For sufficiency part, Let $E[X_{r+1:n} \mid X_{r:n} = x] = \phi(x) = x + \omega(x)$ (33)

where, $\omega(x) = \dfrac{\Gamma[n-r+1,(n-r)(\theta^2+\alpha+\theta\alpha x)]}{\theta\alpha(n-r)^{n-r+1}(\theta^2+\alpha+\theta\alpha x)^{n-r} e^{-\frac{(n-r)}{\alpha}(\theta^2+\alpha+\theta\alpha x)}}$.

From (33), we have

$$\phi(x) = x + \omega(x) .$$

Differentiating above equation with respect to x, we have

$$\phi'(x) = 1 + \omega'(x) ..$$

Then, it is seen that

$$\int \frac{\phi'(t)}{\phi(t)-t} dt = \int \frac{1+\omega'(t)}{\omega(t)} dt = \int \frac{dt}{\omega(t)} + \int \frac{\omega'(t)}{\omega(t)} dt .$$

Using (33) in above equation, we have

$$\int \frac{\phi'(t)}{\phi(t)-t} dt = -\ln\Gamma[n-r+1,(n-r)(\theta^2+\alpha+\theta\alpha t)] + \ln\left(\frac{\Gamma[n-r+1,(n-r)(\theta^2+\alpha+\theta\alpha t)]}{\theta\alpha(n-r)^{n-r+1}(\theta^2+\alpha+\theta\alpha t)^{n-r} e^{-\frac{(n-r)}{\alpha}(\theta^2+\alpha+\theta\alpha t)}} \right)$$

$$= -\ln\left[\theta\alpha(n-r)^{n-r+1}(\theta^2+\alpha+\theta\alpha t)^{n-r} e^{-\frac{(n-r)}{\alpha}(\theta^2+\alpha+\theta\alpha t)} \right]. \tag{34}$$

Taking the limits of integral in (34) from 0 to x, we have

$$\int_0^x \frac{\phi'(t)}{\phi(t)-t} dt = \ln\left[\left\{ \frac{\theta^2+\alpha}{(\theta^2+\alpha+\theta\alpha x)e^{-\theta x}} \right\}^{n-r} \right]. \tag{35}$$

Using the result given by [16], we have

$$[1-F(x)]^{n-r} = \exp\left[-\ln\left[\left\{\frac{\theta^2+\alpha}{(\theta^2+\alpha+\theta\alpha x)e^{-\theta x}}\right\}^{n-r}\right]\right] = \left(\frac{\theta^2+\alpha+\theta\alpha x}{\theta^2+\alpha}e^{-\theta x}\right)^{n-r}$$

which reduces to

$$F(x) = 1-\frac{\theta^2+\alpha+\theta\alpha x}{\theta^2+\alpha}e^{-\theta x}$$

and hence the theorem.

Remark 5: Putting $\alpha = \theta$ in Theorem 3, we get the characterization result for Lindley distribution as obtained by [7].

Remark 6: Putting $\alpha = 0$ in Theorem 3, we get the characterization result for exponential (θ) distribution.

3. Applications

A probability distribution can be characterized in many ways and the method under study here is one of them. We have used here the relation between truncated moments and failure rate functions as well as conditional expectation of order statistics conditioned on an adjacent order statistics to characterize the new quasi Lindley distribution. That is, we have characterized the new quasi Lindley distribution if the regression equation truncated from both sides is given, i. e. the data are truncated from left side at x and truncated from right side at y. In real practice, several times we get the data of which observations are missing either in beginning or in the end. In such type of data we can use the result of this paper.

References

[1] Hamedani, G.G. (2011). Characterizations of the Shakil-Kibria-Singh distribution. *Austrian Journal of Statistics*, 40 (3), 201–207.

[2] Hamedani, G.G. and Ghosh, I. (2015). Kumaraswamy-Half-Cauchy distribution: characterizations and related results. *International Journal of Statistics and Probability*, 4 (3): 94-100.

[3] Ahsanullah, M. and Shakil, M. (2015). Characterizations of continuous probability distributions occurring in physics and allied sciences by truncated moment. *International Journal of Advanced Statistics and Probability*, 3 (1): 100-114.

[4] Ahsanullah, M., Shakil, M. and Golam Kibria, B.M. (2016). Characterizations of continuous distributions by truncated moment. *Journal of Modern Applied Statistical Methods*, 15 (1): 316-331.

[5] Shakil, M., Golam Kibria, B.M. and Singh, J. N. (2016). Characterization of a new class of generalized Pearson distribution by truncated moment. *International Journal of Computational and Theoretical Statistics*, 3 (2): 91-100.

[6] Ahsanullah, M., Ghitany, M.E. and Al-Mutairi, D.K. (2017). Characterization of Lindley distribution by truncated moments. *Communications in Statistics - Theory and Methods*, 46 (12): 6222-6227.

[7] Kilany, N.M. (2017). Characterization of Lindley distribution based on truncated moments of order statistics. *Journal of Statistics Applications and Probability*, 6(2): 1-6.

[8] Kilany, N.M. and Hassanein, W.A. (2018). Characterization of Benktander type II distribution via truncated moments of order statistics. *International Journal of Probabilityand Statistics*, 7(4): 106–113.

[9] Athar, H. and Abdel-Aty, Y. (2020). Characterization of general class of distributions by truncated moment. *Thailand Statistician*, 18(2): 95–107.

[10] Athar, H., Abdel-Aty, Y. and Ali, M.A. (2021). Characterization of generalized distribution by doubly truncated moment. *Statistica*, LXXXI(1): 5-44.

[11] Lindley D.V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal*

Statistical Society Series B, 20(1): 102-107.

[12] Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its applications. *Mathematics and Computers in Simulation*, 78 (4**)**: 493 – 506.

[13] Shanker, R. and Mishra, A. (2013). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research*, 6 (4): 64-71.

[14] Shanker, R. and Ghebretsadik, A.H. (2013). A new quasi Lindley distribution. *International Journal of Statistics and Systems*, 8(2): 43-156.

[15] Gradshteyn, I.S. and Ryzhik, I.M., Table of Integrals, Series, and Products. Academic Press, New York, 2007.

[16] Khan, A.H., and Abu-Salih, M.S. (1989). Characterizations of probability distributions by conditional expectation of order statistics. *Metron*, 47: 171–181.

RT&A, No 4 (76)
Volume 18, December 2023

ANALYSIS OF AN $M/M/1/K$ FEEDBACK WORKING VACATION QUEUE WITH RENEGING

Krishan[1], Neetu Gupta*

•

*,[1]Department of Mathematics, J. C. Bose University of Science and Technology,
YMCA, Faridabad 121006, India
[1]krishan20kashyap16@gmail.com,
*neetuymca@yahoo.co.in

Abstract

The analysis of an $M/M/1/N$ feedback working vacation queueing system with reneging is presented in this paper. Customers may become impatient and even disappointed when they see a long line. In the literature on queueing, customer dissatisfaction caused on by unsatisfactory service is referred as feedback. In the case of feedback, customers retry services after receiving unsatisfactory or incomplete. First, we create the equations for the steady-state probabilities using the Markov process method. The steady-state probabilities are then solved by the matrix method. We then provide some system performance measures. We create a cost model using performance analysis. Finally, we give some numerical examples to show how the various model parameters affect the system's behaviour.

Keywords: reneging, performance measure, steady-state probabilities, cost model

1. Introduction

Because of the wide range of applications in practical situations such as inventory systems, hospital emergency rooms dealing with critical patients, computer and communication systems, impatient telephone switchboard customers, and manufacturing and production systems, many researchers have been drawn to performance modelling of Markovian queues with different customer behaviour. In queueing literature, feedback is defined as consumers' dissatisfaction due to poor service quality.

Takacs [1] studied feedback queue with a single server and determined the queue size distribution. Dhosi [2] give an overview of some general decomposition results and the methodology used to obtain these results for two vacation models. The analysis of a $M/M/1$ queue with multiple and single working vacations was presented by Vijaya Laxmi et al. [3]. Wortman et al., [4] investigated $M/GI/1$ Bernoulli feedback queue, where the server goes on vacation according to Bernoulli schedules. Zhang et al. [5] presented an analysis for an $M/M/1/N$ queueing system with balking, reneging and server vacations.

Kumar et al. [6] studied optimization of an $M/M/1/N$ feedback queue with retention of reneged customer. Rakesh and sumeet [7] researched a multi-server Markovian feedback queueing model with finite capacity that includes balking, reneging, and keeping reneged customers. Jewkes and Buzacott, [8] examined queueing system of a K class $M/G/1$ with feedback. Vijaya Laxmi and Jyothsna [9] worked on the effect of balking and reneging in a single server queue under variant working vacation policy.Bouchentouf et al. [10] analyzed and dealt with variant multiple working vacations and impatience timers of Bernoulli feedback queueing system that is dependent on the server states.

Sundar et al. [11] talks about a single server queuing system where units come in batches of different sizes under a Poisson stream. In k stations, the server offers services. Each station's service times are distributed randomly. The system uses the feedback, vacation, and reneging principles. Van den Berg and Boxma, [12] studied feedback mechanism of an $M/M/1$ queue with a general probability. Bouchentouf et al. [13] studied the examination of a Bernoulli feedback, single vacation, waiting server, and impatient customer Markovian queueing system.

With Bernoulli feedback, synchronous multiple and single working vacations, balking and reneging during busy and working vacation periods, Yahiaoui et al. [14] investigated the cost optimization analysis of a discrete-time, finite capacity, multiserver queueing system. Choi et al., [15, 16] considered $M/M/c$ retrial queues with feedback, geometric loss, and multi-class customers and the Bernoulli feedback policy of an $M/G/1$ queue. A single server $M/M/1/N$ feedback queueing system with vacation, balking, reneging, and retention of reneged customers is the subject of a study by Bouchentouf et al. [17]. Kumar et al., [18, 19, 20] studied retrial queue with Bernoulli feedback, control retrial rate with balking and multi-server, as well as a generalized $M/G/1$ feedback queue in which customers are either "positive" or "negative".

Bouchentouf and Guendouzi [21] deal with the study of an $M^X/M/c$ Bernoulli feedback queueing system with two different policies of synchronous vacations and waiting servers. Ke and Chang [22] studied balking and Bernoulli feedback with a general retrial queue, where a modified vacation policy is operated by the server. Markovian queueing system with working vacation, Bernoulli schedule interruption, setup time with feedback, reneging of impatient customers, and retention of reneged customers are all discussed by Gupta [23]. Jeeva and Kumari [24] gave a mathematical method to generate the membership function with non-linear programming of the $M/G/1$ system, feedback, and bulk arrival queues with server vacation facility, in which departure probability, service time, arrival rate, vacation time, and batch size was all considered as fuzzy numbers. Jain [25] presented various schemes under a probability-based model for queue scheduling operating system. M/G/1 feedback queue is also studied by [26, 27, 28].

In this paper, a single server with a finite capacity queueing system is analyzed. Performance analysis of the Markovian feedback working vacation queue with and reneging is also discussed. Customer dissatisfaction as a result of poor service quality is referred to as feedback in queueing literature. Customers retries service after receiving partial or incomplete service in the case of feedback. The model assumptions are described in section 2. The steady-state equations and their matrix solution is calculated in section 3 and 4 respectively. The system's performance measures are discussed in section 5. Numerical results and cost analysis of the system are described in sections 6 and 7 and the further paper concludes in the next section.

2. Model Assumptions

We consider a $M/M/1/K$ queueing model to analyze its performance and effectiveness of feedback working vacation queue with balking and reneging. The basic assumptions underlying the model are as:

1. Customer's arrival follows a Poisson distribution with arrival rate λ.
2. When the server is unoccupied, then customer may join the system with probability q or leave with complementary probability $p = (1 - q)$. probability of not joining. If a customer receives service and is dissatisfied, they may leave the system or return with a probability of $p_1 = 1 - q_1$ depending on whether they were a feedback customer or not.
3. Every customer has to wait for a certain time interval T, before being served. If the service hasn't begun by then, the customer may become irritated and leave the line, the density function is given below with the random variable T is

$$d(t) = \alpha e^{-\alpha t}, \ \ t \geq 0, \alpha > 0.$$

with mean as $\frac{1}{\alpha}$.

4. Since the departure and arrival rate of the impatient customer without service is independent, therefore the average reneging rate $r(k)$ is defined by the function as

$$r(k) = (k-i)\alpha, \;\; i \leq k \leq K, \, i = 0, 1$$

$$r(k) = 0, \;\; k > K$$

5. The server goes on vacation when the system is empty. After a vacation has ended, the server resumes regular service if there are still patrons in line; otherwise, he leaves for another vacation. The server remains operational and offers service to the arriving customers at a different service rate while on vacation. The assumption is that the vacation and service periods will follow a Poisson distribution with parameters ϕ and η, respectively. While on regular duty and while on working vacation, inter-arrival times, vacation times, and service times are all separate from one another.

$$v(t) = \eta e^{-\eta t}, \;\; t \geq 0, \eta > 0.$$

where η is the vacation rate of a server.

3. Steady-state Equation

In this section, we use the Markov process method to derive the steady-state probabilities. Let $\vartheta_{0,k}$ represent the probability that there are k customer in the system when the server is unavailable, When the server is online, $\vartheta_{1,k}$ represents the probability that k customers are present in the system. By using the Markov process theory, we are able to generate the following collection of steady-state equations.

$$0 = -\lambda\vartheta_{0,0} + (\eta p_1 + \alpha)\vartheta_{0,1} + \mu p_1 \vartheta_{1,1} \tag{1}$$
$$0 = \lambda\vartheta_{0,k-1} - (\eta p_1 + k\alpha + \lambda + \phi)\vartheta_{0,k} + (\eta p_1 + (k+1)\alpha)\vartheta_{0,k+1} \tag{2}$$
$$0 = \lambda\vartheta_{0,K-1} - (\eta p_1 + K\alpha + \phi)\vartheta_{0,K} \tag{3}$$
$$0 = -(\lambda + \eta p_1)\vartheta_{1,1} + (\mu p_1 + \alpha)\vartheta_{1,2} + \phi\vartheta_{0,1} \tag{4}$$
$$0 = \lambda\vartheta_{1,k-1} - (\lambda + \mu p_1 + (k-1)\alpha)\vartheta_{1,k} + (\mu p_1 + k\alpha)\vartheta_{1,k+1} + \phi\vartheta_{0,k} \tag{5}$$
$$0 = \lambda\vartheta_{1,K-1} - (\mu p_1 + (K-1)\alpha)\vartheta_{1,K} + \phi\vartheta_{0,K} \tag{6}$$

4. Solution

The steady-state probabilities $\vartheta_{j,k}, j = 0, 1, j \leq k \leq K$, are acquired when a set of equations has been solved (1)-(6) using matrices. Let $\mathbf{Y} = (\mathbf{Y_0}, \mathbf{Y_1})$ be the steady-state probability vector, where $\mathbf{Y_0} = (\vartheta_{0,0}, \vartheta_{0,1}, \cdots, \vartheta_{0,K})$ and $\mathbf{Y_1} = (\vartheta_{1,1}, \vartheta_{1,2}, \cdots, \vartheta_{1,K})$. The equation (1)-(6) the matrix form of which reads as

$$\mathbf{Y\Theta} = \mathbf{0} \tag{7}$$
$$\mathbf{YE} = \mathbf{1} \tag{8}$$

where $\mathbf{\Theta}$ is the block-formed transition rate matrix for the Markov process and $\mathbf{E}$ is a column vector with each component equal to one

$$\mathbf{\Theta} = \begin{pmatrix} \mathbf{B_1} & \mathbf{B_2} \\ \mathbf{B_3} & \mathbf{B_4} \end{pmatrix}$$

The elements of the matrices $\mathbf{B_1}, \mathbf{B_2}, \mathbf{B_3}$ and $\mathbf{B_4}$ are given by

$$\mathbf{B_1} = \begin{cases} -\lambda, & \text{if } k = j = 1, \\ \lambda, & \text{if } k = j-1, j \geq 2, \\ \eta p_1 + j\alpha, & \text{if } k = j+1, \\ u_j, & \text{if } k = j, j \geq 2, \\ 0, & \text{or else} \end{cases}$$

$$\mathbf{B_2} = \begin{cases} \phi, & \text{if } k = j+1, \\ 0, & \text{or else} \end{cases}$$

$$\mathbf{B_3} = \begin{cases} \mu p_1, & \text{if } k = j = 1, \\ 0, & \text{or else} \end{cases}$$

$$\mathbf{B_4} = \begin{cases} v_j, & \text{if } k = j, \\ \lambda, & \text{if } k = j-1, j \geq 2, \\ \mu p_1 + (j-1)\alpha, & \text{if } k = j+1, \\ 0, & \text{or else} \end{cases}$$

where for $1 \leq j \leq K, u_j = -(\eta p_1 + (j-1)\alpha + \lambda); v_j = -(\lambda + \mu p_1 + (j-1)\alpha)$. $\mathbf{B_1}$ is a square matrix of order $K+1$, $\mathbf{B_2}$ is a $(K+1) \times K$, $\mathbf{B_3}$ is a $K \times (K+1)$, $\mathbf{B_4}$ and is a square matrix of order K.
Based on the partition, $\mathbf{Y} = (\mathbf{Y_0}, \mathbf{Y_1})$, equation (7) and (8) can be written as:

$$\mathbf{Y_0 B_1} + \mathbf{Y_1 B_3} = \mathbf{0} \tag{9}$$

$$\mathbf{Y_0 B_2} + \mathbf{Y_1 B_4} = \mathbf{0} \tag{10}$$

$$\mathbf{Y_0 E_0} + \mathbf{Y_1 E_1} = \mathbf{1} \tag{11}$$

where $\mathbf{E_0}, \mathbf{E_1}$ are, respectively, column vectors of orders $K+1$ and K, each with one element.
From equation (9), we have

$$\mathbf{Y_0} = -\mathbf{Y_1 B_3 B_1^{-1}} \tag{12}$$

Using equation (12) in (10) and (11), we get

$$\mathbf{Y_1}(\mathbf{I} - \mathbf{B_3 B_1^{-1} B_2 B_4^{-1}}) = \mathbf{0} \tag{13}$$

$$\mathbf{Y_1}(\mathbf{E_1} - \mathbf{B_3 B_1^{-1} E_0}) = \mathbf{1} \tag{14}$$

The matrices $\mathbf{B_2}$ and $\mathbf{B_3}$ can be written as

$$\mathbf{B_3} = \begin{pmatrix} v_1 & \mathbf{O_1} \\ \mathbf{O_2} & \mathbf{O_3} \end{pmatrix}_{K \times (K+1)}, \mathbf{B_2} = \phi \begin{pmatrix} \mathbf{O_1} \\ \mathbf{I}_{K \times K} \end{pmatrix}_{(K+1) \times K}$$

where $\mathbf{O_1}, \mathbf{O_2}$ and $\mathbf{O_3}$ are zero matrices of order $1 \times K, (K-1) \times 1$ and $(K-1) \times K$, respectively. Let $\mathbf{B_1}^{-1} = [b_{k,j}]_{(K+1)\times(K+1)}$ and $\mathbf{w} = (b_{1,1}, b_{1,2}, \cdots, b_{1,K+1})$ denotes the 1st row of $\mathbf{B_1}^{-1}$.

$$\mathbf{B_3 B_1^{-1}} = \begin{pmatrix} v_1 \mathbf{w} \\ \mathbf{O_4} \end{pmatrix}_{K \times (K+1)} \tag{15}$$

where $\mathbf{O_4}$ is a zero matrix of order $(K-1) \times (K+1)$
Kow,

$$\mathbf{B_2 B_4^{-1}} = \phi \begin{pmatrix} \mathbf{O_1} \\ \mathbf{B_4^{-1}} \end{pmatrix}_{(K+1) \times K} \tag{16}$$

From equation (15) and (16), we have

$$\mathbf{B_3 B_1^{-1} B_2 B_4^{-1}} = v_1 \phi \begin{pmatrix} \mathbf{w_0 B_4^{-1}} \\ \mathbf{O_3} \end{pmatrix}_{(K+1) \times K} \tag{17}$$

where $\mathbf{w_0} = (b_{1,2}, b_{1,3}, \cdots, b_{1,K+1})$
Let us partition $\mathbf{Y_1}$ as $\left[\vartheta_{1,1}, \tilde{\mathbf{Y_1}}\right]$ where $\tilde{\mathbf{Y_1}} = [\vartheta_{1,k}, 2 \leq k \leq K]_{1\times(K-1)}$. From equation (13) and (17), we have

$$\left[\vartheta_{1,1}, \tilde{\mathbf{Y_1}}\right] = \left[\vartheta_{1,1}, \tilde{\mathbf{Y_1}}\right] \begin{pmatrix} v_1\phi\mathbf{w_0}\mathbf{B_4^{-1}} \\ \mathbf{O_3} \end{pmatrix}$$

As a result, the probabilities of the system's length for the regular service period are given by

$$\vartheta_{1,k} the = \vartheta_{11} v_1 \phi \mathbf{w_0} \mathbf{B_4^{-1}} \epsilon_i, 1 \leq k \leq K,$$

where ϵ_i is a column vector whose k^{th} component is unity and the remaining components are zero. From equation(12) and (15), the system length probabilities of the server being on working vacation is given by

$$[\vartheta_{0,0}, \vartheta_{0,1}, \cdots \vartheta_{0,K}] = -\left[\vartheta_{1,1}, \tilde{\mathbf{Y_1}}\right] \begin{pmatrix} v_1\mathbf{w} \\ \mathbf{O_4} \end{pmatrix}$$

Hence,

$$\vartheta_{0,k} = -\vartheta_{1,1} v_1 \mathbf{w} \epsilon_{k+1}, 0 \leq k \leq K$$

By applying the normalization condition $\sum_{j=0}^{k}\sum_{k=j}^{K}\vartheta_{k,j} = 1$ the only unknown $\vartheta_{1,1}$ is obtained as

$$\vartheta_{1,1} = \left(v_1\phi \sum_{k=1}^{K} \mathbf{w_0}\mathbf{B_4^{-1}}\epsilon_i - v_1 \sum_{k=0}^{K} \mathbf{w}\epsilon_{k+1} \right)^{-1} \tag{18}$$

This completes the evaluation of steady-state probabilities.

5. Performance Measures

The steady-state probabilities can then be used to calculate several model performance metrics. The probability that the server is actively performing routine service P_B, the typical number of users in the system R_R, and the likelihood that the server is on vacation P_V are all given by

$$P_B = \sum_{k=1}^{K} \vartheta_{1,k}; \tag{19}$$

$$P_V = \sum_{k=0}^{K} \vartheta_{0,k} = 1 - P_B \tag{20}$$

$$L_Q = \sum_{j=0}^{1}\sum_{k=0}^{K}(k-j)\vartheta_{j,k} \tag{21}$$

$$L_S = \sum_{k=1}^{K} k(\vartheta_{0,k} + \vartheta_{1,k}); \tag{22}$$

$$R_R = \sum_{j=0}^{1}\sum_{k=0}^{K}(k-j)\alpha\vartheta_{j,k} \tag{23}$$

where $(k-j)\alpha$ is the instantaneous reneging rate.

6. Cost Model

In the following subsection, service rates are used as the decision variables to formulate the cost function of total expected cost function per unit of time. The best service rates with the lowest overall expected cost function are what we are trying to determine. The following cost parameters are presumptive:

Table 1: *Cost per unit time*

Cost per unit time	When
C_1	customer waiting for service
C_2	A customer reneges
C_3	busy server
C_4	server on working vacation
C_5	feedback customer during service period
C_6	feedback customer during working vacation period

The total expected cost function (T_{EC}) per unit of time may be defined as follows using the definitions of each of the cost components mentioned above:

$$T_{EC} = C_1 L_S + C_2 R_R + \mu(C_3 + q_1 C_5) + \eta(C_4 + q_1 C_6) \tag{24}$$

where L_S and R_R are given in equation (22) and (23) respectively.

7. Numerical Results

To illustrate how the various model parameters affect the ideal service rate μ, the ideal expected cost of the system $f(\mu)$, and other system performance measures, we present some numerical examples in this subsection. We fix the maximum number of customers in the system $K = 15$ and the cost elements $C_1 = 10$, $C_2 = 12$, $C_3 = 20$, $C_4 = 14$, $C_5 = 12$, $C_6 = 18$.

(i) **Effect of λ on performance measures and cost:** Here, we study the variation of the performance measures defined in equation (22),(23) and (24). Figure 1(a)-(c) display the sensitivity of performance measures with parameters λ for three different values of K, arrival rate λ, $q = 0.02$, $p_1 = 0.3, \mu = 5$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$ are considered. From (a) for the different number of customers K and the arrival rate λ of customer, the length of the system R_R increases. From (b) when the arrival rate λ of customer increases the average reneging rate R_R increases. Figure (c) when the arrival rate λ of customer increases the expected total cost T_{EC} increases lightly.

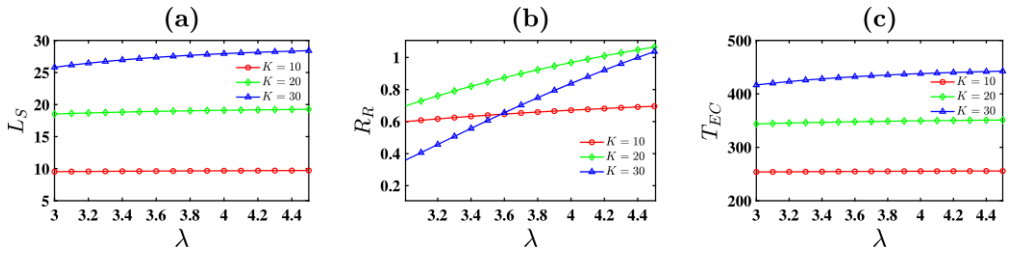


Figure 1: *Variation in performance measures with respect to λ of **(a)** Expected length of the system **(b)** Expected no. of customer loss **(c)** Total expected cost of the system, where $q = 0.02$, $p_1 = 0.3, \mu = 5$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$, K, and $\lambda = 3, 3.1, \cdots, 3.4$ respectively.*

(ii) **Effect of μ on performance measures and cost:** Figure 2(a)-(c) shows the sensitivity of performance measures with parameters μ for different value of K, arrival rate λ, $q = 0.02$, $p_1 = 0.3, \lambda = 3$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$ are considered. From (a) for the different number of customers K and the service rate μ of customer, the length of the system R_S decreases as obvious. From (b) when the service rate μ of customer increases the average customer loss R_R also decreases. Figure (c) when the service rate μ of customer increases the expected total cost T_{EC} increases when system capacity decreases.

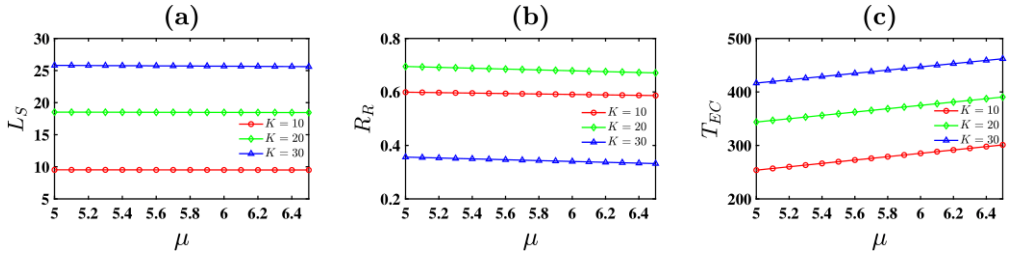


Figure 2: *Variation in performance measures with respect to μ of* ***(a)*** *Expected length of the system* ***(b)*** *Expected no. of customer loss* ***(c)*** *Total expected cost of the system, where $q = 0.02$, $p_1 = 0.3, \lambda = 5$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$, K, and $\mu = 5, 5.1, \cdots, 5.4$ respectively.*

(iii) **Effect of α on performance measures and cost:** The sensitivity of performance measures with parameters α of under three different value of K, arrival rate α, $q = 0.02$, $p_1 = 0.3, \lambda = 3$, $\mu = 5$, $\eta = 0.01$, $\phi = 0.1$ are considered can be viewed in figure 3(a)-(c). From (a) for the different number of customers K and the case parameter α of increases the length of the system R_R decreases. From (b) when the parameter α, the average customer loss L_R also decreases. Figure (c) when the parameter α increases the expected total cost T_{EC} decreases greatly.

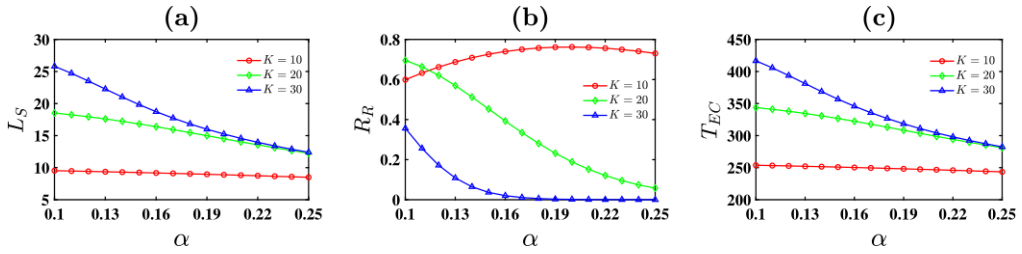


Figure 3: *Variation in performance measures with respect to α of* ***(a)*** *Expected length of the system* ***(b)*** *Expected no. of customer loss* ***(c)*** *Total expected cost of the system, where $q = 0.02$, $p_1 = 0.3$, $\lambda = 3, \mu = 5$, $\eta = 0.01$, $\phi = 0.1$, K, and $\alpha = 0.1, 0.101, \cdots, 0.114$ respectively.*

(iv) **Effect of η on performance measures and cost:** Figure 4(a)-(c) represents the sensitivity of performance measures with parameter η for three different value of K, arrival rate $q = 0.02$, $p_1 = 0.3$, $\lambda = 3$, $\mu = 5$, $\alpha = 0.1$, $\phi = 0.1$ are considered. From (a) for the different number of customer K and the parameter η increases the length of the system R_R decreases lightly. From (b) the average customer loss L_R decreases more when the system capacity is high. Figure (c) when the parameter η increases the expected total cost T_{EC} decreases lightly.

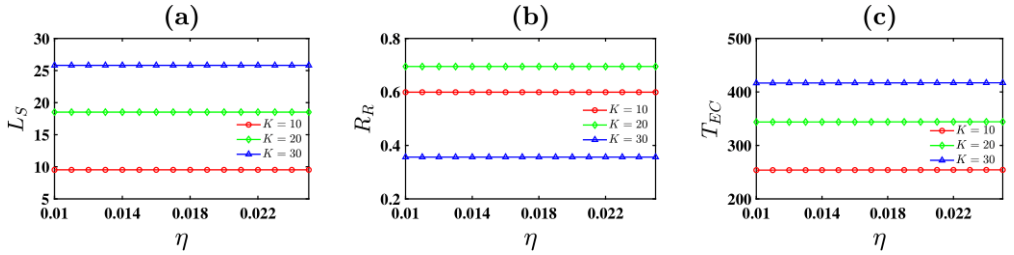


Figure 4: *Variation in performance measures with respect to η of* ***(a)*** *Expected length of the system* ***(b)*** *Expected no. of customer loss* ***(c)*** *Total expected cost of the system, where $q = 0.02$, $p_1 = 0.3, \lambda = 3$, $\mu = 5$, $\alpha = 0.1$, $\phi = 0.1$, K, and $\eta = 0.1, 0.11, \cdots, 0.25$ respectively.*

(v) **Effect on performance measures and cost with respect to k:** Figure 5(a)-(c) display the sensitivity of performance measures with parameters μ under three parameters for different value of K, arrival rate $\lambda = 3, 3.1, \cdots, 3.4$, $\mu = 5, 5.1, \cdots, 5.4$ and $\alpha = 0.1, 0.101, \cdots, 0.114$ $q = 0.02$, $p_1 = 0.3$, $\eta = 0.01$, $\phi = 0.1$ are considered. From (a) when we make the variation in K, then R_R increases greatly w.r.t. λ, but increases very slowly w.r.t. μ and α. From (b) when we make the variation in the parameter α then

the average customer loss is less. Figure (c) total cost of the system is low after making a variation in α and k.

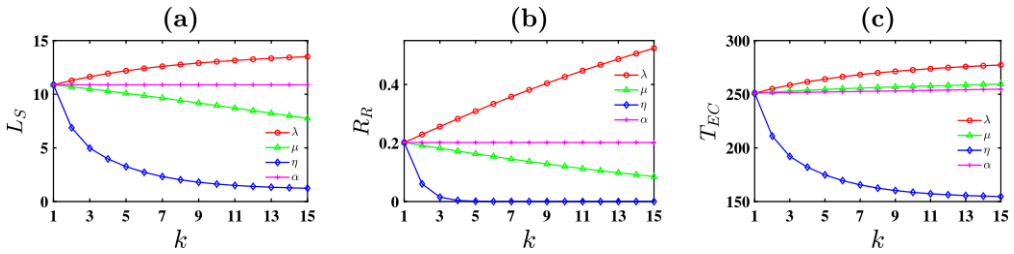


Figure 5: *Variation in performance measures with respect to λ, μ, α of* ***(a)*** *Expected length of the system* ***(b)*** *Expected no. of customer loss* ***(c)*** *Total expected cost of the system, where $q = 0.02$, $p_1 = 0.3$, $\eta = 0.01$, $\phi = 0.1$, and $k = 1, 2, \cdots, 15$, $\lambda = 3, 3.1, \cdots, 3.4$, $\mu = 5, 5.1, \cdots, 5.4$ and $\alpha = 0.1, 0.101, \cdots, 0.114$ respectively.*

(vi) **Effect on system lengh w.r.t. parameters in pair:** Figure 6(a)-(e) display the sensitivity of performance measures with parameters and system capacity $K = 15$ From (a)-(c) making variation in λ and other three parameters μ, α,η, R_R is increasing and decreasing as obvious. From (d) and (e) making variation in μ and other three parameters α and η, R_R is decreasing. From (f) when making variation in α and η, R_R is decreasing when α and η increasing.

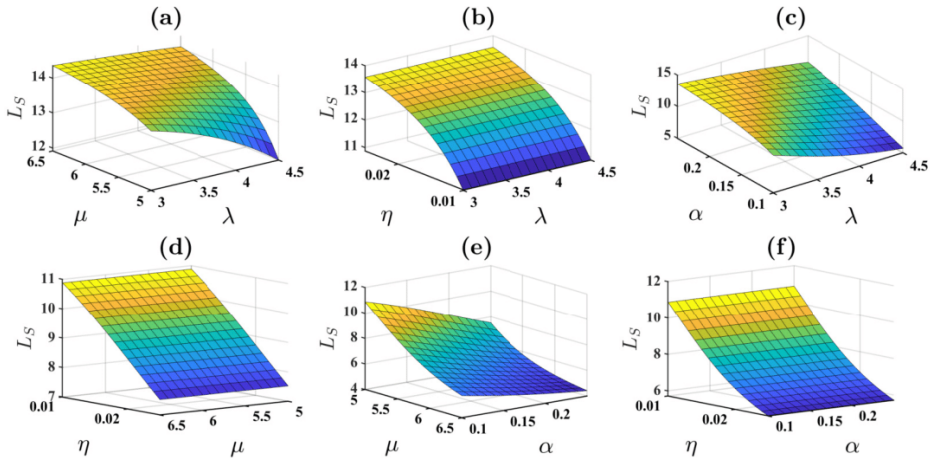


Figure 6: *Variation in length of the system with respect to* ***(a)*** *λ and μ* ***(b)*** *λ and α* ***(c)*** *λ and η* ***(d)*** *μ and α* ***(e)*** *μ and η* ***(f)*** *α and η,where $\lambda = 3, 3.1, \cdots, 6.5$, $\mu = 5, 5.1, \cdots, 8.5$, $\eta = \alpha = 0.1, 0.011, \cdots, 0.045$ else $q = 0.02$, $p_1 = 0.3$, $\lambda = 3$, $\mu = 5$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$, $K = 15$ respectively.*

(vii) **Effect on average reneging rate w.r.t. parameters in pair:** Figure 7(a)-(e) shows the sensitivity of performance measures with parameters and system capacity $K = 15$ From (a)-(c) making variation in λ and other three parameters μ, α, η, R_R is increasing and decreasing as obvious. From (d) and (e) making variation in μ and other three parameters α and η, L_R is decreasing. From (f) when making variation in α and η, L_R is decreasing when α and η increasing.

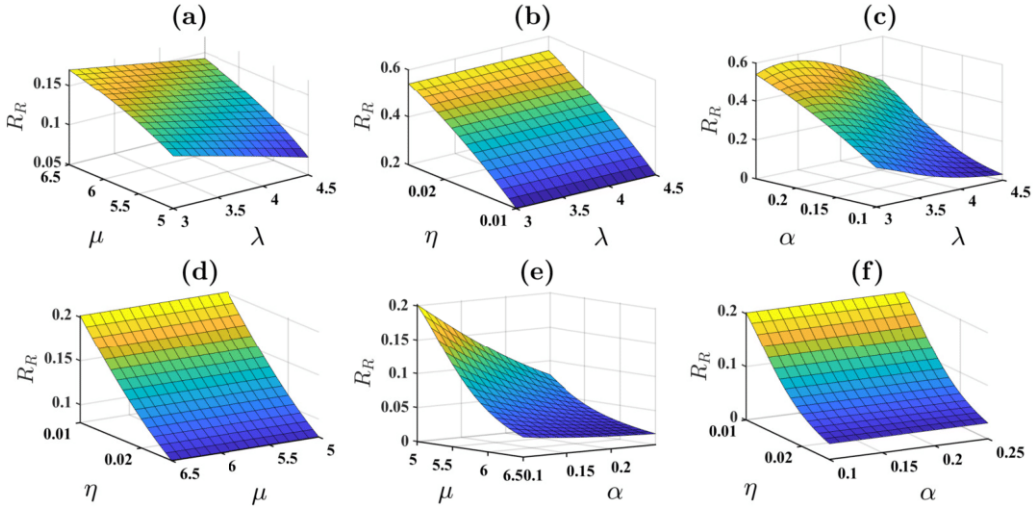


Figure 7: *Variation in average customer loss with respect to* ***(a)*** λ *and* μ ***(b)*** λ *and* α ***(c)*** λ *and* η ***(d)*** μ *and* α ***(e)*** μ *and* η ***(f)*** α *and* η*,where* $\lambda = 3, 3.1, \cdots, 6.5$, $\mu = 5, 5.1, \cdots, 8.5, \eta = \alpha = 0.1, 0.011, \cdots, 0.045$ *else* $q = 0.02$, $p_1 = 0.3, \lambda = 3$, $\mu = 5$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$, $K = 15$ *respectively.*

(viii) **Effect on total expected cost w.r.t. parameters in pair:** Figure 8(a)-(e) display the sensitivity of performance measures with parameters and system capacity $K = 15$ From (a)-(c) making variation in λ and other three parameters μ, α,η, T_{EC} is increasing and decreasing as obvious. From (d) and (e) making variation in μ and other three parameters α and η, T_{EC} is decreasing. From (f) when making variation in α and η, T_{EC} is decreasing when α and η increasing.

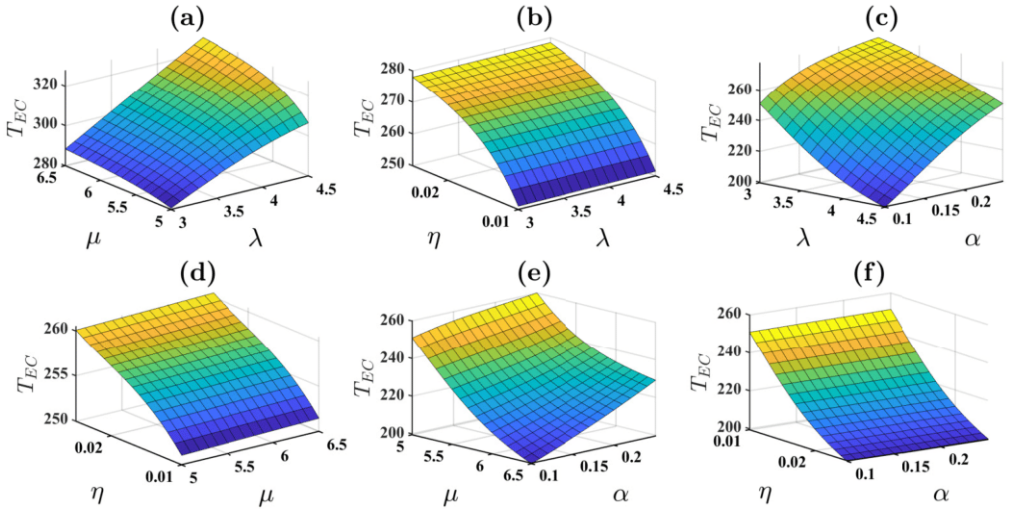


Figure 8: *Variation in total expected cost with respect to* ***(a)*** λ *and* μ ***(b)*** λ *and* α ***(c)*** λ *and* η ***(d)*** μ *and* α ***(e)*** μ *and* η ***(f)*** α *and* η*,where* $\lambda = 3, 3.1, \cdots, 6.5$, $\mu = 5, 5.1, \cdots, 8.5, \eta = \alpha = 0.1, 0.011, \cdots, 0.045$ *else* $q = 0.02$, $p_1 = 0.3, \lambda = 3$, $\mu = 5$, $\eta = 0.01$, $\alpha = 0.1$, $\phi = 0.1$, $K = 15$ *respectively.*

1

Conclusion

In this study, we investigated Markovian feedback queue with reneging and working vacations. We have obtained the steady-state probabilities and solve them using the matrix technique. The model results may be useful in modeling various production and service processes involving feedback and impatient customers. The analysis of the model is restricted to a fixed size. The model's unrestricted size case can also be investigated. Furthermore, to obtain time-dependent results model can be solved in a transient state.

Acknowledgement

Krishan's work is supported by the Council of Scientific and Industrial Research (CSIR), Govt. of India (Award no. 09/1256(11794)/2021-EMR-I).

References

[1] Takács, L. (1963). A single-server queue with feedback. *Bell system Technical journal*, 42(2):505-519.
[2] Doshi, B.T. (1986). Queueing Systems with Vacations, a Survey. *Queueing Systems*, 1:29-66.
[3] Laxmi, P. V., Goswami, V. and Jyothsna, K. (2013). Analysis of finite buffer Markovian queue with balking, reneging and working vacations. *International Journal of Strategic Decision Sciences*, 4(1):1-24.
[4] Wortman, M. A., Disney, R. L. and Kiessler, P. C. (1991). The M/GI/1 Bernoulli feedback queue with vacations. *Queueing Systems*, 9.
[5] Zhang, Y., Yue, D. and Yue, W. (2005). Analysis of an M/M/1/N queue with balking, reneging and server vacations. *Proceedings of the 5th International Symposium on OR and its Applications*, 37-47.
[6] Kumar, R., Jain, N. K. and Som, B. K. (2014). Optimization of an M/M/1/N feedback queue with retention of reneged customers. Operations Research and Decisions, 24(3), 45-58.
[7] Kumar, R. and Sharma, S. K. (2014). A multi-server Markovian feedback queue with balking, reneging and retention of reneged customers. AMO-Advanced Modeling and Optimization, 16(2), 395-406.
[8] Van den berg, J. L. and Boxma, O. J. (1991). The M/G/1 queue with processor sharing and its relation to a feedback queue. *Queueing Systems*, 9.
[9] Laxmi, P. V. and Jyothsna, K. (2014). Performance analysis of variant working vacation queue with balking and reneging. *International Journal of Mathematics in Operational Research*, 6(4):505-521.
[10] Bouchentouf, A. A. and Guendouzi, A. (2020). The $M^X/M/c$ Bernoulli feedback queue with variant multiple working vacations and impatient customers: performance and economic analysis. *Arabian Journal of Mathematics*, 9(2):309-327.
[11] Ganesan, V. and Rita, S. (2017). Feedback queue with services in different stations under reneging and vacation policies. *Int J Appl Eng Res*, 12:11965-11969.
[12] Jewkes, E. M. and Buzacott, J. A. (1991). Flow time distributions in a k class M/G/1 priority feedback queue. *queueing systems*, 8.
[13] Bouchentouf, A. A., Guendouzi, A. and Kandouci, A. (2018). Performance and economic analysis of Markovian Bernoulli feedback queueing system with vacations, waiting server and impatient customers. Acta Universitatis Sapientiae, *Mathematica*, 10(2):218-241.
[14] Yahiaoui, L., Bouchentouf, A. A. and Kadi, M. (2019). Optimum cost analysis for an Geo/Geo/c/N feedback queue under synchronous working vacations and impatient customers. *Croatian Operational Research Review*, 211-226.
[15] Choi, B. D., Kim, Y. C. and Lee, Y. W. (1998). The M/M/c Retrial Queue with Geometric Loss and Feedback. *Computers Math. Applic*, 36(6).
[16] Choi, B. D., Kim, B. and Choi, S. H. (2000). On the M/G/1 Bernoulli feedback queue with multi-class customers. *Computers & Operations Research*, 27.
[17] Bouchentouf, A. A. and Guendouzi, A. (2019). Cost optimization analysis for an $M^X/M/c$ MX/M/c vacation queueing system with waiting servers and impatient customers. SeMA Journal, 76(2):309-341.
[18] Kumar, B.K., Madheswari, S.P. and Vijayakumar, A. (2002). The M/G/1 retrial queue with feedback and starting failures. Available: www.elsevier.com/locate/apm.
[19] Kumar, B. K., Arivudainambi, D. and Krishnamoorthy, A. (2006). Some results on a generalized M/G/1 feedback queue with negative customers. *Annals of Operations Research*, 143(1):277-296.
[20] Kumar, B. K. and Raja, J. (2006). On multiserver feedback retrial queues with balking and control retrial rate. Annals of Operations Research, 141(1):211-232.
[21] Ke, J. C. and Chang, F. M. (2009). Modified vacation policy for M/G/1 retrial queue with balking and feedback. Computers and Industrial Engineering, 57(1):433-443.
[22] Ke, J. C. and Chang, F. M. (2009). Modified vacation policy for M/G/1 retrial queue with balking and feedback. *Computers and Industrial Engineering*, 57(1):433-443.
[23] Gupta, R. (2022). Cost Optimization of Queueing System with Working Vacation, Setup, Feedback, Reneging, and Retention of Reneged Customers. *Journal of Scientific Research*, 14(1):257-268.
[24] Jeeva, M. and Rathna kumari, E. (2012). Bulk Arrival Single Server, Bernoulli Feedback Queue with Fuzzy Vacations and Fuzzy Parameters. *ARPN Journal of Science and Technology*, 2(5).

[25] Jain, S. (2016). Probability-Based Analysis to Determine the Performance of Multilevel Feedback Queue Scheduling Operating System may work on different types of CPU scheduling algorithms with different mechanisms and concepts. The Multilevel Feedback Queue (MLFQ) Scheduling manages. *Int. J. Advanced Networking and Applications*, 3044-3069.

[26] Saravanarajan, M. C., and Chandrasekaran, V. M. (2014). Analysis of M/G/1 feedback queue with two types of services, Bernoulli vacations, and random breakdowns. *International Journal of Mathematics in Operational Research*, 6(5):567-588.

[27] Shanmuga Sundaram, S., and Sivaram, S. (2021). Analysis of M/G/1 Feedback Queue under Steady State When Catastrophes Occur. *Turkish Journal of Computer and Mathematics Education*, 12(1).

[28] Sundari, S. M. and Srinivasan, S. (2012). Analysis of M/G/1 Feedback Queue with Three Stage and Multiple Server Vacation. *In Applied Mathematical Sciences*, 6(125).

RT&A, No 4 (76)
Volume 18, December 2023

PERFORMANCE CHARACTERIZATION OF TWO-SERVER BATCH SERVICE QUEUE WITH SECOND OPTIONAL SERVICE

Andwilile Abrahamu George

•

Department of Mathematics and ICT, College of Business Education,
Dodoma, Tanzania.
gandwilile@gmail.com

P. Vijaya Laxmi

•

Department of Applied Mathematics, Andhra University,
Visakhapatnam, India.

Abstract

In this paper, we analyze the performance of a finite capacity two-server Markovian batch service queueing model with second optional service. The servers provide two kinds of services, the first essential service (FES), which is provided to all incoming customers and the second optional service (SOS) to those who demand it after completing FES. The service times of the two servers are identical and are exponentially distributed. Matrix-decomposition method is used to obtain the steady-state probabilities of the model. Numerical results and discussion are presented to demonstrate the impact of the model parameters on the system behavior. Furthermore, the cost model optimization is developed to determine the optimal service rates using the Quasi-Newton method to minimize the expected cost. Finally, the findings of this work show that the blocking probability is monotonically decreases as finite buffer size increases and approaches its minimum value of zero when finite buffer is sufficiently large.

Keywords: Two-server, Batch service queue, First essential service, Second optional service.

Mathematical Subject Classification: 60k25, 90B22.

I. Introduction

In everyday life, there are various queueing circumstances where all incoming customers demand the FES and only some may demand the SOS provided by the same server. Currently, this has taken a major consideration by various researchers such as Wang and Kuki [23] wherein they analysed the performance of retrial queueing system with SOS. They obtained the queue length, waiting time and busy period using the method of discrete transformation. SOS with a single server fixed batch service queueing system during repeated vacations has been studied by Ayyappan et al., [3]. They analyzed the model using the probability generating function and Rouche's theorem to obtain the probability of the number of customers present in the queue while the server is busy or on vacation.

Multi-server retrial queue with SOS has been presented by Ke et al., [13]. They derived stationary probabilities using matrix analytic approach. An $M/M/1$ queue with SOS and working breakdown has been analysed by Yang and Chen [24] who derived the stationary

probability distribution of the system size using the matrix geometric method. Other extensive studies conducted on an assortment of queueing models with SOS are found in Gupur [10]; Kalyanaraman and Murugan [12]; Ke et al., [14]; Uma and Punniyamoorthy [22], etc.
For the most part, customers get the service individually. However, this rule may not work in all circumstances, since in some places, the server provide service in batch (groups) of customers instead of serving individually. Batch services are more useful in telecommunication, where data bundle is transmitted in the accumulated large entities (batches), in the field of transportation, in a manufacturing system, in a smart city crowdsourcing application for mobile, etc.

Batch service queue has been broadly studied by numerous researchers. Goswami and Samanta [8] presented the two heterogeneous servers with a discrete-time bulk service queueing system and derived a closed-form expression of the stationary probabilities at arbitrary epoch. The manufacturing bulk service queues using Bayesian hierarchical models are presented by Armero and Conesa [2]. They developed the inferential method for the parameter of the governing model using hierarchical models. Barbhuiya and Guptal [6] analysed the $GI/M^Y/1$ queue, wherein the closed form explicit formulations of the system distribution at pre arrival and arbitrary epochs in terms of roots of the characteristic equation generated were derived using the supplementary variable approach and the difference equation method. The queue for bulk service with delayed vacation has been analyzed by Krishna Reddy and Anitha [15]. They obtained stationary probabilities and waiting time distribution of an incoming customer. Ayyappan et al., [4] studied the single server fixed batch service queueing system. The server serves exactly k items at a time. If he finds less than k items in the queue, he takes a multiple vacation of length a. Batch service with multi-server queue model has been presented by Ghimire et al., [7], wherein they computed the stationary probability, queue length size and the amount of waiting time in the queue.

A Markovian single server queueing system has been widely studied by numerous researchers for the last few decades. Due to increasing demand for services, single server operations can lead to congestion in the system such as manufacturing and production, medical clinics, in telecommunication system, etc. Multi-servers have also been studied by various researchers to reduce congestion, Hwang et al., [11]. The queueing system with two servers that are heterogeneous and a variant vacation policy has been studied by Yue and Tian [25], where the two servers take vacations together whenever the system becomes empty. They obtained the explicit expressions of the stationary distribution of the system size. Krishnamoorthy and Sreenivasan [16] investigated a queueing system using two heterogeneous servers, one of which is regularly accessible while the other goes on vacation whenever there are no customers waiting for service and service times of the two servers are exponentially distributed with different service rates. They analyzed the model using matrix geometric method to obtain the mean waiting time in the steady state regime. Kumar et al., [17] investigated the two server queue with heterogeneous servers that were vulnerable to catastrophes wherein any arriving item requires exactly one server for the service and the service rates are different. They obtained both transient and stationary probabilities of number of items in the system using probability generating function and the modified Bessel function of the first kind. Similar studies are found in Ammar [1] and Kumar and Sharma [19] with impatient behavior wherein the service times are independently and exponentially distributed with different service rate. They obtained the explicit transient and steady state probabilities by using probability generating function. Recently, Kumar et al.,[18] generalized the work of Kumar and Sharma [19] by introducing the concept of feedback and reverse balking. They used an iterative approach to obtain the probabilistic measures of the model.

Queues with limited waiting space are more realistic in real-life circumstances. For example, routers servicing incoming packets with varying speeds in a network are examples of this scenario. If the server is busy, the incoming items wait in the queue, but the incoming items are deemed lost when the waiting space is full. One of the major considerations of a system designer in this situation is to provide enough buffer space to keep the blocking probability to a minimum. The limited buffer size queues have been researched by various authors such as Sikdar and Gupta [21],

who analyzed a queue for batch arrival and batch service with a limited buffer size with single and repeated vacations. They obtained stationary probability distributions of the number of items in the queue at departure and pre-arrival epochs. Moreover, they presented the average queue length, waiting time and the blocking probability as the performance measures. Gupta et al.,[9] analyzed a queue for bulk services on a single server with finite buffers and vacation. Using the supplementary variable and embedded Markov chain methods, they obtained the stationary joint distribution of the queue length and the kind of vacation that the server has taken at different epochs. Banerjee [5] presented a queue with a limited capacity and workload-dependent service. They used the embedded Markov chain technique and the supplementary variable approach to obtain stationary probabilities at a departure and arbitrary epoch.

In view of the growing demands for service in practice, our objective here is to investigate the performance of finite buffer two-server batch queue with SOS. In existing literature reviews, no work has been reported with a combination of two-server batch queue with SOS. This inspires us to investigate a two-server batch queue with SOS. Consideration of SOS with varying batch size service makes the model more versatile. The main contributions of this paper are as follows.

- First, we derive the stationary probabilities by using matrix-decomposition method.
- Second, we develop a cost function to optimize service rates for both FES and SOS using Quasi-Newton method so that the expected total cost is minimized.

The remaining part of the paper is structured as follows: Mathematical model description and its formulation are presented in Section 2. In Section 3, we discuss steady state probabilities where the servers are busy or idle in both FES and SOS. The performance measures are discussed in Section 4. The presentation of numerical analysis and discussion is in Section 5. The cost model optimization is developed in Section 6 and in Section 7, we conclude the paper.

II. Model Description and Mathematical Formulation

We consider a finite buffer $M/M^b/2/N$ queue model with SOS, where arrival occurs according to a Poisson process with parameter λ. The servers provide two kinds of service, FES, which is provided to all incoming items and SOS to those whose demand it after completing FES. The item opts the SOS with probability r $(0 \leq r \leq 1)$ or leave the system after completion FES with the probability $(1-r)$. Furthermore, It is assumed that the service times of two servers are independently, identically and exponentially distributed, each with service rate μ_1 in FES and μ_2 in SOS. The queue has limited waiting space of size N. The arriving customers are blocked from entering the queue whenever the queue size is N. The average rate of service for FES is given by

$$\mu_{i1} = \begin{cases} \mu_1, & \text{for } i = 1, \\ 2\mu_1, & \text{for } i = 2. \end{cases}$$

The average rate of service for SOS is given by

$$\mu_{j2} = \begin{cases} \mu_2, & \text{for } j = 1, \\ 2\mu_2, & \text{for } j = 2, \end{cases}$$

where i and j are the number of servers in FES and SOS, respectively.
The servers process a partial batch up to a maximum capacity size of b. If there are less than b in the queue, one of the servers begins service to those customers. If there are more than $2b$ waiting in the queue, then the servers take a batch of size b each based on the order of arrival, while others in excess of $2b$ customers, wait in the queue.

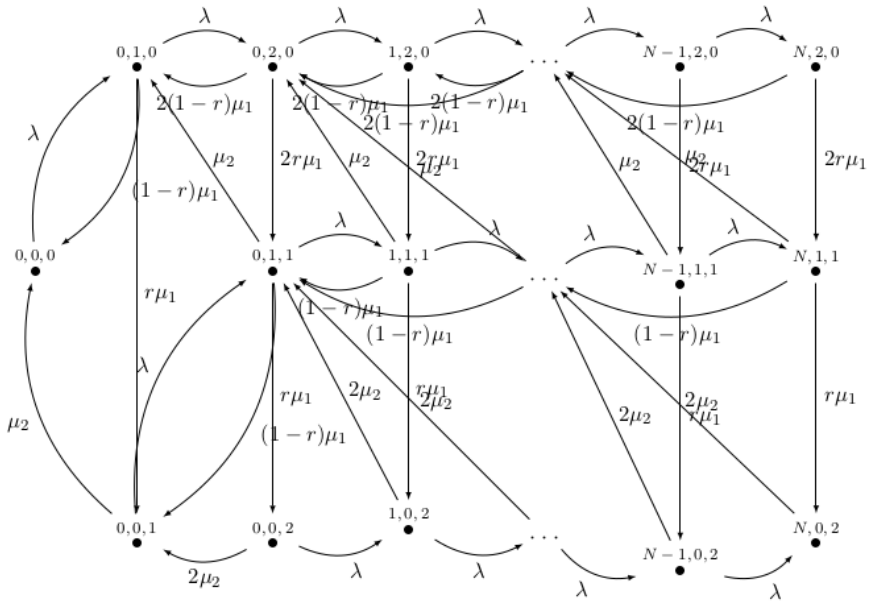


Figure 1: *Transition rate diagram for* $b = 2$

Figure 1 is the state transition diagram showing the various transition states of the queueing model.

I. Formulation of Mathematical Model

At time $t \geq 0$, let $N(t)$ be the number of items in the queue, $S(t)$ be the number of servers in FES and $J(t)$ be the number of servers in SOS, where $S(t)$ and $J(t)$ are defined as follows:

$$S(t) = \ i, \qquad 0 \leq i \leq 2, \ \text{ and } J(t) = \ j, \qquad 0 \leq j \leq 2.$$

We observe that, $\{(N(t), S(t), J(t)); t \geq 0\}$ defines a three dimensional Markov process in continuous time with state space

$$E = \{(n, i, j);\ 0 \leq n \leq N;\ 0 \leq i \leq 2;\ 0 \leq j \leq 2,\ i + j \leq 2\}.$$

Let the transient probabilities are defined as

$$P_{n,i,j}(t) = Pr\{N(t) = n,\ S(t) = i,\ J(t) = j\},$$
$$0 \leq n \leq N;\ 0 \leq i \leq 2;\ 0 \leq j \leq 2,\ i + j \leq 2,$$

where $P_{0,i,j}(t)$ is the transient probability that i and j servers are busy with FES or SOS with no customer waiting in the queue, where $0 \leq i \leq 2,\ 0 \leq j \leq 2,\ i + j \leq 2$.
$P_{n,i,j}(t)$ is the transient probability that all servers are busy with FES or SOS with n customer waiting in the queue, where $0 \leq i \leq 2,\ 0 \leq j \leq 2,\ i + j = 2,\ 1 \leq n \leq N$.
The above set of probabilities at steady state are denoted by $P_{0,i,j}$ and $P_{n,i,j}$, respectively. The following set of equations are obtained using the probabilistic arguments.

$$
\begin{aligned}
\lambda P_{0,0,0} &= (1-r)\mu_1 P_{0,1,0} + \mu_2 P_{0,0,1}, && (1)\\
(\lambda+\mu_1)P_{0,1,0} &= \lambda P_{0,0,0} + 2(1-r)\mu_1 P_{0,2,0} + \mu_2 P_{0,1,1}, && (2)\\
(\lambda+2\mu_1)P_{0,2,0} &= \lambda P_{0,1,0} + 2(1-r)\mu_1 \sum_{i=1}^{b} P_{i,2,0} + \mu_2 \sum_{i=1}^{b} P_{i,1,1}, && (3)\\
(\lambda+2\mu_1)P_{n,2,0} &= \lambda P_{n-1,2,0} + 2(1-r)\mu_1 P_{n+b,2,0} + \mu_2 P_{n+b,1,1},\ 1 \le n \le N-b, && (4)\\
(\lambda+2\mu_1)P_{n,2,0} &= \lambda P_{n-1,2,0},\ \ N-b+1 \le n \le N-1, && (5)\\
2\mu_1 P_{N,2,0} &= \lambda P_{N-1,2,0}, && (6)\\
(\lambda+\mu_2)P_{0,0,1} &= (1-r)\mu_1 P_{0,1,1} + 2\mu_2 P_{0,0,2} + r\mu_1 P_{0,1,0}, && (7)\\
(\lambda+2\mu_2)P_{0,0,2} &= r\mu_1 P_{0,1,1}, && (8)\\
(\lambda+2\mu_2)P_{n,0,2} &= \lambda P_{n-1,0,2} + r\mu_1 P_{n,1,1},\ \ 1 \le n \le N-1, && (9)\\
2\mu_2 P_{N,0,2} &= \lambda P_{N-1,0,2} + r\mu_1 P_{N,1,1}, && (10)\\
(\lambda+\mu_1+\mu_2)P_{0,1,1} &= \lambda P_{0,0,1} + (1-r)\mu_1 \sum_{i=1}^{b} P_{i,1,1} + 2\mu_2 \sum_{i=1}^{b} P_{i,0,2} + 2r\mu_1 P_{0,2,0}, && (11)\\
(\lambda+\mu_1+\mu_2)P_{n,1,1} &= \lambda P_{n-1,1,1} + (1-r)\mu_1 P_{n+b,1,1} + 2\mu_2 P_{n+b,0,2} + 2r\mu_1 P_{n,2,0},\\
&\quad 1 \le n \le N-b, && (12)\\
(\lambda+\mu_1+\mu_2)P_{n,1,1} &= \lambda P_{n-1,1,1} + 2r\mu_1 P_{n,2,0},\ \ N-b+1 \le n \le N-1, && (13)\\
(\mu_1+\mu_2)P_{N,1,1} &= \lambda P_{N-1,1,1} + 2r\mu_1 P_{N,2,0}. && (14)
\end{aligned}
$$

The steady state equations (1)-(14) can be represented as matrix-form

$$\mathbf{PQ=0}, \tag{15}$$

where **0** denotes the column vector with all elements equal to zero, and

$$\mathbf{Q} = \begin{pmatrix}
-\lambda & \mathbf{M}_{12} & \mathbf{M}_{13} & 0 & \mathbf{M}_{15} & \mathbf{M}_{16} & \mathbf{M}_{17} \\
\mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \mathbf{M}_{24} & \mathbf{M}_{25} & \mathbf{M}_{26} & \mathbf{M}_{27} \\
\mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} & \mathbf{M}_{34} & \mathbf{M}_{35} & \mathbf{M}_{36} & \mathbf{M}_{37} \\
0 & \mathbf{M}_{42} & \mathbf{M}_{43} & \mathbf{M}_{44} & \mathbf{M}_{45} & \mathbf{M}_{46} & \mathbf{M}_{47} \\
\mathbf{M}_{51} & \mathbf{M}_{52} & \mathbf{M}_{53} & \mathbf{M}_{54} & \mathbf{M}_{55} & \mathbf{M}_{56} & \mathbf{M}_{57} \\
\mathbf{M}_{61} & \mathbf{M}_{62} & \mathbf{M}_{63} & \mathbf{M}_{64} & \mathbf{M}_{65} & \mathbf{M}_{66} & \mathbf{M}_{67} \\
\mathbf{M}_{71} & \mathbf{M}_{72} & \mathbf{M}_{73} & \mathbf{M}_{74} & \mathbf{M}_{75} & \mathbf{M}_{76} & \mathbf{M}_{77}
\end{pmatrix}$$

is a $(3N+6)\times(3N+6)$ square matrix. The entries of the matrix **Q** are listed below:

$$\mathbf{M}_{12} = \begin{pmatrix} \lambda & 0 \end{pmatrix}_{1\times 2},\ \mathbf{M}_{13} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}_{1\times N},\ \mathbf{M}_{15} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}_{1\times N},$$

$$\mathbf{M}_{27} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{2\times N},\ \mathbf{M}_{17} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}_{1\times N},\ \mathbf{M}_{21} = \begin{pmatrix} (1-r)\mu_1 \\ 0 \end{pmatrix}_{2\times 1},$$

$$\mathbf{M}_{23} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \lambda & 0 & \cdots & 0 \end{pmatrix}_{2\times N},\ \mathbf{M}_{24} = \begin{pmatrix} 0 \\ 2r\mu_1 \end{pmatrix}_{2\times 1},\ \mathbf{M}_{25} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{2\times N},$$

$$\mathbf{M}_{16} = \begin{pmatrix} 0 & 0 \end{pmatrix}_{1\times 2},\ \mathbf{M}_{26} = \begin{pmatrix} r\mu_1 & 0 \\ 0 & 0 \end{pmatrix}_{2\times 2},\ \mathbf{M}_{31} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{N\times 1},\ \mathbf{M}_{32} = \begin{pmatrix} 0 & r_{i2} \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \end{pmatrix}_{N\times 2},$$

$$\mathbf{M}_{51} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{N\times 1},\ \mathbf{M}_{33} = \begin{pmatrix}
\alpha & \lambda & 0 & \cdots & 0 & 0 \\
A_{10} & \alpha & \lambda & \cdots & 0 & 0 \\
A_{20} & A_{21} & \alpha & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A_{N-20} & A_{N-21} & A_{N-22} & \cdots & \alpha & \lambda \\
A_{N-10} & A_{N-11} & A_{N-12} & \cdots & A_{N-1N-2} & -2\mu_1
\end{pmatrix}_{N\times N},$$

where

$$\alpha = -(\lambda + 2\mu_1),\ r_{i2} = 2(1-r)\mu_1,\ \text{for } i = 1,\ 2,\ \cdots, b,\ \text{ and } b \leq N,$$

$$A_{ij} = \begin{cases} 2(1-r)\mu_1, & \text{if } \frac{i-j}{b} = 1 \text{ for } 1 \leq i \leq N-1,\ 0 \leq j \leq N-2,\ b \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

$$\mathbf{M}_{34} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{N\times 1},\ \mathbf{M}_{35} = \begin{pmatrix} 2r\mu_1 & 0 & \cdots & 0 \\ 0 & 2r\mu_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2r\mu_1 \end{pmatrix}_{N\times N},\ \mathbf{M}_{44} = -(\lambda + \mu_1 + \mu_2),$$

$$\mathbf{M}_{37} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{N\times N},\ \mathbf{M}_{42} = \begin{pmatrix} \mu_2 & 0 \end{pmatrix}_{1\times 2},\ \mathbf{M}_{43} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}_{1\times N},$$

$$\mathbf{M}_{45} = \begin{pmatrix} \lambda & 0 & \cdots & 0 \end{pmatrix}_{1\times N},\ \mathbf{M}_{46} = \begin{pmatrix} (1-r)\mu_1 & r\mu_1 \end{pmatrix}_{1\times 2},\ \mathbf{M}_{47} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}_{1\times N},$$

$$\mathbf{M}_{52} = \begin{pmatrix} 0 & s_{i2} \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \end{pmatrix}_{N\times 2},\ \mathbf{M}_{53} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ B_{10} & 0 & 0 & \cdots & 0 & 0 \\ B_{20} & B_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{N-20} & B_{N-21} & B_{N-22} & \cdots & 0 & 0 \\ B_{N-10} & B_{N-11} & B_{N-12} & \cdots & B_{N-1N-2} & 0 \end{pmatrix}_{N\times N},$$

where

$$s_{i1} = (1-r)\mu_1 \text{ and } s_{i2} = \mu_2,\ \text{for } i = 1,\ 2,\ \cdots, b,\ \text{ and } b \leq N,$$

$$B_{ij} = \begin{cases} \mu_2, & \text{if } \frac{i-j}{b} = 1 \text{ for } 1 \leq i \leq N-1,\ 0 \leq j \leq N-2,\ b \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

$$\mathbf{M}_{55} = \begin{pmatrix} \lambda' & \lambda & 0 & \cdots & 0 & 0 \\ C_{10} & \lambda' & \lambda & \cdots & 0 & 0 \\ C_{20} & C_{21} & \lambda' & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{N-20} & C_{N-21} & C_{N-22} & \cdots & \lambda' & \lambda \\ C_{N-10} & C_{N-11} & C_{N-12} & \cdots & C_{N-1N-2} & -(\mu_1 + \mu_2) \end{pmatrix}_{N\times N},\ \mathbf{M}_{61} = \begin{pmatrix} \mu_2 \\ 0 \end{pmatrix}_{2\times 1},$$

where $\lambda' = -(\lambda + \mu_1 + \mu_2)$.

$$C_{ij} = \begin{cases} (1-r)\mu_1, & \text{if } \frac{i-j}{b} = 1 \text{ for } 1 \leq i \leq N-1,\ 0 \leq j \leq N-2,\ b \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

$$\mathbf{M}_{22} = \begin{pmatrix} -(\lambda + \mu_1) & \lambda \\ 2(1-r)\mu_1 & -(\lambda + 2\mu_1) \end{pmatrix}_{2\times 2},\ \mathbf{M}_{57} = \begin{pmatrix} r\mu_1 & 0 & \cdots & 0 \\ 0 & r\mu_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r\mu_1 \end{pmatrix}_{N\times N},$$

$$\mathbf{M}_{62} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}_{2\times 2},\ \mathbf{M}_{63} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{2\times N},\ \mathbf{M}_{67} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \lambda & 0 & \cdots & 0 \end{pmatrix}_{2\times N},$$

$$\mathbf{M}_{65} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{2\times N}, \ \mathbf{M}_{66} = \begin{pmatrix} -(\lambda+\mu_2) & 0 \\ 2\mu_2 & -(\lambda+2\mu_2) \end{pmatrix}_{2\times 2}, \ \mathbf{M}_{64} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}_{2\times 1},$$

$$\mathbf{M}_{71} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{N\times 1}, \ \mathbf{M}_{72} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}_{N\times 2}, \ \mathbf{M}_{73} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{N\times N},$$

$$\mathbf{M}_{74} = \begin{pmatrix} w_{i1} \\ \vdots \\ 0 \\ 0 \end{pmatrix}_{N\times 1}, \ \mathbf{M}_{75} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ D_{10} & 0 & 0 & \cdots & 0 & 0 \\ D_{20} & D_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ D_{N-20} & D_{N-21} & D_{N-22} & \cdots & 0 & 0 \\ D_{N-10} & D_{N-11} & D_{N-12} & \cdots & D_{N-1N-2} & 0 \end{pmatrix}_{N\times N},$$

where

$$w_{i1} = 2\mu_2, \text{ for } i = 1,\ 2,\ \cdots, b, \text{ and } b \le N,$$

$$D_{ij} = \begin{cases} 2\mu_2, & \text{if } \frac{i-j}{b} = 1 \text{ for } 1 \le i \le N-1,\ 0 \le j \le N-2,\ b \le N, \\ 0, & \text{otherwise.} \end{cases}$$

$$\mathbf{M}_{54} = \begin{pmatrix} s_{i1} \\ \vdots \\ 0 \\ 0 \end{pmatrix}_{N\times 1}, \ \mathbf{M}_{36} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}_{N\times 2}, \ \mathbf{M}_{76} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}_{N\times 2}, \ \mathbf{M}_{56} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}_{N\times 2},$$

$$\mathbf{M}_{77} = \begin{pmatrix} -(\lambda+2\mu_2) & \lambda & 0 & \cdots & 0 & 0 \\ 0 & -(\lambda+2\mu_2) & \lambda & \cdots & 0 & 0 \\ 0 & 0 & (-\lambda+2\mu_2) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -(\lambda+2\mu_2) & \lambda \\ 0 & 0 & 0 & \cdots & 0 & -2\mu_2 \end{pmatrix}_{N\times N},$$

In the following sequel, we use a matrix decomposition approach to obtain the stationary probabilities in a recursive manner. Let $P_{0,0,0}$, P_1, P_2, $P_{0,1,1}$, P_3, P_4, P_5 be the vectors of stationary probabilities where $P_1 = (P_{0,1,0},\ P_{0,2,0})$, $P_2 = (P_{1,2,0},\ P_{2,2,0}, \cdots,\ P_{N,2,0})$, $P_3 = (P_{1,1,1},\ P_{2,1,1}, \cdots,\ P_{N,1,1})$, $P_4 = (P_{0,0,1},\ P_{0,0,2})$, $P_5 = (P_{1,0,2},\ P_{2,0,2}, \cdots,\ P_{N,0,2})$.

From equation (15) it follows that

$$\begin{aligned} -\lambda P_{0,0,0} + P_1 M_{21} + P_4 M_{61} &= 0, && (16) \\ P_{0,0,0} M_{12} + P_1 M_{22} + P_2 M_{32} + P_{0,1,1} M_{42} + P_3 M_{52} &= 0, && (17) \\ P_1 M_{23} + P_2 M_{33} + P_3 M_{53} &= 0, && (18) \\ P_1 M_{24} - P_{0,1,1} M_{44} + P_3 M_{54} + P_4 M_{64} + P_5 M_{74} &= 0, && (19) \\ P_2 M_{35} + P_{0,1,1} M_{45} + P_3 M_{55} + P_5 M_{75} &= 0, && (20) \\ P_1 M_{26} + P_{0,1,1} M_{46} + P_4 M_{66} &= 0, && (21) \\ P_3 M_{57} + P_4 M_{67} + P_5 M_{77} &= 0. && (22) \end{aligned}$$

Equation (21) yields

$$P_4 = P_1 \Phi_1 + P_{0,1,1} \Phi_2, \quad (23)$$

where $\Phi_1 = -\mathbf{M}_{26}\mathbf{M}_{66}^{-1}$ and $\Phi_2 = -\mathbf{M}_{46}\mathbf{M}_{66}^{-1}$.
Using equation (23) into equation (16), we obtain

$$P_{0,0,0} = \frac{1}{\lambda}\left[P_1\Phi_3 + P_{0,1,1}\Phi_4\right], \tag{24}$$

where $\Phi_3 = \mathbf{M}_{21} + \Phi_1\mathbf{M}_{61}$ and $\Phi_4 = \Phi_2\mathbf{M}_{61}$.
From equation (22), we obtain

$$P_5 = P_3\Phi_5 + P_1\Phi_6 + P_{0,1,1}\Phi_7, \tag{25}$$

where $\Phi_5 = -\mathbf{M}_{57}\mathbf{M}_{77}^{-1}$, $\Phi_6 = -\Phi_1\mathbf{M}_{67}\mathbf{M}_{77}^{-1}$ and $\Phi_7 = -\Phi_2\mathbf{M}_{67}\mathbf{M}_{77}^{-1}$.
Substituting equation (25) into equation (20), we get

$$P_3 = (P_2\Phi_8 + P_{0,1,1}\Phi_9 + P_1\Phi_{10})\Phi_{11}^{-1}, \tag{26}$$

where $\Phi_8 = -\mathbf{M}_{35}$, $\Phi_9 = -(\mathbf{M}_{45} + \Phi_7\mathbf{M}_{75})$, $\Phi_{10} = -\Phi_6\mathbf{M}_{75}$
and $\Phi_{11} = \mathbf{M}_{55} + \Phi_5\mathbf{M}_{75}$.
Using equation (26) into equation (18), we get

$$P_2 = (P_1\Phi_{12} + P_{0,1,1}\Phi_{13})\Phi_{14}^{-1}, \tag{27}$$

where $\Phi_{12} = -(\mathbf{M}_{23} + \Phi_{10}\mathbf{M}_{53}\Phi_{11}^{-1})$, $\Phi_{13} = -\Phi_9\mathbf{M}_{53}\Phi_{11}^{-1}$
and $\Phi_{14} = \mathbf{M}_{33} + \Phi_8\mathbf{M}_{53}\Phi_{11}^{-1}$.
Using equations (27) into equation (26), we have

$$P_3 = P_1\Phi_{15} + P_{0,1,1}\Phi_{16}, \tag{28}$$

where $\Phi_{15} = \Phi_{12}\Phi_{14}^{-1}\Phi_8\Phi_{11}^{-1} + \Phi_{10}\Phi_{11}^{-1}$ and $\Phi_{16} = \Phi_{13}\Phi_{14}^{-1}\Phi_8\Phi_{11}^{-1} + \Phi_9\Phi_{11}^{-1}$.
Substituting equations (24), (27) and (28) into equation (17), we have

$$P_1 = P_{0,1,1}\Phi_{17}, \tag{29}$$

and Φ_{17} is given by

$$\begin{aligned} \Phi_{17} &= -AB^{-1}, \\ \text{where } A &= \Phi_4\mathbf{M}_{12} + \lambda\Phi_{13}\Phi_{14}^{-1}\mathbf{M}_{32} + \lambda\mathbf{M}_{42} + \lambda\Phi_{16}\mathbf{M}_{52}, \\ B &= \Phi_3\mathbf{M}_{12} + \lambda\mathbf{M}_{22} + \lambda\Phi_{12}\Phi_{14}^{-1}\mathbf{M}_{32} + \lambda\Phi_{15}\mathbf{M}_{52}. \end{aligned}$$

Substituting the value of P_1 into equations (23), (24), (25), (27) and (28), respectively, we obtain

$$P_4 = P_{0,1,1}(\Phi_{17}\Phi_1 + \Phi_2), \tag{30}$$

$$P_{0,0,0} = P_{0,1,1}\frac{1}{\lambda}(\Phi_{17}\Phi_3 + \Phi_4), \tag{31}$$

$$P_5 = P_{0,1,1}(\Phi_{17}\Phi_{15}\Phi_5 + \Phi_{16}\Phi_5 + \Phi_{17}\Phi_6 + \Phi_7), \tag{32}$$

$$P_2 = P_{0,1,1}(\Phi_{17}\Phi_{12}\Phi_{14}^{-1} + \Phi_{13}\Phi_{14}^{-1}), \tag{33}$$

$$P_3 = P_{0,1,1}(\Phi_{17}\Phi_{15} + \Phi_{16}). \tag{34}$$

Now all probabilities have been expressed as a function of $P_{0,1,1}$. The normalization condition is

$$P_{0,0,0} + P_{0,1,1} + P_1e_1 + P_2e_2 + P_3e_2 + P_4e_1 + P_5e_2 = 1, \tag{35}$$

where e_1 and e_2 are vectors with all of the entries equal to one of dimensions (2×1) and $(N \times 1)$, respectively.

Substituting equations (29), (30), (31), (32), (33) and (34) into (35), we get

$$P_{0,1,1} = \frac{\lambda}{\Phi_{18}}, \tag{36}$$

where

$$\begin{aligned}\Phi_{18} &= \Phi_{17}\Phi_3 + \Phi_4 + \lambda(1 + \Phi_{17}e_1 + \Phi_{17}\Phi_{12}\Phi_{14}^{-1}e_2 + \Phi_{13}\Phi_{14}^{-1}e_2 \\ &+ \Phi_{17}\Phi_{15}e_2 + \Phi_{16}e_2 + \Phi_{17}\Phi_1e_1 + \Phi_2e_2 + \Phi_{17}\Phi_{15}\Phi_5e_2 \\ &+ \Phi_{16}\Phi_5e_2 + \Phi_{17}\Phi_6e_2 + \Phi_7e_2).\end{aligned}$$

The derivation is complete for all stationary probabilities, which can be used to find the measures of performance of the model.

III. Measures of Performance

In this section, we list out some key measures of performance that reflect the bahaviour of the model.

- Let L_s denote the mean number of customers in the system when the servers are busy

$$L_s = \sum_{l=1}^{b}\sum_{m=1}^{b}\sum_{n=0}^{N}(n+l+m)P_{n,i,j},\ 0 \le i \le 2,\ 0 \le j \le 2,\ i+j=2.$$

- Let L_q denote the mean number of customers in the queue when the servers are busy

$$L_q = \sum_{n=1}^{N} nP_{n,i,j},\ 0 \le i \le 2,\ 0 \le j \le 2,\ i+j=2. \tag{37}$$

- The arriving customers are blocked from entering the queue whenever the queue size is N. In this case the blocking probability (P_{block}) is given by

$$P_{block} = P_{N,2,0} + P_{N,1,1} + P_{N,0,2}. \tag{38}$$

- The effective arrival rate is given by

$$\lambda_{eff} = \lambda(1 - P_{block}) = \lambda\left(\sum_{n=0}^{N-1} P_{n,i,j}\right),\ 0 \le i \le 2,\ 0 \le j \le 2,\ i+j=2.$$

  The expected waiting time in the system using Little's law, we get

$$W_s = \frac{L_s}{\lambda_{eff}}.$$

- The expected waiting time in the queue using Little's law, we get

$$W_q = \frac{L_q}{\lambda_{eff}}. \tag{39}$$

- Percentage Variation (P.V.) in waiting time (W_q) is defined as

$$P.V. = \frac{|(W_q)_{b2} - (W_q)_{b1}|}{(W_q)_{b1}} \times 100\%,$$

  where $(W_q)_{b1}$ and $(W_q)_{b2}$ are the waiting time of two values of b.

I. Practical Application of the Model

The queueing model described in this paper has an application in hospital management systems. The model can be applied to situations where the outpatients request an appointment for HIV testing in the clinic centre. The clinic officer monitors the length of booking windows for appointments of the outpatients, but since, in most cases, there is a limited number of doctors, it leads to an unbalanced ratio between the number of outpatients and the doctors. This situation leads to an increase in the length of the booking window and brings the long waiting for appointments. Therefore, to shorten outpatients waiting time, we limit the size of the booking window and assume the clinic centre has a limited slots capacity during a limited length of the booking window. The clinic centre capacity can be divided into sessions/periods of an equal amount of slots. The slots are termed as the maximum size of the outpatient appointments per session. In this scenario, we consider the pooled testing being used to screen the blood of outpatients to detect HIV infections. The slots are taken into service as the primary test for blood sample testing, which is tested in groups. The slots that test positive results opt for a secondary test for further testing. In contrast, all outpatients in the slots are cleared and exempted from further testing if slots test negative. In this situation, appointment, doctors, pooled test (testing in group), limited length of the booking window, primary test and secondary test correspond to the arrivals, servers, batch testing, limited queue capacity, FES and SOS, respectively, in queueing terminology. In the next section, this model application in practice is analyzed with a numerical investigation.

IV. Numerical Investigation

In this section, we present the applicability of the solutions obtained using matrix-decomposition method. We compute the model numerically by taking the model parameters that have close incidence with the practical situations as $N = 14,\ b = 3, \lambda = 0.6,\ \mu_1 = 0.4,\ \mu_2 = 0.3,\ r = 0.4$, with the assumption that $b \leq N$, where

- $\lambda =$ Appointment rate of outpatients to the clinic centre,
- $\mu_1 =$ Service rate during primary test,
- $\mu_2 =$ Service rate during secondary test,
- $r =$ Probability of opting secondary test,
- $b =$ Maximum number of appointment of outpatients per session,
- $N =$ Maximum capacity of the outpatients at the clinic centre.

Table 1: *Variation in different measures of performance with the change in service rate (μ_1 and μ_2)*

μ_1	L_q	W_q	P_{block}	μ_2	L_q	W_q	P_{block}
0.1	3.745370	6.480820	0.038017	0.1	6.193580	10.32330	0.105915
0.2	1.727790	2.884150	0.001948	0.2	1.423120	2.371980	0.001781
0.3	1.116660	1.861360	0.000308	0.3	0.836311	1.393880	0.000123
0.4	0.836311	1.393880	0.000123	0.4	0.655543	1.092590	0.000029
0.5	0.678974	1.131630	0.000078	0.5	0.569728	0.949556	0.000014
0.6	0.580261	0.967104	0.000061	0.6	0.520022	0.866710	0.000009
0.7	0.513680	0.856134	0.000053	0.7	0.487796	0.812998	0.000008
0.8	0.466371	0.777284	0.000048	0.8	0.465319	0.775536	0.000007
0.9	0.431387	0.718978	0.000045	0.9	0.448812	0.748024	0.000006
1.0	0.404684	0.674474	0.000042	1.0	0.436214	0.727027	0.000005

Table 1 shows the effect of the service rate of primary test (μ_1) and secondary test (μ_2) on some performance measures. We observe that increasing $\mu_1(\mu_2)$ decreases L_q, W_q and P_{block}. This is because as $\mu_1(\mu_2)$ increases, outpatients are served faster so that the queue length and the waiting time decrease. Moreover, P_{block} tends to zero due to faster services.

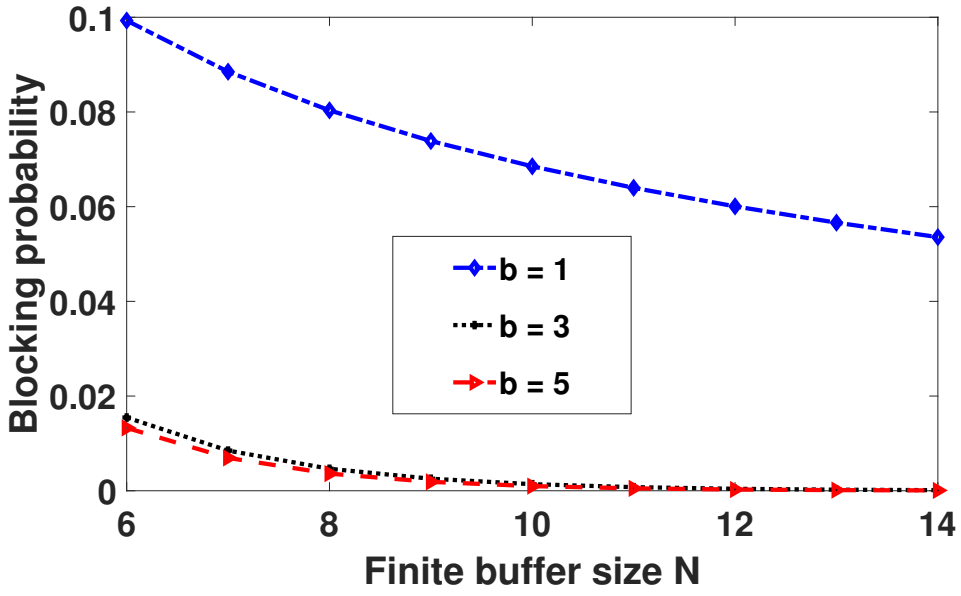


Figure 2: *Effect of buffer size on the blocking probability*

Figure 2 shows the impact of buffer size (N) on the blocking probability (P_{block}) for different batch sizes b. We observe that P_{block} monotonically decreases as N increases and reaches its minimum value zero as N is sufficiently larger. Moreover, when N is kept fixed, P_{block} decreases with the increase of b, as we expect. This is because as b incrases, more number of outpatients are taken for service in a batch, which leads to a decrease in size of the queue. Hence P_{block} decreases.

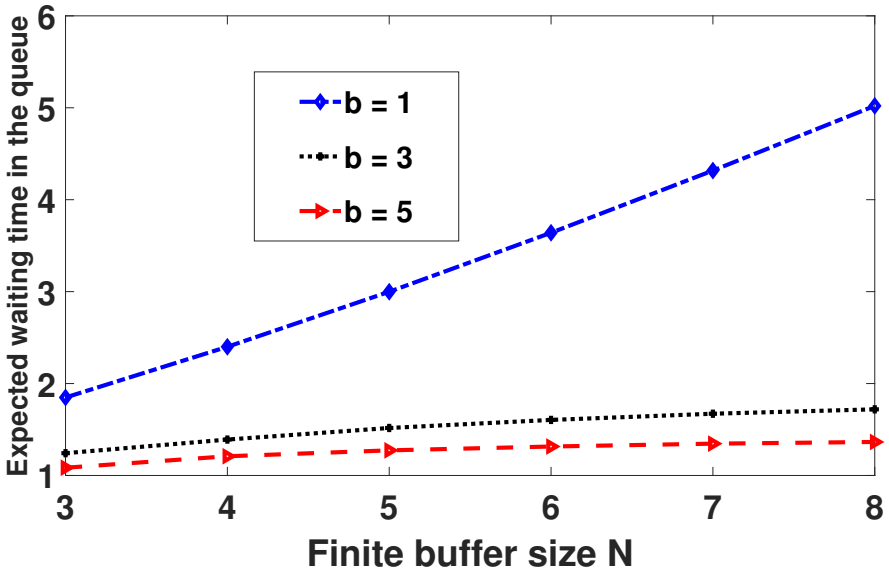


Figure 3: *Effect of finite buffer size on the expected waiting time in the queue* (W_q)

In Figure 3, we observe that for fixed N, W_q decreases as b increases. This is because, as b increases, more outpatients are served in a batch at a time, as a result, W_q decreases. Further, for fixed b, except $b = 1$, the waiting time is more prominently increasing when N lies in $[3, 6]$. However, when the buffer size increases beyond 6, the waiting time is not much affected by increasing N, since the arrival rate is constant and the doctors serve the outpatients in batches $b > 1$. When $b = 1$, the waiting time increases monotonically as N increases. The reason is that by increasing the buffer size, more outpatients accumulate in the queue, and the doctor serves one outpatient at a time, this leads to an increase in the queue length. Hence, the waiting time increases compared to $b = 3$ ($b = 5$).

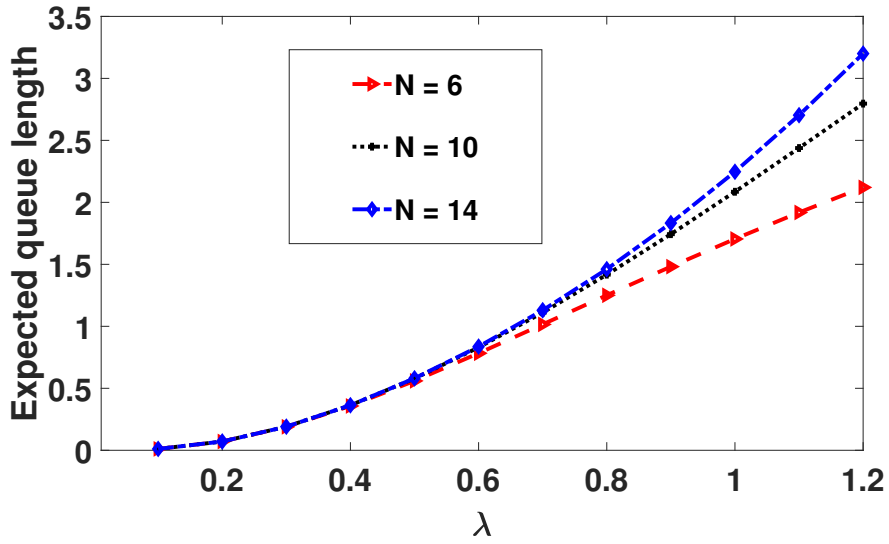


Figure 4: *Impact of the arrival rate on the expected queue length*

Figure 4 shows the impact of λ on the expected queue length (L_q) for different values of N. It is clear that as λ increases, L_q increases for all values of N, as it should be. Moreover, for a fixed λ, as N increases, more outpatients accumulate in the queue thereby an increasing trend can be seen in L_q.

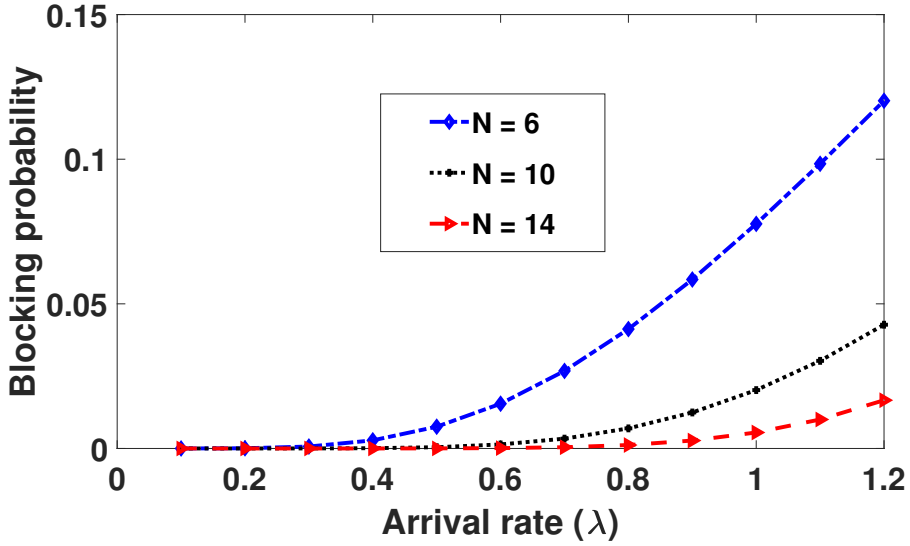


Figure 5: *Impact of arrival rate on the blocking probability*

Figure 5 shows the impact of λ on P_{block}. As intuitively expected, P_{block} increases with the increase of λ. Furthermore, for a fixed λ, P_{block} is high for smaller values of N, which is true.

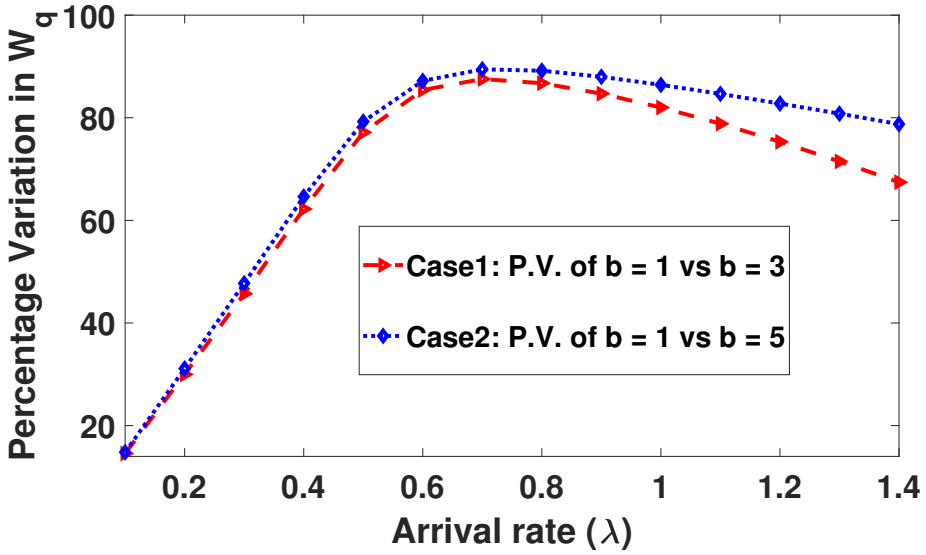


Figure 6: *Impact of arrival rate on the percentage variation in expected waiting time*

Figure 6 shows the impact of arrival rate λ on the P.V. in W_q in two cases, case 1: P.V. of $b = 1$ vs $b = 3$ and case 2: $b = 1$ vs $b = 5$. As λ increases, the P.V. in W_q for case 1 and case 2 initially increases up to $\lambda = 0.7$ and for $\lambda > 0.7$ it slightly decreases. Moreover, the P.V. in W_q varies widely as λ increases. This means that as arrival rate increases, there is a high probability of blocking outpatients to enter the clinic centre.

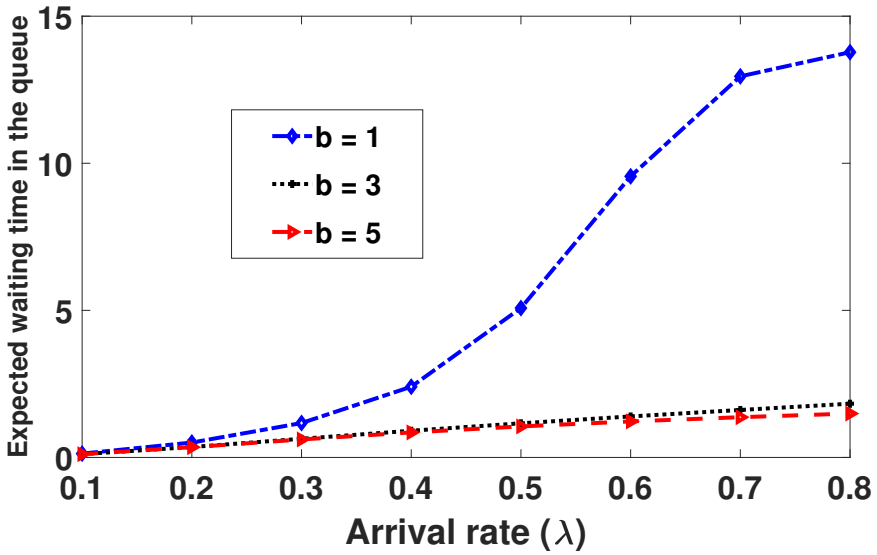


Figure 7: *Impact of arrival rate on the anticipated waiting time*

For different values of b, we show the impact of λ on the anticipated waiting time (W_q) in Figure 7. As λ increases, the inflow of outpatients to the clinic centre increases, which tends to a longer queue. Hence, W_q increases. Further, W_q shows an opposite trend with the increase of batch size taken for the service, as intuitively expected.

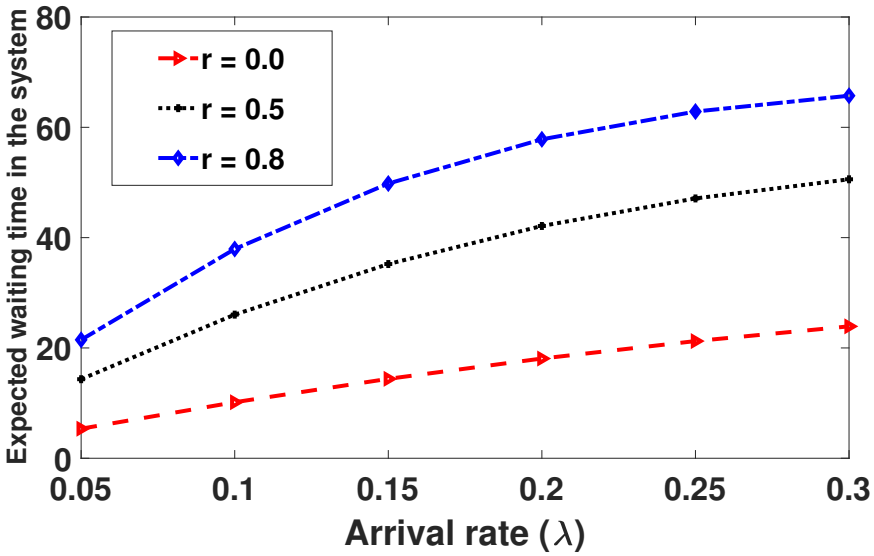


Figure 8: *Impact of arrival rate on the anticipated waiting time in the system*

Figure 8 demonstrates the impact of λ on the anticipated waiting time in the system (W_s) for different values of r. For a fixed r, it is observed that W_s increases with the increase of λ, as we expect. Furthermore, for a fixed λ, W_s is smaller in the absence of secondary test ($r = 0.0$), and as r increases, more number of outpatients are tend to secondary test (SOS), which leads to an increase in W_s.

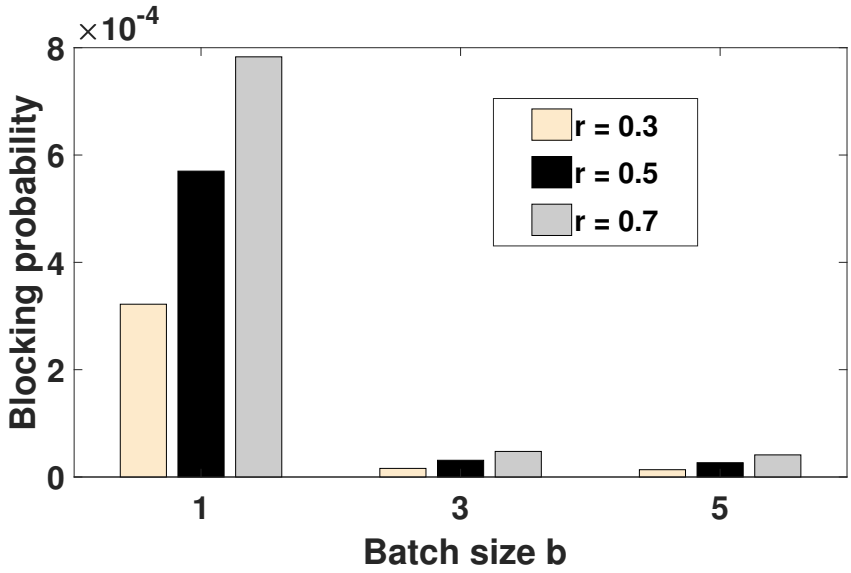


Figure 9: *Effect of batch size on blocking probability*

Figure 9 shows the effect of b on the blocking probability (P_{block}) for different r. As b increases, more outpatients are served in batch and leave the queue, which results in a decrease of P_{block}. Furthermore, for fixed b, P_{block} increases as r increases. The reason is that the outpatients opting for secondary test increase the expected waiting time of other outpatients to be served.

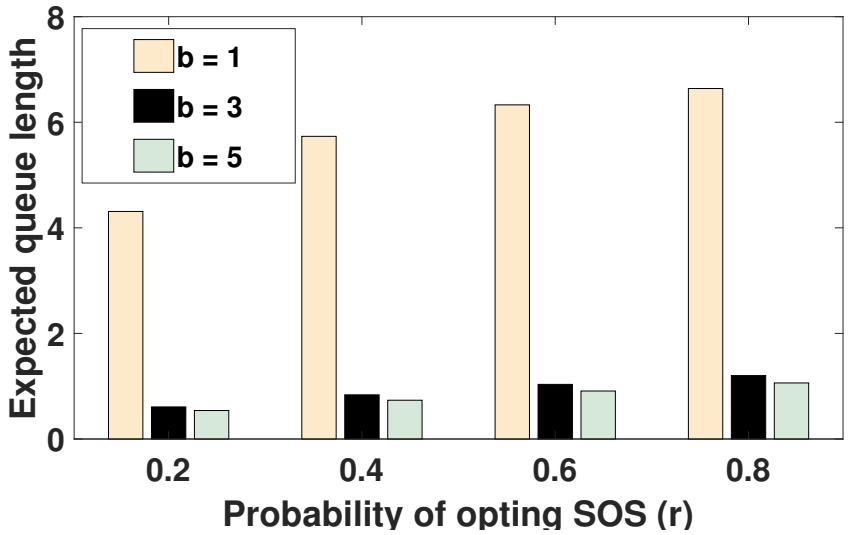


Figure 10: *Effect of probability r on the expected queue length*

Figure 10 depicts the effect of r on the anticipated queue length (L_q). We observe that as r increases, outpatients opt for secondary test, thereby increasing the expected queue length. Also, for a fixed r, as batch size increases, more outpatients are served at a time, resulting in a decrease in L_q.

V. Cost Model Optimization

In this section, we present the total anticipated cost function per item per unit time. The main goal is to figure out the optimal value of service rates μ_1 and μ_2 for primary and secondary tests, respectively, so that the cost is minimized. We apply the Quasi-Newton technique to derive the optimal values of the service rates.

I. Quasi-Newton Method

Quasi-Newton method is reliable and efficient for finding a minimizer of a nonlinear function by estimating the curvature along a sequence of search directions with some fixed tolerance (say ϵ). Let μ^i denote the current point at iteration $i = 0,\ 1,\ 2, \cdots$ The gradient of f at μ^i is denoted $\overrightarrow{\nabla} f(\mu^i)$ and abbreviated to $\overrightarrow{\nabla} f(\mu)$. We use $H(\mu^i)$ to denote a positive definite matrix which is an estimate of the inverse Hessian $\overrightarrow{\nabla}^2 f(\mu^i)^{-1}$. It is important to note that if f is not differentiable at μ^{i+1} we say that the algorithm breaks down (in theory) and if $\overrightarrow{\nabla} f(\mu) = 0$ we say it terminates at a smooth stationary point (for more details the reader may refer **[?]**).

II. Algorithm for Quasi-Newton method

The steps of Quasi-Newton method can be described as follows:

Step 1: Pick any starting trial solution for $\mu^0 = (\mu_1^0, \mu_2^0)$ and compute $f(\mu_1^0, \mu_2^0)$.

Step 2: While $|\frac{\partial f}{\partial \mu_1}| > \epsilon$ or $|\frac{\partial f}{\partial \mu_2}| > \epsilon$; do step $3 - 5$.

Step 3: Compute the cost gradient $\overrightarrow{\nabla} f(\mu) = [\frac{\partial f}{\partial \mu_1}, \frac{\partial f}{\partial \mu_2}]^T$ and the cost Hessian matrix $H(\mu) = \begin{bmatrix} \frac{\partial^2 f}{\partial \mu_1^2} & \frac{\partial^2 f}{\partial \mu_1 \partial \mu_2} \\ \frac{\partial^2 f}{\partial \mu_2 \partial \mu_1} & \frac{\partial^2 f}{\partial \mu_2^2} \end{bmatrix}$ at point $\overrightarrow{\mu}^{(i)}$, provided that $\frac{\partial^2 f}{\partial \mu_1^2}(\frac{\partial^2 f}{\partial \mu_2^2}) - (\frac{\partial^2 f}{\partial \mu_1 \partial \mu_2})^2 \neq 0$, which enables the

existence of inverse of $H(\mu)$.

Step 4: Update the new trial solution $\mu^{(i+1)} = \mu^{(i)} - [H(\mu^{(i)})]^{-1}\overrightarrow{\nabla} f(\mu^{(i)})$.

Step 5: Set $i = i + 1$ and return to step 2. If the gradient is sufficiently smaller than ϵ, then stop.

To implement the above algorithm, we propose the cost function $f(\mu_1, \mu_2)$ per item per unit time as

$$f(\mu_1, \mu_2) = c_q L_q + c_1\mu_1 + c_2\mu_2 + 2c_3. \tag{40}$$

Let us consider the following optimization problem:

$$f(\mu_1^*, \mu_2^*) = \underset{s.t \quad \mu_1 > \mu_2}{\text{Minimize}} f(\mu_1, \mu_2),$$

where the various cost components are defined as follows:
c_q = cost per unit per item present in the queue,
c_1 = cost per unit time when the servers are in FES,
c_2 = cost per unit time when the servers are in SOS,
c_3 = fixed cost for purchase of one server.
The goal of using this cost model is to give emphasis on the service rates in order to have a cost benefit and less congestion at the queueing system.
We apply the Quasi-Newton technique to obtain the optimal value of service rates (μ_1, μ_2). Assuming the values to the coefficients of the expected total cost function in (40) as c_q = \$600, c_1 = \$300, c_2 = \$400, c_3 = \$200, $\epsilon = \epsilon_1 = \epsilon_2 = 10^{-7}$ and $\lambda = 0.6$, $b = 3$, $r = 0.4$ and $N = 14$. Starting with the initial values $(\mu_1^0, \mu_2^0) = (0.4, 0.3)$, the following Table values have been calculated.

Table 2: *Quasi-Newton Method.*

Iteration	0	1	2	3
μ_1	0.40000	0.53331	0.64672	0.69485
μ_2	0.30000	0.37052	0.43434	0.46377
$f(\mu_1, \mu_2)$	1141.79	1010.26	974.409	971.328
L_q	0.83631	0.50344	0.34443	0.29561
$\frac{\partial f}{\partial \mu_1}$	-914.739	-308.569	-68.0794	-2.25317
$\frac{\partial f}{\partial \mu_2}$	-1275.38	-437.473	-99.3126	-3.79388
Hessian	$\begin{bmatrix} 6861.83 & 2.13400 \\ 2.13400 & 18086.7 \end{bmatrix}$	$\begin{bmatrix} 2720.75 & 2.82643 \\ 2.82643 & 6854.12 \end{bmatrix}$	$\begin{bmatrix} 1414.49 & 2.10218 \\ 2.10218 & 3374.53 \end{bmatrix}$	$\begin{bmatrix} 1098.71 & 1.79815 \\ 1.79815 & 2530.27 \end{bmatrix}$
Iteration	4	5	6	7
μ_1	0.69690	0.69667	0.69669	**0.69669**
μ_2	0.46530	0.46513	0.46515	**0.46515**
$f(\mu_1, \mu_2)$	971.324	971.324	971.324	**971.324**
L_q	0.29358	0.29378	0.29376	0.29376
$\frac{\partial f}{\partial \mu_1}$	0.25677	-0.02454	0.00303	-0.00029
$\frac{\partial f}{\partial \mu_2}$	0.34083	-0.04239	0.00402	-0.00050
Hessian	$\begin{bmatrix} 1087.19 & 1.78470 \\ 1.78470 & 2494.97 \end{bmatrix}$	$\begin{bmatrix} 1088.50 & 1.78605 \\ 1.78605 & 2498.18 \end{bmatrix}$	$\begin{bmatrix} 1088.38 & 1.78590 \\ 1.78590 & 2497.79 \end{bmatrix}$	$\begin{bmatrix} 1088.39 & 1.78592 \\ 1.78592 & 2497.83 \end{bmatrix}$

From Table 2, we find that the minimum cost per unit time is $f(\mu_1^*, \mu_2^*)$ = \$971.324 at $(\mu_1^*, \mu_2^*) = (0.69669, 0.46515)$ achieved at seventh iteration.

Table 3: *The optimal service rates* $(\mu_1,\ \mu_2)$ *and cost function* $f(\mu_1, \mu_2)$ *obtained in variation of b and r.*

		μ_1^*	μ_2^*	L_q	$f(\mu_1^*, \mu_2^*)$
	$r = 0.2$	0.80571	0.572062	0.234602	1011.3
$b = 1$	$r = 0.5$	0.728169	-	-	-
	$r = 0.8$	0.691397	-	-	-
	$r = 0.2$	0.687462	0.357727	0.243295	895.306
$b = 3$	$r = 0.5$	0.698797	0.505399	0.314466	1000.48
	$r = 0.8$	0.700593	0.598757	0.367520	1070.19
	$r = 0.2$	0.670994	0.327201	0.250341	882.383
$b = 5$	$r = 0.5$	0.687876	0.475721	0.321209	989.377
	$r = 0.8$	0.697033	0.570477	0.373146	1061.19
	$r = 0.2$	0.667993	0.319684	0.252881	880.000
$b = 7$	$r = 0.5$	0.685428	0.469603	0.323526	987.585
	$r = 0.8$	0.695602	0.565179	0.375082	1059.80
	$r = 0.2$	0.667434	0.317632	0.25369	879.497
$b = 9$	$r = 0.5$	0.684913	0.468318	0.32414	987.285
	$r = 0.8$	0.695243	0.564205	0.375558	1059.59

In Table 3, we observe that

- For fixed b, as r increases, we observe that μ_1^* shows an increasing trend. However, for $b = 1$, as r increases, the reverse trend is observed in μ_1^*. This is necessary in order to balance the system in the profitable manner. On the other hand, it is obvious that $\mu_2^*,\ L_q$ and $f(\mu_1^*, \mu_2^*)$ increase with the increase of r.
- Similarly, for a fixed r, as b increases, $\mu_1^*,\ \mu_2^*$ and $f(\mu_1^*, \mu_2^*)$ all decreases except L_q which increases as b increase. L_q increase because the service rates decrease in order to balance the system profitably.
- For $b = 1$ and $r = (0.5)0.8$, the service rate in secondary test (SOS) does not satisfy the condition $\mu_1 > \mu_2$. Hence, we exclude those values.

In general, we observe the following features of the queueing system:

- As service rates increases, blocking probability, expected waiting time, and queue size decrease.
- The blocking probability is higher for smaller values of buffer size.
- An increase in arrival rate increases the blocking probability, expected waiting time, and queue size.
- An increase in batch size decreases the blocking probability, expected waiting time, queue size, and optimum cost.
- An increase of r probability increases the optimum cost.

VI. Conclusion

A two-server batch service queueing model with SOS is studied using matrix-decomposition method. The cost model optimization was also developed to determine the optimal service rates to minimize the cost. Performance measures and numerical illustrations discussed in this paper provide valuable insights about the functioning of clinic centre in providing the services

of primary and secondary test. The clinic offices will benefit from the simplified numerical simulations, which will help them increase system efficiency.

In future work, we will incorporate a two-server batch service queue with SOS, adding the concepts of working vacations and vacation interruption. Also, we will consider the transient state in the current model with batch arrival.

References

[1] Ammar, S. (2014). Transient analysis of a two-heterogeneous servers queue with impatient behavior. *Egyptian Mathematical Society*, 22(1), 90–95.

[2] Armero, C. and Conesa, D. (2006). Bayesian hierarchical models in manufacturing bulk service queues. *Journal of Statistical Planning and Inference*, 136(2), 335–354.

[3] Ayyappan, G., Devipriya, G., and Subramanian, A. (2014). Analysis of single server batch service queueing system under multiple vacations with second optional service. *International Journal of Mathematical Archive*, 5(2), 126–144.

[4] Ayyappan, G., Devipriya, G., and Subramanian, A. (2014). Analysis of single server fixed batch service queueing system under multiple vacations with gated service. *International Journal of Computer Applications*, 89(5),15–19.

[5] Banerjee, A. (2016). Analysis of finite buffer queue with state dependent service and correlated customer arrivals. *Egyptian Mathematical Society*, 24(2), 295–302.

[6] Barbhuiya, F. and Guptal, U. (2019). A difference equation approach for analysing a batch service queue with the batch renewal arrival process. *Journal of Industrial and Production Engineering*, 25(2), 233–242.

[7] Ghimire, S., Ghimire, R., Thapa, G. B., and Fernandes, S. (2017). Multi-server batch service queueing model with variable service rates. *International Journal of Applied Mathematics & Statistical Sciences*, 6(4), 43–54.

[8] Goswami, V. and Samanta, S. (2009). Discrete time bulk service queue with two heterogeneous servers. *Computers and Industrial Engineering*, 56(4), 1348–1356.

[9] Gupta, G. K., Banerjee, A., and Gupta, U. C. (2019). On finite-buffer batch-size-dependent bulk service queue with queue-length dependent vacation. *Quality Technology and Quantitative Management*, 17(6), 501–527.

[10] Gupur, G. (2014). On the eigenvalues of the $m/m/1$ queueing model with second optional service. *Journal of Pseudo-Differential Operators and Applications*, 5(3), 359–410.

[11] Hwang, J., Gao, L., and Jang, W. (2010). Joint demand and capacity management in a restaurant system. *European journal of operational research*, 207(1), 465–472.

[12] Kalyanaraman, R. and Murugan, S. (2008). A single server queue with additional optional service in batches and server vacation. *Applied mathematical sciences*, 2(56), 2765–2776.

[13] Ke, J. C., Wu, C. H., and Pearn, W. L. (2011). Multi-server retrial queue with second optional service: Algorithmic computation and optimisation. *International journal of systems science*, 42(10), 1755–1769.

[14] Ke, J. C., Wu, C. H., and Pearn, W. L. (2013). Infinite capacity multi-server queue with second optional service channel. *Journal of physics. Conference series*, 410, 1–4.

[15] Krishna Reddy, G. and Anitha, R. (1998). Markovian bulk service queue with delayed vacations. *Computers and Operations Research*, 25(12), 1159–1166.

[16] Krishnamoorthy, A. and Sreenivasan, C. (2012). An $m/m/2$ queueing system with heterogeneous servers including one with working vacation. *International Journal of Stochastic Analysis*, 2012, 1–16.

[17] Kumar, B. S., Madheswari, S., and Venkatakrishnan, K. S. (2007). Transient solution of an $m/m/2$ queue with heterogeneous servers subject to catastrophes. *International Journal of Information and Management Sciences*, 18(1), 63–80.

[18] Kumar, B. S., Sharma, V., and Seth, A. (2020). An $m/m/2$ heterogeneous service markovian feedback queuing model with reverse balking, reneging and retention of reneged customers. *Advances in Computing and Intelligent Systems, Algorithms for Intelligent System*, pages 291–296.

[19] Kumar, R. and Sharma, S. (2017). Transient solution of a two-heterogeneous server's queuing system with retention of reneging customers. *Bulletin of the Malaysian Mathematical Sciences Society*, 42(4), 1–18.

[20] Lewis, A. S. and Overton, M. L. (2013). Nonsmooth optimization via quasi-newton methods. *Mathematical Programming, Series B*, 141, 135–166.

[21] Sikdar, K. and Gupta, U. (2008). On the batch arrivals batch service queue with finite buffer under server's vacation: $m^X/g^Y/1/n$. *Computers and Mathematics with Applications*, 56(11), 2861–2873.

[22] Uma, S. and Punniyamoorthy, K. (2016). Single server bulk queue with second optional service, balking and compulsory vacation. *International journal of engineering research and applications*, 6(10), 15–20.

[23] Wang, J., W. F. S. J. and Kuki, A. (2017). Finite source retrial queues with two phase service. *International Journal of Operational Research*, 30(4), 421–440.

[24] Yang, D. Y. and Chen, Y. H. (2018). Computational and optimization of a working breakdown queue with second optimal service. *Journal of Industrial and Production Engineering*, 35(3), 81–88.

[25] Yue, D., Y. W. and Tian, R. (2010). Analysis of two-server queues with a variant vacation policy. *JThe Ninth International Symposium on Operations Research and Its Applications (ISORA), Chengdu-Jiuzhaigou, China, August 19-23*, pages 483–491.

PROPERTIES OF QUADRASOPHIC FUZZY SET AND ITS APPLICATIONS

G. Aruna [1], J. Jesintha Rosline[2]

•

[1,2] PG and Research Department of Mathematics, Auxilium College (Autonomous) ,
Affiliated to Thiruvalluvar University, Vellore Dist., Tamil Nadu, India
anu9117@gmail.com[1] , jesi.simple@gmail.com[2]

Abstract

Fuzzy set theory is a distinctive way of approaching ambiguous information. In this artifact, we introduce a new extension of fuzzy set known as Quadrasophic Fuzzy set and its properties. The Quadrasophic Fuzzy set has four parameters. The attributes and operations of the Quadrasophic Fuzzy sets are defined with pertinent examples. The arithmetic aggregator operators with a redefined level of 0.5 are introduced. The theorems of aggregator operators of Quadrasophic Fuzzy sets are explained using mathematical formulations. Suitable results and examples are provided to enlighten the proposed method. The arithmetic aggregator operators of the proposed method have been used in decision- making to get the optimal solution with supplementary statistics. Additionally, the selection of appropriate fertilizer in farming is demonstrated using the operators of the suggested model. A decision making approach is also used to develop the proposed method in order to identify the ideal solution. An illustration is provided to examine the unique feature of the proposed method to resolve the decision-making problems with a perfect solution.

Keywords: Quadrasophic Fuzzy set, Operations, Aggregated operator, Decision making , QF-VIKOR

1. Introduction

The most crucial factor in analyzing the vagueness state is the fuzzy set. The theory of fuzzy sets and its numerous extensions help in solving decision-making issues and are useful in real-world circumstances that involve uncertain data. To address the newly introduced non-membership parameter the Intuitionistic fuzzy set was introduced by Atanassov [1] . R. Yager [7] develops the idea of the Pythagorean fuzzy set (PFS) in order to address the IFS deficiency. To tackle a bipolar environment, Zhang [8] proposed the concept of a bipolar fuzzy set. Cuong proposed the idea of a Picture fuzzy set to manage the neutral environment. The Neurosophic fuzzy set was a concept that Florentin Smarandache [2] presented to handle three new parameters .

There is a great deal of unclear information available today. Poor choices can cause a person to suffer in life. A successful existence depends on making decisions. Every field relies heavily on decisions to move forward in the right direction. For the examination of decision-making, fuzzy set theory is useful. The authors devised a variety of strategies to find the MCDM problem's ideal answer. Some of the operators used in MCDM problems comprise the fuzzy weighted technique, arithmetic and geometric operators, power operators, Yager's operator, Dombi's operator, weighted ordered arithmetic and geometric, and hybrid operator. To obtain precise outcomes, authors today use cutting-edge techniques to tackle decision-making problems. And also, in order to obtain the results in an appropriate manner, several authors have incorporated the current models into new methodologies. Also, various extensions of the fuzzy set are used in decision-making environments to produce the best results. The use of MDCM

is widespread, including applications in areas like business management, operations research, neural-network, and medical science.

C. Jana et al. [5] analyzed the role of bipolar fuzzy Dombi aggregation operators in the multi-attribute decision-making process. K. Mohana and R. Jansi [6] developed the weighted arithmetic operator and applied it to the MCDM problem in a bipolar-Pythagorean environment. The VIKOR technique [10] for decision-making was employed by several authors.

A new idea in fuzzy set design, known as a Quadrasophic fuzzy set (QFS), employs two new parameters. Restricted membership values are a QFS exclusive feature that helps it effectively manage uncertain situations. QFS has four parameters: positive membership functions, restricted positive, restricted negative, and negative. The system's advantages and disadvantages are revealed by the restricted value. As a result, QFS consists of techniques with built-in values that make it simpler to understand the situation and deliver accurate results.

In this artifact, Section 2 provides the preliminary definitions. Basic operations of QFS and results are given in Section 3, certain properties, operations and theorems of QFS are presented in Section 4 with pertinent examples. The comparison study with other current models is presented in Section 5 to corroborate our findings. Section 6 illustrates the use of QFS to analyze fertilizer in agricultural decision-making. The novel application of the QF-VIKOR technique in decision-making is shown in Section 7 and Section 8 supplies the conclusion with its scope for further research.

2. Preliminaries

Intuitionistic Fuzzy sets [4]: An intuitionistic fuzzy set I in the non-empty set X is defined as $I = \{(x, \mu(x), \nu(x); x \in X)\}$ where the function $\mu(x), \nu(x) : X \to [0,1]$ represents the membership degree and the non- membership degree value with the condition $0 \leq \mu(x) + \nu(x) \leq 1$. The value $\pi(x) = 1 - \mu(x) - \nu(x)$ is named as the degree of indeterminacy $\forall x \in X$.

Pythagorean Fuzzy set [7] : A Pythagorean fuzzy set P, is defined as $P = \{(x, \mu(x), \nu(x); x \in X\}$ where the function $\mu(x), \nu(x)$:$X \to [0,1]$ represents the membership degree and the non- membership degree value with the condition $0 \leq \mu_P^2(x) + \nu_P^2(x) \leq 1$. The value $\pi_P(x) = \sqrt{1 - (\mu_P^2(x) + \nu_P^2(x))}$ is named as the degree of indeterminacy for $\forall x \in X$.

Picture fuzzy sets [3]: A picture fuzzy set (PFS) on the universe U is defined as $P = \{(x, \mu(x), \eta(x), \nu(x))\}$ where $\mu(x) \in [0,1]$ is the positive membership degree of x in P, $\eta(x) \in [0,1]$ is the neutral membership degree and $\nu(x) \in [0,1]$ is the negative membership degree.

Bipolar fuzzy sets [8] : Let X be the non-empty fuzzy set. The bipolar fuzzy set $B = \{B^-, B^+\}$ is defined in X. $B^+ \in [0,1]$ is the satisfaction degree of x in B, $B^- \in [-1,0]$ is the satisfaction degree of the implicit counter property of x in B.

Properties of Pythagorean fuzzy set [7]: Let $PF_1, PF_2 \in PFS(x)$ then the result as follows,

$$PF_1 + PF_2 = \{x, \sqrt{(\mu_{PF_1}(x))^2 + (\mu_{PF_2}(x))^2 - (\mu_{PF_1}(x))^2(\mu_{PF_2}(x))^2}, \nu_{PF_1}(x)\nu_{PF_2}(x)\}, \forall x \in X.$$

3. Quadrasophic Fuzzy Set

Quadrasophic fuzzy set: The Quadrasophic fuzzy set (QFS) on the universal set X is defined as

$$Q = \{(x, \eta(x), \lambda_\eta(x), \lambda_\mu(x), \mu(x)) | x \in X\}$$

where $\mu(x) : X \to [0,1]$ is the degree of high positive membership of x in Q, $\eta(x) : X \to [-1,0]$ is the degree of high negative membership of x in Q, $\lambda_\mu(x) : X \to [0,0.5]$ is the degree of restricted positive membership of x in Q,$\lambda_\eta(x) : X \to [-0.5,0]$ is the degree of restricted negative membership of x in Q. And it satisfies the following condition: for all $x \in X$, $-1 \leq \mu(x) + \eta(x) \leq 1$, $-0.5 \leq \lambda \leq 0.5$ and $0 \leq \mu^2 + \eta^2 + \lambda^2 \leq 3$ where λ =Length of $(\lambda_\mu, \lambda_\eta)$. Let $QFS(x)$ denotes the collection of all Quadrasophic fuzzy set on X.

3.1. Operations of Quadrasophic Fuzzy Set

Intersection of QFS : The intersection of two quadrasophic fuzzy set Q_1 and Q_2 in QFS is defined as:

$$Q_1 \cap Q_2 = \{max(\eta_{Q_1}(x), \eta_{Q_2}(x)), max(\lambda_{\eta_{Q_1}}(x), \lambda_{\eta_{Q_2}}(x)), \\ min(\lambda_{\mu_{Q_1}}(x), \lambda_{\mu_{Q_2}}(x)), min(\mu_{Q_1}(x), \mu_{Q_2}(x))\} \forall x \in X.$$

Union of QFS: The union of Q_1 and Q_2 in Quadrasophic fuzzy set is defined as:

$$Q_1 \cup Q_2 = \{min(\eta_{Q_1}(x), \eta_{Q_2}(x)), min(\lambda_{\eta_{Q_1}}(x), \lambda_{\eta_{Q_2}}(x)), \\ max(\lambda_{\mu_{Q_1}}(x), \lambda_{\mu_{Q_2}}(x)), max(\mu_{Q_1}(x), \mu_{Q_2}(x))\} \forall x \in X.$$

Subset: Let $Q_1, Q_2 \in Q$ defined on the non - empty set X then Q_1 is the subset of Q_2 denoted by $Q_1 \subseteq Q_2$, if for each $x \in X$; $\eta_{Q_1}(x) \geq \eta_{Q_2}(x)$, $\lambda_{\eta Q_1}(x) \geq \lambda_{\eta Q_2}(x)$, $\lambda_{\mu Q_1}(x) \leq \lambda_{\mu Q_2}(x)$, $\mu_{Q_1}(x) \leq \mu_{Q_2}(x)$.

Complement of QFS: The complement of the set $Q_1 \in Q$ in X is represented as Q_1^C and is defined as $Q_1^C = (\eta^C, \lambda_\eta^C, \lambda_\mu^C, \mu^C)$, where $\eta^C = -1 - \eta$, $\lambda_\eta^C = -0.5 - \lambda_\eta$, $\lambda_\mu^C = 0.5 - \lambda_\mu$ and $\mu^C = 1 - \mu$.

Equal Set: Let $Q_1, Q_2 \in Q$ be defined on the non empty set X then Q_1 is equal set to Q_2 denoted by $Q_1 = Q_2$ if for each $x \in X$;

$$\eta_{Q_1}(x) = \eta_{Q_2}(x), \lambda_{\eta_{Q_1}}(x) = \lambda_{\eta_{Q_2}}(x), \lambda_{\mu_{Q_1}}(x) = \lambda_{\mu_{Q_2}}(x), \mu_{Q_1}(x) = \mu_{Q_2}(x).$$

Distance Metric of QFS : The normalized Hamming distance between any QFS set $Q_1, Q_2 \in Q(x)$ is defined as,

$$d_{Qh}(Q_1, Q_2) = \frac{1}{2n} \sum_{i=1}^{n} [\left|\eta_{Q_1}(x_i))^2 - (\eta_{Q_2}(x_i))^2\right| + \left|(\lambda_{\eta_{Q_1}}(x_i))^2 - (\lambda_{\eta_{Q_2}}(x_i))^2\right| \\ + \left|(\lambda_{\mu_{Q_1}}(x_i))^2 - (\lambda_{\mu_{Q_2}}(x_i)))^2\right| + \left|(\mu_{Q_1}(x_i))^2 - \mu_{Q_2}(x_i))^2\right|].$$

The normalized Euclidean distance between any QFS set $Q_1, Q_2 \in Q(x)$ is defined as,

$$d_{Qh}(Q_1, Q_2) = \sqrt{\frac{\frac{1}{2n} \sum_{i=1}^{n} [\eta_{Q_1}(x_i))^2 - (\eta_{Q_2}(x_i))^2]^2 + [](\lambda_{\eta_{Q_1}}(x_i))^2 - (\lambda_{\eta_{Q_2}}(x_i))^2]^2}{+[(\lambda_{\mu_{Q_1}}(x_i))^2 - (\lambda_{\mu_{Q_2}}(x_i)))^2]^2 + [(\mu_{Q_1}(x_i))^2 - \mu_{Q_2}(x_i))^2]^2}}$$

Proposition 1. The QFS is not the simplification of bipolar fuzzy set.

Proof. Let $Q_1 = (\eta_{Q_1}, \lambda_{\eta Q_1}, \lambda_{\mu Q_1}, \mu_{Q_1})$ be the set in QFS and $B_1 = \{b_1^-, b_1^+\}$ be the set in BFS. Although positive (b_1^+) and negative (b_1^-) degree of membership exists in BFS, there is no restricted level. It does not give the information about the partial belongingness or level of influential. In QFS, separate parameter is fixed to trace the level of restricted. Hence, QFS differs from BFS. ■

Remark 1. If both membership grade of positive and negative restricted value is zero. Then, QFS is equal to BFS.

4. Certain Properties of Quadrasophic Fuzzy Set

In this segment, the operations and Quadrasophic fuzzy set weighted arithmetic operator $(QFWA)$ is defined and proved with theorems and illustrated with model.
Algebraic Sum : Let $Q_1, Q_2 \in Q$
i) If $Q_1 = (\eta_{Q_1}, \lambda_{\eta Q_1}, \lambda_{\mu Q_1}, \mu_{Q_1})$ and $Q_2 = (\eta_{Q_2}, \lambda_{\eta Q_2}, \lambda_{\mu Q_2}, \mu_{Q_2})$ be any QFS, then the algebraic sum is defined as:

$$Q_1 \oplus Q_2 = \{-(-\eta_{Q_1})(-\eta_{Q_2}), -(-\lambda_{\eta_{Q_1}})(-\lambda_{\eta_{Q_2}}), \sqrt{0.5\lambda_{\mu Q_1}^2 + 0.5\lambda_{\mu Q_2}^2 - (0.5\lambda_{\mu Q_1}^2)(0.5\lambda_{\mu Q_2}^2)}, \sqrt{\mu_{Q_1}^2 + \mu_{Q_2}^2 - \mu_{Q_1}^2\mu_{Q_2}^2}\}$$

Then ,

$$Q_1 \oplus Q_2 = \{-(-\eta_{Q_1})(-\eta_{Q_2}), -(-\lambda_{\eta_{Q_1}})(-\lambda_{\eta_{Q_2}}), ((1-(1-0.5\lambda_{\mu Q_1}^2) + (1-(1-0.5\lambda_{\mu Q_2}^2) - (1-(1-0.5\lambda_{\mu Q_1}^2).(1-(1-0.5\lambda_{\mu Q_2}^2))^{\frac{1}{2}}, ((1-(1-\mu_{Q_1}^2) + (1-(1-\mu_{Q_2}^2) - (1-(1-\mu_{Q_1}^2).(1-(1-\mu_{Q_2}^2))^{\frac{1}{2}}\}$$

ii) For $\alpha \geq 0$,

$$\alpha Q_1 = \{-(-\eta_{Q_1})^{\alpha}, -(-\lambda_{\eta Q_1})^{\alpha}, (1-(1-0.5\lambda_{\mu Q_1}^2)^{\alpha})^{\frac{1}{2}}, (1-(1-\mu_{Q_1}^2)^{\alpha})^{\frac{1}{2}}\}.$$

Theorem 1. If $Q_1 = (\eta_{Q_1}, \lambda_{\eta Q_1}, \lambda_{\mu Q_1}, \mu_{Q_1})$ and $Q_2 = (\eta_{Q_2}, \lambda_{\eta Q_2}, \lambda_{\mu Q_2}, \mu_{Q_2})$ be any two set in QFS defined in the non-empty set X and for $\alpha_1, \alpha_2 \geq 0$ the results follow:
i) $Q_1 \oplus Q_2 = Q_2 \oplus Q_1$

$$\begin{aligned} Q_1 \oplus Q_2 &= \{-(-\eta_{Q_1})(-\eta_{Q_2}), -(-\lambda_{\eta_{Q_1}})(-\lambda_{\eta_{Q_2}}), \\ &\quad \sqrt{0.5\lambda_{\mu Q_1}^2 + 0.5\lambda_{\mu Q_2}^2 - (0.5\lambda_{\mu Q_1}^2)(0.5\lambda_{\mu Q_2}^2)}, \sqrt{\mu_{Q_1}^2 + \mu_{Q_2}^2 - \mu_{Q_1}^2\mu_{Q_2}^2}\} \\ &= \{-(-\eta_{Q_2})(-\eta_{Q_1}), -(-\lambda_{\eta_{Q_2}})(-\lambda_{\eta_{Q_1}}), \\ &\quad \sqrt{0.5\lambda_{\mu Q_2}^2 + 0.5\lambda_{\mu Q_1}^2 - (0.5\lambda_{\mu Q_2}^2)(0.5\lambda_{\mu Q_1}^2)}, \sqrt{\mu_{Q_2}^2 + \mu_{Q_1}^2 - \mu_{Q_2}^2\mu_{Q_1}^2}\} \\ &= Q_2 \oplus Q_1 \end{aligned}$$

ii) $\alpha(Q_1 \oplus Q_2) = (\alpha Q_1 \oplus \alpha Q_2)$

$$\begin{aligned} \alpha(Q_1 \oplus Q_2) &= \alpha\{-(-\eta_{Q_1})(-\eta_{Q_2}), -(-\lambda_{\eta_{Q_1}})(-\lambda_{\eta_{Q_2}}), \\ &\quad (1-(1-(0.5\lambda_{\mu Q_1}^2 + 0.5\lambda_{\mu Q_2}^2 - (0.5\lambda_{\mu Q_1}^2 \times 0.5\lambda_{\mu Q_2}^2))))^{\frac{1}{2}}, \\ &\quad (1-(1-(\mu_{Q_1}^2 + \mu_{Q_2}^2 - \mu_{Q_1}^2.\mu_{Q_2}^2)))^{\frac{1}{2}}\} \\ &= \{-(-\eta_{Q_1})^{\alpha}(-\eta_{Q_2})^{\alpha}, -(-\lambda_{\eta_{Q_1}})^{\alpha}(-\lambda_{\eta_{Q_2}})^{\alpha}, \\ &\quad (1-(1-(0.5\lambda_{\mu Q_1}^2 + 0.5\lambda_{\mu Q_2}^2 - (0.5\lambda_{\mu Q_1}^2 \times 0.5\lambda_{\mu Q_2}^2)))^{\alpha})^{\frac{1}{2}}, \\ &\quad (1-(1-(\mu_{Q_1}^2 + \mu_{Q_2}^2 - \mu_{Q_1}^2.\mu_{Q_2}^2))^{\alpha})^{\frac{1}{2}}\} \end{aligned}$$

$$
\begin{aligned}
(\alpha Q_1 \oplus \alpha Q_2) &= \{-(-\eta_{Q_1})^{\alpha}, -(-\lambda_{\eta_{Q_1}})^{\alpha}, (1-(1-0.5\lambda_{\mu Q_1}^2)^{\alpha})^{\frac{1}{2}}, (1-(1-\mu_{Q_1}^2)^{\alpha})^{\frac{1}{2}}\} \\
&\oplus \{-(-\eta_{Q_2})^{\alpha}, -(-\lambda_{\eta_{Q_2}})^{\alpha}, (1-(1-0.5\lambda_{\mu Q_2}^2)^{\alpha})^{\frac{1}{2}}, (1-(1-\mu_{Q_2}^2)^{\alpha})^{\frac{1}{2}}\} \\
&= \{-(-\eta_{Q_1})^{\alpha}(-\eta_{Q_2})^{\alpha}, -(-\lambda_{\eta_{Q_1}})^{\alpha}(-\lambda_{\eta_{Q_2}})^{\alpha}, \\
&(1-(1-0.5\lambda_{\mu Q_1}^2)^{\alpha}(1-0.5\lambda_{\mu Q_2}^2)^{\alpha})^{\frac{1}{2}}, (1-(1-\mu_{Q_1}^2)^{\alpha}(1-\mu_{Q_2}^2)^{\alpha})^{\frac{1}{2}}\} \\
&= \{-(-\eta_{Q_1})^{\alpha}(-\eta_{Q_2})^{\alpha}, -(-\lambda_{\eta_{Q_1}})^{\alpha}(-\lambda_{\eta_{Q_2}})^{\alpha}, \\
&(1-(1-(0.5\lambda_{\mu Q_1}^2+0.5\lambda_{\mu Q_2}^2-(0.5\lambda_{\mu Q_1}^2\times 0.5\lambda_{\mu Q_2}^2)))^{\alpha})^{\frac{1}{2}}, \\
&(1-(1-(\mu_{Q_1}^2+\mu_{Q_2}^2-\mu_{Q_1}^2.\mu_{Q_2}^2)))^{\alpha})^{\frac{1}{2}}\}
\end{aligned}
$$

Hence, $\alpha(Q_1 \oplus Q_2) = (\alpha Q_1 \oplus \alpha Q_2)$

iii) $\alpha_1 Q_1 \oplus \alpha_2 Q_1 = (\alpha_1 \oplus \alpha_2) Q_1$

$$
\begin{aligned}
\alpha_1 Q_1 \oplus \alpha_2 Q_1 &= \alpha_1\{(-\eta_{Q_1}), (-\lambda_{\eta_{Q_1}}), (1-(1-0.5\lambda_{\mu Q_1}^2))^{\frac{1}{2}}, (1-(1-(\mu_{Q_1}^2)))^{\frac{1}{2}}\} \\
&\oplus \alpha_2\{(-\eta_{Q_1}), (-\lambda_{\eta_{Q_1}}), (1-(1-0.5\lambda_{\mu Q_1}^2))^{\frac{1}{2}}, (1-(1-(\mu_{Q_1}^2)))^{\frac{1}{2}}\} \\
&= \{-(-\eta_{Q_1})^{\alpha_1}(-\eta_{Q_1})^{\alpha_2}, -(-\lambda_{\eta_{Q_1}})^{\alpha_1}(-\lambda_{\eta_{Q_1}})^{\alpha_2}, \\
&(1-(1-0.5\lambda_{\mu Q_1}^2)^{\alpha_1}(1-0.5\lambda_{\mu Q_1}^2)^{\alpha_2})^{\frac{1}{2}}, (1-(1-\mu_{Q_1}^2)^{\alpha_1}(1-\mu_{Q_1}^2)^{\alpha_2})^{\frac{1}{2}}\} \\
&= \{-(-\eta_{Q_1})^{\alpha_1+\alpha_2}, -(-\lambda_{\eta_{Q_1}})^{\alpha_1+\alpha_2}, \\
&(1-(1-0.5\lambda_{\mu Q_1}^2)^{\alpha_1+\alpha_2})^{\frac{1}{2}}, (1-(1-\mu_{Q_1}^2)^{\alpha_1+\alpha_2})^{\frac{1}{2}}\}
\end{aligned}
$$

$$
\begin{aligned}
(\alpha_1 \oplus \alpha_2) Q_1 &= (\alpha_1 \oplus \alpha_2)\{(-\eta_{Q_1}), (-\lambda_{\eta_{Q_1}}), (1-(1-0.5\lambda_{\mu Q_1}^2))^{\frac{1}{2}}, (1-(1-(\mu_{Q_1}^2)))^{\frac{1}{2}}\} \\
&= \{-(-\eta_{Q_1})^{\alpha_1+\alpha_2}, -(-\lambda_{\eta_{Q_1}})^{\alpha_1+\alpha_2}, \\
&(1-(1-0.5\lambda_{\mu Q_1}^2)^{\alpha_1+\alpha_2})^{\frac{1}{2}}, (1-(1-\mu_{Q_1}^2)^{\alpha_1+\alpha_2})^{\frac{1}{2}}\}
\end{aligned}
$$

Hence, $\alpha_1 Q_1 \oplus \alpha_2 Q_1 = (\alpha_1 \oplus \alpha_2) Q_1$

Example 1: Consider $Q_1 = (-0.5, -0.2, 0.4, 0.7)$ and $Q_2 = (-0.6, -0.4, 0.5, 0.8)$
i) Since, $-0.5 \geq -0.6$, $-0.2 \geq -0.4$, $0.4 \leq 0.5$, $0.7 \leq 0.8$
$\Longrightarrow Q_1 \subseteq Q_2$.
Assume $n = 0.5$ then $-0.25 \geq -0.3$, $-0.1 \geq -0.2$, $0.2 \leq 0.25$, $0.35 \leq 0.4$
$\Longrightarrow nQ_1 \subseteq nQ_2$(for any positive integer n).
ii) The union of Q_1 and Q_2 is: $Q_1 \cup Q_2 = (-0.6, -0.4, 0.5, 0.8)$
Assume $n = 0.5$ then $n(Q_1 \cup Q_2) = (-0.3, -0.2, 0.25, 0.4)$
$nQ_1 = (-0.25, -0.1, 0.2, 0.35)$, $nQ_2 = (-0.3, -0.2, 0.25, 0.4)$
$\Longrightarrow nQ_1 \cup nQ_2 = (-0.3, -0.2, 0.25, 0.4)$
Thus, $n(Q_1 \cup Q_2) = nQ_1 \cup nQ_2$.
iii) The intersection of Q_1 and Q_2 is: $Q_1 \cap Q_2 = (-0.5, -0.2, 0.4, 0.7)$
If $n = 0.5$, then $n(Q_1 \cap Q_2) = (-0.25, -0.1, 0.2, 0.35)$
$\Longrightarrow nQ_1 \cap nQ_2 = (-0.25, -0.1, 0.2, 0.35)$
Thus, $n(Q_1 \cap Q_2) = nQ_1 \cap nQ_2$.

Theorem 2. Let $Q_1 = (\eta_{Q_1}, \lambda_{\eta Q_1}, \lambda_{\mu Q_1}, \mu_{Q_1})$ and $Q_2 = (\eta_{Q_2}, \lambda_{\eta Q_2}, \lambda_{\mu Q_2}, \mu_{Q_2})$ be any two QFS set defined in the non-empty set X and for $n \geq 0$ the results follow:
i) If $Q_1 \subseteq Q_2$ then $nQ_1 \subseteq nQ_2$.
ii) $n(Q_1 \cup Q_2) = nQ_1 \cup nQ_2$.
iii) $n(Q_1 \cap Q_2) = nQ_1 \cap nQ_2$.

Proof. The proof is obvious by Example 1.

■

QFWA Operator: Let $Q_s = (\eta_s, \lambda_{\eta_s}, \lambda_{\mu_s}, \mu_s)\ (s = 1, 2, \ldots, n)$ be the set of QFS. Then $QFWA$ (Quadrasophic Fuzzy Weighted Arithmetic operator) with respect to α_i=$(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is the weight vector, where $\alpha_i \in [0, 1]$ such that $\sum_{i=1}^{n} \alpha_i = 1$ is a function defined from $Q^n \to Q$. Then $QFWA$ is defined as:

$$\begin{aligned} QFWA(Q_1, Q_2, \ldots, Q_n) &= \oplus_{i=1}^{n} \alpha_i Q_i \\ &= (\alpha_1 Q_1 + \alpha_2 Q_2 + \cdots + \alpha_n Q_n) \end{aligned}$$

Theorem 3. Let $Q_s = (\eta_s, \lambda_{\eta_s}, \lambda_{\mu_s}, \mu_s)\ (s = 1, 2, 3 \ldots, n)$ be the set of QFS defined in the non-empty set X. Then $QFWA$ operator of QFS is defined as

$$\begin{aligned} QFWA(Q_1, Q_2, \ldots, Q_s) &= (-\prod_{k=1}^{s}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{s}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, \\ &(1 - \prod_{k=1}^{s}(1 - 0.5\lambda_{\mu_{Q_k}}^2)^{\alpha_k})^{\frac{1}{2}}, (1 - \prod_{k=1}^{s}(1 - \mu_{Q_k}^2)^{\alpha_k})^{\frac{1}{2}}) \end{aligned} \tag{1}$$

Proof. The proof follows the method of mathematical induction. Assume that $n = 2$ then,

$$QFWA(Q_1, Q_2) = \alpha_1 Q_1 \oplus \alpha_2 Q_2 = \alpha_1(\eta_{Q_1}, \lambda_{\eta_{Q_1}}, \lambda_{\mu_{Q_1}}, \mu_{Q_1}) \oplus \alpha_2(\eta_{Q_2}, \lambda_{\eta_{Q_2}}, \lambda_{\mu_{Q_2}}, \mu_{Q_2})$$

$$\begin{aligned} \alpha_1 Q_1 \oplus \alpha_2 Q_1 &= \alpha_1\{(-\eta_{Q_1}), (-\lambda_{\eta_{Q_1}}), (1 - (1 - 0.5\lambda_{\mu Q_1}^2))^{\frac{1}{2}}, (1 - (1 - (\mu_{Q_1}^2)))^{\frac{1}{2}}\} \\ &\oplus \alpha_2\{(-\eta_{Q_1}), (-\lambda_{\eta_{Q_1}}), (1 - (1 - 0.5\lambda_{\mu Q_1}^2))^{\frac{1}{2}}, (1 - (1 - (\mu_{Q_1}^2)))^{\frac{1}{2}}\} \\ &= \{-(-\eta_{Q_1})^{\alpha_1}(-\eta_{Q_1})^{\alpha_2}, -(-\lambda_{\eta_{Q_1}})^{\alpha_1}(-\lambda_{\eta_{Q_1}})^{\alpha_2}, \\ &([1 - (1 - 0.5\lambda_{\mu Q_1}^2)^{\alpha_1}] + [1 - (1 - 0.5\lambda_{\mu Q_2}^2)^{\alpha_2}] \\ &- [1 - (1 - 0.5\lambda_{\mu Q_1}^2)^{\alpha_1}][1 - (1 - 0.5\lambda_{\mu Q_2}^2)^{\alpha_2}])^{\frac{1}{2}}, \\ &([1 - (1 - \mu_{Q_1}^2)^{\alpha_1}] + [1 - (1 - \mu_{Q_2}^2)^{\alpha_2}] \\ &- [1 - (1 - \mu_{Q_1}^2)^{\alpha_1})[1 - (1 - \mu_{Q_2}^2)^{\alpha_2}])^{\frac{1}{2}}\} \\ &= \{-\prod_{k=1}^{2}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{2}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, (1 - (1 - 0.5\lambda_{\mu_{Q_1}}^2)^{\alpha_1})(1 - 0.5\lambda_{\mu_{Q_2}}^2)^{\alpha_2})^{\frac{1}{2}}, \\ &(1 - (1 - \mu_{Q_1}^2)^{\alpha_1}(1 - \mu_{Q_2}^2)^{\alpha_2})^{\frac{1}{2}}\} \end{aligned}$$

$$\begin{aligned} \alpha_1 Q_1 \oplus \alpha_2 Q_2 = \{&-\prod_{k=1}^{2}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{2}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, (1 - \prod_{k=1}^{2}(1 - 0.5\lambda_{\mu_{Q_k}}^2)^{\alpha_k})^{\frac{1}{2}}, \\ &(1 - \prod_{k=1}^{2}(1 - \mu_{Q_k}^2)^{\alpha_k})^{\frac{1}{2}}\} \end{aligned}$$

For $n = s$ assume Equation 1 is true. Thus, the result follows:

$$\begin{aligned} QFWA(Q_1, Q_2, \ldots, Q_s) &= (-\prod_{k=1}^{s}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{s}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, \\ &(1 - \prod_{k=1}^{s}(1 - 0.5\lambda_{\mu_{Q_k}}^2)^{\alpha_k})^{\frac{1}{2}}, (1 - \prod_{k=1}^{s}(1 - \mu_{Q_k}^2)^{\alpha_k})^{\frac{1}{2}}) \end{aligned}$$

For $n = s+1$,

$$\begin{aligned}
QFWA(Q_1, Q_2, \ldots, Q_{s+1}) &= \oplus_{k=1}^{s}(\alpha_k Q_k) \oplus (\alpha_{s+1} Q_{s+1}) \\
&= (-\prod_{k=1}^{s}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{s}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, \\
&(1-\prod_{k=1}^{s}(1-0.5\lambda^2_{\mu_{Q_k}})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{s}(1-\mu^2_{Q_k})^{\alpha_k})^{\frac{1}{2}}) \\
&\oplus ((-\eta_{Q_{s+1}})^{\alpha_{s+1}}, (-\lambda_{\eta_{Q_{s+1}}})^{\alpha_{s+1}}, \\
&(1-(1-0.5\lambda^2_{\mu_{Q_{s+1}}})^{\alpha_{s+1}})^{\frac{1}{2}}, (1-(1-\mu^2_{Q_{s+1}})^{\alpha_{s+1}})^{\frac{1}{2}})
\end{aligned}$$

$$\begin{aligned}
QFWA(Q_1, \ldots, Q_{s+1}) &= (-\prod_{k=1}^{s}(-\eta_{Q_k})^{\alpha_k} \oplus (-\eta_{Q_{s+1}})^{\alpha_{s+1}}, \\
&-\prod_{k=1}^{s}(-\lambda_{\eta_{Q_k}})^{\alpha_k} \oplus (-\lambda_{\eta_{Q_{s+1}}})^{\alpha_{s+1}}, \\
&(1-\prod_{k=1}^{s}(1-0.5\lambda^2_{\mu_{Q_k}})^{\alpha_k})^{\frac{1}{2}} \oplus (1-(1-0.5\lambda^2_{\mu_{Q_{s+1}}})^{\alpha_{s+1}})^{\frac{1}{2}}, \\
&(1-\prod_{k=1}^{s}(1-\mu^2_{Q_k})^{\alpha_k})^{\frac{1}{2}} \oplus (1-(1-\mu^2_{Q_{s+1}})^{\alpha_{s+1}})^{\frac{1}{2}}) \\
&= \{(-\prod_{k=1}^{s+1}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{s+1}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, \\
&(1-\prod_{k=1}^{s}(1-0.5\lambda^2_{\mu Q_k})^{\alpha_k})^{\frac{1}{2}} + (1-(1-0.5\lambda^2_{\mu Q_{s+1}})^{\alpha_{s+1}})^{\frac{1}{2}} \\
&-(1-\prod_{k=1}^{s}(1-0.5\lambda^2_{\mu Q_k})^{\alpha_k})^{\frac{1}{2}}(1-(1-0.5\lambda^2_{\mu Q_{s+1}})^{\alpha_{s+1}}])^{\frac{1}{2}}, \\
&(1-\prod_{k=1}^{s}(1-\mu^2_{Q_k})^{\alpha_k})^{\frac{1}{2}} + (1-(1-\mu^2_{Q_{s+1}})^{\alpha_{s+1}})^{\frac{1}{2}} \\
&-(1-\prod_{k=1}^{s}(1-\mu^2_{Q_k})^{\alpha_k})^{\frac{1}{2}}(1-(1-\mu^2_{Q_{s+1}})^{\alpha_{s+1}})^{\frac{1}{2}}\} \\
QFWA(Q_1, Q_2, \ldots, Q_s) &= (-\prod_{k=1}^{s+1}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{s+1}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, \\
&(1-\prod_{k=1}^{s+1}(1-0.5\lambda^2_{\mu_{Q_k}})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{s+1}(1-\mu^2_{Q_k})^{\alpha_k})^{\frac{1}{2}})
\end{aligned}$$

Hence 1 is true for $n = s+1$. By, the method of mathematical induction we conclude 1 is true for any $n > 0$. ■

Example 2: Let $Q_k = (\eta_k, \lambda_{\eta_k}, \lambda_{\mu_k}, \mu_k)$, $k = 1,2,3,4$ be any four Quadrasophic fuzzy set.Consider $Q_1 = (-0.7, -0.3, 0.3, 0.5)$, $Q_2 = (-0.5, -0.2, 025, 0.7)$, Q_3=$(-0.8, -0.4, 0.3, 0.7)$, and $Q_4 = (-0.6, -0.4, 0.5, 0.9)$. We assume the weight vector is $\alpha = (0.2, 0.3, 0.4, 0.1)$. Then

$$\begin{aligned}
QFWA(Q_1, Q_2, Q_3, Q_4) &= \oplus_{k=1}^{4}(\alpha_k Q_k) \\
&= 0.2(-0.7, -0.3, 0.3, 0.5) + 0.3(-0.5, -0.2, 025, 0.7) \\
&+ 0.4(-0.8, -0.4, 0.3, 0.7) + 0.1(-0.6, -0.4, 0.5, 0.9) \\
&= (-0.657, -0.307, 0.221, 0.708)
\end{aligned}$$

Theorem 4 (Idem-potency Property)**.** If $Q_k = (\eta_k, \lambda_{\eta_k}, \lambda_{\mu_k}, \mu_k)$, k=1, 2, 3, . . . , n be the set of QFS and $Q_k = Q\ \forall k$. Then $QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) = Q$.

Proof. Consider

$$\begin{aligned}
QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) &= \oplus_{k=1}^{n}(\alpha_k Q_k) \\
QFWA(Q_1, Q_2, \ldots, Q_n) &= (-\prod_{k=1}^{n}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{n}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, \\
&\quad (1-\prod_{k=1}^{n}(1-0.5\lambda_{\mu_{Q_k}}^2)^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{n}(1-\mu_{Q_k}^2)^{\alpha_k})^{\frac{1}{2}}) \\
&\quad (since\, Q_k = Q) \\
&= ((-\eta_Q), (-\lambda_{\eta_{Q_k}}), (1-(1-0.5\lambda_{\mu_Q}^2))^{\frac{1}{2}}, \\
&\quad (1-(1-\mu_Q^2))^{\frac{1}{2}}) \\
QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) &= Q.
\end{aligned}$$

■

Theorem 5 (Monotonicity)**.** Let Q_k and $Q_k^{'}$ be the pair of sets in QFS. If $Q_k < Q_k^{'}\ \forall k$, then $QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) = QFWA(Q_1^{'}, Q_2^{'}, Q_3^{'}, \ldots, Q_n^{'})$.

Proof. If $Q_k, Q_k^{'} \in QFS$ and $Q_k < Q_k^{'}\ \forall k$, then

$$\begin{aligned}
&(-\prod_{k=1}^{n}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{n}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, (1-\prod_{k=1}^{n}(1-0.5\lambda_{\mu_{Q_k}}^2)^{\alpha_k})^{\frac{1}{2}}, \\
&(1-\prod_{k=1}^{n}(1-\mu_{Q_k}^2)^{\alpha_k})^{\frac{1}{2}}) \leq (-\prod_{k=1}^{n}(-\eta_{Q_k}^{'})^{\alpha_k}, -\prod_{k=1}^{n}(-\lambda_{\eta_{Q_k}}^{'})^{\alpha_k}, \\
&(1-\prod_{k=1}^{n}(1-(0.5\lambda_{\mu_{Q_k}}^{'})^2)^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{n}(1-(\mu_{Q_k}^{'})^2)^{\alpha_k})^{\frac{1}{2}})
\end{aligned}$$

$$QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) = QFWA(Q_1^{'}, Q_2^{'}, Q_3^{'}, \ldots, Q_n^{'}).$$

■

Example 3: Consider $Q_1 = (-0.7, -0.3, 0.4, 0.8)$ and $Q_2 = (-0.6, -0.2, 0.5, 0.7)$ be two sets in QFS. The weight vector is $\alpha = (0.5, 0.5)$.

Solution: Consider (s) for smaller and (l) for larger value of membership.
We know that,

$$\begin{aligned}
&\prod_{k=1}^{2}(-\eta_{Q_{k_{(s)}}})^{\alpha_k} < \prod_{k=1}^{2}(-\eta_{Q_k})^{\alpha_k} < \prod_{k=1}^{2}(-\eta_{Q_{k_{(l)}}})^{\alpha_k} \\
&= (-0.7)^{0.5}(-0.7)^{0.5} < (-0.7)^{0.5}(-0.6)^{0.5} < (-0.6)^{0.5}(-0.6)^{0.5} \\
&= -0.7006 < -0.6481 < -0.600.
\end{aligned}$$

Also,

$$\begin{aligned}
&\prod_{k=1}^{2}(-\lambda_{\eta_{Q_{k_{(s)}}}})^{\alpha_k} < \prod_{k=1}^{2}(-\lambda_{\eta_{Q_k}})^{\alpha_k} < \prod_{k=1}^{2}(-\lambda_{\eta_{Q_{k_{(l)}}}})^{\alpha_k} \\
&= -0.299 < -0.244 < -0.199
\end{aligned}$$

Also,

$$\begin{aligned}
&(1-\prod_{k=1}^{2}(1-0.5\lambda_{\mu_{Q_{k_{(s)}}}}^2)^{\alpha_k})^{\frac{1}{2}} < (1-\prod_{k=1}^{2}(1-0.5\lambda_{\mu_{Q_k}}^2)^{\alpha_k})^{\frac{1}{2}} < (1-\prod_{k=1}^{2}(1-0.5\lambda_{\mu_{Q_{k_{(l)}}}}^2)^{\alpha_k})^{\frac{1}{2}} \\
&= 0.282 < 0.320 < 0.353
\end{aligned}$$

Also,

$$(1-\prod_{k=1}^{2}(1-\mu_{Q_{k_{(s)}}}^{2})^{\alpha_k})^{\frac{1}{2}} < (1-\prod_{k=1}^{2}(1-\mu_{Q_k}^{2})^{\alpha_k})^{\frac{1}{2}} < (1-\prod_{k=1}^{2}(1-\mu_{Q_{k_{(l)}}}^{2})^{\alpha_k})^{\frac{1}{2}}$$

$$= 0.700 < 0.756 < 0.800$$

$$\{-\prod_{k=1}^{2}(-\eta_{Q_{k_{(s)}}})^{\alpha_k}, -\prod_{k=1}^{2}(-\lambda_{\eta_{Q_{k_{(s)}}}})^{\alpha_k}, (1-\prod_{k=1}^{2}(1-0.5\lambda_{\mu_{Q_{k_{(s)}}}}^{2})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{2}(1-\mu_{Q_{k_{(s)}}}^{2})^{\alpha_k})^{\frac{1}{2}}\}$$
$$< \{-\prod_{k=1}^{2}(-\eta_{Q_k})^{\alpha_k}), -\prod_{k=1}^{2}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, (1-\prod_{k=1}^{2}(1-0.5\lambda_{\mu_{Q_k}}^{2})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{2}(1-\mu_{Q_k}^{2})^{\alpha_k})^{\frac{1}{2}}\}$$
$$< \{-\prod_{k=1}^{2}(-\eta_{Q_{k_{(l)}}})^{\alpha_k}, -\prod_{k=1}^{2}(-\lambda_{\eta_{Q_{k_{(l)}}}})^{\alpha_k}, (1-\prod_{k=1}^{2}(1-0.5\lambda_{\mu_{Q_{k_{(l)}}}}^{2})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{2}(1-\mu_{Q_{k_{(l)}}}^{2})^{\alpha_k})^{\frac{1}{2}}\}$$

And thus, the set of smaller membership grade is lesser than actual membership grade which is lesser than the larger membership grade.

Table 1: *Comparative study between QFS and Bipolar environment*

Type of operator	Environment	Results
Dombi Weighted [5]	Bipolar Fuzzy environment	$Q_4 > Q_1 > Q_2 > Q_5 > Q_3$ $0.706 > 0.694 > 0.661 > 0.600 > 0.583$
Proposed method	Quadrasophic Fuzzy Set	$Q_4 > Q_1 > Q_2 > Q_5 > Q_3$ $0.134 > 0.124 > 0.121 > 0.057 > 0.017$

Table 2: *Comparative study between QFS and Bipolar Pythagorean environment*

Type of operator	Environment	Results
Weighted Average[6]	Bipolar Pythagorean environment	$Q_1 > Q_2 > Q_3 > Q_4$ $0.383 > 0.189 > -0.031 > -0.102$
Proposed method	Quadrasophic Fuzzy set	$Q_1 > Q_2 > Q_4 > Q_3$ $0.256 > 0.084 > -0.045 > -0.052$

Table 3: *Types, ratio and duration of fertilizer*

Fertilizer types	Composition in ratio	Duration
f_1	120:40:40	Short
f_2	150:60:50	Medium
f_3	150:80:50	Long

Theorem 6 (Bounded property)**.** Consider $Q_k = (\eta_k, \lambda_{\eta_k}, \lambda_{\mu_k}, \mu_k)$, k=1, 2, 3, . . . , n be the set of QFS. If
$Q_{(s)} = min(Q_1, Q_2, Q_3, \ldots, Q_n) = (\eta_{k_{(s)}}, \lambda_{\eta_{k_{(s)}}}, \lambda_{\mu_{k_{(s)}}}, \mu_{k_{(s)}})$ and
$Q_{(l)} = max(Q_1, Q_2, Q_3, \ldots, Q_n) = (\eta_{k_{(l)}}, \lambda_{\eta_{k_{(l)}}}, \lambda_{\mu_{k_{(l)}}}, \mu_{k_{(l)}}))$.
Then $Q_{(s)} < QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) < Q_{(l)}$.

Proof. By the example 3, the proof is obvious.

$$
\begin{aligned}
&(-\prod_{k=1}^{n}(-\eta_{Q_{k_{(s)}}})^{\alpha_k}, -\prod_{k=1}^{n}(-\lambda_{\eta_{Q_{k_{(s)}}}})^{\alpha_k},\\
&(1-\prod_{k=1}^{n}(1-0.5\lambda^2_{\mu_{Q_{k_{(s)}}}})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{n}(1-\mu^2_{Q_{k_{(s)}}})^{\alpha_k})^{\frac{1}{2}})\\
&< (-\prod_{k=1}^{n}(-\eta_{Q_k})^{\alpha_k}, -\prod_{k=1}^{n}(-\lambda_{\eta_{Q_k}})^{\alpha_k}, (1-\prod_{k=1}^{n}(1-0.5\lambda^2_{\mu_{Q_k}})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{n}(1-\mu^2_{Q_k})^{\alpha_k})^{\frac{1}{2}})\\
&< (-\prod_{k=1}^{n}(-\eta_{Q_{k_{(l)}}})^{\alpha_k}, -\prod_{k=1}^{n}(-\lambda_{\eta_{Q_{k_{(l)}}}})^{\alpha_k},\\
&(1-\prod_{k=1}^{n}(1-0.5\lambda^2_{\mu_{Q_{k_{(l)}}}})^{\alpha_k})^{\frac{1}{2}}, (1-\prod_{k=1}^{n}(1-\mu^2_{Q_{k_{(l)}}})^{\alpha_k})^{\frac{1}{2}})
\end{aligned}
$$

$$\implies Q_{(s)} < QFWA(Q_1, Q_2, Q_3, \ldots, Q_n) < Q_{(l)}.$$ ■

Table 4: *Decision Matrix of Quadrasophic Fuzzy Set*

Fertilizer	s_1	s_2	s_3
f_1	(-0.8,-0.4,0.3,0.6)	(-0.7,-0.3,0.5,0.8)	(-0.6,-0.2,0.3,0.7)
f_2	(-0.7,-0.3,0.2,0.4)	(-0.5,-0.3,0.5,0.7)	(-0.4,-0.2,0.4,0.8)
f_3	(-0.4,-0.2,0.4,0.8)	(-0.7,-0.3,0.4,0.8)	(-0.6,-0.3,0.3,0.6)

5. Comparative study

In this segment, to corroborate our method, a comparison study is performed to prove that the proposed method generates better and more accurate results than the other existing methodologies. The Table 1 and 2 gives the study of comparison.

Table 5: *Normalized decision matrix of QFS*

Fertilizer	s_1	s_2	s_3
f_1	(-0.2,-0.1,0.2,0.4)	(-0.7,-0.3,0.5,0.8)	(-0.6,-0.2,0.3,0.7)
f_2	(-0.3,-0.2,0.3,0.6)	(-0.5,-0.3,0.5,0.7)	(-0.4,-0.2,0.4,0.8)
f_3	(-0.6,-0.3,0.1,0.2)	(-0.7,-0.3,0.4,0.8)	(-0.6,-0.3,0.3,0.6)

Table 6: *QFWA value of QFS*

Fertilizer	QFWA value
f_1	(-0.4519,-0.183,0.25,0.683)
f_2	(-0.3923,-0.225,0.28,0.724)
f_3	(-0.628,-0.299,0.2096,0.6259)

6. Application of Quadrasophic Fuzzy Set in agricultural field

The QFS has a special function that allows the difficulties to be solved perfectly. In this segment, we design a technique to deal with the decision-making assessment based on the QFWA

Table 7: *Decision Matrix of Quadrasophic Fuzzy Set*

	M_1	M_2	M_3
B_1	(-0.7,-0.4,0.2,0.6)	(-0.6,-0.3,0.4,0.7)	(-0.5,-0.3,0.1,0.3)
B_2	(-0.5,-0.2,0.1,0.3)	(-0.3,-0.2,0.4,0.8)	(-0.2,-0.1,0.3,0.7)
B_3	(-0.3,-0.1,0.4,0.7)	(-0.6,-0.3,0.2,0.7)	(-0.5,-0.3,0.2,0.5)

Table 8: *Score Value of Quadrasophic Fuzzy Set*

	M_1	M_2	M_3
B_1	-0.1	0.06	-0.13
B_2	-0.1	0.23	0.23
B_3	0.23	0	-0.03

operator and score function of QFS.

Crop yields are significantly influenced by fertilizers. Several types of fertilizers are applied to the soil to improve cultivation. The two primary types of fertilizer used in agriculture are organic and inorganic materials. For various purposes during the cultivation process, three primary plant nutrients: nitrogen, potassium, and phosphorus are used in different ratios. Suppose that the farmer wants to get better yields from the cultivation in medium term. The following Table 3 lists the type, ratio, and length of time employed in the cultivation for better results. Let $F = \{f_1, f_2, f_3\}$ be the set of alternative fertilizers used on agricultural land.

Membership grade of QFS in Agriculture :
$\eta-$ refers to the side effects of using fertilizer
$\lambda_\eta-$refers to the rate of soil pollution
$\lambda_\mu-$ refers to the rate of soil potential
$\mu-$ refers to the rate of fertilizer absorption level
Let $s = \{s_1, s_2, s_3\}$ be the group of criteria satisfied by the fertilizers used in the cultivation land. where, s_1- to upgrade the soil fertility, s_2- to promote the cultivation and s_3- healthy crop.
Step 1: The Table 4 provides the QFS decision matrix.
Step 2: The normalized QFS decision matrix is presented in the Table 5 by considering s_1 as the cost factor.
Step 3: Assume that $\alpha = (0.3, 0.3, 0.4)$ is the weight vector of each criteria.
Step 4: Use the QFWA operator to find the aggregated value. The Table 6 presents the QFWA value of a normalized QFS decision matrix.
Step 5: Apply the formula $sv(Q) = \frac{\mu(x)+\lambda_\mu(x)+\eta(x)+\lambda_\eta(x)}{3}$ to find the score value of QFSDM.

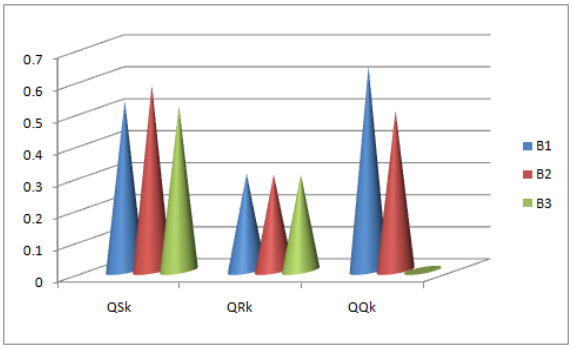


Figure 1: *Values of QS_k, QR_k, and QQ_k*

Table 9: *QFS positive and negative ideal solutions*

	$s_l^{(*)}$	$s_l^{(-)}$
M_1	(-0.3,-0.1,0.4,0.7)	(-0.7,-0.4,0.1,0.3)
M_2	(-0.3,-0.2,0.4,0.8)	(-0.6,-0.3,0.2,0.7)
M_3	(-0.2,-0.1,0.3,0.7)	(-0.5,-0.3,0.1,0.3)

Table 10: *Value of QS_k, QR_k, QQ_k*

	QS_k	QR_k	QQ_k
B_1	0.5324	0.3028	0.6378
B_2	0.5802	0.3000	0.5
B_3	0.51421	0.3000	0
Ranking	B_3, B_1, B_2	$B_3 = B_2, B_1$	B_3, B_2, B_1

$sv(f_1) = 0.0993$, $sv(f_2) = 0.1289$, and $sv(f_3) = -0.0306$.
Step 6: Rank the score value: $sv(f_2) > sv(f_1) > sv(f_3)$.
The fertilizer f_2 is therefore the ideal option for the best production in a medium period of time. Evidently, a medium-duration cultivation technique uses the nutrient ratio 150:60:50. The application gives sufficient justification for QFS to be able to determine the most appropriate decision.

7. A novel MCDM QF-VIKOR technique

A VIKOR technique focuses on ranking options and identifying compromises that are closest to the ideal answer. To define and identify the optimal solution of the Quadrasophic Fuzzy Set, we employ the QF-VIKOR technique in this segment.Consider a businessman who wants to boost profits from their investment in the share marketing sector. Let $M = \{M_1, M_2, M_3\}$ be the set of investment criteria that leads to the alternative businesses B_1, B_2, and B_3.
Step 1: The QFS decision-making matrix is shown in Table 7.
Step 2: The QFS Score matrix is shown in Table 8.
Step 3: The Table 9 shows the values of the QFS ideal solutions, both positive and negative.
Step 4: Assume that the weight vector for each criterion is $\alpha = \{0.3, 0.4, 0.3\}$
Step 5: The values of QS_k, QR_k, QQ_k are provided in the Table 10 using the distance metric 3.1. Rank the alternatives as well. where, [9] $QS_k = \sum_{l=1}^{n} \alpha_l \frac{d(s_l^{(*)}, s_{kl})}{d(s_l^{(*)}, s_l^{(-)})}$, $QR_k = max_l\, \alpha_l \frac{d(s_l^{(*)}, s_{kl})}{d(s_l^{(*)}, s_l^{(-)})}$
$QQ_k = \frac{\beta(QS_k - QS^{(*)})}{(QS^{(-)} - QS^{(*)})} + \frac{(1-Q\beta)(QR_k - QR^{(*)})}{(QR^{(-)} - QR^{(*)})}$ and $QS^{(*)} = minQS_l$, $QS^{(-)} = maxQS_l$, $QR^{(*)} = minQR_l$, and $QR^{(*)} = maxQR_l$, $Q\beta \in [0, 1]$.
Step 6: Figure 1 shows that B_3 is the minimum value. We evaluate how effectively the compromise solution of QQ_k accepts B_3 and B_2. $QQ(B_2) - QQ(B_3) \geq \frac{1}{n-1} \implies 0.5 \geq \frac{1}{3-1} = 0.5 \geq 0.5$ Similarly for QS_k and QR_k .

Hence, the greatest alternative is B_3, whereas B_2 and B_3 are the compromise solutions. Therefore, the ideal firm to invest in for increased profit is B_3.

8. Conclusion

The definitions, characteristics, and some Quadrasophic Fuzzy set operations are defined in this artifact. Theorems and results are also demonstrated using pertinent examples and remarks. To validate our technique, comparative research was conducted in several fuzzy environments. The use of Quadrasophic Fuzzy set in the field of agriculture to identify appropriate

fertilizer has been explored. Further,the QF-VIKOR approach is introduced with its new decision-making application. The use of Quadrasophic Fuzzy set is grounded in the notion that it also works in artificial intelligence, neural-networks, and medicine. Other aggregation operators and their uses in various fields will be examined in further work. Due to its unique parameter classification, Quadrasophic Fuzzy set will also be functional in many domains, including corporate management and psychology.

References

[1] Atanassov, Krassimir T., and S. Stoeva.(1986).Intuitionistic fuzzy sets. *Fuzzy sets and Systems* 20, no. 1 : 87-96.

[2] Broumi, S., A. Bakali, M. Talea, F. Smarandache, V. Ulucay, M. Sahin, A. Dey et al.(2018). Neutrosophic sets: An overview, new trends in neutrosophic theory and applications,*F. Smarandache, S. Pramanik (Editors)* 2 : 388-418.

[3] Cuong, Bui Cong, and V. Kreinovich. (2014). Picture fuzzy sets. *Journal of computer science and cybernetics* 30, no. 4: 409-420.

[4] Ejegwa, P. A., S. O. Akowe, P. M. Otene, and J. M. Ikyule.(2014). An overview on intuitionistic fuzzy sets. *Int. J. Sci. Technol. Res 3*, no. 3 : 142-145.

[5] Jana, Chiranjibe, Madhumangal Pal, and Jian-qiang Wang.(2019).Bipolar fuzzy Dombi aggregation operators and its application in multiple-attribute decision-making process. *Journal of Ambient Intelligence and Humanized Computing* 10 : 3533-3549.

[6] Mohana, K., and R. Jansi. (2018). Bipolar Pythagorean fuzzy sets and their application based on multi-criteria decision-making problems. *International Journal of Research Advent in technology* 6 : 3754-3764.

[7] Yager, Ronald R. (2013).Pythagorean fuzzy subsets. *In 2013 joint IFSA world congress and NAFIPS annual meeting (IFSA/NAFIPS)*, pp. 57-61. IEEE.

[8] Zhang, W-R. (1998). (Yin)(Yang) bipolar fuzzy sets. *In 1998 IEEE international conference on fuzzy systems proceedings. IEEE world congress on computational intelligence (Cat. No. 98CH36228)*, vol. 1, pp. 835-840. IEEE.

[9] Shanthi, S. Anita, and Prathipa Jayapalan.(2019). A VIKOR method based on bipolar intuitionistic fuzzy soft set. *Advances in mathematics* 9, no. 4 : 1511-1519.

[10] Alsolame, Bashayer, and Noura Omair Alshehri.(2020).Extension of VIKOR method for MCDM under bipolar fuzzy set, *International Journal of Analysis and Applications* 18, no. 6 : 989-997.

RT&A, No 4 (76)
Volume 18, December 2023

A PENTAGONAL FUZZY-BASED SOLUTION OF MULTIPLE OBJECTIVE LPP

Junaid Basha M, *Nandhini S, Nur Aisyah Abdul Fataf

•

Department of mathematics, Vellore Institute of Technology, Vellore, 632014, India
*Department of mathematics, Vellore Institute of Technology, Vellore, 632014, India
Cyber Security and Digital Industrial Revilution Centre, Universiti Pertahanan Nasional Malaysia, Malaysia
bashajunaid432@gmail.com
*nandhini.s@vit.ac.in
n.aisyah@upnm.edu.my

Abstract

In this paper, the researchers compare proposed approach and Excel solver. In this proposed technique, the researchers converted fuzzy multiple objective linear programming problems (FMOLPP) into multiple objective linear programming problems (MOLPP) with the help of Defuzzified mean of maxima method. Before that, the researchers changed the pentagonal fuzzy numerical valuation to a triangular fuzzy value by using the proposed theorem. Further, the crisp value of MOLPP is solved using standard simplex algorithms. Then the outcomes of the optimal solutions are compared with both the results.

Keywords: pentagonal fuzzy number, triangular fuzzy number, MOLPP, Defuzzified mean of maxima

1. Introduction

In a 2014 study [1], the Pentagonal Fuzzy Number (PFN) and its mathematical operations were described in this article. These operations were used to resolve a few cases. The problem with five points of approximation could be resolved with a pentagonal fuzzy number. In the 2017 study [2], this paper aims to describe the fundamental idea of a PFN. The researchers examine canonical pentagonal fuzzy numbers (CPFN) utilizing internal arithmetic operations and α-cut procedures. Internal arithmetic features had been studied with the idea of PFN. The CPFN and associated arithmetic operations were given special consideration. This research provides a PFN modification. PFNs of many types are created. Here, a specific kind of PFN's arithmetic operation was discussed. This article also discusses the distinction between two pentagonal-valued functions. The abovementioned numbers were used to demonstrate pentagonal fuzzy results for the fuzzy equation [3]. The 2018 concepts [4], portfolio optimization technique builds on the standard Markowitz mean-variance model and uses PFN to represent returns. The valuation and variance of fuzzy numbers were defined using the alpha-level set approach. The proposed model performs more effectively than the conventional mean-variance method. In this study [5], modeling techniques over multiple PFNs were presented. These similarity criteria depend on geometric distance, l_p-metric distance, graded mean integration form, and the perimeter of a PFN. The ideas of symmetric PFNs and quadratic PFNs, as well as their geometrical examples, were introduced in this study [6]. In addition, define the fundamental operations of arithmetic, such as the addition and subtraction of two symmetric PFNs. This work suggests a straightforward

method for solving the fuzzy transportation problem in the context of a fuzzy environment, where the rates of conveyance, the availability of resources at the providers, and the consumption at the targets were all represented as PFNs. With the use of a strong ranking method and innovative fuzzy arithmetic on PFNs, the fuzzy transportation problem was solved without having to transform it to its equivalent crisp formulation [7].
The focus of 2019 work [8] discusses the fuzzy optimum solution to fully fuzzy linear programming problems (FFLPP) using PFNs. It had been suggested that new methods using multiple ranking functions would be used to solve FFLPP with PFNs. In this study [9], various methods of interval-valued pentagonal fuzzy numbers (IVPFN) connected to various membership functions (MF) were investigated, taking into account the prevalence of various interval-valued fuzzy numbers in specialized studies. Additionally, the concept of MF was substantially generalized to nonlinear membership functions for observing the asymmetries and symmetries of pentagonal fuzzy structures. By using the parameters as PFNs, the produced intellects were applied to a game challenge, leading to a new approach for simulating actual issues and a better knowledge of the parameters' uncertainties during the testing process. In the 2020 study, using a ranking function and comparing the results with completely fuzzy LPP, a technique was suggested for solving fuzzy LPP utilizing PFNs. It involves observing that which produces optimum results [10]. The 2021 study, implementing the Leasing Strategy, aims to lower the machine's rental price. A strong ranking approach and fuzzy arithmetic pentagonal fuzzy numbers were utilized to solve the fuzzy flow-shop scheduling problem without translating the processing time into its equivalent crisp results [11]. In this research design [12], fuzzy results were used to represent the pertinent samples of imprecise rainfall data that were gathered in southern and northern India. Moreover, fuzzy numbers were removed by explicit classification using the PFN pivot points more extensively than the PFN. After the fuzzy phase, a pertinent statistical technique was applied to assess the hypotheses, allowing for better decision-making. In a 2022 study, three steps make up the suggested study technique. With the mean approach of α-cut, the coefficients are first defuzzified. In the second step, a crisp multi-objective quadratic fractional programming model (MOQFP) was built to create a non-fractional system based on an iterative parametric technique. Then, for the final step, the Σ-constraint approach was used to turn this multi-objective, non-fractional model into one with a single objective [13].
The remaining research is organized as follows: 2. Preliminary and Theorem, 3. Proposed algorithm, 4. Numerical example, and 5. Conclusions.

2. Preliminary and Theorem

2.1. Fuzzy set

Let Q be a non-empty set. A fuzzy set P in Q is identified by its membership function $\mu_{\tilde{P}}(y) : Q \rightarrow [0,1]$ and $\mu_{\tilde{P}}(y)$ is described as the degree to which an element is such a member y in fuzzy set P for each $y \in Q$. Then a fuzzy set P in Q is a collection of ordered pairs [14].

$$\tilde{P} = \{(y, \mu_{\tilde{P}}(y))/y \in Q\}$$

Theorem 1. In this theorem, we convert the pentagonal fuzzy number into a triangular fuzzy number ($\sum_{i=1}^{5}\tilde{Q}_i$ are convert into $\sum_{i=1}^{3}\tilde{Q}_i$). By finding the results of pentagonal gradient points? And then obtain the three-tuples fuzzy process transformed as a Defuzzified mean of the maxima of the crisp process.

Proof. Case 1: Let us take the pentagonal fuzzy number

$$\begin{aligned}\tilde{S} &= \sum_{i=1}^{5}\tilde{Q}_i \\ &= (\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3, \tilde{Q}_4, \tilde{Q}_5; \tilde{\omega}_1, \tilde{\omega}_2)\end{aligned}$$

Since, $(\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3, \tilde{Q}_4, \tilde{Q}_5; \tilde{\omega}_1, \tilde{\omega}_2)$ are the pentagonal fuzzy numbers and also we known that $(\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3, \tilde{Q}_4, \tilde{Q}_5)$ are real numbers.
In this case the pentagonal fuzzy number of $\tilde{\omega}_1$ and $\tilde{\omega}_2$ is the graded point of $\tilde{Q}_2$ and $\tilde{Q}_4$ [15].
So, the value of $\tilde{\omega}_1 = \tilde{\omega}_2 = 0$ means the graded points of $\tilde{Q}_2 = \tilde{Q}_4 = 0$
Thus, the values of pentagonal fuzzy number transformed as triangular fuzzy numbers.
So, we can write as, $\tilde{S} = (\tilde{Q}_1, \tilde{Q}_3, \tilde{Q}_5)$ (or) $\tilde{S} = (\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3)$.
As a result, we can obtain three tuples of fuzzy number value, which we refer to as triangular fuzzy number.
Hence, the triangular fuzzy number is $\tilde{S} = \sum_{i=1}^{3} \tilde{Q}_i$
Case 2: Let us take three-tuples triangular fuzzy number

$$\tilde{S} = \sum_{i=1}^{3} \tilde{Q}_i = (\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3)$$

Now, we are converting the three-tuples fuzzy process to the Defuzzified mean of the maxima (MoM) crisp process.

The Defuzzified process of $MoM = \frac{\sum_{\tilde{Q}_s \in D} \tilde{Q}_s}{|D|}$, Where $|D|$ is the number of counted values of $\tilde{Q}_s$.
Since triangular fuzzy numbers have three tuples only, we submit the three tuples in the MoM process.
Thus, the value of $MoM = \frac{\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3}{3} = R_s$, Here we take three-tuples, so the value of $|D| = 3$ and where R_s is the crisp value.
Therefore, the numerical result has been changed from a fuzzy to a crisp process as per the above format. ■

3. Proposed algorithm

- We take that first pentagonal numerical example; convert it into triangular values, and then Defuzzified as per 1 of the mean of maxima.
- A three-tuples of fuzzy MOLPP utilized in the form of as usual simplex algorithm and the transformed into Defuzzified mean of maxima MOLPP crisp values.
- Here, we used the usual simplex algorithm to solve the objective function and constraints.
- Discover the valuation of each of the particular objective functions that are to be maximized or minimized.
- Test the feasibility of the solution in step 3. If it is feasible, then go to step 6. Otherwise, use the dual simplex method to remove infeasibility.
- To give a name to the optimum valuation of the objective functions as r_w.
- Step 4 should be repeated with $w = 1, 2, 3, 4$ and 5. We are only considering five objective functions in this work.
- Find an optimum solution to each LP problem obtained in step 7.
- Then, LP objective problems with an efficient, optimal solution will receive it. Otherwise, we have to follow the same procedure.

4. Numerical example

$$\begin{aligned}
\text{Max } \tilde{w}_1 &= (1,0,2,0,3)\tilde{x}_1 + (7,0,8,0,9)\tilde{x}_2 \\
\text{Max } \tilde{w}_2 &= (10,10,11,11,12)\tilde{x}_1 + (12,13,13,14,14)\tilde{x}_2 \\
\text{Max } \tilde{w}_3 &= (20,0,21,22,22)\tilde{x}_1 + (16,16,17,0,18)x_2\tilde{x}_2 \\
\text{Min } \tilde{w}_4 &= (6,6,7,7,8)\tilde{x}_1 + (2,2,3,4,4)\tilde{x}_2 \\
\text{Min } \tilde{w}_5 &= (25,0,26,0,27)\tilde{x}_1 + (13,0,14,0,15)\tilde{x}_2
\end{aligned}$$

S.to

$$(2,0,3,0,4)\tilde{x}_1 + (11,0,12,0,13)\tilde{x}_2 \leq (4,0,5,0,6)$$
$$(1,1,2,2,3)\tilde{x}_1 + (6,7,7,8,8)\tilde{x}_2 \leq (9,9,10,11,11)$$
$$(5,0,6,6,7)\tilde{x}_1 + (3,3,4,0,5)\tilde{x}_2 \leq (14,14,15,16,16)$$
$$\tilde{x}_2 1, \tilde{x}_2 \geq \tilde{0} \quad (1)$$

By 1 using as per Theorem1 (Case 1) to convert pentagonal fuzzy numerical values into triangular fuzzy values.

$$\text{Max } \tilde{w}_1 = (1,2,3)\tilde{x}_1 + (7,8,9)\tilde{x}_2$$
$$\text{Max } \tilde{w}_2 = (10,11,12)\tilde{x}_1 + (12,13,14)\tilde{x}_2$$
$$\text{Max } \tilde{w}_3 = (20,21,22)\tilde{x}_1 + (16,17,18)\tilde{x}_2$$
$$\text{Min } \tilde{w}_4 = (6,7,8)x_1\tilde{x}_1 + (2,3,4)\tilde{x}_2$$
$$\text{Min } \tilde{w}_5 = (25,26,27)\tilde{x}_1 + (13,14,15)\tilde{x}_2$$

S.to

$$(2,3,4)\tilde{x}_1 + (11,12,13)\tilde{x}_2 \leq (4,5,6)$$
$$(1,2,3)\tilde{x}_1 + (6,7,8)\tilde{x}_2 \leq (9,10,11)$$
$$(5,6,7)\tilde{x}_1 + (3,4,5)\tilde{x}_2 \leq (14,15,16)$$
$$\tilde{x}_1, \tilde{x}_2 \geq \tilde{0} \quad (2)$$

By 2 using as per Theorem1 (Case 2) to convert three-tuples of fuzzy numerical values into crisp values.

$$\text{Max } w_1 = 2x_1 + 8x_2$$
$$\text{Max } w_2 = 11x_1 + 13x_2$$
$$\text{Max } w_3 = 21x_1 + 17x_2$$
$$\text{Min } w_4 = 7x_1 + 3x_2$$
$$\text{Min } w_5 = 26x_1 + 14x_2$$

S.to

$$3x_1 + 12x_2 \leq 5$$
$$2x_1 + 7x_2 \leq 10$$
$$6x_1 + 4x_2 \leq 15$$
$$x_1, x_2 \geq 0 \quad (3)$$

As per the algorithm used, the usual simplex techniques and Excel solver approaches.
First objective function:

$$\text{Max } w_1 = 2x_1 + 8x_2$$

S.to

$$3x_1 + 12x_2 + x_3 = 5$$

$$2x_1 + 7x_2 + x_4 = 10$$
$$6x_1 + 4x_2 + x_5 = 15$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

Max $w_1 = 3.3336$

Table 1: *Excel solver*

Decision variables	x_1	x_2	Objective value (Max w_1)		
Value	0	0.416667	3.333333333		
Coefficients	2	8			
			LHS		RHS
Constraints 1	3	12	5	$\leq$	5
Constraints 2	2	7	3.3333333333	$\leq$	10
Constraints 3	6	4	107	$\leq$	15

Second objective function:

$$\text{Max } w_2 = 11x_1 + 13x_2$$

S.to

$$3x_1 + 12x_2 + x_3 = 5$$
$$2x_1 + 7x_2 + x_4 = 10$$
$$6x_1 + 4x_2 + x_5 = 15$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

Max $w_2 = 18.337$

Table 2: *Excel solver*

Decision variables	x_1	x_2	Objective value (Max w_2)		
Value	1.666667	0	18.33333333		
Coefficients	11	13			
			LHS		RHS
Constraints 1	3	12	5	$\leq$	5
Constraints 2	2	7	3.333333333	$\leq$	10
Constraints 3	6	4	10	$\leq$	15

Third objective function:

$$\text{Max } w_3 = 21x_1 + 17x_2$$

S.to

$$3x_1 + 12x_2 + x_3 = 5$$
$$2x_1 + 7x_2 + x_4 = 10$$
$$6x_1 + 4x_2 + x_5 = 15$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

Max $w_3 = 35.007$

Table 3: *Excel solver*

Decision variables	x_1	x_2	Objective value (Max w_3)		
Value	1.666667	0	35		
Coefficients	21	17			
			LHS		RHS
Constraints 1	3	12	5	$\leq$	5
Constraints 2	2	7	3.333333333	$\leq$	10
Constraints 3	6	4	10	$\leq$	15

Fourthd objective function:

$$\text{Min } w_4 = 7x_1 + 3x_2$$

S.to

$$3x_1 + 12x_2 + x_3 = 5$$
$$2x_1 + 7x_2 + x_4 = 10$$
$$6x_1 + 4x_2 + x_5 = 15$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

Min $w_4 = -11.667$

Table 4: *Excel solver*

Decision variables	x_1	x_2	Objective value (Min w_4)		
Value	1.666667	0	-11.66666667		
Coefficients	-7	-3			
			LHS		RHS
Constraints 1	3	12	5	$\leq$	5
Constraints 2	2	7	3.333333333	$\leq$	10
Constraints 3	6	4	10	$\leq$	15

Fifth objective function:

$$\text{Min } w_5 = 26x_1 + 14x_2$$

S.to

$$3x_1 + 12x_2 + x_3 = 5$$
$$2x_1 + 7x_2 + x_4 = 10$$
$$6x_1 + 4x_2 + x_5 = 15$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

Min $w_5 = -43.342$

Table 5: *Excel solver*

Decision variables	x_1	x_2	Objective value (Min w_5)		
Value	1.666667	0	-43.33333333		
Coefficients	-26	-14			
			LHS		RHS
Constraints 1	3	12	5	$\leq$	5
Constraints 2	2	7	3.333333333	$\leq$	10
Constraints 3	6	4	10	$\leq$	15

Table 6: *Final table*

Comparison of w		
S.no	Excel solver	Proposed technique
1	3.3333	3.3336
2	18.333	18.337
3	35	35.007
4	-11.667	-11.667
5	-43.333	-43.342

5. Conclusion

In this work, the researchers first take a pentagonal triangular fuzzy number and then convert it into a triangular fuzzy number with the use of Theorem 2.1. Furthermore, the triangular or tri-tuple fuzzy valuations are changed into Defuzzified crisp values. The algorithm employs the mean maxima method of crisp values to solve the standard linear programming simplex method. As a comparison result of our proposed LPP simplex method to Excel's solver simplex method, our proposed results are more optimal than Excel's solver.

References

[1] Pathinathan, T. and Ponnivalavan, K. (2014). Pentagonal fuzzy number. *International journal of computing algorithm*, 3:1003–1005.

[2] Kamble, A. J. (2017). Some notes on pentagonal fuzzy numbers. *Int J fuzzy math arch*, 13(2):113–121.

[3] Mondal, S. P. and Mandal M. (2017). Pentagonal fuzzy number, its properties and application in fuzzy equation. *Future Computing and Informatics Journal*, 2(2):110–117.

[4] Ramli, S. and Jaaman, S. H. (2018). Optimal solution of fuzzy optimization using pentagonal fuzzy numbers. *In AIP Conference Proceedings*, 1974(1).

[5] Pathinathan, T. and Mike Dison, E. (2018). Similarity measures of pentagonal fuzzy numbers. *Int J Pure Appl Math*, 119(9):165–175.

[6] Rosline, J. J. and Dison, E. M. (2018). Symmetric pentagonal fuzzy numbers. *Int. J. Pure Appl. Math*, 119:245–253.

[7] Maheswari, P. U., Ganesan, K. (2018). Solving fully fuzzy transportation problem using pentagonal fuzzy numbers. *In Journal of physics: conference series* 1000(1).

[8] Dinagar, D. S. and Jeyavuthin, M. M. (2019). Distinct Methods for Solving Fully Fuzzy Linear Programming Problems with Pentagonal Fuzzy Numbers. *Journal of Computer and Mathematical Sciences*, 10(6):1253–1260.

[9] Chakraborty, A. Mondal, S. P. Alam, S. Ahmadian, A. Senu N. De, D. and Salahshour, S. (2019). The pentagonal fuzzy number: its different representations, properties, ranking, defuzzification and application in game problems. *Symmetry*, 11(2):248.

[10] Siddi, S. (2020). Solving fuzzy LPP for pentagonal fuzzy number using ranking approach. *Mukt Shabd Journal*, 9.

[11] Alharbi, M. G. and El-Wahed Khalifa, H. A. (2021). On a flow-shop scheduling problem with fuzzy pentagonal processing time. *Journal of Mathematics*, 1–7.

[12] Keerthika, K. S. and Parthiban, S. (2021). A Fuzzy Approach To The Test Of Hypothesis Using Pentagonal Fuzzy Number. *NVEO-NATURAL VOLATILES and ESSENTIAL OILS Journal| NVEO*, 3641–3649.

[13] Goyal, V. Rani, N. and Gupta, D. (2022). An algorithm for quadratically constrained multi-objective quadratic fractional programming with pentagonal fuzzy numbers. *Operations Research and Decisions*, 32(1):49–71.

[14] Basha, J. and Nandhini, S. (2023). A fuzzy based solution to multiple objective LPP. *AIMS Mathematics*, 8(4):7714-7730.

[15] Bisht, M. Beg, I. and Dangwal, R. (2023). Optimal solution of pentagonal fuzzy transportation problem using a new ranking technique. *Yugoslav Journal of Operations Research*.

DIFFERENT ESTIMATION METHODS FOR THE PARAMETER OF XGAMMA DISTRIBUTION AND THEIR COMPARISON

Sukanta Pramanik[1*] and Sandipan Maiti[2]

•

[1]Department of Statistics, Siliguri College
Siliguri, India
skantapramanik@gmail.com
[2]Data Science and Analytic Division, MMA, Ipsos Research Pvt. Ltd.
Bangaluru, India
sandipanone@gmail.com

Abstract

The xgamma distribution is vital in reliability/survival analysis and biomedical research. In this article, different estimation methods are proposed for the parameter of this distribution. The distribution is a unique finite mixture of exponential distribution and gamma distribution. Some further properties of the distribution that are not available in the earlier literature are studied. We consider the maximum likelihood estimator, least squares estimator, weighted least squares estimator, percentile estimator, the maximum product spacing estimator, the minimum spacing absolute distance estimator, the minimum spacing absolute log-distance estimator, Cramér von Mises estimator, Anderson Darling estimator, right-tailed Anderson Darling estimator, and compare them using a comprehensive simulation study. For comparison purposes, the estimators' bias and mean squared error are considered. A real data example is also a part of this work. Some model selection techniques are used to choose the best fitting of the distribution to the data.

Keywords: Bootstrap confidence intervals; classical methods of estimation; entropy; mixture distribution; stress-strength reliability.

1. Introduction

The xgamma distribution was introduced by Sen et al. [14]. It is a mixture distribution of $F_1(x) \sim Exp(\theta)$ and $F_2(x) \sim Gamma(3,\theta)$ with their mixing proportions $\pi_1 = \theta/(1+\theta)$ and $\pi_2 = 1-\pi_1$ respectively. The probability density function (pdf) and cumulative distribution function (cdf) of the xgamma distribution are, respectively, given by

$$f(x;\theta) = \frac{\theta^2}{(1+\theta)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}, \qquad x>0, \theta>0 \tag{1}$$

and

$$F(x;\theta) = 1 - \frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}, \qquad x>0, \theta>0. \tag{2}$$

The distribution has gained widespread popularity using reliability, survival analysis and biomedical research. The distribution does not belong to the regular exponential family of distributions; hence, the statistical inferential aspects are not used for the exponential family. The present study aims to estimate the parameter of the xgamma distribution with seven different methods. From the literature survey, there is little attempt made in this direction, and this article is an effort to fill the gap. For this reason, the maximum likelihood estimator (MLE), least squares estimator (LSE), weighted least squares estimator (WLSE), Crmer"von Mises estimator (CvME), Maximum product of spacings estimator (MPSE), Anderson-Darling estimator (ADE), and Right-tail Anderson-Darling estimator (RADE) have been considered for estimation.
The article is organized as follows. Section 2 introduces some further properties of the xgamma distribution. In Section 3, we introduce seven different methods of estimation. A comprehensive Monte Carlo simulation study is presented to evaluate the performances of these estimators concerning bias and mean squared error (MSE) criteria in Section 4. In Section 5, we consider a real data illustration. The concluding remarks are made in Section 6.

2. New Properties

This section discusses some new statistical properties that have yet to be available in earlier literature.

2.1. Incomplete Moments, Mean Deviations, and Lorenz and Benferroni Curves

The r^{th} incomplete moment, say, $m_r^I(t)$, of the xgamma distribution is given by

$$\begin{aligned} m_r^I(t) &= \int_0^t x^r f(x)dx \\ &= \frac{\gamma(r+1,\theta t)}{(1+\theta)\theta^{r-1}} + \frac{\gamma(r+3,\theta t)}{2(1+\theta)\theta^r}. \end{aligned}$$

Apart from range and standard deviation, mean deviation about the mean, δ_1 and median, δ_2 are used as measures of spread in a population. Incomplete moments are used to define $\delta_1 = 2\mu_1^{'}F(\mu_1^{'}) - 2m_1^I(\mu_1^{'})$ and $\delta_2 = \mu_1^{'} - 2m_1^I(\mu_e)$,respectively. Here, $\mu_1^{'} = E(X)$ is to be obtained from r^{th} moment of xgamma distribution with $r = 1$, $F(\mu_1^{'})$ is to be calculated from (2), $m_1^I(\mu_1^{'})$ is the first incomplete function obtained from the above equation with $r = 1$.

The Lorenz and Benferroni curves are defined by $L(p) = m_1^I(x_p)/\mu_1^{'}$ and $B(p) = \frac{m_1^I(x_p)}{(p\mu_1^{'})}$, respectively, where $x_p = F^{-1}(p)$ can be computed numerically by the quantile function with $u = p$. These curves are significantly used in economics, reliability, demography, insurance, and medicine. We refer to Pundir, Arora, and Jain[28] and the references cited therein for details on this aspect.

2.2. Entropies

The entropy measures the variation of the uncertainty of X, a random variable. A popular entropy measure is Renyi entropy [13]. If X has the pdf, f(x), then Renyi entropy is defined by

$$H_R(\beta) = \frac{1}{1-\beta}\ln\left\{\int_0^\infty f^\beta(x)dx\right\} \tag{3}$$

where $\beta > 0$ and $\beta \neq 1$. Suppose X has the pdf in (2). Then, the Renyi entropy of xgamma distribution is

$$H_R(\beta) = \frac{1}{1-\beta}\ln\left\{\sum_{i=0}^{\beta}\binom{\beta}{i}\frac{\theta^{2\beta-i-1}}{2^i(1+\theta)^\beta}\frac{\Gamma(2i+1)}{\beta^{2i+1}}\right\}$$

Shannon measure of entropy is defined as

$$\begin{aligned} H(f) &= E[-\ln f(X)] \\ &= \left\{\frac{\theta+3}{\theta+1}\right\} - ln\left\{\frac{\theta^2}{1+\theta}\right\} - \sum_{i=1}^{\infty}(-1)^i \frac{\Gamma(2i+1)}{2^i(1+\theta)\theta^{i-1}} - \sum_{i=1}^{\infty}(-1)^i \frac{\Gamma(2i+3)}{2^{i+1}(1+\theta)\theta^i} \end{aligned}$$

2.3. Stress-Strength Reliability

The Stress-Strength model is the life of a component with a random strength X subjected to a random stress Y. When a component experiences stress greater than its capacity to withstand, i.e., strength, it breaks and works well when $X > Y$. So, Stress-Strength Reliability is $R = P(Y < X)$. Let $X \sim xgamma(\theta_1)$ and $Y \sim xgamma(\theta_2)$ be independent random variables. Then Stress-Strength Reliability

$$\begin{aligned} R &= P(Y < X) \\ &= \int_0^{\infty} G_y(x) f(x) dx \\ &= 1 - \frac{\theta_2^2}{(\theta_1+\theta_2)(1+\theta_1)(1+\theta_2)}. \\ &\quad \left[(1+\theta_1) + \frac{\theta_1}{(\theta_1+\theta_2)} + \frac{\theta_1^2+\theta_1\theta_2+\theta_2}{(\theta_1+\theta_2)^2} + \frac{3\theta_1\theta_2}{(\theta_1+\theta_2)^3} + \frac{6\theta_1^2\theta_2}{(\theta_1+\theta_2)^4}\right]. \end{aligned}$$

Also if $\theta_1 = \theta_2 = \theta$, then

$$R = \frac{1-\theta}{2}.$$

2.4. Moments of the residual life

The residual life function is essential in reliability/survival analysis, social studies, bio-medical sciences, economics, population study, the insurance industry, maintenance and product quality control, and product technology. Let X denote the lifetime of a unit at age t, then $X_t = X - t \mid X > t$ is the remaining lifetime beyond that age t.

The cdf $F(x)$ is uniquely determined by the r^{th} moment of the residual life of X (for $r = 1, 2, ...$) [Navarro, Franco, and Ruiz [10], and it is given by

$$\begin{aligned} m_r(t) &= \frac{1}{\bar{F}(t)} \int_t^{\infty} (x-t)^r dF(x) \\ &= \frac{(1+\theta)e^{\theta x}}{(1+\theta+\theta x+\frac{\theta^2 x^2}{2})} \left[\sum_{i=0}^{r}(-1)^i \binom{r}{i} t^i \frac{\Gamma(r+1-i,\theta t)}{(1+\theta)\theta^{r-i-1}} + \sum_{i=0}^{r}(-1)^i \binom{r}{i} t^i \frac{\Gamma(r+3-i,\theta t)}{2(1+\theta)\theta^{r-i}}\right]. \end{aligned}$$

In particular, if $r = 1$, then $m_1(t)$ represents an important function called the mean residual life (MRL) function, representing the average life length for a unit alive at age t.

2.5. Moments of the reversed residual life

In some real-life situations, uncertainty is not only related to the future but can also refer to the past. Consider a system whose state is observed only at a specific preassigned inspection time t. If the system is inspected for the first time and found to be 'down', failure relies on the past, i.e. on which instant in $(0, t)$ it has failed. So, the study of a dual notion of the residual life that deals with the past time seems worthwhile [see Di Crescenzo and Longobardi [6]]. If X, a random variable denotes the lifetime of a unit is down at age t, then $\bar{X}_t = t - X \mid X < t$ indicates the idle time or inactivity time or reversed residual life of the unit at age t.

In the case of forensic science, people may be interested in estimating $\bar{X}_t$ to ascertain a person's exact time of death. In the Insurance industry, it represents the period that remained unpaid by a policyholder due to death. The r^{th} moment of $\bar{X}_t$ (for $r = 1,\ 2,\ ...$) is given by

$$\begin{aligned}\bar{m}_r(t) &= \frac{1}{F(t)}\int_0^t (t-x)^r dF(x)\\ &= \frac{1}{\left\{1-\frac{(1+\theta+\theta x+\frac{\theta^2x^2}{2})e^{-\theta x}}{(1+\theta)}\right\}}\left[\sum_{i=0}^{r}(-1)^i\binom{r}{i}t^{r-i}\frac{\gamma(i+1,\theta t)}{(1+\theta)\theta^{i-1}}+\sum_{i=0}^{r}(-1)^i\binom{r}{i}t^{r-i}\frac{\gamma(i+3,\theta t)}{2(1+\theta)\theta^{i}}\right]\end{aligned}$$

In particular, if $r = 1$, then $\bar{m}_1(t)$ represents a function called the mean idle time or inactivity time (MIT) or reversed residual life (MRRL) function that indicates the mean inactive life length for a unit which is first observed down at age t. The properties of the MIT function have been explored by Ahmad, Kayid, and Pellerey [1] and Kayid and Ahmad [7].

3. Estimation on distribution parameter

In this section, we describe seven estimation methods, namely, MLE, LSE, WLSE, CvME, MPSE, ADE and RADE, to obtain the estimators of the parameter θ of the xgamma distribution.

3.1. Maximum likelihood estimator

Let $(X_1, X_2, \ldots, X_n)$ is a random sample from the distribution in (1). Then, the log-likelihood function is given by

$$\ell(\theta) = 2n\log\theta - n\log(1+\theta) + \sum_{i=1}^{n}\log(1+\frac{\theta}{2}X_i^2) - \theta\sum_{i=1}^{n}X_i. \tag{4}$$

The derivative of the log-likelihood function is

$$\frac{d\ell(\theta)}{d\theta} = \frac{2n}{\theta} - \frac{n}{1+\theta} - \sum_{i=1}^{n}X_i + \sum_{i=1}^{n}\frac{\frac{X_i^2}{2}}{1+\frac{\theta}{2}X_i^2}. \tag{5}$$

Equating this to zero does not yield a closed-form solution for the MLE; thus, a numerical method, like Newton Raphson, is used to solve this equation.

3.2. Ordinary and weighted least squares estimator

The ordinary least squares and weighted least squares estimators were proposed by Swain et al. [16] to estimate the parameters of Beta distributions. Suppose $F(X_{i:n}|\theta)$ denotes the cumulative distribution function of the ordered random variables $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ of size n from a distribution function $F(\cdot|\theta)$. Therefore, in this case, the **least square** estimator of θ, say, $\hat{\theta}_{LSE}$ can be obtained by minimizing the function

$$S(\theta) = \sum_{i=1}^{n}\left[F\left(X_{i:n}|\theta\right) - \frac{i}{n+1}\right]^2$$

with respect to θ, where $F(\cdot|\theta)$ is the cdf, given in Eqn. 2. Equivalently, this can be obtained by solving:

$$\sum_{i=1}^{n}\left[F\left(X_{i:n}\mid\theta\right) - \frac{i}{n+1}\right]\eta_1\left(X_{i:n}\mid\theta\right) = 0,$$

where,

$$\eta_1\left(X_{i:n} \mid \theta\right)=\frac{1}{1+\theta}\left[\theta+(\theta^2-1)(1+X_{i:n})+\theta\left(\frac{\theta^2}{2}-1\right)X_{i:n}^2\right]e^{-\theta X_{i:n}}. \tag{6}$$

The **weighted least squares** estimator of θ, say, $\hat{\theta}_{WLSE}$, can be obtained by minimizing

$$W\left(\theta\right)=\sum_{i=1}^{n}\frac{(n+1)^2\,(n+2)}{i\,(n-i+1)}\left[F\left(X_{i:n} \mid \theta\right)-\frac{i}{n+1}\right]^2.$$

This estimator can also be obtained by solving the following:

$$\sum_{i=1}^{n}\frac{(n+1)^2\,(n+2)}{i\,(n-i+1)}\left[F\left(X_{i:n} \mid \theta\right)-\frac{i}{n+1}\right]\eta_1\left(X_{i:n} \mid \theta\right)=0$$

where $\eta_1\left(\cdot \mid \theta\right)$ is given in Equation (6).

3.3. Cramèr-von-Mises estimator

To motivate our choice of Cramèr-von-Mises type minimum distance estimators, MacDonald [8] provided empirical evidence that the estimator's bias is smaller than the other minimum distance estimators. Thus, the Cramèr-von-Mises estimator of θ, say $\hat{\theta}_{CvME}$ can be obtained by minimizing

$$C(\theta)=\frac{1}{12n}+\sum_{i=1}^{n}\left[F\left(X_{i:n}|\theta\right)-\frac{2i-1}{2n}\right]^2$$

with respect to θ. This estimator can also be obtained by solving the non-linear equations

$$\sum_{i=1}^{n}\left[F\left(X_{i:n} \mid \theta\right)-\frac{2i-1}{2n}\right]\eta_1\left(X_{i:n} \mid \theta\right)=0$$

where $\eta_1\left(\cdot \mid \theta\right)$ is given in Equation (6).

3.4. Maximum product of spacings estimator

The maximum product spacing method was introduced by Cheng and Amin [4] as an alternative to MLE to estimate the unknown parameters of continuous univariate distributions. The maximum product spacing method was also derived independently by Ranneby [12] as an approximation to the Kullback-Leibler measure of information. To motivate our choice, Cheng and Amin [5] proved that this method is as efficient as the MLE estimators and consistent under more general conditions. We define the uniform spacings of a random sample from the xgamma distribution as:

$$D_i(\theta)=F\left(X_{i:n} \mid \theta\right)-F\left(X_{i-1:n} \mid \theta\right), \qquad i=1,2,\ldots,n,$$

where $F(X_{0:n} \mid \theta)=0$ and $F(X_{n+1:n} \mid \theta)=1$. Clearly $\sum_{i=1}^{n+1}D_i(\theta)=1$. The maximum product of spacings estimator $\hat{\theta}_{MPSE}$, of the parameter θ is obtained by maximizing, with respect to θ, the geometric mean of the spacings:

$$G\left(\theta\right)=\left[\prod_{i=1}^{n+1}D_i(\theta)\right]^{\frac{1}{n+1}} \tag{7}$$

or, equivalently, by maximizing the function

$$H\left(\theta\right)=\frac{1}{n+1}\sum_{i=1}^{n+1}\log D_i(\theta) \tag{8}$$

The estimator $\hat{\theta}_{MPSE}$ of the parameter θ can be obtained by solving the non-linear equation

$$\frac{d}{d\theta}H(\theta) = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\theta)}\left[\eta_1(X_{i:n}|,\theta) - \eta_1(X_{i-1:n}|,\theta)\right] \quad = \quad 0$$

where, $\eta_1(\cdot \mid \theta)$ is given in Equation (6).

3.5. Anderson-Darling and Right-tail Anderson-Darling estimators

The Anderson-Darling (AD) test [see Anderson and Darling [2]] is an alternative to other statistical tests for detecting sample distribution's departure from normality. Specifically, the Anderson-Darling test converges very quickly towards the asymptote [see Anderson and Darling [3], Pettitt [11] and Stephens [15]]. The **Anderson-Darling estimator** $\hat{\theta}_{ADE}$ of the parameter θ are obtained by minimizing, with respect to θ, the function:

$$A(\theta) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log F(X_{i:n} \mid \theta) + \log \overline{F}(X_{n+1-i:n} \mid \theta)\right\}. \tag{9}$$

This estimator can also be obtained by solving the non-linear equations:

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1(X_{i:n} \mid \theta)}{F(X_{i:n} \mid \theta)} - \frac{\eta_1\left(X_{n+1-i:n} \mid \theta\right)}{\overline{F}(X_{n+1-i:n} \mid \theta)}\right] \quad = \quad 0$$

where, $\eta_1(\cdot \mid \theta)$ is defined in Equation (6).

The **right-tail Anderson-Darling estimator** $\hat{\theta}_{RADE}$ of the parameter θ is obtained by minimizing, with respect to θ, the function:

$$R(\theta) = \frac{n}{2} - 2\sum_{i=1}^{n}F(X_{i:n} \mid \theta) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\overline{F}(X_{n+1-i:n} \mid \theta). \tag{10}$$

These estimators can also be obtained by solving the non-linear equations:

$$-2\sum_{i=1}^{n}\eta_1(X_{i:n} \mid \theta) + \frac{1}{n}\sum_{i=1}^{n}(2i-1)\frac{\eta_1\left(X_{n+1-i:n} \mid \theta\right)}{\overline{F}(X_{n+1-i:n} \mid \theta)} \quad = \quad 0$$

where, $\eta_1(\cdot \mid \theta)$ is defined in Equations (6).

4. Simulation Study

In this section, we have carried out a Monte Carlo simulation study to assess the performance of the proposed estimators (MLE, LSE, WLSE, CvME, MPSE, ADE and RADE) of the parameter θ for the xgamma distribution. First, we generate random data from the xgamma distribution where we can use the fact that the xgamma distribution is a special mixture of the exponential(θ) and gamma$(3, \theta)$ distributions. To generate random data X_i, i = 1, 2, 3, ... n, from the xgamma distribution with parameter θ, we can use the following algorithm:

1. Generate $U_i \sim$ uniform(0, 1), i = 1, 2, 3, ... n
2. Generate $V_i \sim$ exponential(θ), i = 1, 2, 3, ... n
3. Generate $W_i \sim$ gamma(3, θ), i = 1, 2, 3, ... n
4. If $U_i \leq \theta/(1+\theta)$, then set $Z_i = V_i$. Otherwise, set $Z_i = W_i$

Table 1: *True value of θ and the average bias of the different estimation procedures for xgamma distribution*

θ	n	MLE	LSE	WLSE	CvME	MPSE	ADE	RADE
0.1	10	0.003704	0.002111	0.001874	0.002768	0.001992	0.001884	0.001008
	20	0.001981	0.001151	0.001088	0.001505	0.001511	0.001123	0.000638
	40	0.000934	0.000556	0.000529	0.000735	0.001200	0.000531	0.000310
	70	0.000544	0.000318	0.000318	0.000421	0.000845	0.000302	0.000156
	100	0.000392	0.000223	0.000233	0.000296	0.000657	0.000219	0.000120
0.5	10	0.025356	0.016170	0.014216	0.021056	0.009505	0.014093	0.008516
	20	0.011770	0.007034	0.006230	0.009586	0.009145	0.006256	0.003748
	40	0.005585	0.003072	0.002902	0.004403	0.007156	0.002951	0.001367
	70	0.003144	0.002050	0.002021	0.002805	0.005178	0.001866	0.000982
	100	0.001619	0.000868	0.000826	0.001398	0.003546	0.000768	0.000143
1	10	0.066489	0.051949	0.046934	0.063382	0.021548	0.041516	0.027305
	20	0.027416	0.017687	0.015768	0.023747	0.016970	0.014983	0.008669
	40	0.017020	0.012473	0.012031	0.015611	0.013533	0.011465	0.007898
	70	0.008023	0.004010	0.004044	0.005801	0.011038	0.003737	0.002373
	100	0.005219	0.002871	0.003082	0.004134	0.008104	0.002752	0.001560
1.5	10	0.107576	0.079362	0.071569	0.097780	0.039168	0.062468	0.040920
	20	0.051824	0.037330	0.033998	0.047283	0.029162	0.032012	0.021111
	40	0.024598	0.015411	0.014735	0.020502	0.023889	0.013955	0.009203
	70	0.014383	0.009096	0.009045	0.012023	0.017182	0.008117	0.005517
	100	0.011339	0.008878	0.008546	0.010944	0.011182	0.008025	0.005675

A Monte Carlo simulation study was carried out considering N = 5000 times for selected values of n, θ. For the first simulation, samples of sizes 10, 20, 40, 70 and 100 were considered, and values of θ were taken as 0.1, 0.5, 1.0, 1.5. The required numerical evaluations are carried out using R 3.1.1 software. The following two measures were computed:

1. Average bias of the simulated estimate $\hat{\theta}$, for i=1, 2, 3, ...,N is $\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)$
2. Average Mean Square Error (MSE) of the simulated estimate $\hat{\theta}$, for i=1, 2, 3, ..., N is $\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2$

In Table 1, we have calculated the average bias of the parameter θ using MLE, LSE, WLSE, CvME, MPSE, ADE and RADE.
In Table 2, we have calculated the average MSEs of the parameter θ using MLE, LSE, WLSE, CvME, MPSE, ADE and RADE.
Table 1 shows that
(i) Bias decreases as n increases.
(ii) Bias decreases as the values of θ increases.
Table 2 shows that
(i) MSE decreases as n increases.
(ii) MSE decreases as the values of θ increases.
Comparing the Tables $1-2$ and Figures $1-2$, even though the MLE is comparatively easy to calculate, the ADE or RADE is preferable from the bias and MSE point of view.

5. Data Analysis

The data set is given by Murthy et al. [9] and represents the failure time of 20 components. The data are 0.072, 4.763, 8.663, 12.089, 0.477, 5.284, 9.511, 13.036, 1.592, 7.709, 10.636, 13.949, 2.475, 7.867, 10.729, 16.169, 3.597, 8.661, 11.501 and 19.809. A summary of these data is: n = 20, $\bar{x} = 8.42945$, s = 5.322056, skewness = 0.1769692, kurtosis = 2.430915. The box plot and the Total Time on Test (TTT) plot of these observations are displayed in Figure 3. The box plot indicates

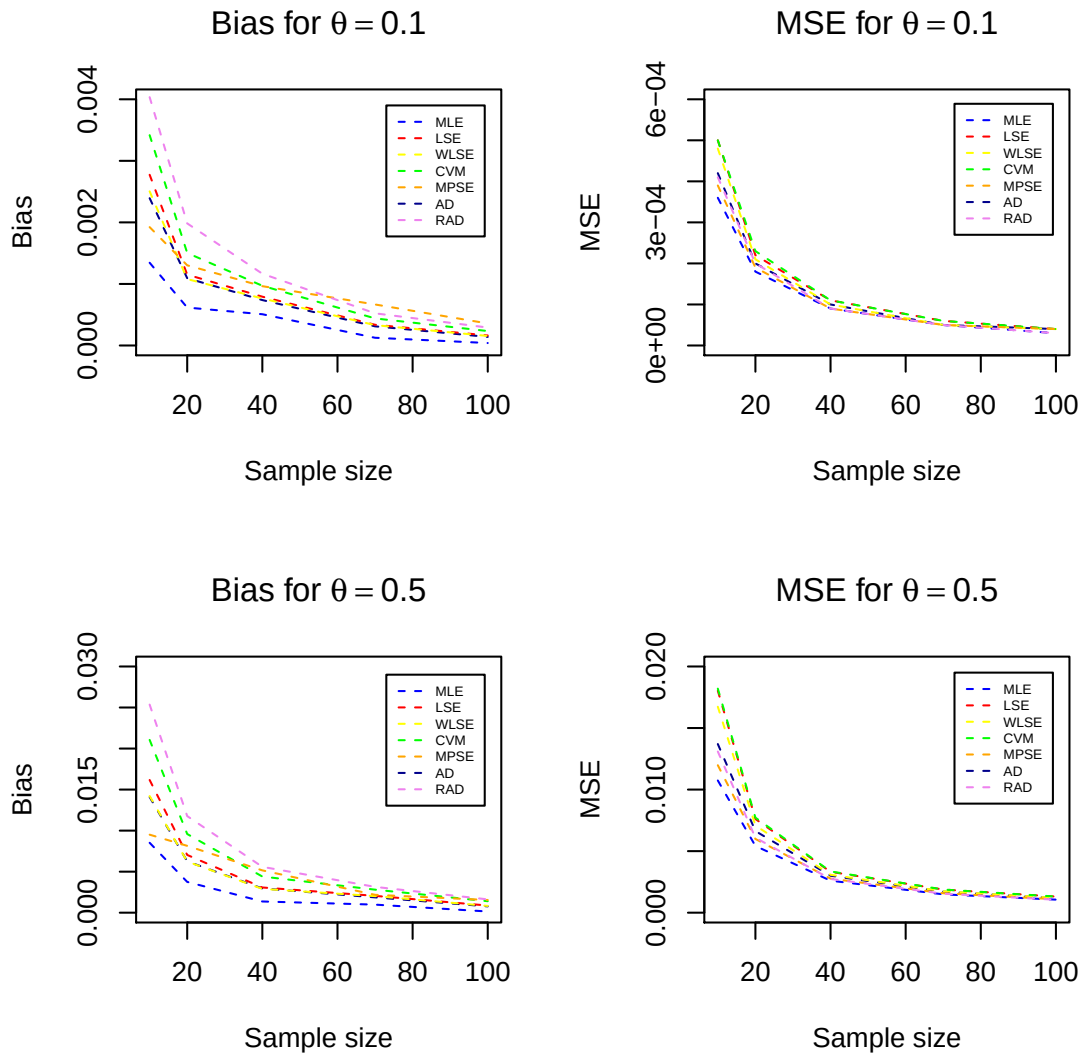


Figure 1: *Average bias and MSE of the estimator of xgamma distribution for different estimation procedures*

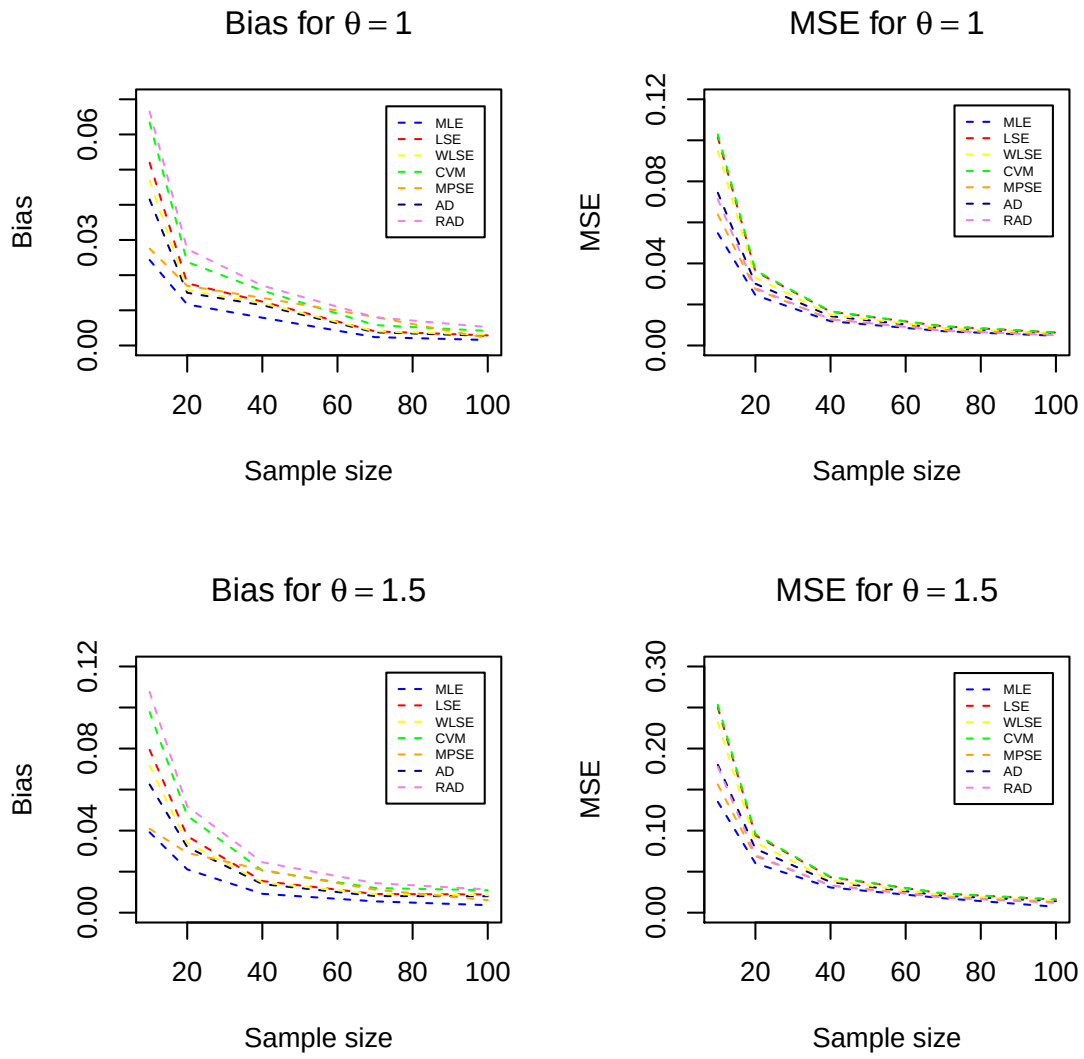


Figure 2: *Average bias and MSE of the estimator of xgamma distribution for different estimation procedures*

Table 2: *True value of θ and the average MSEs of the different estimation procedures for xgamma distribution*

θ	n	MLE	LSE	WLSE	CvME	MPSE	ADE	RADE
0.1	10	0.000439	0.000502	0.000483	0.000506	0.000381	0.000443	0.000414
	20	0.000199	0.000229	0.000217	0.000230	0.000184	0.000209	0.000197
	40	0.000091	0.000106	0.000100	0.000106	0.000088	0.000097	0.000093
	70	0.000052	0.000061	0.000057	0.000062	0.000050	0.000056	0.000053
	100	0.000036	0.000042	0.000039	0.000042	0.000036	0.000039	0.000037
0.5	10	0.013628	0.016936	0.015879	0.017242	0.011157	0.013783	0.012303
	20	0.005962	0.007405	0.006890	0.007501	0.005361	0.006442	0.005854
	40	0.002634	0.003267	0.003026	0.003294	0.002502	0.002930	0.002677
	70	0.001539	0.001940	0.001776	0.001949	0.001497	0.001739	0.001597
	100	0.001069	0.001308	0.001205	0.001312	0.001057	0.001189	0.001106
1	10	0.069448	0.100324	0.093373	0.101749	0.053900	0.072068	0.062812
	20	0.028016	0.034550	0.031993	0.035101	0.024756	0.029757	0.027279
	40	0.012726	0.015990	0.014680	0.016163	0.011771	0.014040	0.012747
	70	0.007001	0.008904	0.008123	0.008952	0.006748	0.007925	0.007184
	100	0.004781	0.006223	0.005606	0.006245	0.004748	0.005514	0.004999
1.5	10	0.178790	0.256766	0.239198	0.260282	0.136610	0.181424	0.156670
	20	0.073687	0.097266	0.089498	0.098933	0.063242	0.080475	0.071668
	40	0.032243	0.041780	0.037965	0.042180	0.029828	0.036235	0.032962
	70	0.017389	0.022625	0.020359	0.022757	0.016632	0.019824	0.018033
	100	0.012620	0.016351	0.014757	0.016432	0.007166	0.014528	0.013152

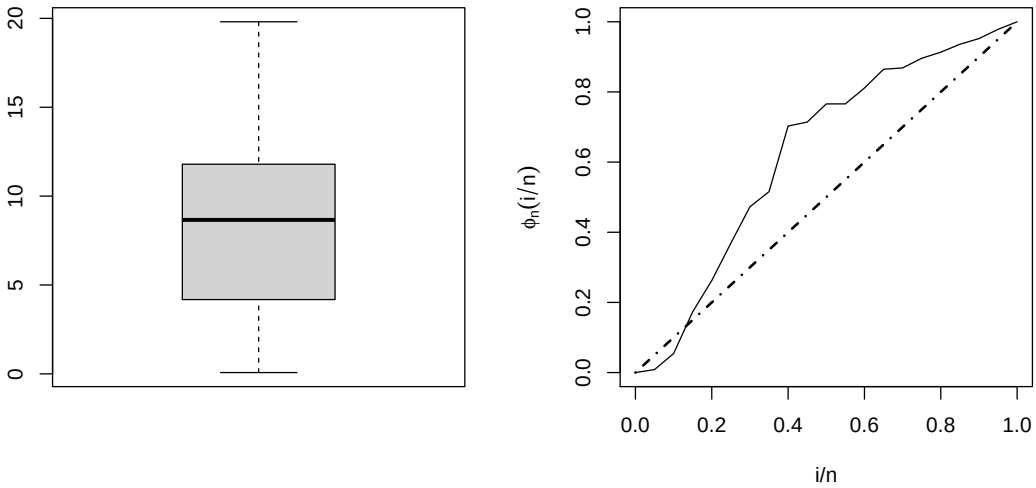


Figure 3: *Box plot and TTT plot for the failure time data*

Table 3: *Summarized results of fitting different distributions for failure time data set*

	xgamma	Akash	Exponential	Lindley	Shankar	Sujatha
ℓ	-60.0189	-61.6744	-62.6346	-61.3792	-62.2797	-61.8345
AIC	122.0378	125.3488	127.2693	124.7583	126.5595	125.6689
BIC	123.0336	126.3445	128.2650	125.7541	127.5552	126.6647
$CAIC$	122.5357	125.5710	127.4915	124.9805	126.7817	125.8912
$HQIC$	122.2322	125.5431	127.4636	124.9527	126.7538	125.8633

Table 4: *Estimated values with SEs for fitting different distributions for failure time data set*

	xgamma	Akash	Exponential	Lindley	Shankar	Sujatha
$\hat{\theta}_{MLE}$	0.3005	0.3427	0.1186	0.2162	0.2352	0.3293
(SE)	(0.0105)	(0.0254)	(0.0315)	(0.0202)	(0.0237)	(0.0247)
$\hat{\theta}_{LSE}$	0.2503	0.3150	0.0954	0.1930	0.2067	0.3026
(SE)	(0.0112)	(0.0235)	(0.0235)	(0.0164)	(0.0200)	(0.0225)
$\hat{\theta}_{WLSE}$	0.2610	0.3192	0.1041	0.2002	0.2133	0.3073
(SE)	(0.0124)	(0.0237)	(0.0265)	(0.0176)	(0.0208)	(0.0229)
$\hat{\theta}_{CvME}$	0.2512	0.3309	0.1003	0.2019	0.2179	0.3174
(SE)	(0.0113)	(0.0245)	(0.0252)	(0.0178)	(0.0214)	(0.0237)
$\hat{\theta}_{MPSE}$	0.3024	0.3171	0.0964	0.1945	0.2083	0.3047
(SE)	(0.0178)	(0.0236)	(0.0239)	(0.0167)	(0.0201)	(0.0227)
$\hat{\theta}_{ADE}$	0.2690	0.3206	0.1071	0.2134	0.2166	0.3208
(SE)	(0.0133)	(0.0238)	(0.0275)	(0.0197)	(0.0212)	(0.0240)
$\hat{\theta}_{RADE}$	0.2381	0.3268	0.1034	0.1969	0.2284	0.3285
(SE)	(0.0107)	(0.0243)	(0.0263)	(0.0170)	(0.0227)	(0.0246)

that the distribution is right-skewed. The TTT plot suggests an increasing failure rate; thus, the xgamma distribution could be appropriate for modeling the current data. Figure 4 shows the fitted probability distribution and empirical distribution function of the xgamma distribution based on different estimates of the parameter to the data set. Table 3 summarises the results of fitting different distributions. Based on the results listed in the table, we conclude that the xgamma distribution provides the best fit with the lowest values of model selection criteria. The xgamma model provides the closest fit to the data. In Table 4, we have presented different estimates of θ under various distributional assumptions to the data and their corresponding standard error (SE). It is also noticed that the SE is the least for the assumption of xgamma distribution, and the MLE and RADE are efficient estimates.

6. Concluding remarks

In this paper, different estimation procedures of the parameter of the xgamma distribution have been studied. Simulation studies are carried out for seven different initial values. As the sample size increases, the MSEs and biases of all estimators decrease and become close to each other. In a small sample situation, the MSEs of the ADE and RADE are smaller than the others. A real data application is conducted to show the appropriateness of xgamma distribution in practical data modelling. The xgamma distribution was compared with some known distributions and presented the estimates according to various parameter estimators. The xgamma distribution

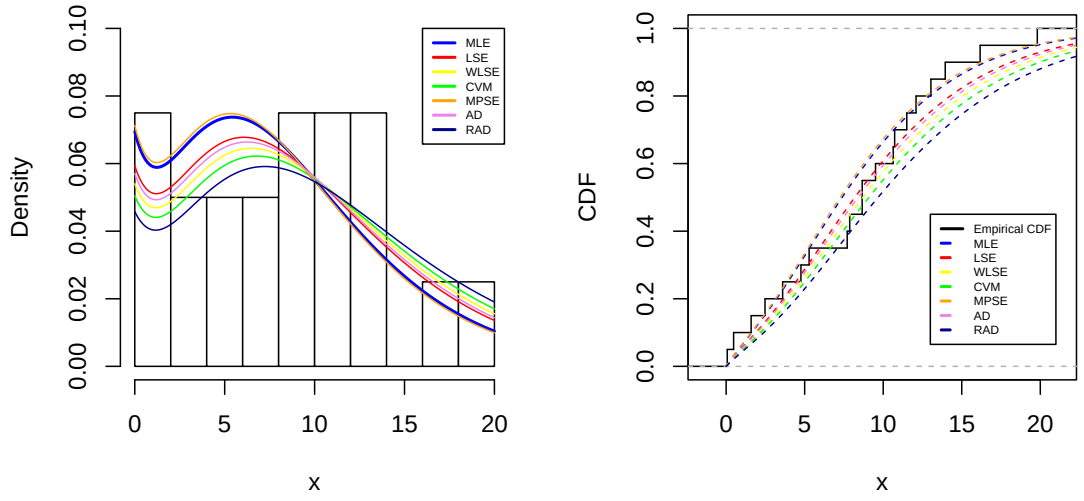


Figure 4: *The fitted pdfs and cdfs of xgamma distribution for different methods of estimation based on failure time data*

is the best-fitting model for some failure time data, and the ADE and RADE are preferable for estimating the parameter even though the MLE has computational ease.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

[1] Ahmad I. A., Kayid M. and Pellerey F. (2005). Further results involving the MIT order and the IMIT class. *Probability in the Engineering and Informational Sciences*, 19, 377-395.

[2] Anderson T. W. and Darling D. A. (1952). Asymptotic theory of certain goodness of fit criteria based on stochastic processes, *The Annals of Mathematical Statistics*, 3(2), 193-212.

[3] Anderson T. W. and Darling D. A. (1954). A test of goodness of fit, *Journal of the American Statistical Association*, 49(268), 765-769.

[4] Cheng R. C. H. and Amin N. A. K. (1979). Maximum product-of-spacings estimation with applications to the log-normal distribution. *Math Report*, 79.

[5] Cheng R. C. H. and Amin N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(3), 394-403.

[6] Di Crescenzo A. and Longobardi M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *Journal of Applied Probability*, 39, 434-440.

[7] Kayid M. and Ahmad I. A. (2004). On the mean inactivity time ordering with reliability applications. *Probability in the Engineering and Informational Sciences*, 18(03), 395-409.

[8] MacDonald P. D. M. (1971). Comment on an estimation procedure for mixtures of distributions by Choi and Bulgren. *Journal of Royal Statistical Society: Series B*, 33(2), 326-329.

[9] Murthy D. N. P., Xie M. and Jiang R. (2004). Weibull Models, *New York: Wiley*.

[10] Navarro J., Franco M. and Ruiz J. M. (1998). Characterization through moments of the residual life and conditional spacings. *Sankhya*, A 60, 36-48.

[11] Pettitt A. N. (1976). A two-sample Anderson-Darling rank statistic, *Biometrika*, 63(1), 161-168.

[12] Ranneby B. (1984). The maximum spacing method. an estimation method related to the maximum likelihood Method. Scandinavian *Journal of Statistics*, 11(2), 93-112.

[13] Renyi A. (1961). On Measures of Entropy and Information. In: Proceedings of the 4^{th} Berkeley Symposium on Mathematical. *Statistics and Probability, University of California Press, Berkeley*.
[14] Sen S., Maiti S. S. and Chandra N. (2016). The xgamma distribution: Statistical properties and application. *Journal of Applied Statistical Methods*, 15(1), 774-788.
[15] Stephens. M. A. (1974). EDF statistics for goodness of fit and some comparisons, *Journal of the American Statistical Association*, 69(347), 730-737.
[16] Swain J., Venkatraman S. and Wilson J. (1988). Least squares estimation of distribution function in Johnsons translation system. *Journal of Statistical Computation and Simulation*, 29, 271-297.

CONFIDENCE INTERVAL USING MAXIMUM LIKELIHOOD ESTIMATION FOR THE PARAMETERS OF POISSON TYPE LENGTH BIASED EXPONENTIAL CLASS MODEL

[1]Rajesh Singh, [2]Preeti A. Badge, [3]Pritee Singh

[1,2] Department of Statistics, S. G. B. Amravati University, Amravati, India
[3]Department of Statistics, Institute of Science, Nagpur, India.
rsinghamt@hotmail.com
preetibadge10@gmail.com
priteesingh25@gmail.com

Abstract

In this research paper, Confidence interval using Maximum likelihood estimation is obtained for Poisson type Length biased exponential class for the parameters. Failure intensity, mean time to failure and likelihood function for the parameter is obtained. Confidence interval has been derived for parameters using maximum likelihood estimation. To study the performance of confidence interval, average length and coverage probability are calculated using Monte Carlo simulation technique. From the obtained intervals it is concluded that Confidence interval for the parameter perform better for appropriate choice of execution time and certain values of parameters.

Keywords: Length biased exponential distribution, Software reliability growth model, Maximum likelihood estimation (MLE), Confidence interval using MLE, Average length and coverage probability.

I. Introduction

In this research paper Poisson type length biased exponential class model is considered according to Musa and Okumoto [9] classification scheme. Seenoi et al [12] proposed length biased exponentiated invented weibull distribution including some probability functions and moments of this distribution. Mir et al [6] introduced a length biased Beta distribution and also given a test for detection of length biasedness of beta distribution. The exponential exponentiated distribution proposed by Gupta and Kundu [4] which is special case of the exponentiated Weibull family. Mudholkar and Shrivastava [7] proposed the exponentiated weibull distribution as an extension of the weibull family obtained by adding the second shape parameter. Gupta and Keating [3] developed relationship between the survival function, the failure rate and mean residual life of exponential distribution and its length biased form.

Mudholkar et al [8] applied exponentiated weibull distribution to serve survival data and showed those hazard rates are increasing, decreasing bathtub shape and unimodal. Neppala et al [11] proposed Pareto type II based software reliability growth model with interval domain data using maximum likelihood estimation to estimate the parameter. Singh et al [13] proposed Bayes estimators for length biased distribution compared with ML estimators. Cunha and Rao [1] estimated credible interval and confidence interval through MLE for lognormal distribution also compared average length and coverage probability of the calculated interval. In the field of software reliability most of the work done on point estimation which give single guess value. Interval estimation with confidence interval gives more information than a point estimate. Confidence interval will be derived for both finite and infinite failure type models. Interval estimate indicates the error related to point estimate by the extent of its range and by probability of the true population parameter lying within that range. Thus, the purpose of this research paper is to study confidence interval using maximum likelihood estimation.

The association of the paper is such that section II presents length biased exponential model and derivation of failure intensity and expected number of failures using Length biased exponential distribution. Section III presents Likelihood function and derivation of maximum likelihood estimates of the parameters. Section IV contains derivation of confidence limits for the parameters θ_0 and θ_1 using maximum likelihood estimation. Results are discussed in the section V while concluding remarks are provided in section VI.

II. Model Formulation and Evaluation

Consider that software is tested for its performance and observed the time of failure occurs during software system performance. Let the number of failures present in software be θ_0,and t_e be the execution time i.e. time during which CPU is busy and m_e be the number of failures observed up to execution time t_e. Consider that time between the failures t_i (i=1,2,……m_e) follows the exponential distribution with parameter θ_1. The length biased exponential distribution is given as

$$f^*(t) = \begin{cases} t\theta_1^2 e^{-\theta_1 t} & , t > 0, \theta_1 > 0, E[t] \neq 0 \\ 0 & otherwise \end{cases} \quad (1)$$

Where $f^*(t)$ denotes the length biased exponential distribution.

The failure intensity function is obtained by using equation (1)

$$\lambda(t) = \theta_0 t\theta_1^2 e^{-\theta_1 t},\ t > 0,\ \theta_0 > 0 \quad (2)$$

Where, θ_0 express the number of failures and θ_1 express the for failure rate.

The mean failure function i.e. expected number of failures at time t_e can be obtained by using equations (2) and given by:

$$\mu(t_e) = \theta_0\theta_1^2 I_1 \quad (3)$$

Where, $I_1 = \int_0^{t_e} t_i e^{-\theta_1 t_i}\, dt$ and by solving (see Gradshteyn and Ryzhik [2]) we get,

$$\mu(t_e) = \theta_0\left[1-(1+\theta_1 t_e)e^{-\theta_1 t_e}\right] \ , t>0, \theta_0>0, \theta_1>0 \tag{4}$$

The study of behavior of failure intensity and expected number of failure of length biased exponential class model has been done by Singh et al [13]. They have compared the MLE's and Bayesian estimators on the basis of risk efficiencies.

III. Maximum Likelihood Estimation

Maximum likelihood estimation is most preferable because of its easy computation, greater efficiency and better numerical stability. It requires likelihood function for estimation. The likelihood function of parameters $\theta_0$and θ_1 is obtained with the help of failure intensity (2) and expected number of failures (4) (see for details Musa et al [10]) given by:

$$L(\theta_0,\theta_1) = \theta_0^{m_e}\theta_1^{2m_e}\left[\prod_{i=1}^{m_e} t_i\right]e^{-T\theta_1}e^{-\theta_0\left[1-(1+\theta_1 t_e)e^{-\theta_1 t_e}\right]} \tag{5}$$

Where, $\sum_{i=1}^{m_e} t_i = T$

After taking the logarithm of both sides of above equation and applying the procedure of obtaining the MLE's for parameters $\theta_0$andθ_1, the MLE's are

$$\hat{\theta}_{m0} = \left[\frac{m_e}{\left(1-(1+\hat{\theta}_{m1}t_e)e^{-\hat{\theta}_{m1}t_e}\right)}\right] \tag{6}$$

and

$$\hat{\theta}_{m1} = \left[\frac{(2m_e - T\theta_{m1})e^{\hat{\theta}_{m1}t_e}}{\hat{\theta}_{m0}t_e^2}\right]^{1/2} \tag{7}$$

respectively.

The values of $\hat{\theta}_{m0}$ and $\hat{\theta}_{m1}$ can be obtained by solving simultaneous equations (6) and (7) using any available standard numerical method viz. Bisection Method, Newton Rapson method. Singh et al [13] obtained maximum likelihood estimates for parameters of length biased exponential model. They compared maximum likelihood estimates and Bayes estimates on the basis of risk efficiencies and concluded that Bayes estimates preferred over maximum likelihood estimates.

IV. Confidence Interval using maximum likelihood estimation

Now to obtain confidence interval for both the parameter, it requires variance-covariance matrix for Σ all the MLE. Variance-covariance matrix is derived using Fisher information matrix. For asymptotic variance we can calculate Fisher information matrix which is negative second partial derivative of log likelihood function (see for details Kale [5]). Second derivative of log likelihood function can be given as follows:

$$\text{Var}\,(\hat{\theta}_0) = \frac{\hat{\theta}_0^2}{me} \tag{8}$$

$$\text{Var}\,(\hat{\theta}_1) = \left[(1/2\text{m}_\text{e} + \theta_0\theta_1^2 e^{-\hat{\theta}_1 t_e}t_e^2 - \theta_0\theta_1^3 e^{-\hat{\theta}_1 t_e}t_e^3)\right] \tag{9}$$

Using equation (8) confidence limits for parameter θ_0 is given by:

$$\hat{\theta}_{0L} = \hat{\theta}_0 + Z_{\alpha/2}(\hat{\theta}_0/(m_e)^{1/2}) \quad (10)$$

$$\hat{\theta}_{0U} = \hat{\theta}_0 - Z_{\alpha/2}(\hat{\theta}_0/(m_e)^{1/2}) \quad (11)$$

For parameter θ_1 confidence limits using equation (9) are given by:

$$\hat{\theta}_{1L} = \hat{\theta}_1 - Z_{\alpha/2}[(1/2m_e + \hat{\theta}_0\hat{\theta}_1^2 e^{-\hat{\theta}_1 t_e} t_e^2 - \hat{\theta}_0\hat{\theta}_1^3 e^{-\hat{\theta}_1 t_e} t_e^3)]^{1/2} \quad (12)$$

$$\hat{\theta}_{1U} = \hat{\theta}_1 + Z_{\alpha/2}[(1/2m_e + \hat{\theta}_0\hat{\theta}_1^2 e^{-\hat{\theta}_1 t_e} t_e^2 - \hat{\theta}_0\hat{\theta}_1^3 e^{-\hat{\theta}_1 t_e} t_e^3)]^{1/2} \quad (13)$$

By substituting tabulated values of Z_α, 95% confidence interval can be obtained.

V. Discussion and Results

Here, 95% confidence interval using maximum likelihood estimation is obtained for the parameters θ_0 i.e. total number of failures and θ_1. Confidence interval for the proposed length biased exponential model is calculated as defined equation (1). To study the performance of confidence interval, sample size m_e was generated up to execution time t_e and it was repeated 1000 times from the length biased exponential distribution for distinct values of θ_0 and θ_1.Using Monte Carlo simulation technique 95% confidence interval has been obtained. The values of average length and coverage probability have been obtained by assuming execution time $t_e(= 15,20,25,30,35)$, and parameters $\theta_0(= 24(2)32)$, and $\theta_1(= 0.02(0.02)0.1)$. Average length and coverage probability obtained for confidence interval has been summarized in the tables 1to 10

Table 1 to 5 represents the 95% confidence interval for the parameter θ_0. From the table, it is seen that as values of parameter θ_0 increases the calculated average length decreases and there is slight increase in calculated average length as values of parameter θ_1 increases. From table it can be seen that values of coverage probability decreases as θ_0 increases and as θ_1 increases coverage probability increases. Similarly, it can also be observed that as execution time t_e increases average length increases. As average length decreases it effects on coverage probability. Coverage probability decreases as average length decreases and it increases as execution time increases.

The table 6 to 10 represents the 95% confidence interval for the parameter θ_0. From the table, it is seen that the values of average length increases as θ_0 increases and average length increases slightly as θ_1 increases. From table it can be seen that values of coverage probability increases as θ_0 increases and as θ_1 increases coverage probability also increases. Similarly, it can also be observed that execution time t_e increases average length decreases and coverage probability also decreases as execution time increases.

Table 1: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 15*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.16894 (0.995)	56.38039 (0.995)	55.55027 (0.994)	54.65604 (0.994)	53.83833 (0.992)
0.04	57.222518 (0.995)	56.39313 (0.995)	55.63409 (0.995)	54.70139 (0.994)	54.12473 (0.993)
0.06	57.233046 (0.995)	56.42863 (0.994)	55.65794 (0.995)	54.90752 (0.994)	54.16633 (0.993)
0.08	57.233615 (0.995)	56.45414 (0.995)	55.67368 (0.995)	54.92776 (0.994)	54.23347 (0.993)
0.1	57.243652 (0.995)	56.50778 (0.995)	55.71389 (0.995)	54.93576 (0.994)	54.27069 (0.993)

*The values in the parenthesis are coverage probability.

Table 2: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 20.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.605773 (0.995)	57.18747 (0.995)	56.78515 (0.994)	56.38402 (0.994)	55.94543 (0.992)
0.04	57.612651 (0.995)	57.21467 (0.995)	56.84414 (0.994)	56.41711 (0.994)	56.02504 (0.993)
0.06	57.61924 (0.995)	57.23135 (0.995)	56.85961 (0.994)	56.44942 (0.9944)	56.07487 (0.993)
0.08	57.62179 (0.995)	57.2444 (0.995)	56.89678 (0.994)	56.46993 (0.994)	56.09547 (0.993)
0.1	57.63312 (0.995)	57.26667 (0.995)	56.90905 (0.995)	56.53445 (0.995)	56.18467 (0.994)

*The values in the parenthesis are coverage probability.

Table 3: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 25.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.687913 (0.995)	57.35914 (0.995)	57.05328 (0.995)	56.72293 (0.994)	56.39379 (0.993)
0.04	57.694764 (0.995)	57.38959 (0.995)	57.05003 (0.995)	56.73697 (0.994)	56.42135 (0.994)
0.06	57.70429 (0.996)	57.38601 (0.995)	57.07483 (0.995)	56.77586 (0.994)	56.47019 (0.994)
0.08	57.725421 (0.996)	57.38709 (0.995)	57.09107 (0.995)	56.77883 (0.994)	56.48683 (0.994)
0.1	57.728669 (0.996)	57.40079 (0.995)	57.11992 (0.995)	56.78598 (0.995)	56.50296 (0.994)

*The values in the parenthesis are coverage probability.

Table 4: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 30.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.757765 (0.996)	57.47327 (0.995)	57.18907 (0.995)	56.92065 (0.994)	56.70381 (0.994)
0.04	57.760038 (0.996)	57.48198 (0.995)	57.19682 (0.995)	56.95059 (0.994)	56.72307 (0.994)
0.06	57.760261 (0.996)	57.48631 (0.995)	57.21524 (0.995)	56.96878 (0.995)	56.73889 (0.994)
0.08	57.765491 (0.996)	57.49554 (0.995)	57.23952 (0.995)	57.01431 (0.995)	56.75868 (0.994)
0.1	57.765513 (0.997)	57.53795 (0.996)	57.27481 (0.995)	57.04687 (0.995)	57.02576 (0.995)

*The values in the parenthesis are coverage probability.

Table 5: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 35.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.784104 (0.997)	57.53507 (0.996)	57.33528 (0.995)	57.04729 (0.995)	56.85084 (0.994)
0.04	57.786394 (0.997)	57.54886 (0.996)	57.34892 (0.995)	57.06216 (0.995)	56.87181 (0.994)
0.06	57.793615 (0.997)	57.55129 (0.996)	57.35886 (0.995)	57.07193 (0.995)	56.87427 (0.994)
0.08	57.797047 (0.997)	57.57595 (0.996)	57.35927 (0.995)	57.08127 (0.995)	56.88975 (0.994)
0.1	57.798991 (0.998)	57.58077 (0.996)	57.37618 (0.995)	57.13901 (0.995)	57.00542 (0.995)

*The values in the parenthesis are coverage probability.

Table 6: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 15.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.023096 (0.996)	0.023270 (0.996)	0.023425 (0.996)	0.023611 (0.997)	0.023803 (0.997)
0.04	0.023100 (0.996)	0.023273 (0.996)	0.023426 (0.996)	0.023648 (0.997)	0.023842 (0.997)
0.06	0.023103 (0.996)	0.023279 (0.996)	0.023448 (0.996)	0.023651 (0.997)	0.023856 (0.997)
0.08	0.023103 (0.996)	0.023297 (0.996)	0.023461 (0.996)	0.023671 (0.997)	0.023865 (0.997)
0.1	0.023107 (0.996)	0.023298 (0.996)	0.023488 (0.996)	0.023691 (0.997)	0.023871 (0.998)

*The values in the parenthesis are coverage probability.

Table 7: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 20.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.019317 (0.995)	0.019425 (0.995)	0.019531 (0.995)	0.019644 (0.996)	0.019751 (0.996)
0.04	0.019321 (0.995)	0.019431 (0.995)	0.019535 (0.995)	0.019657 (0.996)	0.019779 (0.996)
0.06	0.019321 (0.995)	0.019435 (0.995)	0.019546 (0.995)	0.019664 (0.996)	0.019786 (0.996)
0.08	0.019323 (0.995)	0.019440 (0.995)	0.019551 (0.995)	0.019679 (0.996)	0.019802 (0.996)
0.1	0.019325 (0.995)	0.019448 (0.995)	0.019568 (0.995)	0.019685 (0.996)	0.019826 (0.996)

*The values in the parenthesis are coverage probability.

Table 8: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 25.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.016918 (0.994)	0.017003 (0.995)	0.017076 (0.995)	0.017164 (0.995)	0.017245 (0.995)
0.04	0.016919 (0.994)	0.017006 (0.995)	0.017083 (0.995)	0.017166 (0.995)	0.017262 (0.995)
0.06	0.016924 (0.994)	0.017008 (0.995)	0.017088 (0.995)	0.017167 (0.995)	0.017269 (0.995)
0.08	0.016927 (0.995)	0.017010 (0.995)	0.017094 (0.995)	0.017177 (0.995)	0.017276 (0.994)
0.1	0.016928 (0.995)	0.017013 (0.995)	0.017093 (0.995)	0.017181 (0.995)	0.017284 (0.996)

*The values in the parenthesis are coverage probability.

Table 9 : *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 30.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.015232 (0.994)	0.015288 (0.994)	0.015349 (0.994)	0.015427 (0.994)	0.015466 (0.994)
0.04	0.015235 (0.994)	0.015298 (0.994)	0.015363 (0.994)	0.015431 (0.994)	0.015473 (0.994)
0.06	0.015236 (0.994)	0.015301 (0.994)	0.015364 (0.994)	0.015434 (0.994)	0.015477 (0.994)
0.08	0.015237 (0.994)	0.015303 (0.994)	0.015368 (0.994)	0.015442 (0.994)	0.015481 (0.994)
0.1	0.015256 (0.994)	0.015328 (0.994)	0.015372 (0.994)	0.015467 (0.994)	0.015489 (0.994)

*The values in the parenthesis are coverage probability.

Table 10 : *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 35.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.013966 (0.993)	0.01402 (0.993)	0.014061 (0.993)	0.014124 (0.993)	0.014167 (0.994)
0.04	0.013967 (0.993)	0.014014 (0.993)	0.014064 (0.993)	0.014129 (0.993)	0.014171 (0.994)
0.06	0.013968 (0.993)	0.014023 (0.993)	0.014067 (0.993)	0.014136 (0.993)	0.014176 (0.994)
0.08	0.013969 (0.993)	0.014029 (0.993)	0.014076 (0.993)	0.014141 (0.993)	0.014184 (0.994)
0.1	0.013972 (0.993)	0.014033 (0.993)	0.014086 (0.993)	0.014153 (0.994)	0.014186 (0.994)

*The values in the parenthesis are coverage probability.

VI. Conclusion

Confidence interval through MLE is derived in this paper, the confidence interval suggested for the parameters of Poisson type length biased exponential class as SRGM. From the computation and above discussion it is concluded that proposed confidence interval maintained high coverage probability for different values of parameters for fixed execution time. Confidence interval for parameter θ_0 can be preferred for small execution time whereas Confidence interval for parameter θ_1 can be preferred for large execution time.

References:

[1] D'Cunha, J. G. and Rao, K. A. (2016). Frequentist Comparison of the Bayesian Credible and Maximum Likelihood Confidence interval for the Median of the Lognormal Distribution for the Censored Data.*International Journal of Scientific and Research Publications*, 6:61-66.

[2] Gradshteyn, I. S and Ryzhik, I. M. Table of Integrals, Series, and Products, Alan Jeffrey (editor) 5th Ed., Academic Press, New York, (1994).

[3] Gupta, R. C. and Keating, J. P. (1986).Relations for reliability measures under length biased sampling. *Scand Journal of Statistics*, 13: 49-56.

[4] Gupta, R. D. and Kundu, D. (2009). A new class of weighted exponential distribution. Statistics, 43:621 - 634.

[5] Kale, B. K., A First Course on Parametric Inference (Second Edition). Narosa, (1999).

[6] Mir, K. A., Ahmad, A. Reshi, J. A. (2013). Structural properties of length biased Beta distribution of first kind. *American Journal of Engineering Research*, 2:1-6.

[7] Mudholkar, G.S. and Shrivastava, D. K. (1993). ExponentiatedWeibull Family for Analyzing Bathtub Failure Rate Data.*IEEE Transanctions on Reliability*, 42:299-302.

[8] Mudholkar, G.S., Shrivastava, D. K., Kollia G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data.*Journal of the American Statistical Association*, 91:1575-1585.

[9] Musa, J.D. and Okumoto, K.(1984). A logarithmic Poisson execution time model for software reliability measurement. *Proc. 7th International Conference on Software Engineering*, Orlando, Florida, 230–238.

[10] Musa, J. D., Lannino, A. and Okumoto, K. Software Reliability: Measurement, Prediction, Application, New York, McGraw-Hill, (1987).

[11] Neppala, G. Satya Prasad, R. Kantam, R.R.L. (2011).Software Reliability Growth Model using Interval Domain Data. *International Journal of Computer Applications*, 34: 5-9.

[12] Seenoi, P. Supapakorn, T., Bodhiswan, W. (2014). The length-biased exponetiated inverted Weibull distribution. *International Journal of Pure and Applied Mathematics*, 92:191-206.

[13] Singh, R. Singh, P. and Kale, K. (2016). Bayes Estimators for the parameters of Poisson type length biased exponential class model using non- informative priors. *Journal of Reliability and Statistical Studies*, 9: 21-28.

A TWO NON IDENTICAL UNITS COLD STANDBY SYSTEM WITH CORRELATED FAILURE TIME AND REPAIR MACHINE FAILURE

Alka Chaudhary, Shivali Sharma and Anika Sharma
•
Principal, KLPG College, Meerut, 250002
Department of Statistics, Meerut College, Meerut, 250004
Ch. Charan singh university, Meerut, 250002
alkachaudhary527@gmail.com , sharma14shivali@gmail.com , ash27sharma@gmail.com

Abstract

The paper deals with a system composed of two-non identical units (unit-1 and unit-2). Initially, unit-1 is operative and unit-2 is kept in cold standby. The cold standby unit can't fail in its standby mode. Each unit of the system has two possible modes: Normal (N) and total failure (F). When the unit-1 fails the cold standby (unit-2) becomes operative instantaneously with the help of a perfect and instantaneous switching device. A single repairman is always available with the system to repair a failed unit and failed RM. Unit-1 gets priority in operation and repair over unit-2. However, the RM gets priority in repair over any of the units. The RM machine is good initially and can't fail unless it becomes operative. The system failure occurs when both the units are in total failure mode. The joint distribution of failure and repair times for each unit is taken bivariate exponential distribution. Each repaired unit works as good as new. Using regenerative point technique, various important measures of system effectiveness have been obtained.

Keywords: Transition probabilities, mean sojourn time, bi-vairate exponential distribution, reliability, MTSF, availability, expected busy period of repairman, net expected profit.

1. Introduction

Two units standby system models have been investigated by a large number of authors including A. Kumar, D. Pawar and S.C. Malik [11], P. Chaudhary, A. Sharma and R.Gupta [3], P. Chaudhary and A. Sharma [2], N.Kumar and N. Nandal [12], P. Gupta and P. Vinodiya [9], R. Gupta and P.Bhardwaj [5], R. Gupta and A. Tyagi [8], N. Kumar, S.C. Malik and N. Nandal [10], P. Chaudhary and S. Masih [1], P. Chaudhary and L. Tyagi[4] by using the concepts of warm standby with common cause failure and human error, correlated failure and repairs, two types of repairmen, two priority units warm standby with preparation for repair, two unit priority standby with repair, two unit cold standby with two operating modes.

In the analysis of above system models it has been assumed that a failed unit is always repairable manually and after repair the unit becomes as good as new. There are many situations where a repair machine (RM) is needed to repair a failed unit and the RM may also fail during the

repair of a failed unit .In this case the RM is first taken up for repair and the failed unit waits for getting repair.

For example, in case of nuclear reactors, marine equipment etc. the robots are used for the repair of such type of systems. It is evident that a robot being a Machine, may fail while performing its intended task. In this case the repairman will repair the RM first and then begins the repair of the failed unit.

Keeping above fact in view, the present chapter deals with the analysis of a two non-identical units cold standby system model with constant failure and general repair rates assuming that the first unit gets priority in operation and repair both. The RM may also fail during the repair of a unit. The failure rate of RM is taken as constant and its repair rate as general.

The objective of the present paper is to provide the analysis of a two non-identical unit standby system with correlated failure time and repair machine failure. The joint distributions of failure and repair times of each unit are taken to be bivariate exponential distribution with p.d.f. of the type-

$$f_i(x,y) = \alpha_i\beta_i(1-r_i)e^{-\alpha_i x-\beta_i y} I_0\left(2\sqrt{\alpha_i\beta_i r_i xy}\right)$$
$$; i = 1,2;\ x, y, \alpha_i, \beta_i > 0; 0 \le r_i < 1$$

$$\text{Where, } I_0 = \sum_{k=0}^{\infty} \frac{(z/2)^{2k}}{(k!)^2}$$

is the modified Bessel function of type-I and order zero. Gupta et al. [6] and Gupta and Shivakar [7] have analyzed some of two unit redundant system models by taking the joint distribution of failure and repair as bivariate exponential having the above form of pdf.

Using regenerative point technique, the following measures of system effectiveness are obtained-

i. Transition probabilities and mean sojourn times in various states.
ii. Reliability and Mean time to system failure (MTSF).
iii. Point-wise and steady-state availabilities of the system as well as expected up time of the system during time interval (0, t).
iv. Expected busy period of repairman in the repair of unit-1 and unit-2 during time interval (0, t).
v. Net expected profit earned by the system in time interval (0, t).

2. System description and assumptions

1. The system consists of two non-identical units (unit-1 and unit-2). Initially, unit-1 is operative and unit-2 is kept in cold standby. The cold standby unit can't fail in its standby mode.
2. Each unit of the system has two possible modes: Normal (N) and total failure (F).
3. When the unit-1 fails the cold standby (unit-2) becomes operative instantaneously with the help of a perfect and instantaneous switching device.
4. A single repairman is always available with the system to repair a failed unit and failed RM.
5. Unit-1 gets priority in operation and repair over unit-2. However, the RM gets priority in repair over any of the units.
6. The RM machine is good initially and can't fail unless it becomes operative.
7. The system failure occurs when both the units are in total failure mode.

8. The joint distribution of failure and repair times for each unit is taken bivariate exponential with density function given by ,

$$f_i(x,y) = \alpha_i\beta_i(1-r_i)e^{-\alpha_i x-\beta_i y} I_0\left(2\sqrt{\alpha_i\beta_i r_i x y}\right)$$

$$; i = 1,2;\ x, y, \alpha_i, \beta_i > 0; 0 \le r_i < 1$$

Where, $I_0 = \sum_{k=0}^{\infty} \frac{(z/2)^{2k}}{(k!)^2}$

9. Each repaired unit works as good as new.

3. Notations and states of the system

3.1. Notations:

E	:	Set of regenerative states.
X_i, Y_i	:	Random variable denoting the failure and repair time for unit-1and unit-2 respectively ;(i=1, 2)
$f_i(x,y)$	:	Joint probability density function of (X_i, Y_i) ;(i=1, 2) $= \alpha_i\beta_i(1-r_i)e^{-\alpha_i x-\beta_i y} I_0(2\sqrt{\alpha_i\beta_i r_i xy})dx$; $x, y, \alpha_i, \beta_i > 0;\ 0 \le r_i < 1$ Where, $I_0(2\sqrt{\alpha_i\beta_i r_i xy}) = \sum_{k=0}^{\infty} \frac{(\alpha_i\beta_i r_i xy)^k}{(k!)^2}$
$g_i(x)$	:	Marginal p.d.f. of $X_i = x$; (i=1, 2) $= \alpha_i(1-r_i)e^{-\alpha_i(1-r_i)x}$
$k_i(y\|X_i = x)$	:	Conditional p.d.f. of Y_i given $X_i = x$; (i=1, 2) $= \beta_i e^{-(\beta_i y+\alpha_i r_i x)} \sum_{j=0}^{\infty} \frac{(\alpha_i\beta_i r_i xy)^j}{(j!)^2}$
$g_{ij}(\bullet), g_{ij}^{(k)}(\bullet)$	:	P.d.f. of transition time from state S_i to S_j and S_i to S_j via S_k .
$p_{ij}(\bullet), p_{ij}^{(k)}(\bullet)$	:	Steady-state transition probabilities from state S_i to S_j and S_i to S_j via S_k .
$p_{ij\|x}(\bullet), p_{ij\|x}^{(k)}(\bullet)$	:	Steady-state transition probabilities from state S_i to S_j and S_i to S_j via S_k when it is known that the unit has worked for time x before its failure.
*	:	Symbol for Laplace Transform i.e. $g_{ij}^*(s) = \int e^{-st} q_{ij}(t)dt$
~	:	Symbol for Laplace Stieltjes Transform i.e. $\tilde{Q}_{ij}(s) = \int e^{-st} dQ_{ij}(t)t$
©	:	Symbol for ordinary convolution i.e. $A(t)©B(t) = \int_0^t A(u)B(t-u)du$
θ	:	waiting time of unit-1.

3.2. Symbols for the states of the system

N_{1o}, N_{2o}	:	Unit-1and Unit-2 is in N-mode and operative.
N_{2s}	:	Unit-2 is in N-mode and kept into cold standby.
F_{1r}, F_{2r}	:	Unit-1 and unit-2 is in F-mode and under repair.

F_{1w}, F_{2w}	:	Unit -1 and Unit -2 is in F-mode and waiting for repair.
RM_g	:	Repair machine is good.
RM_o	:	Repair machine is operative.
RM_r	:	Repair machine is failed and under repair.

Considering the above symbols in view of the assumptions stated earlier, we have the following states of the system:

Up States

$$S_0 \equiv \begin{pmatrix} N_{1o}, N_{2s} \\ RM_g \end{pmatrix}$$

$$S_1 \equiv \begin{pmatrix} F_{1r}, N_{2o} \\ RM_o \end{pmatrix}$$

$$S_2 \equiv \begin{pmatrix} F_{1w}, N_{2o} \\ RM_r \end{pmatrix}$$

$$S_5 \equiv \begin{pmatrix} N_{1o}, F_{2w} \\ RM_r \end{pmatrix}$$

$$S_6 \equiv \begin{pmatrix} N_{1o}, F_{2w} \\ RM_r \end{pmatrix}$$

Failed States

$$S_3 \equiv \begin{pmatrix} F_{1r}, F_{2w} \\ RM_o \end{pmatrix}$$

$$S_4 \equiv \begin{pmatrix} F_{1w}, F_{2w} \\ RM_r \end{pmatrix}$$

The transition diagram of the system model along with the transition rate or transition time c.d.f. is shown in Fig.1. The epochs of the transition into state S_4 from S_2 and S_6 from S_5 are non-regenerative.

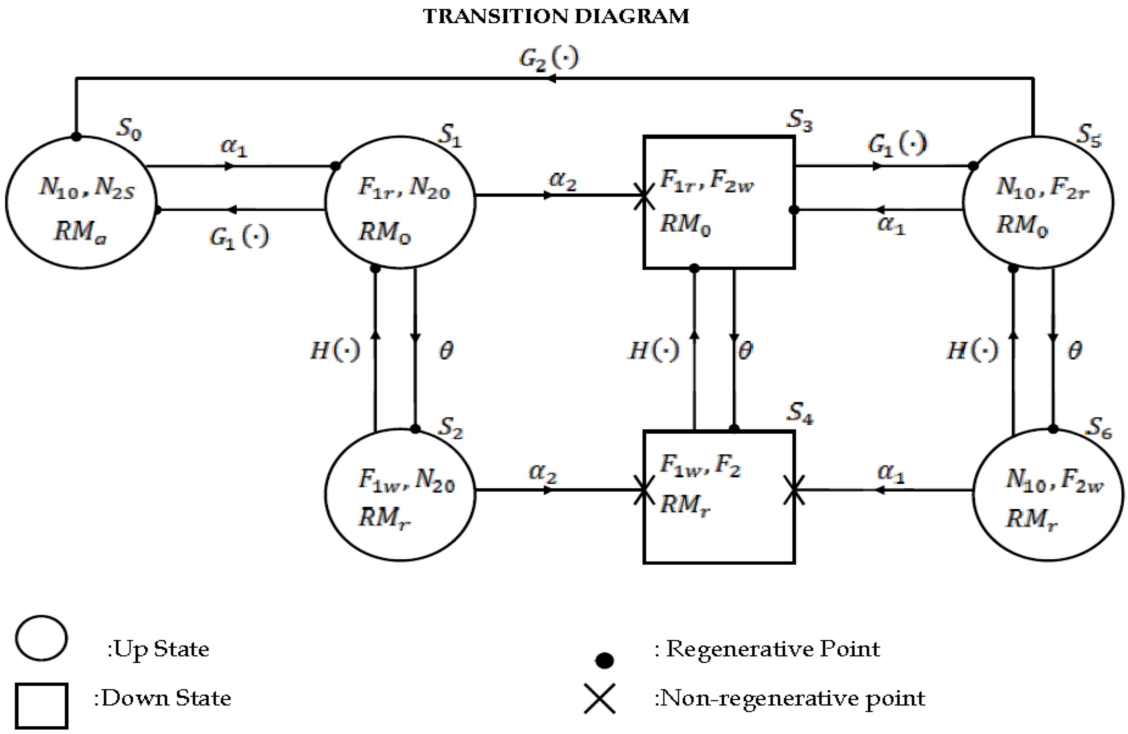


Figure 1: *Correlation Model*

[†]The limits of integration are 0 to ∞ whenever they are not mentioned.

4. Transition Probabilities and Mean Sojourn Times

Let X(t) be the state of the system at epoch t, then $\{X(t); t \geq 0\}$ constitutes a continuous parametric Markov-Chain with state space $E = \{S_0 \text{to} S_6\}$ .The various measures of system effectiveness are obtained in terms of steady-state transition probabilities and mean sojourn times in various states.

$$p_{01} = 1, \qquad p_{21} = \frac{1}{(\theta + \alpha_2(1-r_2))}$$

$$p_{43} = \int dG(t) = 1, \qquad p_{65} = \frac{1}{(\theta + \alpha_1(1-r_1))}$$

$$\begin{aligned} p_{10|x} &= \int \alpha_2(1-r_2) e^{-\alpha_2(1-r_2)t} e^{-\theta t} dK_1(t \mid x) \\ &= \widetilde{K}_1[\theta + \alpha_2(1-r_2) \mid x] \\ &= k_1^*[\theta + \alpha_2(1-r_2) \mid x], \end{aligned}$$

$$\begin{aligned} p_{12|x} &= \theta \int e^{-[\theta+\alpha_2(1-r_2)]t} d\overline{K_1}(t \mid x) \\ &= \theta \int e^{-[\theta+\alpha_2(1-r_2)]t} dt\left[1 - K_1(t \mid x)\right] \\ &= \left(\frac{\theta}{\theta + \alpha_2(1-r_2)}\right)\left(1 - k_1^*[\theta + \alpha_2(1-r_2) \mid x]\right) \end{aligned}$$

$$\begin{aligned} p_{13|x} &= \int \alpha_2(1-r_2) e^{-\alpha_2(1-r_2)t} e^{-\theta t} dt\overline{K_1}(t \mid x) \\ &= \alpha_2(1-r_2) \int e^{-[\theta+\alpha_2(1-r_2)]t} dt\left[1 - K_1(t \mid x)\right] \\ &= \left(\frac{\alpha_2(1-r_2)}{\theta + \alpha_2(1-r_2)}\right)\left(1 - k_1^*[\theta + \alpha_2(1-r_2) \mid x]\right) \end{aligned}$$

$$\begin{aligned} p_{34|x} &= \int \theta e^{-\theta t} dt\overline{K_1}(t \mid x) \\ &= \int \theta e^{-\theta t} dt\left[1 - K_1(t \mid x)\right] \\ &= 1 - k_1^*(\theta \mid x), \end{aligned}$$

$$\begin{aligned} p_{35|x} &= \int e^{-\theta t} dK_1(t \mid x) \\ &= \widetilde{K}_1[\theta \mid x] \\ &= k_1^*(\theta \mid x) \end{aligned}$$

$$\begin{aligned} p_{50|x} &= \int e^{-\theta t} dK_2(t \mid x) \\ &= \widetilde{K}_2\left[\theta + \alpha_1(1-r_1) \mid x\right] \\ &= k_2^*(\theta + \alpha_1(1-r_1) \mid x) \end{aligned}$$

$$\begin{aligned} p_{53|x} &= \int \alpha_1(1-r_1) e^{-\alpha_1(1-r_1)t} e^{-\theta t} dt\overline{K_2}(t \mid x) \\ &= \alpha_1(1-r_1) \int e^{-[\theta+\alpha_1(1-r_1)]t} dt\left[1 - K_2(t \mid x)\right] \\ &= \left(\frac{\alpha_1(1-r_1)}{\theta + \alpha_1(1-r_1)}\right)\left(1 - k_2^*[\theta + \alpha_1(1-r_1) \mid x]\right) \end{aligned}$$

$$\begin{aligned} p_{56|x} &= \theta \int e^{-[\theta+\alpha_1(1-r_1)]t} dt\left[1 - K_2(t \mid x)\right] \\ &= \left(\frac{\theta}{\theta + \alpha_1(1-r_1)}\right)\left(1 - k_2^*[\theta + \alpha_1(1-r_1) \mid x]\right) \end{aligned}$$

$$p_{23}^{(4)} = \int \left(1 - e^{-\alpha_2 (1-r_2)t}\right) dG(t)$$

$$= 1 - \frac{1}{\theta + \alpha_2 (1-r_2)},$$

$$p_{63}^{(4)} = \int \left(1 - e^{-\alpha_1 (1-r_1)t}\right) dG(t)$$

$$= 1 - \frac{1}{\theta + \alpha_1 (1-r_1)}$$

It can be easily verified that

$$p_{01} = p_{45|x} = 1, \qquad p_{12} + p_{13} = 1, \qquad p_{34} = 1$$

$$p_{20|x} + p_{25|x}^{(4)} = 1, \qquad p_{50|x} + p_{52|x}^{(6)} = 1 \qquad (1\text{-}5)$$

From the conditional steady state transition probabilities, the unconditional steady state transition probabilities can be obtained by using the result-

$$p_{ij} = \int p_{ij|x}\, g(x)\, dx$$

Thus,

$$p_{10} = \frac{\beta_1 (1-r_1)}{[\alpha_2 (1-r_2) + \beta_1 (1-r_1) + \theta]}$$

$$p_{12} = \frac{\theta}{\theta + \alpha_2 (1-r_2)} \left[1 - \frac{\beta_1 (1-r_1)}{[\alpha_2 (1-r_2) + \beta_1 (1-r_1) + \theta]}\right]$$

$$p_{13} = \frac{\alpha_2 (1-r_2)}{\theta + \alpha_2 (1-r_2)} \left[1 - \frac{\beta_1 (1-r_1)}{[\alpha_2 (1-r_2) + \beta_1 (1-r_1) + \theta]}\right]$$

$$p_{34} = \left[1 - \frac{\beta_1 (1-r_1)}{[\beta_1 (1-r_1) + \theta]}\right],$$

$$p_{35} = \frac{\beta_1 (1-r_1)}{[\beta_1 (1-r_1) + \theta]}$$

$$p_{50} = \frac{\beta_2 (1-r_2)}{\left[\theta + \beta_2 (1-r_2) + \alpha_1 (1-r_1)\right]}$$

$$p_{53} = \frac{\alpha_1 (1-r_1)}{\theta + \alpha_1 (1-r_1)} \left[1 - \frac{\beta_2 (1-r_2)}{[\alpha_1 (1-r_1) + \beta_2 (1-r_2) + \theta]}\right]$$

$$p_{56} = \frac{\theta}{\theta + \alpha_1 (1-r_1)} \left[1 - \frac{\beta_2 (1-r_2)}{[\alpha_1 (1-r_1) + \beta_2 (1-r_2) + \theta]}\right]$$

Thus, we have

$$p_{01} = p_{43} = 1, \qquad p_{10} + p_{12} + p_{13} = 1 \qquad p_{34} + p_{35} = 1$$

$$p_{50} + p_{53} + p_{56} = 1 \qquad p_{21} + p_{23}^{(4)} = 1, \qquad p_{65} + p_{63}^{(4)} = 1 \qquad (6\text{-}11)$$

5. Mean Sojourn Time

The mean sojourn time ψ_i in state S_i is defined as the expected time taken by the system in state S_i before transiting into any other state. If random variable U_i denotes the sojourn time in state S_i then,

$$\psi_i = \int P[U_i > t] dt$$

The mean sojourn times in various states are as follows-

$$\psi_0 = \frac{1}{\alpha_1(1-r_1)}, \qquad \psi_1 = \frac{1}{\theta+\alpha_2(1-r_2)+\beta_1(1-r_1)}$$

$$\psi_2 = \frac{1}{\theta+\alpha_2(1-r_2)}, \qquad \psi_3 = \frac{1}{\theta+\beta_1(1-r_1)}$$

$$\psi_4 = \int \overline{G}(t)\,dt = 1, \qquad \psi_5 = \frac{1}{\theta+\alpha_1(1-r_1)+\beta_2(1-r_2)}$$

$$\psi_6 = \frac{1}{\theta+\alpha_1(1-r_1)} \tag{12-18}$$

6. Analysis of Characteristics

6.1. Reliability and MTSF

Let $R_i(t)$ be the probability that the system operates during (0, t) given that at t=0 system starts from $S_i \in E$. To obtain it we assume the failed states S_3 and S_4 as absorbing. By simple probabilistic arguments, the value of $R_0(t)$ in terms of its Laplace Transform (L.T.) is given by

$$R_0^*(s) = \frac{\left(1-q_{56}^* q_{65}^*\right)\left[\left(1-q_{12}^* q_{21}^*\right)Z_0^* + q_{01}^* Z_1^* + q_{01}^* q_{12}^* Z_2^*\right]}{\left(1-q_{56}^* q_{65}^*\right)\left(1-q_{12}^* q_{21}^* - q_{01}^* q_{10}^*\right)} \tag{1}$$

We have omitted the argument's from $q_{ij}^*(s)$ and $Z_i^*(s)$ for brevity. $Z_i^*(s)$; i = 0, 1, 2, 5, 6 are the L. T. of

$$Z_0(t) = e^{-\alpha_1(1-r_1)t}$$

$$Z_1(t) = \int e^{-(\theta+\alpha_2(1-r_2))t}\,\overline{K}_1(t \mid x)\,dt, \qquad Z_2(t) = \int e^{-\alpha_2(1-r_2)t}\,\overline{G}(t)\,dt$$

$$Z_5(t) = e^{-(\theta+\alpha_1(1-r_1))t}, \qquad Z_6(t) = \int e^{-\alpha_1(1-r_1)t}\,\overline{G}(t)\,dt \tag{2-6}$$

Taking the Inverse Laplace Transform of (1), we can get the reliability of the system when system initially starts from state S_0 .

The MTSF is given by,

$$E(T_0) = \int R_0(t) = \lim_{s\to 0} R_0^*(s) = \frac{\left(1-p_{56}p_{65}\right)\left[\left(1-p_{12}p_{21}\right)\psi_0 + \psi_1 + p_{12}\psi_2\right]}{\left(1-p_{56}p_{65}\right)\left[1-p_{12}p_{21}-p_{10}\right]} \tag{7}$$

6.2. Availability Analysis

Let $A_i(t)$ be the probability that the system is up at epoch t, when initially it starts operation from state $S_i \in E$. Using the regenerative point technique and the tools of Laplace transform, one can obtain the value of $A_0(t)$ in terms of its Laplace transforms i.e. $A_0^*(s)$ given as follows-

$$A_0^*(s) = \frac{N_1(s)}{D_1(s)} \tag{8}$$

Where,

$$N_1(s) = \left(1-q_{25}^{(4)*} q_{52}^{(6)*}\right)\left[Z_0^* + q_{01}^* Z_1^*\right] + Z_2^* q_{01}^*\left[q_{12}^* + q_{13}^* q_{34}^* q_{45}^* q_{52}^{(6)*}\right] + Z_5^* q_{01}^*\left[q_{12}^* q_{25}^{(4)*} + q_{13}^* q_{34}^* q_{45}^*\right]$$

and

$$D_1(s) = 1 - q_{25}^{(4)*} q_{52}^{(6)*} - q_{01}^* q_{12}^* q_{20}^* - q_{01}^* q_{13}^* q_{20}^* q_{34}^* q_{45}^* q_{52}^{(6)*} - q_{01}^* q_{12}^* q_{50}^* q_{25}^{(4)*} - q_{01}^* q_{13}^* q_{34}^* q_{45}^* q_{50}^* \tag{9}$$

Where, $Z_i(t)$, i=0,1,2,5,6 are same as given in section 6.1.

The steady-state availability of the system is given by

$$A_0 = \lim_{t\to\infty} A_0(t) = \lim_{s\to 0} s\, A_0^*(s) \tag{10}$$

We observe that

$$D_1(0) = 0$$

Therefore, by using L. Hospital's rule the steady state availability is given by

$$A_0 = \lim_{s\to 0} \frac{N_1(s)}{D_1'(s)} = \frac{N_1}{D_1'} \tag{11}$$

Where,

$$N_1(0) = \left(1 - p_{25}^{(4)} p_{52}^{(6)}\right)\left[\psi_0 + \psi_1\right] + \psi_2\left[1 - p_{13}p_{50}\right] + \psi_5\left[1 - p_{12}p_{20}\right]$$

and

$$D_1' = \left(\psi_0 + \psi_1\right)\left(1 - p_{25}^{(4)} p_{52}^{(6)}\right) + \psi_3 p_{13}\left(1 - p_{25}^{(4)} p_{52}^{(6)}\right) + n_1\left(1 - p_{13}p_{20}p_{52}^{(6)}\right) + n_2\left(1 - p_{12}p_{20}\right) \tag{12}$$

The expected up time of the system in interval (0, t) is given by

$$\mu_{up}(t) = \int_0^t A_0(u)\, du$$

So that, $\mu_{up}^*(s) = \dfrac{A_0^*(s)}{s}$ (13)

6.3. Busy Period Analysis

Let $B_i^1(t)$ and $B_i^2(t)$ be the respective probabilities that the repairman is busy in the repair of unit-1 failed due to first repair with priority of unit-1 and unit-2 failed due to second repair at epoch t, when initially the system starts operation from state $S_i \in E$. Using the regenerative point technique and the tools of L. T., one can obtain the values of above two probabilities in terms of their L. T. i.e. $B_i^{1*}(s)$ and $B_i^{2*}(s)$ as follows-

$$B_i^{1*}(s) = \frac{N_2(s)}{D_1(s)}, \qquad B_i^{2*}(s) = \frac{N_3(s)}{D_1(s)} \tag{14-15}$$

Where,

$$N_2(s) = Z_1^* q_{01}^*\left[\left(1 - q_{56}^* q_{65}^*\right)\left(1 - q_{34}^* q_{43}^*\right) + q_{35}^* q_{43}^*\left(q_{56}^* q_{63}^{(4)*} + q_{53}^*\right)\right] + Z_3^* q_{01}^*\left(1 - q_{56}^* q_{65}^*\right)\left(q_{12}^* q_{23}^{(4)*} + q_{13}^*\right) \tag{16}$$

and

$$N_3(s) = -q_{01}^* q_{35}^*\left(q_{13}^* + q_{12}^* q_{23}^{(4)*}\right) Z_5^*$$

and $D_1(s)$ is same as defined by the expression (9) of section 6.2.

The steady state results for the above two probabilities are given by-

$$B_0^1 = \lim_{s\to 0} s\, B_0^{1*}(s) = N_2 \setminus D_1' \quad \text{and} \quad B_0^2 = \lim_{s\to 0} s\, B_0^{2*}(s) = N_3 \setminus D_1' \tag{17-18}$$

Where,

$$N_2(0) = \psi_1\left[\left(1 - p_{56}p_{65}\right)\left(1 - p_{34}\right) + p_{35}\left(p_{56}p_{63}^{(4)} + p_{53}\right)\right] + \psi_3\left(1 - p_{56}p_{65}\right)\left(p_{12}p_{23}^{(4)} + p_{13}\right) \tag{19}$$

$$N_4(0) = -p_{35}\left(p_{13} + p_{12}p_{23}^{(4)}\right)\psi_5 \tag{20}$$

and D_1' is same as given in the expression (12) of section 6.2.

The expected busy period in repair of unit-1 failed due to first repair with priority of unit-1 and unit-2 failed due to second repair during time interval (0, t) are respectively given by-

$$\mu_b^1(t) = \int_0^t B_0^1(u)\,du, \qquad \text{and} \qquad \mu_b^2(t) = \int_0^t B_0^2(u)\,du$$

So that,

$$\mu_b^{1*}(s) = \frac{B_0^{1*}(s)}{s} \qquad \text{and} \qquad \mu_b^{2*}(s) = \frac{B_0^{2*}(s)}{s} \tag{21-22}$$

7. Profit Function Analysis

The net expected total cost incurred in time interval (0, t) is given by

P (t) = Expected total revenue in (0, t) - Expected cost of repair in (0, t)

$$= K_0\mu_{up}(t) - K_1\mu_b^1(t) - K_2\mu_b^2(t) \tag{23}$$

Where, K_0 is the revenue per- unit up time by the system during its operation. K_1 and K_2 are the amounts paid to the repairman per-unit of time when the system is busy in repair of unit-1 failed due first repair with priority of unit-1 and unit-2 failed due to second repair respectively.

The expected total profit incurred in unit interval of time is $P = K_0A_0 - K_1B_0^1 - K_2B_0^2$

8. Particular Case

Let, $\overline{G}(t) = e^{-\lambda t}$

In view of above, the changed values of transition probabilities and mean sojourn times.

$$p_{21} = \frac{1}{\lambda + \alpha_2(1-r_2)}, \qquad p_{65} = \frac{1}{\lambda + \alpha_1(1-r_1)}$$

$$p_{23}^{(4)} = 1 - \frac{1}{\lambda + \alpha_2(1-r_2)}, \qquad p_{63}^{(4)} = 1 - \frac{1}{\lambda + \alpha_1(1-r_1)}$$

$$\psi_2 = \frac{1}{\lambda + \alpha_2(1-r_2)}, \qquad \psi_6 = \frac{1}{\lambda + \alpha_1(1-r_1)}$$

9. Graphical Study of Behaviour and Conclusions

For a more clear view of the behaviour of system characteristics with respect to the various parameters involved, we plot curves for MTSF and profit function in Fig. 2 and Fig. 3 w.r.t. $\alpha_{1=}$for three different values of correlation coefficient β_1=0.15, 0.45, 0.85 and two different values of repair parameter r_2 =0.8, 0.9 while the other parameters are $\theta = 0.9$,=$\beta_{2=}$=0.4, r_1 =0.7, $\alpha_{2=}$=0.8, λ=0.9. It is clearly observed from Fig. 2 that MTSF increases uniformly as the value of $\alpha_{1=}$and $r_2$increase and it decrease with the increase in $\alpha_{1=}$. Further, to achieve MTSF at least 80 units we conclude for smooth curves that the values of $\alpha_{1=}$must be less than 0.1, 0.13 and 0.4 respectively for β_1=0.15, 0.45, 0.85 when r_2 =0.8. Whereas from dotted curves we conclude that the values of $\alpha_{1=}$must be less than 0.1, 0.11 and 0.22 for β_1=0.15, 0.45, 0.85 when r_2 =0.9.

Similarly, Fig.3 reveals the variations in profit (P) with respect to $\boldsymbol{\alpha_1}$for three different values of β_1 =0.15, 0.45, 0.75 and two different values of r_2 =0.01, 0.03, when the values of other parameters $\theta = 0.8$,=$\beta_{2=}$=0.98, r_1 =0.3, $\alpha_{2=}$=0.6, λ=0.8, K_0=95, K_1=250 and K_2=225. Here also the same

trends in respect of α_1 , β_1 and r_2 are observed in case of MTSF. Moreover, we conclude from the smooth curves that the system is profitable only if α_1 is less than 0.1, 0.19 and 0.59 respectively for β_1 =0.15, 0.45, 0.75 when r_2 =0.01. From dotted curves, we conclude that the system is profitable only if α_1 is less than 0.1, 0.25 and 0.32 respectively for β_1 =0.15, 0.45, 0.75 when r_2 =0.03.

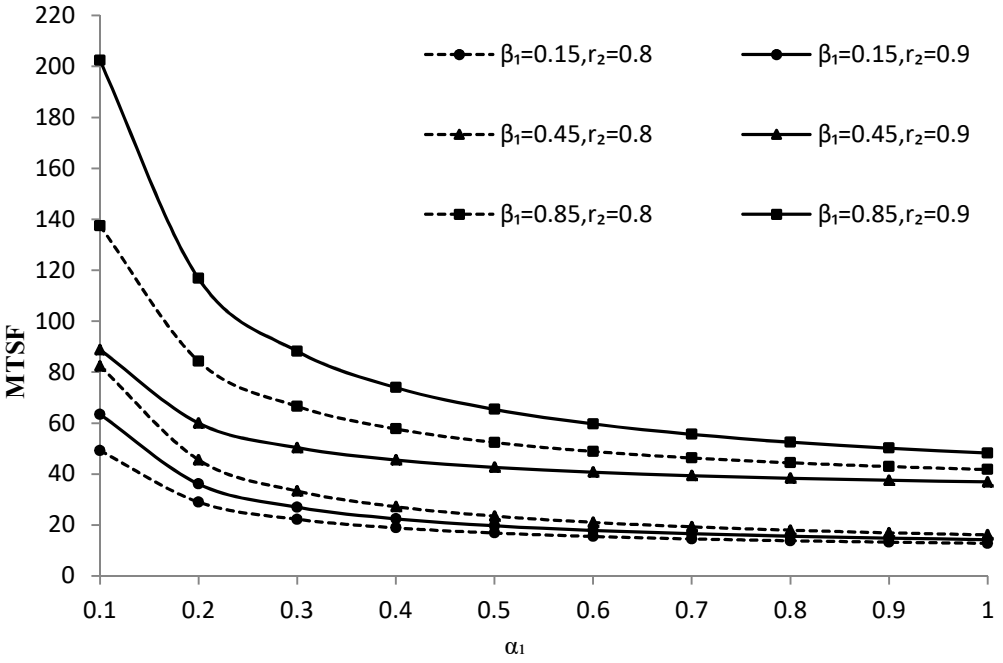


Figure 2: *Behaviour of MTSF with respect to α_1, β_1 and r_2*

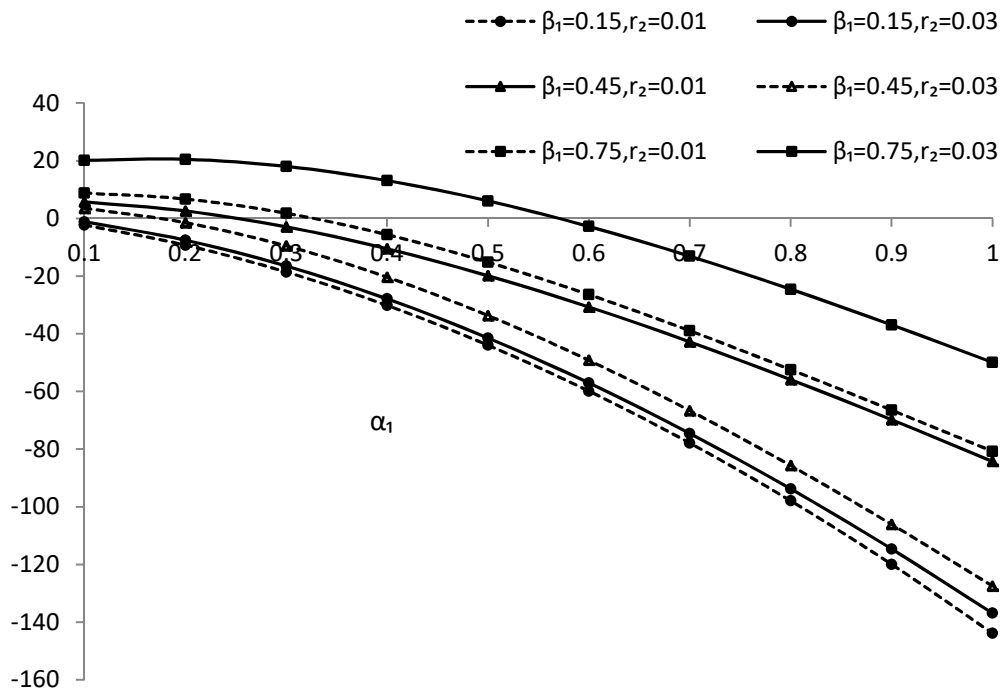


Figure 3: *Behaviour of PROFIT (P) with respect to α_2, β_1 and r_2*

References

[1] Chaudhary, P. and Masih, S. (2020).Analysis of a standby system with three modes of priority unit and correlated failure and repair times. *StatisticsandReliabilityEngineering,* 7(1):123-130.

[2] Chaudhary, P. and Sharma, A. (2021). A two dissimilar unit standby system with waiting time of repairman and correlated failure and repair times. *Statistics and Reliability Engineering,* 8(3):343-349.

[3] Chaudhary, P., Sharma, A. and Gupta, R. (2022). A Discrete Parametric Markov-Chain Model of a two non-identical units warm standby system with two types of failure. *Reliability: Theory& Applications,* 17(2(68)):21-30.

[4] Chaudhary, P. and Tyagi, L. (2020).A two non-identical unit standby system with minor and major repairs in priority unit and correlated failure and repair. *Agriculture and Statistical Science*, 16(1):243-250.

[5] Gupta, R. and Bhardwaj, P. (2019). A Discrete Parametric Markov-Chain Model of a two-unit cold standby system with appearance and disappearance of repairman. *Reliability: Theory & Applications,* 14(1(52)):13-22.

[6] Gupta, R. and kumar, K. (2008). A two unit complex system with correlated failure and repair times. *PureandappliedMathekaScience LXVII,*(1-2):23-34.

[7] Gupta, R. and Shivakar. (2010). Cost-Benefit analysis of a two unit parallel system with correlated failure and repair times. *IAPQRTransactions,*(35(2)):117-140.

[8] Gupta, R. and Tyagi, A. (2019). A Discrete Parametric Markov-Chain Model of a two-unit cold standby system with repair efficiency depending on environment. *Reliability: Theory & Applications,* 14(1(52)):23-33.

[9] Gupta, P. and Vinodiya, P. (2018). Analysis of reliability of a two non-identical units cold standby repairable system has two types of failure.*ComputerScienceandEngineering,* 8(3):343-349.

[10] Kumar, N., Malik, S.C. and Nandal, N. (2022). Stochastic analysis of a repairable system of non-identical units with priority and conditional failure of repairman. *Reliability: Theory & Applications,* 17(1(67)):123-133.

[11] Kumar, A., Panwar, D. and Malik, S.C. (2019). Profit analysis of a warm standby non-identical units system with single server subject to preventive maintenance. A*griculture and Statistical Science*, 15(1):261-269.

[12] Kumar, N. and Nandal, N. (2020). Stochastic modeling of a system of a two non-identical units with priority for operation and repair to main unit subject to conditional failure of repairman. *Statistics and Reliability Engineering*,7(1):114-122.

Decision Making Through Fuzzy Linear Programming Approach

Pandit U. Chopade
●
Research Supervisor, Department of Mathematics
D. S. M's Arts Commerce and Science College, Jintur
chopadepu@rediffmail.com
Mahesh M. Janolkar
●
Department of First Year Engineering
Prof. Ram Meghe College of Engineering & Management, Badnera-Amravati
maheshjanolkar@gmail.com
Kirankumar L. Bondar
●
P. G. Department of Mathematics, Govt Vidarbh Institute of Science and Humanities, Amravati
klbondar_75@rediffmail.com

Abstract

In this study a real world industrial MPS problem is addressed using the SMF approach. A decision maker, analyst and implementer, all play significant roles in making judgements in an uncertain environment, which is where this difficulty arises in the chocolate manufacturing business. As analysts our goal is to identify a solution with a higher LOS that will enable the decision maker to reach a conclusion. Because all the coefficients including the goals, technical and resource factors are well defined. The MPS problem is taken into consideration. With 24 constraints and 6 variables, this is regarded as one of the sufficiently large problem, which LOV is appropriate for getting satisfactory OR can be determined by a decision maker. To increase the satisfactory income, the decision maker can also advice to the analyst some feasible modification to FI. This collaborative process between the analyst, decision maker and implementer must continue until the best possible solution is found and put into action.

Keywords: Linear Programming problem, S-curve membership function, Uncertainty, Mix-Product selection, Decision maker.

Abbreviations

MPS	:	mix-product selection
FLP	:	fuzzy linear programming
MF	:	membership function
SMF	:	s-curve membership function
FO	:	fuzzy outcome
FS	:	fuzzy system
FI	:	fuzzy interval
UOP	:	units of product
LOS	:	level of satisfaction

LOV : level of vagueness
OR : optimal revenue
OF : objective function

1. Introduction

A non-linear MF, referred to as the SMF has been used in problems involving interactive FS. The modified SMF can be applied and tested for its suitability through an applied problem. In this problem, the SMF was applied to reach a decision, when all three coefficients such as OF, technical coefficients and resources of MPS were fuzzy. The solution thus obtained is suitable to be given to decision maker and implementer for final implementation. The problem illustrated in this paper is only one of six cases of MPS problems which occur in real life applications. It will be interesting to investigate the fuzzy solution patterns of these above MPS problem. Non-SMF conversion function is used for problems related to FLP. The function S can be applied and tested for its effectiveness by applied pressure. In this example, the S function is applied to make a decision after binary, such as the number of technologies and equipment, of which MPS is complex. Solutions thus obtained can provide the decision maker and the coordinator for the final implementation. The wording described in this article is just one of the three FPS words that actually have an application. The above FPS term is considered to be the real-life situation when it comes to making chocolate. Data for this problem are provided in the database of Choco man Inc, USA. Choco man manufactures chocolate bars, candies and wafer using a variety of ingredients and formulas. The goal is to use the modified S-function as a system to get the best UOP through the FLP system [1-3]. Compared with this FLP system. The recommended method is based on its relationship with the decision maker, developer and researcher to find satisfactory solutions for the FLP problem. In the decision-making process using the FLP model, modifications and source software can be complex, rather than providing exact numbers as in the net LP model. For example, machine hours, work, requirements, etc. and manufacturing, which is not always good, due to insufficient information and uncertainty among potential importers in the environment. Therefore, they should be considered as non-essential components and the FLP problem can be solved by using the FLP method. The problem of non-compliant MPS has been described. The aim of this article is to find the best UOP with high satisfaction and nonsense as the main thing. This problem is considered because all the parameters such as technology and hardware changes are uncertain. This is considered to be a major overall problem that includes 29 barriers and 8 barriers. Since there are so many decisions to make, the best UOP table is described for uncertainty and satisfaction to find a solution. with the highest UOP level and the highest satisfaction. It should be borne in mind that a high UOP does not mean it will lead to a high level of satisfaction. The best UOP was calculated at the satisfaction level using the FLP method. OF indicates that a high UOP will not lead to a high level of satisfaction. The results of this work suggest that the best decision is based on the negative impact on the FS of the MPS. In addition, high levels of UOP are obtained when blur is low in the system [4-25].

2. FLP Model

A general model of classical LP is formulated as,

$Max(w) = dy$; The standard formulation subject to

$$B \le c; y \ge 0 \tag{2.1}$$

Where d and y are the m-part vector, B is $n \times m$ matrix. Since we live in an uncertain

environment, the number of objective functions (d), the number of matrix technologies (B) and the variability of assets (d) are complex. Therefore, an infinite number can be displayed, so that the problem can be solved by the FLP system. FLP problems are designed as follows:

$$Max(w) = d^* y \text{ ; The fuzzy formulation subject to}$$
$$B^* y \le c^*; y \ge 0 \tag{2.2}$$

Where w is the vector of the decision change, $B^*, c^* \& d^*$ are zero numbers. The function of addition and multiplication is explained by fact that in-depth numbers are derived from the extension principles of Li [26]. Njikọ Inequalities are provided by some relationship and work objectives, w, must take into account the given LP problem. The approach of Mohammed [27] is being considered to solve the problem of FLP 2 depletion., which means that the solution will probably be to some satisfaction. First, design the team function for the zero parameter of $B^*, c^* \& d^*$. Here, non-existent team functions, such as logic, are used. Here vb_{kl} represents the work of members; $vc_k \& vd_l$ are the numerical functions for matrix B for $k = 1,2,3...n \& l = 1,2,3,...m$. c_k is the numerical variable for $k = 1,2,3...n$ and d_l are the integers of purpose point w for $l = 1,2,3,...m$.

Then, with the appropriate change in the concept of agreement between the non- $b^* k_l$ numbers; $c^* k$ and $d^* k_l$ & l, words for $b^* k_l$, $c^* k$ and $d^* l$ will be obtained. When an agreement between $b^* k_l$; The solution $c^* k$ and $d^* l$ will be [28];

$$V = v_{dl} = vb_{kl} = vc_k \text{ for all}$$
$$k = 1,2,3...n \& l = 1,2,3,...m \tag{2.3}$$

Therefore, we can obtain,

$$D = pd(v); B = pb(v) \& C = pc(v) \tag{2.4}$$

Where $v \in [0,1]$ in $pd, pb \& pc$ are distinct functions [29], of $vd, vB \& vc$ resp. Equation (2.2) would be;

$$Max(w) = [pd(v)]y \text{ ; fuzzy formulation subject to}$$
$$[pb(v)]y \le pc(v); y \ge 0 \tag{2.5}$$

First, create a group function for the complex part of $B^* \& c^*$. Here, non-uniform functions are used as S-curve function [30]. vb_{kl} represents the work of members; where b_{kl} is the coefficient of matrix B for $k = 1,2,3...,29$ and $l = 1,2,3,...8$, c_k is the material variable for $k = 1,2,3...,29$. Group function is also obtained for b_k and beard time, $c_{kb} \& c_{kc}$ for c_k^*. Similarly, we can create team work for a number of non-core technologies and their production [31]. Due to the high cost of production and the need to meet certain production and demand conditions, the problem of inefficiency arises in the manufacturing process. This problem also arises in the production of chocolate when deciding on the combination of ingredients to create different types of products. This is called here the choice of product mix [32].

3. The Fuzzy MPS

There are products that can be made by mixing different ingredients and using k type processing. It is expected that the infrastructure will be massive. There are also some restrictions by the retail department, such as the requirement for the product mix, the requirement of the main product line, as well as the minimum and maximum query for each product. Not everything that is needed in these circumstances is obvious. It is important to achieve maximum UOP and satisfaction using the FLP method. Since the number of technologies and equipment changes is running high, the results of the UOP would be foolish. FLP problem, customized in size. 2 can be written:

$Max(w) = \sum_{l=1}^{8} y_l$, subject to

$$\sum_{l=1}^{8} b_{kl}^{*} y_l \le c_k^{*}, \text{ where } y_l \ge 0, l = 1,2,...8. \tag{3.1}$$

where $b_{kl}^{*} \,\&\, c_k^{*}$ are fuzzy parameters.

3.1 Fuzzy Resource Variable vc_k

For an interval $c_k^b \le c_k \le c_k^c$,

$$b_{c_k} = \frac{c}{1 + De^{b\left(c_k - c_k^b / c_k^c - c_k^b\right)}}$$

$$\frac{\beta\left(c_k - c_k^b\right)}{e\left(c_c^k - c_k^b\right)} = \frac{1}{D}\left(\frac{c}{\theta_{c_k}} - 1\right) \tag{3.1.1}$$

$$\frac{\beta\left(c_k - c_k^b\right)}{\left(c_c^k - c_k^b\right)} = \ln\left(\frac{1}{D}\left(\frac{c}{\theta_{c_k}} - 1\right)\right)$$

$$c_k = c_k^a + \frac{1}{D}\left(\frac{\left(c_c^k - c_k^b\right)}{\beta}\right)\ln\left(\frac{1}{D}\left(\frac{c}{\theta_{c_k}} - 1\right)\right)$$

Since c_k is a non-trivial material change therefore, from (3.2)

$$cc_k^{*} = c_k^b + \left(\frac{\left(c_c^k - c_k^b\right)}{\beta}\right)\ln\left(\frac{1}{D}\left(\frac{c}{\theta_{c_k}} - 1\right)\right) \tag{3.1.2}$$

3.2 Fuzzy Constraints

The products, materials and equipment requirements are shown in Tables 1 as well as 2, respectively. Tables 3 as well as 4 provide the mix size and use the required material to make each product.

Table 1: *Product's Demand*.

Item	Fuzzy Interval $(\times 1000 units)$
Milk Chocolate, (200 gram)	[450-575) Gram
Milk Chocolate, (50 gram)	[750-950) Gram
Crunchy Chocolate, (200 gram)	[350-450) Gram
Crunchy Chocolate, (50 gram)	[550-700) Gram
Chocolate with Nuts (200 gram)	[250-325) Gram
Chocolate with Nuts (50 gram)	[450-575) Gram
Chocolate Candy	[150-200) Gram
Wafer	[350-450) Gram

Table 2: *Material and Ease of Access*

Raw Material	Fuzzy Interval (x1000 units)
Coco (Kilo Gram)	[75-125) Kilo Gram
Milk (Kilo Gram)	[90-150) Kilo Gram
Nuts (Kilo Gram)	[45-75) Kilo Gram
Sugar (Kilo Gram)	[150-450) Kilo Gram
Flour (Kilo Gram)	[15-25) Kilo Gram
Aluminum Foil (Kilo Gram)	[375-625) Kilo Gram
Paper (Per Feet Square)	[375-625) Per Feet Square
Plastic (Per Feet Square)	[375-625) Per Feet Square
Cooking (Ton per H)	[750-1250) Ton Per H
Mixing (Ton per H)	[150-250) Ton Per H
Forming (Ton per H)	[1125-1875) Ton Per H
Grinding (Ton per H)	[150-250) Ton Per H
Wafer Making (Ton per H)	[75-125) Ton Per H
Cutting (H)	[300-350) H
Packaging 1 (H)	[300-500) H
Packaging 2 (H)	[900-1500) H
Labor (H)	[750-1250) H

There are two unclear barriers such as access to the equipment and restrictions on the capacity of the equipment. These barriers are inevitable for any object and property depending on the consumption of the property, to trade and acquire property. These selections are based on the FLP resolution of Chocoman Inc. Decision changes for the FPSP are defined as:

$y_1 =$ 250 grams of chocolate milk to be produced (in 1000)

$y_2 =$ 100 grams of chocolate milk to be produced (per 1000)

$y_3 =$ Chocolate Crispy of 250 grams to be produced (in 1000)

$y_4 =$ 100 grams of Chocolate Crispy to be produced (in 1000)

$y_5 =$ Chocolate with 250 grams of fruit to produce (en1000)

$y_6 =$ Chocolate contains 100 grams per gram to produce (in 1000)

$y_7 =$ Chocolate candies will be produced (in 1000 packages)

$y_8 =$ Chocolate wafer production (in 1000 packages)

The Chocoman Marketing Department has issued the following restrictions:
Product mix required. Large product (250 grams) of any kind should not exceed 60% (uncertain value) of small product (100 grams)

$$y_1 \le 0.6 y_2 \quad (3.2.1)$$

$$y_3 \le 0.6 y_4 \quad (3.2.2)$$

$$y_5 \le 0.6 y_6 \quad (3.2.3)$$

The required product line is key. Total sales of confectionery products and wafers should not exceed 15% (uncertain value) of total confectionery product.

Table 3: *Mixing Proportions*

Materials Required Per 1000 Units	Product types (fuzzy interval)							
	AMC 150	AMC 50	ACC 150	ACC 50	ACN 150	ACN 50	Candy	Wafer
Coco (Kilo Gram)	[60-90)	[20-45)	[105-130)	[25-60)	[150-250)	[0-0)	[1200-1400)	[150-300)
Milk (Kilo Gram)	[0-0)	[0-0)	[60-90)	[0-0)	[78-101)	[35-80)	[230-500)	[0-0)
Nuts (Kilo Gram)	[325-456)	[78-105)	[230-280)	[34-87)	[0-0)	[0-0)	[110-230)	[73-130)
Sugar (Kilo Gram)	[172-201)	[0-0)	[78-99)	[0-0)	[321-436)	[103-120)	[0-0)	[54-90)
Flour (Kilo Gram)	[0-0)	[0-0)	[120-150)	[0-0)	[450-487)	[245-298)	[1001-1200)	[540-670)
Aluminum Foil (Kilo Gram)	[110-165)	[78-95)	[0-0)	[0-0)	[330-420)	[110-154)	[0-0)	[0-0)
Paper (Per Feet Square)	[156-185)	[0-0)	[190-245)	[0-0)	[100-150)	[56-89)	[0-0)	[0-0)
Plastic (Per Feet Square)	[0-0)	[0-0)	[170-240)	[40-82)	[510-725)	[120-179)	[0-0)	[0-0)

Table 4: *Facility Usage*

Facility Usage Required Per 1000 Units	Product types (fuzzy interval)							
	AMC 150	AMC 50	ACC 150	ACC 50	ACN 150	ACN 50	Candy	Wafer
Cooking (Ton per H)	[0.60-0.90)	[0.20-0.45)	[0.105-0.130)	[0.25-0.60)	[0.150-0.250)	[0-0)	[0.1200-0.1400)	[0.150-0.300)
Mixing (Ton per H)	[0-0)	[0-0)	[0.60-0.90)	[0-0)	[0.78-0.101)	[0.35-0.80)	[0.230-0.500)	[0-0)
Forming (Ton per H)	[0.325-0.456)	[0.78-0.105)	[0.230-0.280)	[0.34-0.87)	[0-0)	[0-0)	[0.110-0.230)	[0.73-0.130)
Grinding (Ton per H)	[0.172-0.201)	[0-0)	[0.78-0.99)	[0-0)	[0.321-0.436)	[0.103-0.120)	[0-0)	[0.54-0.90)
Wafer Making (Ton per H)	[0-0)	[0-0)	[0.120-0.150)	[0-0)	[0.450-0.487)	[0.245-0.298)	[0.1001-0.1200)	[0.540-0.670)
Cutting (H)	[0.110-0.165)	[0.78-0.95)	[0-0)	[0-0)	[0.330-0.420)	[0.110-0.154)	[0-0)	[0-0)
Packaging 1 (H)	[0.156-0.185)	[0-0)	[0.190-0.245)	[0-0)	[0.100-0.150)	[0.56-0.89)	[0-0)	[0-0)
Packaging 2 (H)	[0-0)	[0-0)	[0.170-0.240)	[0.40-0.82)	[0.510-725)	[0.120-0.179)	[0-0)	[0-0)
Labor (H)	[0.325-0.456)	[0.78-0.105)	[0.230-0.280)	[0.34-0.87)	[0-0)	[0-0)	[0.110-0.230)	[0.73-0.130)

Table 5: *OS with S-curve MF for* $\theta = 14.120$.

Number	Satisfaction degree ($\boldsymbol{\theta}$)	Optimal UOP (w*)	Number	Satisfaction degree ($\boldsymbol{\theta}$)	Optimal UOP (w*)
1	7.562	2438.54	11	50.0115	2965.11
2	14.076	2500.51	12	52.1911	3001.89
3	15.2145	2615.83	13	52.8741	3057.48
4	16.1148	2651.25	14	59.6383	3152.55
5	18.057	2701.67	15	63.3374	3160.55
6	24.8497	2845.48	16	63.538	3180.37
7	28.9782	2848.79	17	64.8241	3204.67
8	30.3968	2889.39	18	70.4424	3250.39
9	31.7572	2923.44	19	85.5813	3277.92
10	42.6513	2955.9	20	95.4286	3344.58

4. Results

The FPS problem is solved by using MATLAB and its LP application. It provides complexity and a degree of satisfaction. The LP application has two extras in addition to the non-existent. There is an output w *, the best UOP.

Table 6: *The Vagueness* β *as well as objective value* w^* *with* $\theta = 50\%$

Vagueness β	UOP w*	Vagueness β	UOP w*
1	2465.54	21	3037.45
3	2533.72	23	3080.78
5	2568.99	25	3223.61
7	2631.09	27	3239.79
9	2730.54	29	3282.03
11	2740.35	31	3352.45
13	2778.95	33	3368.74
15	2784.04	35	3438.1
17	2833.00	37	3446.69
19	3011.15		

Table 7: *Optimal UOP* w^*

w^*	Vagueness β				w^*	Vagueness β			
θ	1	3	5	7	θ	1	3	5	7
7.562	2421.27	2478.47	2594.46	2488.84	42.6513	2957.06	2847.5	3230.2	2810.63
14.076	2514.88	2502.54	2673.13	2509.44	50.0115	2960.57	3010.7	3234.95	2838.32
15.2145	2638.86	2623.91	2765.32	2574.27	52.1911	2981.24	3017.36	3248.8	2843.2
16.1148	2639.8	2632.57	2780.56	2604.7	52.8741	3078.7	3080.9	3297.06	3039.16
18.057	2668.82	2675.98	2797.33	2618.06	59.6383	3079.57	3086.95	3298.37	3157.71
24.8497	2686.3	2680.99	2919.95	2621.45	63.3374	3132.07	3162.39	3334.88	3206.49
28.9782	2753.94	2747.67	2930.67	2652.31	63.538	3273.09	3202.78	3415.55	3315.88
30.3968	2827.54	2773.03	3028.05	2723.29	64.8241	3443.79	3348.41	3426.19	3411.56

31.7572	2870.88	2807.2	3189.58	2753.75	70.4424	3479.39	3434.25	3470.15	3476.37

Different standards of Chocolate production are transferred to the toolbox. The answer can be listed in the following tables. From Table 5, it can be seen that a high level of satisfaction provides a high UOP. But the best solution to the above problem is at a satisfaction rate of 50%, or 2833 minutes. From the tables below, we conclude that within the objective, w^* is an ever-increasing function [33].

Table 8: *Optimal UOP* w^*

w^*	Vagueness β				w^*	Vagueness β			
θ	9	11	13	15	θ	9	11	13	15
7.562	2517.93	2511.75	2700.82	2626.7	42.6513	3006.57	3238.42	3211.28	3082.57
14.076	2555.17	2562	2817.03	2713.6	50.0115	3106.2	3252.29	3236.27	3155.49
15.2145	2610.27	2712.45	2818.6	2730.28	52.1911	3110.49	3312.54	3276.6	3166.6
16.1148	2694.71	2735.65	2917.06	2735.94	52.8741	3155.25	3326.07	3285.56	3215.15
18.057	2704.95	2778.61	3015.94	2814.01	59.6383	3206.75	3341.22	3292.6	3306.44
24.8497	2768.05	2785.92	3017.65	2843.42	63.3374	3367.82	3383.69	3312.35	3339.97
28.9782	2803.52	2982.47	3019.4	2857.43	63.538	3432.71	3393.02	3319.99	3353.86
30.3968	2912.9	3162.64	3200.54	2919.49	64.8241	3461.5	3394.43	3341.83	3462.87
31.7572	2959.22	3205.75	3210.48	2936.06	70.4424	3478.85	3435.72	3421.66	3493.17

Table 9: *Optimal UOP* w^*

w^*	Vagueness β				w^*	Vagueness β			
θ	17	19	21	23	θ	17	19	21	23
7.562	2560.71	2591.74	2598.75	2569.53	42.6513	3279.76	3093.95	3025.39	3012.8
14.076	2577.5	2681.47	2671.48	2712.04	50.0115	3289.08	3100.34	3089.09	3119.28
15.2145	2827.45	2695.28	2725.3	2774.99	52.1911	3329.94	3206.97	3105.94	3133.89
16.1148	2857.61	2745.12	2898.84	2857.97	52.8741	3339.61	3249.02	3118.94	3212.27
18.057	2877.99	2760.14	2919.28	2910.07	59.6383	3343.42	3287.02	3159.21	3267.98
24.8497	3081.74	2770.16	2962.64	2962.97	63.3374	3362.92	3361.71	3185.11	3331.74
28.9782	3093.67	2858.84	2989.96	2977.2	63.538	3373.1	3417.77	3275.53	3457.72
30.3968	3157.45	3063.62	3018.63	2983.99	64.8241	3440.06	3434.14	3397.49	3486.65
31.7572	3202.92	3087.9	3020.53	2988.83	70.4424	3492.01	3471.26	3495.27	3498.94

Table 10: *Optimal UOP* w^*

w^*	Vagueness β				w^*	Vagueness β			
θ	23	25	27	29	θ	23	25	27	29
7.562	2557.26	2509.77	2624.58	2522.45	42.6513	3110.12	2866.61	3012.12	3001.32
14.076	2639.95	2531.72	2637.73	2547.82	50.0115	3128.99	2880.25	3060.57	3044.8
15.2145	2727.12	2561.53	2645.54	2584.66	52.1911	3139.91	2957.15	3075.73	3135.83
16.1148	2785.23	2610.31	2745.36	2750.06	52.8741	3240.09	3012.5	3126.45	3297.11
18.057	2845.05	2680.12	2766.93	2756.62	59.6383	3259.24	3066.82	3170.93	3305.56
24.8497	2879.51	2758.1	2778.77	2762.94	63.3374	3263.83	3118.69	3292.42	3313.34
28.9782	2937.4	2800.6	2817.91	2832.69	63.538	3378.55	3132.87	3296.45	3384.03
30.3968	2967.17	2840.55	2893.03	2886.01	64.8241	3422.86	3324.07	3375.38	3404.9
31.7572	3057.98	2846.94	2961.62	2938.18	70.4424	3483.18	3350.47	3470.84	3428.67

Table 11: *Optimal UOP* w^*

w^*	Vagueness β				w^*	Vagueness β			
θ	31	33	35	37	θ	31	33	35	37
7.562	2522.48	2523.96	2533.43	2519.95	42.6513	3144.28	2901.63	3220.44	3041.08
14.076	2532.12	2608.62	2618.64	2611.46	50.0115	3183.95	2934.68	3236.11	3068.4
15.2145	2571.52	2618.64	2717.62	2615.81	52.1911	3202.9	3052.3	3264.69	3102
16.1148	2712.13	2739.13	2749.95	2652.37	52.8741	3213.79	3204.34	3330.91	3109.29
18.057	2916.79	2771.39	2778.74	2857.52	59.6383	3342.85	3264.08	3393.05	3214.24
24.8497	2943.77	2797.06	2979.54	2891.37	63.3374	3361.04	3270.6	3426.9	3242.07
28.9782	3088.17	2828.98	3023.91	2963.05	63.538	3403.39	3377.37	3432.62	3352.56
30.3968	3126.97	2886.21	3082.34	3010.27	64.8241	3406.28	3467.32	3455.09	3392.32
31.7572	3130.92	2887.8	3171.68	3020.85	70.4424	3435.75	3483.32	3461.04	3459.68

4.1 UOP of w^* for different vagueness values

Reasonable solutions and some uncertainties in the zero parameter of the technical rate and the hardware change are equal to 50%. Thus, the result of the 50% satisfaction level for $1 \le \beta \le 37$ and the principle corresponding to w^* are shown in Table 6. OF's of UOP reduce β imprecision and increase of the non-linear parameter of the number of technologies and asset exchange. This is clearly shown in Table 6. Table 6 is very important for the decision maker when choosing UOP, so that the result is at perfect level.

4.2 Output for θ, β & w^*

The result in the table below shows that when the inaccuracy of the increase results in a small UOP.

Table 12: w^* *with resp. to* β & θ

Satisfaction Degree (θ)	Vagueness (β)	Optimal UOP (w^*)	Satisfaction Degree (θ)	Vagueness (β)	Optimal UOP (w^*)
7.562	1	2500.51	50.0115	21	3001.89
14.076	3	2615.83	52.1911	23	3057.48
15.2145	5	2651.25	52.8741	25	3152.55
16.1148	7	2701.67	59.6383	27	3180.37
18.057	9	2845.48	63.3374	29	3204.67
24.8497	11	2848.79	63.538	31	3250.39
28.9782	13	2889.39	64.8241	33	3277.92
30.3968	15	2923.44	70.4424	35	3338.54
31.7572	17	2955.9	83.3374	37	3344.58
42.6513	19	2965.11			

It is also seen that SMF has a variety of standards that provide possible solutions with some satisfaction. Also, the link between w^* & θ is provided in Tables 7, 8, 9, 10 and 11. This is clearly shown in Table 6. Table 6 is very important for the decision maker when choosing UOP, so that the result is a perfect level. From Tables 7, 8, 9, 10 and 11, we find that for each type of satisfaction θ, the optimal UOP w^* decreases as the endpoint increases between 1 and 37. Similarly, with any positive value, the optimal UOP increases as the degree of satisfaction increases. Table 12 is the result of the diagonal pattern of w^* in Table 6. This result shows that, when the inaccuracies are low at $\beta = 1,3 \& 5$, UOP w * is best and reached the lowest satisfaction level, $\theta = 7.5\%, 14.1\% \& 15.2\%$. When the odds are high at $\beta = 33,35 \& 37$, UOP w^* is best reached with high satisfaction level, i.e., $\theta = 64.8\%, 70.4\% \& 83.3\%$.

5. Selection of Parameter β and Decision Making

In order for the decision maker to get the best results for the UOP w^*, the researcher creates a production table. From the table above, the decision maker can select the negative value according to his preference. Hair volume is divided into w^* in three parts, namely short, medium and high. It can be slightly modified if the input data for the number of technologies and hardware changes. It can be called a bunch of empty vanities. The decision can be made by the decision maker by choosing the best UOP for w^* and providing solutions for its implementation.

5.1 Discussion

The results show that the UOP minimum is 2,755.4 with a maximum of 3,034.9. It can be seen that when the understanding is between 0 and 1, the maximum value of w^* 3 034.9 is obtained by the minimum value. Similarly, when over 39, the minimum gain of w^* 2,755.4 and the maximum gain are obtained. Since the solution for MPS nonsense is the most satisfying solution with a high satisfaction degree, it is important to choose a blur between the minimum value and the maximum value of w^*.

6. Conclusion

The purpose of this research project was to find the most effective POU for MPS problems that have not been identified. SMF was recently developed as a framework for the task of solving the above problems effectively. The decision-making process and its implementation will be easier if the decision maker and consultant can work with the analyst to get the best and most satisfactory results. There are two more cases to consider in future work of the running technology that is not negative and that the dynamic assets are running and not complicated. FS mathematical relationships can be developed for MPS problems to find satisfying solutions. The decision maker, researcher and practitioner can apply their knowledge and experience to get the best results.

References:

[1] Azadeh A, Raoofi Z, Zarrin M. A multi-objective fuzzy linear programming model for optimization of natural gas supply chain through a greenhouse gas reduction approach. Journal of Natural Gas Science and Engineering. 2015;26:702-10.

[2] Chandrawat RK, Kumar R, Garg B, Dhiman G, Kumar S, editors. An analysis of modeling and optimization production cost through fuzzy linear programming problem with symmetric and right angle triangular fuzzy number. Proceedings of Sixth International Conference on Soft Computing for Problem Solving; 2017: Springer.

[3] Wan S-P, Wang F, Lin L-L, Dong J-Y. An intuitionistic fuzzy linear programming method for logistics outsourcing provider selection. Knowledge-Based Systems. 2015;82:80-94.

[4] Kumar D, Rahman Z, Chan FT. A fuzzy AHP and fuzzy multi-objective linear programming model for order allocation in a sustainable supply chain: A case study. International Journal of Computer Integrated Manufacturing. 2017;30(6):535-51.

[5] Rani D, Gulati T, Garg H. Multi-objective non-linear programming problem in intuitionistic fuzzy environment: Optimistic and pessimistic view point. Expert Systems with Applications. 2016;64:228-38.

[6] Govindan K, Sivakumar R. Green supplier selection and order allocation in a low-carbon paper industry: integrated multi-criteria heterogeneous decision-making and multi-objective linear programming approaches. Annals of operations research. 2016;238(1-2):243-76.

[7] Abdel-Basset M, Gunasekaran M, Mohamed M, Smarandache F. A novel method for solving the fully neutrosophic linear programming problems. Neural computing and applications. 2019;31(5):1595-605.

[8] Liao H, Jiang L, Xu Z, Xu J, Herrera F. A linear programming method for multiple criteria decision making with probabilistic linguistic information. Information Sciences. 2017;415:341-55.

[9] Yang X-P, Zhou X-G, Cao B-Y. Latticized linear programming subject to max-product fuzzy relation inequalities with application in wireless communication. Information Sciences. 2016;358:44-55.

[10] Edalatpanah S. A direct model for triangular neutrosophic linear programming. International journal of neutrosophic science. 2020;1(1):19-28.

[11] Rodger JA, George JA. Triple bottom line accounting for optimizing natural gas sustainability: A statistical linear programming fuzzy ILOWA optimized sustainment model approach to reducing supply chain global cybersecurity vulnerability through information and communications technology. Journal of cleaner production. 2017;142:1931-49.

[12] Talaei M, Moghaddam BF, Pishvaee MS, Bozorgi-Amiri A, Gholamnejad S. A robust fuzzy optimization model for carbon-efficient closed-loop supply chain network design problem: a numerical illustration in electronics industry. Journal of cleaner production. 2016;113:662-73.

[13] Alavidoost M, Babazadeh H, Sayyari S. An interactive fuzzy programming approach for

bi-objective straight and U-shaped assembly line balancing problem. Applied Soft Computing. 2016;40:221-35.

[14] Ebrahimnejad A. An improved approach for solving fuzzy transportation problem with triangular fuzzy numbers. Journal of intelligent & fuzzy systems. 2015;29(2):963-74.

[15] Garg H. A linear programming method based on an improved score function for interval-valued Pythagorean fuzzy numbers and its application to decision-making. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 2018;26(01):67-80.

[16] Mirzaee H, Naderi B, Pasandideh SHR. A preemptive fuzzy goal programming model for generalized supplier selection and order allocation with incremental discount. Computers & Industrial Engineering. 2018;122:292-302.

[17] Singh SK, Yadav SP. Efficient approach for solving type-1 intuitionistic fuzzy transportation problem. International Journal of System Assurance Engineering and Management. 2015;6(3):259-67.

[18] Darbari JD, Kannan D, Agarwal V, Jha P. Fuzzy criteria programming approach for optimising the TBL performance of closed loop supply chain network design problem. Annals of operations research. 2019;273(1-2):693-738.

[19] Xu Y, Xu A, Wang H. Hesitant fuzzy linguistic linear programming technique for multidimensional analysis of preference for multi-attribute group decision making. International Journal of Machine Learning and Cybernetics. 2016;7(5):845-55.

[20] Li M, Fu Q, Singh VP, Ma M, Liu X. An intuitionistic fuzzy multi-objective non-linear programming model for sustainable irrigation water allocation under the combination of dry and wet conditions. Journal of Hydrology. 2017;555:80-94.

[21] Paydar MM, Saidi-Mehrabad M. Revised multi-choice goal programming for integrated supply chain design and dynamic virtual cell formation with fuzzy parameters. International Journal of Computer Integrated Manufacturing. 2015;28(3):251-65.

[22] Tirkolaee EB, Goli A, Weber G-W, editors. Multi-objective aggregate production planning model considering overtime and outsourcing options under fuzzy seasonal demand. International Scientific-Technical Conference Manufacturing; 2019: Springer.

[23] Zaidan A, Atiya B, Bakar MA, Zaidan B. A new hybrid algorithm of simulated annealing and simplex downhill for solving multiple-objective aggregate production planning on fuzzy environment. Neural computing and applications. 2019;31(6):1823-34.

[24] Nematian J. An Extended Two-stage Stochastic Programming Approach for Water Resources Management under Uncertainty. Journal of Environmental Informatics. 2016;27(2).

[25] Gholamian N, Mahdavi I, Tavakkoli-Moghaddam R. Multi-objective multi-product multi-site aggregate production planning in a supply chain under uncertainty: fuzzy multi-objective optimisation. International Journal of Computer Integrated Manufacturing. 2016;29(2):149-65.

[26] Li D-F. Linear programming models and methods of matrix games with payoffs of triangular fuzzy numbers: Springer; 2015.

[27] Mohammed A, Harris I, Soroka A, Nujoom R. A hybrid MCDM-fuzzy multi-objective programming approach for a G-resilient supply chain network design. Computers & Industrial Engineering. 2019;127:297-312.

[28] Mahmoudi A, Liu S, Javed SA, Abbasi M. A novel method for solving linear programming with grey parameters. Journal of intelligent & fuzzy systems. 2019;36(1):161-72.

[29] Oliveira C, Coelho D, Antunes CH. Coupling input–output analysis with multiobjective linear programming models for the study of economy–energy–environment–social (E3S) trade-offs: a review. Annals of operations research. 2016;247(2):471-502,

[30] Garg H. Non-linear programming method for multi-criteria decision making problems under interval neutrosophic set environment. Applied Intelligence. 2018;48(8):2199-213, https://link.springer.com/article/10.1007/s10489-017-1070-5

[31] Chen S-M, Huang Z-C. Multiattribute decision making based on interval-valued intuitionistic fuzzy values and linear programming methodology. Information Sciences. 2017;381:341-51, doi: 10.1016/j.ins.2016.11.010

[32] Afzali A, Rafsanjani MK, Saeid AB. A fuzzy multi-objective linear programming model based on interval-valued intuitionistic fuzzy sets for supplier selection. International Journal of Fuzzy Systems. 2016;18(5):864-74, doi.org/10.1007/s40815-016-0201-1

[33] Subulan K, Taşan AS, Baykasoğlu A. A fuzzy goal programming model to strategic planning problem of a lead/acid battery closed-loop supply chain. Journal of Manufacturing Systems. 2015;37:243-64, doi.org/10.1016/j.jmsy.2014.09.001

DENSITY BY MODULI AND LACUNARY STATISTICAL CONVERGENCE OF DOUBLE SEQUENCES

[1] A. G. K. Ali, [2]A. M. Brono and [3]A. Masha

[1,3]Department of Mathematics and Computer Science, Borno State University Maiduguri, Nigeria
[2]Department of Mathematical Sciences, University of Maiduguri, Borno State, Nigeria
Email: akachallaali@gmail.com, bronoahmadu@unimaid.edu.ng,
mashafantami@gmail.com

Abstract

In this paper, we introduced and studied the concept of lacunary statistical convergence of double sequence with respect to modulus function where the modulus function is an unbounded double sequence. We also introduced the concept of lacunary strong convergence of double sequence via modulus function. We further characterized those lacunary convergence of double sequence for which the lacunary statistically convergent of double sequence with respect to modulus function equals statistically convergent of double sequence with respect to modulus function. Finally, we established some inclusion relations between these two lacunary methods and proved some essential analogue for double sequence.

Keywords: modulus function, statistical convergence, lacunary strong convergence, lacunary statistical convergence, double sequence.

1. Introduction

The concept of statistical convergence was formally introduced by [1] and [2] independently. Although statistical convergence was introduced over fifty years ago, it has become an active area of research in recent years. It has been applied in various areas such as summability theory [3] and [4], topological groups [5] and [6], topological spaces [7], locally convex spaces [8], measure theory [9], [10 and [11], Fuzzy Mathematics [12] and [13]. In recent years generalization of statistical convergence has appeared in the study of strong summability and the structure of ideals of bounded functions, [14]. Extension of the notion of statistical convergence of single sequence to double sequences by proposed by [15]. The concept of lacunary statistical convergence of single sequence was introduced by [16]. The extension of the concept of lacunary statistical of single sequence to double sequences was proposed by [17]. The notion of modulus function was introduced by [18]. Following [19] and [20], we recall that a function $f: [0,\infty) \to [0,\infty)$ is said to be a modulus function if it satisfies the following properties

(1) $f(x) = 0$ if and only if $x = 0$

(2) $f(x+y) \leq f(x) + f(y)$ for $x \geq 0, y \geq 0$,

(3) f is increasing,

(4) f is continuous from the right at 0.

It follows that f is continuous on $[0,\infty)$. The modulus function may be bounded or unbounded. For example, if we take $f(x)=\frac{x}{x+1}$, then $f(x)$ is bounded. But, $0< p < 1$, $f(x)=x^p$ is not bounded.

The definition of a new concept of density with help of an unbounded modulus function was proposed by [21], as a consequence, they obtained a new concept of non-matrix convergence, namely, f-statistical convergence, which is intermediate between the ordinary convergence and statistical and agrees with the statistical convergence when the modulus function is the identity mapping.

Quite recently, [22] and [23] have introduced and studied the concepts of f-statistical convergence of order α and f-statistical boundedness, respectively, by using approach of [21]. Quite recently, [24] introduced and studied the concept of f-lacunary statistical convergence and the concept of strong lacunary statistical convergence with respect to modulus function. We further extended and introduced some analogues results of double in line with that of [24].

Definition 1.1: ([15]): A real double sequence $x=\left(x_{jk}\right)$ is statistically convergent to a number l if for each $\varepsilon>0$, the set

$$\left\{(j,k), j\le n \text{ and } k\le m: \left|x_{jk}-l\right|\ge\varepsilon\right\} \tag{1}$$

has double natural density zero. In this case we write $st_2-\lim_{jk} x_{jk}=l$ and we denote the set of all statistically convergent double sequences by st_2.

Definition 1.2 ([17]). The double sequence $\theta_{r,s}=(j_r,k_s)$ is called double lacunary if there exist two increasing sequences of integers such that $j_0=0, h_r=j_r-j_{r-1}\to\infty$ as $r\to\infty$ and $k_0=0, h_s=k_s-k_{s-1}\to\infty$ as $s\to\infty$. Let $j_{r,s}=j_r k_s, h_{r,s}=h_r\overline{h_s}$ and $\theta_{r,s}$ is determined by $I_{r,s}=\{(j,k): j_{r-1}<j\le j_r \text{ and } k_{s-1}<k\le k_s\}, q_r=\frac{j_r}{j_{r-1}}$, $\overline{q_s}=\frac{k_s}{k_{s-1}}$ and $q_{r,s}=q_r\overline{q_s}$.

Definition 1.3 ([17]): Let $\theta_{r,s}$ be a double lacunary sequence, the double number sequence x is double lacunary statistical convergent to L provided that for every $\varepsilon>0$,

$$\lim_{r,s}\frac{1}{h_{r,s}}\left|\left\{(j,k)\in I_{r,s}: \left|x_{j,k}-L\right|\ge\varepsilon\right\}\right|=0. \tag{2}$$

Throughout this paper s, L_2^{∞} and c will denote the spaces of all, bounded and convergent double sequences of real numbers, respectively.

Now in this paper we introduce the concept of $f_{j,k}$-lacunary statistical convergence of double sequence, where $f_{j,k}$ is an unbounded modulus functions of double sequence.

Definition 1.4: Let $f_{j,k}$ be an unbounded modulus functions of double sequence. Let $\theta_{r,s}=(j_r,k_s)$ be double lacunary sequence. A double sequence $x=(x_{jk})$ is said to be $f_{j,k}$-lacunary statistically convergent of double sequence to L or $S_{\theta_{r,s}}^{f_{j,k}}$- convergent to L, if, for each $\varepsilon>0$,

$$\lim_{r,s\to\infty}\frac{1}{f_{j,k}(h_{r,s})}f_{j,k}\left(\left|\left\{(j,k)\in I_{r,s}: \left|x_{jk}-L\right|\ge\varepsilon\right\}\right|\right)=0. \tag{3}$$

In this case we write

$$S_{\theta_{r,s}}^{f_{j,k}}-\lim x_{jk}=L \text{ or } x_{jk}\to L\left(S_{\theta_{r,s}}^{f_{j,k}}\right).$$

For a given double lacunary sequence $\theta_{r,s}=(j_r,k_s)$ and unbounded modulus function $f_{j,k}$, by $S_{\theta_{r,s}}^{f_{j,k}}$ we denote the set of all $f_{j,k}$-lacunary statistically convergent of double sequences.

2. Methods

2.1 $f_{j,k}$-Lacunary Statistical Convergence of Double Sequence

We begin by establishing elementary connections between convergence of double sequence, $f_{j,k}$-lacunary statistical convergence of double sequence and double lacunary statistical convergence.

Theorem 2.1: Every convergent double sequence is $f_{j,k}$-lacunary statistically convergent of sequence, that is $c \subset S_{\theta_{r,s}}^{f_{j,k}}$ for any unbounded modulus functions f of double sequence and double lacunary statistical convergence sequence $\theta_{r,s}$.

Proof: Let $x = (x_{jk})$ be any convergent double sequence. Then, for each $\varepsilon > 0$, the set

$$\{(j,k) \in \mathbb{N} \times \mathbb{N}: |x_{jk} - L| \geq \varepsilon\} \text{ is finite. Suppose } \{(j,k) \in \mathbb{N} \times \mathbb{N}: |x_{jk} - L| \geq \varepsilon\} = g_0.$$

Now, since $\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\} \subset \{(j,k) \in \mathbb{N} \times \mathbb{N}: |x_{jk} - L| \geq \varepsilon\}$ and $f_{j,k}$ is modulus increasing, therefore

$$\frac{f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right)}{f_{j,k}(h_{r,s})} \leq \frac{f(g_0)}{f_{j,k}(h_{r,s})}.$$

Taking limit as $r, s \to \infty$, on both sides, we get

$$\lim_{r,s\to\infty} \frac{f_{j,k}\left((j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\right)}{f_{j,k}(h_{r,s})} = 0,$$

as $f_{j,k}(h_{r,s}) \to \infty$ as $r, s \to \infty$.

Theorem 2.2: Every $f_{j,k}$-lacunary statistical convergent double sequence is double lacunary statistical convergent.

Proof: Suppose $x = (x_{jk})$ is $f_{j,k}$-lacunary statistically convergent double sequence to L. Then by the definition of limit and the fact that $f_{j,k}$ being modulus is subadditive, for every $p \in \mathbb{N}$, there exist $r_0, s_0 \in \mathbb{N}$ such that, for $r, s \geq r_0, s_0$, we have

$$f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right) \leq \frac{1}{p} f_{j,k}(h_{r,s}) \leq \frac{1}{p} f_{j,k}\left(\frac{h_{r,s}}{p}\right) = f_{j,k}\left(\frac{h_{r,s}}{p}\right)$$

Since $f_{j,k}$ is increasing, we have

$$\frac{1}{h_{r,s}}\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right| \leq \frac{1}{p}.$$

Hence, x= (x_{jk}) is a double lacunary statistically convergent to L.

Remark 2.1: It seems that the inclusion $S_{\theta_{r,s}}^{f_{j,k}} \subset S_{\theta_{r,s}}$ is strict. But right now we are not in a position to give an example of a double sequence which is $S_{\theta_{r,s}}$-convergent but not $S_{\theta_{r,s}}^{f_{j,k}}$-convergent. So it is left as an open problem.

Remark 2.2: From theorem 2.1 and 2.2, we can say that the concept of $f_{j,k}$-lacunary statistical convergence is intermediate between the usual notion of convergence of double sequence and the double lacunary statistical convergence of double sequences.

We now establish a relationship between $f_{j,k}$-lacunary statistical convergence of double sequences and double lacunary strong convergence with respect to modulus functions $f_{j,k}$ of double sequence.

Theorem 2.3 Let $\theta_{r,s}=(j_r,k_s)$ be a double lacunary sequence, then consider the following:

(a) For any unbounded modulus functions f for which $\lim_{t\to\infty}\frac{f(t)}{t}>0$ and there is a positive constant c such that $f(xy)\geq cf(x)f(y)$, for all $x\geq 0, y>0$,

(i) $x_{jk}\to L\left(N_{\theta_{r,s}}^{f_{j,k}}\right)$ implies $x_{jk}\to L\left(S_{\theta_{r,s}}^{f_{j,k}}\right)$,

(ii) $N_{\theta_{r,s}}^{f_{j,k}}$ is a proper subset of $S_{\theta_{r,s}}^{f_{j,k}}$.

(b) $x\in L_\infty^2$ and $x_{jk}\to L\left(N_{\theta_{r,s}}^{f_{j,k}}\right)$ imply $x_{jk}\to L\left(S_{\theta_{r,s}}^{f_{j,k}}\right)$, for any unbounded modulus functions $f_{j,k}$ of double sequence.

(c) $N_{\theta_{r,s}}^{f_{j,k}}\cap L_\infty^2=S_{\theta_{r,s}}^{f_{j,k}}\cap L_\infty^2$ for any unbounded modulus function $f_{j,k}$ of double sequence for which $\lim_{t\to\infty}\frac{f(t)}{t}>0$ and there is a positive constant c such that $f(xy)\geq cf(x)f(y)$, for all $x\geq 0, y\geq 0$.

Proof: (a) (i) For any double sequence $x=(x_{jk})$ and $\varepsilon>0$, by the definition of a modulus function (1) and (3) we have

$$\frac{1}{h_{r,s}}\sum_{j,k\in I_{r,s}}\sum f_{j,k}(|x_{jk}-L|)\geq\frac{1}{h_{r,s}}f_{j,k}\left(\sum_{j,k\in I_{r,s}}\sum|x_{jk}-L|\right)\geq\frac{1}{h_{r,s}}f_{j,k}\left(\sum_{\substack{j,k\in I_{r,s}\\ |x_{jk}-L|\geq\varepsilon}}\sum|x_{jk}-L|\right)$$

$$\geq\frac{1}{h_{r,s}}f_{j,k}(|\{(j,k)\in I_{r,s}:|x_{jk}-L|\geq\varepsilon\}|\varepsilon)\geq\frac{c}{h_{r,s}}f_{j,k}(|\{(j,k)\in I_{r,s}:|x_{jk}-L|\geq\varepsilon\}|)f(\varepsilon)$$

$$=\frac{c}{h_{r,s}}\frac{f_{j,k}(|\{(j,k)\in I_{r,s}:|x_{jk}-L|\geq\varepsilon\}|)}{f_{j,k}(h_{r,s})}f_{j,k}(h_{r,s})f(\varepsilon)$$

From where it follows that $x\in S_{\theta_{r,s}}^{f_{j,k}}$ as $x\in N_{\theta_{r,s}}^{f_{j,k}}$ and $\lim_{r,s\to\infty}\left({f_{j,k}(h_{r,s})}/{h_{r,s}}\right)>0$.

(ii) To show the strictness of inclusion, consider the double sequence $x=(x_{jk})$ such that x_{jk} is to be 1,2, ...,$\lfloor\sqrt{h_{r,s}}\rfloor$ at the first $\lfloor\sqrt{h_{r,s}}\rfloor$ integers in $I_{r,s}$, and x_{jk}=0 otherwise. Note that (x_{jk}) is not bounded. Also, for every $\varepsilon>0$,

$$\frac{1}{f_{j,k}(h_{r,s})}f_{j,k}(|\{(j,k)\in I_{r,s}:|x_{jk}-L|\geq 0\}|)=\frac{f_{j,k}(\sqrt{h_{r,s}})}{f_{j,k}(h_{r,s})}=\frac{f_{j,k}(\sqrt{h_{r,s}})}{\sqrt{(h_{r,s})}}\times\frac{h_{r,s}}{f_{j,k}(h_{r,s})}\times\frac{\lfloor\sqrt{h_{r,s}}\rfloor}{h_{r,s}}\to\infty\text{ as }r\to\infty,$$

Because $\lim_{r,s\to\infty}\left({f_{j,k}(\lfloor\sqrt{h_{r,s}}\rfloor)}/{(\lfloor\sqrt{h_{r,s}}\rfloor)}\right)$, $\lim_{r,s\to\infty}\left({f_{j,k}(h_{r,s})}/{h_{r,s}}\right)$ are positive and $\lim_{r,s\to\infty}\left({f_{j,k}(\lfloor\sqrt{h_{r,s}}\rfloor)}/{(\lfloor\sqrt{h_{r,s}}\rfloor)}\right)=0$.

Thus, $x_{jk}\to 0\left(S_{\theta_{r,s}}^{f_{j,k}}\right)$. On other hand,

$$\frac{1}{h_{r,s}}\sum_{j,k\in I_{r,s}}\sum f_{j,k}(|x_{jk}-L|)=\frac{f_{j,k}(1)+f_{j,k}(2)+\cdots+f_{j,k}(\lfloor\sqrt{h_{r,s}}\rfloor)}{h_{r,s}}\geq\frac{f_{j,k}(1+2+\cdots+\lfloor\sqrt{h_{r,s}}\rfloor}{h_{r,s}}$$

$$=\frac{f_{j,k}\left(\lfloor\sqrt{h_{r,s}}\rfloor\left({(\lfloor\sqrt{h_{r,s}}\rfloor+1)}/{2}\right)\right)}{h_{r,s}}\geq c\frac{f_{j,k}(\lfloor\sqrt{h_{r,s}}\rfloor)f_{j,k}\left({(\lfloor\sqrt{h_{r,s}}\rfloor+1)}/{2}\right)}{h_{r,s}}$$

$$=c\times\frac{f_{j,k}(\lfloor\sqrt{h_{r,s}}\rfloor)}{\lfloor\sqrt{h_{r,s}}\rfloor}\times\frac{f_{j,k}\left({(\lfloor\sqrt{h_{r,s}}\rfloor+1)}/{2}\right)}{{(\lfloor\sqrt{h_{r,s}}\rfloor+1)}/{2}}\times\frac{\lfloor\sqrt{h_{r,s}}\rfloor\left({(\lfloor\sqrt{h_{r,s}}\rfloor+1)}/{2}\right)}{h_{r,s}}>0$$

As c, $\lim_{r,s\to\infty}\left(f_{j,k}\left(\left[\sqrt{h_{r,s}}\right]\right)/\left[\sqrt{h_{r,s}}\right]\right)$, $\lim_{r,s\to\infty}\left(f_{j,k}\left(\left[\sqrt{h_{r,s}}\right]+1\right)/2\right)/\left(\left(\left[\sqrt{h_{r,s}}\right]+1\right)/2\right)$, and

$\lim_{r,s\to\infty}\left(\left[\sqrt{h_{r,s}}\right]\right)\left(\left(\left[\sqrt{h_{r,s}}\right]+1\right)/2\right)/h_{r,s}\right)$ are positive. Hence $x_{jk}\nrightarrow 0\left(N_{\theta_{r,s}}^{f_{j,k}}\right)$.

(b) Suppose that $x_{jk}\to L\left(S_{\theta_{r,s}}^{f_{j,k}}\right)$ and $x\in L_\infty^2$, say $\left|x_{jk}-L\right|\le H$ for all $j,k\in\mathbb{N}$. Given $\varepsilon>0$, we have

$$\frac{1}{h_{r,s}}\sum_{j,k\in I_{r,s}}\sum f_{j,k}\left(\left|x_{jk}-L\right|\right)=\frac{1}{h_{r,s}}\sum_{\substack{j,k\in I_{r,s}\\ |x_{jk}-L|\ge\varepsilon}}\sum f_{j,k}\left(\left|x_{jk}-L\right|\right)+\sum_{\substack{j,k\in I_{r,s}\\ |x_{jk}-L|<\varepsilon}}\sum f_{j,k}\left(\left|x_{jk}-L\right|\right)\le\frac{1}{h_{r,s}}\left|\{(j,k)\in I_{r,s}:\left|x_{jk}-L\right|\ge\varepsilon\}\right|f_{j,k}(H)+\frac{1}{h_{r,s}}h_{r,s}f_{j,k}(\varepsilon).$$

Taking limit on both sides as $r,s\to\infty$, we get $\lim_{r,s\to\infty}\left(\frac{1}{h_{r,s}}\right)\sum_{j,k\in I_{r,s}}\sum f_{j,k}\left(\left|x_{jk}-L\right|\right)=0$, in view of theorem 2.2 and the fact that $f_{j,k}$ is increasing.

(c) This is an immediate consequence of (a) and (b)

Remark 2.3 The example given in part (a) of the above theorem shows that the boundedness condition cannot be omitted from the hypothesis of part (b).

3. Results

3.1 $f_{j,k}$-Lacunary Statistical Convergence of Double Sequence Versus $f_{j,k}$-Statistical Convergence of Double Sequence

In this section we study the inclusion $S_{\theta_{r,s}}^{f_{j,k}}\subset S^{f_{j,k}}$ and $S^{f_{j,k}}\subset S_{\theta_{r,s}}^{f_{j,k}}$ under certain restrictions on $\theta_{r,s}$ and $f_{j,k}$.

Lemma 3.1.1: For any double lacunary sequence $\theta_{r,s}$ and unbounded modulus function $f_{j,k}$ for which $\lim_{t\to\infty}(f(t)/t)>0$ and there is a positive constant c such that $f(xy)\ge cf(x)f(y)$, for all $x\ge 0$, $y\ge 0$, one has $S^{f_{j,k}}\subset S_{\theta_{r,s}}^{f_{j,k}}$ if and only if $\liminf_{r,s} q_{r,s}>1$.

Proof: Sufficiency: If $\liminf_{r,s} q_{r,s}>1$ then there exists $\delta>0$ such that $q_{r,s}\ge 1+\delta$ for sufficiently large r,s. Since $h_{r,s}=k_{r,s}-k_{r-1,s-1}$, we have

$$\frac{h_{r,s}}{k_{r,s}}\ge\left(\frac{\delta}{1+\delta}\right)^2$$

For sufficiently large r,s. If $x_{jk}\to L\left(S^{f_{j,k}}\right)$, then, for given $\varepsilon>0$ and sufficiently large r,s we have

$$\frac{1}{f_{j,k}(j_rk_s)}f_{j,k}\left(\left|\{j\le j_r\ and\ k\le k_s:\left|x_{jk}-L\right|\ge\varepsilon\}\right|\right)\ge\frac{f_{j,k}\left(\left|\{(j,k)\in I_{r,s}:\left|x_{jk}-L\right|\ge\varepsilon\}\right|\right)}{f_{j,k}(j_rk_s)}=\frac{f_{j,k}(h_{r,s})}{f_{j,k}(j_rk_s)}\times\frac{f_{j,k}\left(\left|\{(j,k)\in I_{r,s}:\left|x_{jk}-L\right|\ge\varepsilon\}\right|\right)}{f_{j,k}(h_{r,s})}=\left(\frac{f_{j,k}(h_{r,s})}{h_{r,s}}\right)\cdot\left(\frac{j_rk_s}{f_{j,k}(j_rk_s)}\right)\left(\frac{h_{r,s}}{j_rk_s}\right)\frac{f_{j,k}\left(\left|\{(j,k)\in I_{r,s}:\left|x_{jk}-L\right|\ge\varepsilon\}\right|\right)}{f_{j,k}(j_rk_s)}\ge\left(\frac{f_{j,k}(h_{r,s})}{h_{r,s}}\right)\left(\frac{j_rk_s}{f_{j,k}(j_rk_s)}\right)\left(\frac{\delta}{1+\delta}\right)^2\cdot\frac{f_{j,k}\left(\left|\{(j,k)\in I_{r,s}:\left|x_{jk}-L\right|\ge\varepsilon\}\right|\right)}{f_{j,k}(h_{r,s})}.$$

This proves the sufficiency.

Necessity: Assume that $\liminf_{r,s} q_{r,s}=1$. We can select a subsequence $\left(j_{r(i)}k_{s(j)}\right)$ of $\theta_{r,s}$ satisfying

$$\frac{j_{r(i)}k_{s(j)}}{j_{r(i)-1}k_{s(j)-1}}<1+\frac{1}{ij},\frac{j_{r(i)}k_{s(j)}}{j_{r(i)-1}k_{s(j)-1}}>ij,$$

Where $r(i) \geq r(i-1)+2$ and $s(j) \geq s(j-1)+2$.
Define a bounded double sequence by

$$x_{jk} = \begin{cases} 1 \; if \; j,k \in I_{r,s(i,j)}, for \; some \; i,j = 1,2,3,\ldots \\ 0 \qquad\qquad\qquad\qquad otherwise. \end{cases}$$

It is shown that $x \notin N_{\theta_{r,s}}$ but $x \in \omega$. Thus, we have $x \notin S_{\theta_{r,s}}^{f_{j,k}}$. Hence $S^{f_{j,k}} \not\subset S_{\theta_{r,s}}^{f_{j,k}}$. But this is a contradiction to the assumption that $S^{f_{j,k}} \subset S_{\theta_{r,s}}^{f_{j,k}}$. This contradiction shows that our assumption is wrong. Hence $\liminf_{r,s} q_{r,s} > 1$.

Remark 3.1.1: The double sequence $x = (x_{jk})$, constructed in the necessity part of the above lemma, is an example of $f_{j,k}$-statistically convergent double sequence which is not $f_{j,k}$-lacunary statistically convergent of double sequence.

Lemma 3.1.2: For any double lacunary sequence $\theta_{r,s}$ and unbounded modulus functions $f_{j,k}$ for which $\lim_{t\to\infty}(f(t)/t) > 0$ and there is a positive constant c such that $f(xy) > cf(x)f(y)$, for all $x \geq 0, y \geq 0$ one has $S_{\theta_{r,s}}^{f_{j,k}} \subset S^{f_{j,k}}$ if and only if $\limsup_{r,s} q_{r,s} > 1$.

Proof: Sufficiency: If $\limsup_{r,s} q_{r,s}$, then there is $H > 0$ such that $q_{r,s} < H$ for all r, s. Now, suppose that $x_{jk} \to L\left(S_{\theta_{r,s}}^{f_{j,k}}\right)$ and $\lim_{r,s\to\infty}\left(f(h_{r,s})/h_{r,s}\right) = L'$. Therefore, for given $\varepsilon > 0$, there exist $r_0, s_0 \in \mathbb{N}$ such that for all $r, s > r_0, s_0$

$$\frac{f_{j,k}(h_{r,s})}{h_{r,s}} < L' + \varepsilon,$$

$$\frac{1}{f_{j,k}(h_{r,s})} f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right) < \varepsilon.$$

Let $N_{r,s} = \left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|$. Using this notion, we have

$\frac{f_{j,k}(N_{r,s})}{f_{j,k}(h_{r,s})} < \varepsilon \; \forall \; r, s > r_0, s_0.$

Now, let $M = \max\{f_{j,k}(N_{1,1}), f_{j,k}(N_{2,2}), \ldots, f_{j,k}(N_{r_0,s_0})\}$ and let m, n be integers such that $j_{r-1} < m \leq j_r$ and $k_{s-1} < n < k_s$, then we can write

$$\frac{1}{f_{j,k}(mn)} f_{j,k}\left(\left|\{j \leq m, k \leq n: |x_{jk} - L| \geq \varepsilon\}\right|\right) \leq \frac{1}{f_{j,k}(j_{r-1}k_{s-1})} \cdot f_{j,k}\left(\left|\{j \leq m, k \leq n: |x_{jk} - L| \geq \varepsilon\}\right|\right)$$

$$= \frac{1}{f_{j,k}(j_{r-1}k_{s-1})} f_{j,k}\left(N_{1,1}, N_{2,2} + \cdots + N_{r_0,s_0} + N_{r_0+1,s_0+1} + \cdots + N_{r,s}\right)$$

$$\leq \frac{1}{f_{j,k}(j_{r-1}k_{s-1})}\left(f_{j,k}(N_{1,1}) + f_{j,k}(N_{2,2}) + \cdots + f_{j,k}(N_{r_0,s_0}) + f_{j,k}(N_{r_0+1,s_0+1}) + \cdots + f_{j,k}(N_{r,s})\right)$$

$$\leq \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \left[f_{j,k}(N_{r_0+1,s_0+1}) + \cdots + f_{j,k}(N_{r,s})\right]$$

$$= \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})}\left[\frac{f_{j,k}(h_{r_0+1,s_0+1})}{h_{r_0+1,s_0+1}} \frac{f_{j,k}(N_{r_0+1,s_0+1})}{f_{j,k}(h_{r_0+1,s_0+1})} h_{r_0+1,s_0+1} + \cdots \right.$$

$$\left. + \frac{f_{j,k}(h_{r,s})}{h_{r,s}} \frac{f_{j,k}(N_{r,s})}{f_{j,k}(h_{r,s})} h_{r,s}\right] + \frac{1}{f_{j,k}(j_{r-1}k_{s-1})}\left[(L'+\varepsilon)\varepsilon h_{r_0+1,s_0+1} + \cdots + (L'+\varepsilon)\varepsilon h_{r,s}\right]$$

$$= \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \frac{1}{f_{j,k}(j_{r-1}k_{s-1})}\varepsilon(L'+\varepsilon)\left[h_{r_0+1,s_0+1} + \cdots + h_{r,s}\right]$$
$$= \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \frac{1}{f_{j,k}(j_{r-1}k_{s-1})}\varepsilon(L'+\varepsilon)\left[j_r k_s - j_{r_0}k_{s_0}\right]$$
$$< \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \varepsilon(L'+\varepsilon)\left(\frac{j_r k_s}{f_{j,k}(j_{r-1}k_{s-1})}\right)$$
$$= \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \varepsilon(L'+\varepsilon)\frac{1}{f_{j,k}(j_{r-1}k_{s-1})/j_{r-1}k_{s-1}}\frac{j_r k_s}{j_{r-1}k_{s-1}}$$
$$= \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \varepsilon(L'+\varepsilon)q_{r,s}\frac{1}{f_{j,k}(j_{r-1}k_{s-1})/j_{r-1}k_{s-1}}$$
$$< \frac{r_0 s_0 M}{f_{j,k}(j_{r-1}k_{s-1})} + \varepsilon(L'+\varepsilon)H\cdot\frac{1}{f_{j,k}(j_{r-1}k_{s-1})/j_{r-1}k_{s-1}},$$

From where the sufficiency follows immediately, in view of the above fact that

$$\lim_{r,s\to\infty}\left(f_{j,k}(j_{r-1}k_{s-1})/j_{r-1}k_{s-1}\right) > 0.$$

Necessity: Suppose that $\limsup_{r,s} q_{r,s} = \infty$. We can select a subsequence $\left(j_{r(i)}k_{s(j)}\right)$ of double lacunary sequence $\theta_{r,s}$ such that $q_{r(i),s(j)} > ij$. Define a bounded double sequence $x = (x_{jk})$ by

$$x_{jk} = \begin{cases} 1 \; if \; j_{r(i)}k_{s(j)} < jk \le 2j_{r(i)-1}k_{s(j)-1}, for \; some \; i,j = 1,2,3,\ldots, \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad otherwise. \end{cases}$$

It is shown that $x \in N_\theta$ but $x \notin \omega$. We conclude that $x \in S_{\theta_{r,s}}^{f_{j,k}}$, but $x \notin S^{f_{j,k}}$, for every $f_{j,k}$-statistically convergent of double sequence is statistically convergent double sequence. $S_{\theta_{r,s}}^{f_{j,k}} \not\subset S^{f_{j,k}}$. But this is a contradiction to the assumption that $S_{\theta_{r,s}}^{f_{j,k}} \subset S^{f_{j,k}}$. This contradiction shows that $\limsup_{r,s} q_{r,s} < \infty$.

Remark 3.2.1: The double sequence $x = \left(x_{jk}\right)$, constructed in the necessity part of the above lemma, is an example of $f_{j,k}$-lacunary statistically convergent double sequence which is not $f_{j,k}$-statistically convergent double sequence.

Combining lemma 3.1.1 and 3.2.1 we have the following.

Theorem 3.1.1: For any double lacunary sequence $\theta_{r,s}$ and unbounded modulus functions $f_{j,k}$ for which $\lim_{t\to\infty}(f(t)/t) > 0$ and there is positive constant c such that $f(xy) \ge cf(x)f(y)$, for all $x \ge 0, y \ge 0$, one has $S_{\theta_{r,s}}^{f_{j,k}} = S^{f_{j,k}}$ if and only if $1 < \liminf_{r,s} q_{r,s} < \limsup_{r,s} q_{r,s} < \infty$.

Theorem 3.2.1: For any double lacunary sequence $\theta_{r,s}$ and unbounded modulus functions $f_{j,k}$ for which $\lim_{t\to\infty}(f(t)/t) > 0$ and there is positive constant c such that $f(xy) \ge cf(x)f(y)$, for all $x \ge 0, y \ge 0$, one has

$$S^{f_{j,k}} = \bigcap_{\liminf_{r,s} q_{r,s}>1} S_{\theta_{r,s}}^{f_{j,k}} = \bigcup_{\limsup_{r,s} q_{r,s}<\infty} S_{\theta_{r,s}}^{f_{j,k}}. \tag{4}$$

Proof: In view lemma 3.1, we have $S^{f_{j,k}} \subset \bigcap_{\liminf_{r,s} q_{r,s}>1} S_{\theta_{r,s}}^{f_{j,k}}$. Suppose if possible $x = (x_{jk}) \in \bigcap_{\liminf_{r,s} q_{r,s}>1} S_{\theta_{r,s}}^{f_{j,k}}$ but $x \notin S^{f_{j,k}}$. We have $(x_{jk}) \in S_{\theta_{r,s}}^{f_{j,k}}$ for all $\theta_{r,s} = (j_r, k_s)$ for which $\liminf_{r,s} q_{r,s} > 1$. If we take $\theta_{r,s} = (2^{r+s})$, then, in view theorem 3.1, we have $S_{\theta_{r,s}}^{f_{j,k}} = S^{f_{j,k}}$ and so $x \in S^{f_{j,k}}$, contrary to our assumption. Hence $S^{f_{j,k}} = \bigcap_{\liminf_{r,s} q_{r,s}>1} S_{\theta_{r,s}}^{f_{j,k}}$. The remaining part can be proved similarly and hence is omitted.

Remark 3.3.1: The double sequence $x = (x_{jk})$ constructed in part (a) of theorem 2.1 belongs to $S_{\theta_{r,s}}^{f_{j,k}}$

for every double lacunary sequence $\theta_{r,s}$, as well unbounded modulus functions $f_{j,k}$ for which $\lim_{t\to\infty}(f(t)/t) > 0$ and there is a positive constant c such that $f(xy) \geq cf(x)f(y)$ for all $x \geq 0, y \geq 0$.
Hence $\cap \lim \inf_{r,s} q_{r,s} S_{\theta_{r,s}}^{f_{j,k}} \neq \phi$.

3.2 Inclusion Between two Lacunary Methods of $f_{j,k}$-Statistical Convergence.

Our first results shows that, for certain modulus function $f_{j,k}$, if $\theta'_{r,s}$ is a lacunary refinement of the double lacunary sequence $\theta_{r,s}$ $S_{\theta'_{r,s}}^{f_{j,k}} \subset S_{\theta_{r,s}}^{f_{j,k}}$. To establish this result, we first recall the definition of double lacunary refinement of double sequence.

Definition 3.2.1: The double lacunary sequence $\theta'_{r,s} = (j'_r, k'_s)$ is called a double lacunary refinement of double lacunary sequence $\theta_{r,s} = (j_r, k_s)$ if $(j_r, k_s) \subset (j'_r, k'_s)$.

Theorem 3.2.1: If $\theta'_{r,s} = (j'_r, k'_s)$ is a double lacunary refinement of $\theta_{r,s} = (j_r, k_s)$ and $f_{j,k}$ is an unbounded modulus functions of double sequence such that

$$\left|f_{j,k}(x) - f_{j,k}(y)\right| = f_{j,k}(|x-y|), \forall\, x > 0, y > 0, \tag{5}$$

Then $x \in S_{\theta'_{r,s}}^{f_{j,k}}$ implies $x \in S_{\theta_{r,s}}^{f_{j,k}}$.

Proof: Suppose each $I_{r,s}$ of $\theta_{r,s}$ contains the points $\left(j'_{r(i)}k'_{s(j)}\right)_{i,j=1}^{v(r,s)}$ of $\theta'_{r,s}$ so that

$$j_{r-1}, k_{s-1} < j'_{r,1}, k'_{s,1} < j'_{r,2}, k'_{s,2} < \cdots < j'_{r,v(r)}, k'_{s,v(s)} = j_r, k_s$$

where $I'_{r,s} = \{(j,k): j'_{r-1} < j' \leq j'_{r-1}\ and\ k'_{s-1} < k' \leq k'_s\}$.

Note that, for all $(r,s), v(r,s) \geq 1$ because $(j_r, k_s) \subset (j'_{r-1}, k'_{s-1})$. Let $x_{jk} \to L\left(S_{\theta'_{r,s}}^{f_{j,k}}\right)$. Therefore, for each $\varepsilon > 0$, we have

$$\lim_{\substack{r,s\to\infty \\ 1\leq i,j\leq v(r,s)}} \frac{1}{f_{j,k}\left(h'_{r(i),s(j)}\right)} f_{j,k}\left(\left|\{(j,k) \in I'_{r(i),s(j)}: \left|x_{jk} - L\right| \geq \varepsilon\}\right|\right) = 0,$$

where $h'_{r(i),s(j)} = k'_{r(i),s(j)} - k'_{r(i-1),s(j-1)}$ and $h'_{(r,1),(s,1)} = k'_{(r,1),(s,1)} - k'_{r-1,s-1,}$ whence

$$\lim_{r,s\to\infty} \sum_{\substack{I'_{r(i),s(m)} \subset I_{r(i),s(m)} \\ 1\leq i,j\leq v(r,s)}} \sum \frac{1}{f_{j,k}\left(h'_{r(i),s(j)}\right)} f_{j,k}\left(\left|\{(j,k) \in I'_{r(i),s(j)}: \left|x_{jk} - L\right| \geq \varepsilon\}\right|\right) = 0.$$

For each $\varepsilon > 0$, we have

$$\frac{1}{f_{j,k}\left(h_{r,s}\right)} f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: \left|x_{jk} - L\right| \geq \varepsilon\}\right|\right)$$

$$= \frac{1}{f_{j,k}\left(h_{r,s}\right)} \cdot f_{j,k}\left(\left|\left\{(j,k) \in \bigcup_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} I'_{r(i),s(j)}: \left|x_{jk} - L\right| \geq \varepsilon\right\}\right|\right)$$

$$= \frac{1}{f_{j,k}\left(h_{r,s}\right)} \cdot f_{j,k}\left(\sum_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} \sum \left|\{(j,k) \in I'_{r(i),s(j)}: \left|x_{jk} - L\right| \geq \varepsilon\}\right|\right)$$

$$\leq \frac{1}{f_{j,k}(h_{r,s})} \sum_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} \sum f_{j,k}\left(\{(j,k) \in I'_{r(i),s(j)} : |x_{jk} - L| \geq \varepsilon\}\right)$$

$$= \frac{1}{f_{j,k}(h'_{r(i),s(j)})} \sum_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} \sum f_{j,k}(h'_{r(i),s(j)})$$

$$\cdot \frac{1}{f_{j,k}(h'_{r(i),s(j)})} f_{j,k}\left(\{(j,k) \in I'_{r(i),s(j)} : |x_{jk} - L| \geq \varepsilon\}\right).$$

Also, in view of the choice of unbounded modulus functions f and using the fact that $\theta'_{r,s} = (j'_r, k'_s)$ is increasing, we have

$$\sum_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} \sum f_{j,k}(h'_{(r,i),(s,j)}) = f_{j,k}(h'_{(r,1),(s,1)}) + f_{j,k}(h'_{(r,2),(s,2)}) + \cdots + f_{j,k}(h'_{(r,s),v(r,s)})$$

$$= f_{j,k}(j'_{(r,1)}, k'_{(s,1)}) + f_{j,k}(j'_{(r,2)}, k'_{(s,2)}) + \cdots$$
$$+ f_{j,k}\left((j'_{(r,v(r))}, k'_{(s,v(s))}) - (j'_{(r,v(r)-1)}, k'_{(s,v(s)-1)})\right)$$
$$= f_{j,k}(|(j'_{r,1}, k'_{s,1}) - (j'_{r-1}, k'_{s-1})|) + f_{j,k}(|(j'_{r,2}, k'_{s,2}) - (j'_{r,1}, k'_{s,1})|) + \cdots$$
$$+ f_{j,k}(|(j'_{r,v(r)}, k'_{s,v(s)}) - (j'_{(r,v(r)-1)}, k'_{(s,v(s)-1)})|)$$
$$= |f_{j,k}(j'_{r,1}, k'_{s,1}) - f_{j,k}(j'_{r-1}, k'_{s-1})| + |f_{j,k}(j'_{r,2}, k'_{s,2}) - f_{j,k}(j'_{r,1}, k'_{s,1})| + \cdots$$
$$+ |f_{j,k}(j'_{r,v(r)}, k'_{s,v(s)}) - f_{j,k}(j'_{(r,v(r)-1)}, k'_{(s,v(s)-1)})|$$
$$= f_{j,k}(j'_{r,1}, k'_{s,1}) - f_{j,k}(j'_{r-1}, k'_{s-1}) + f_{j,k}(j'_{r,2}, k'_{s,2}) - f_{j,k}(j'_{r,1}, k'_{s,1}) + \cdots$$
$$+ f_{j,k}(j'_{r,v(r)}, k'_{s,v(s)}) = f_{j,k}(j'_{(r,v(r)-1)}, k'_{(s,v(s)-1)}) = |f_{j,k}(j'_{(r,v(r)-1)}, k'_{(s,v(s)-1)})|$$
$$= f_{j,k}(|j'_{(r,v(r)-1)}, k'_{(s,v(s)-1)}|) = f_{j,k}(|h_{r,s}|) = f_{j,k}(h_{r,s}).$$

Thus, we have

$$\frac{1}{f_{j,k}(h_{r,s})} f_{j,k}(|\{(j,k) \in I_{r,s} : |x_{jk} - L| \geq \varepsilon\}|) \leq \frac{1}{\sum_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} f_{j,k}(h'_{r(i),s(j)})} \sum_{\substack{I'_{r(i),s(j)} \subset I_{r,s} \\ 1\leq i,j\leq v(r,s)}} f_{j,k}(h'_{r(i),s(j)}) t_{r(i),s(j)},$$

(6)

where

$$t_{r(i),s(j)} = \left(f_{j,k}(h'_{r(i),s(j)})\right)^{-1} f_{j,k}(|\{(j,k) \in I'_{r(i),s(j)} : |x_{jk} - L| \geq \varepsilon\}|).$$

Since the term on the right hand of (6) is regular weighted mean transformation of the double sequence $t_{r(i),s(j)}$, which tend to zero as $r, s \to \infty$, therefore the term on the right hand side of (6) also tends to zero as $r, s \to \infty$. Thus,

$\frac{1}{f_{j,k}(h_{r,s})} f_{j,k}(|\{(j,k) \in I_{r,s} : |x_{jk} - L| \geq \varepsilon\}|) \to 0$ as $r, s \to \infty$. Hence $x \in S_{\theta_{r,s}}^{f_{j,k}}$.

Theorem 3.2.2: Let $f_{j,k}$ be an unbounded modulus functions and $\theta'_{r,s} = (j'_m, k'_n)$ is a double lacunary refinement of double lacunary sequence

$\theta_{r,s} = (j_r, k_s)$. Let $I_{r,s} = \{(j,k) : j_{r-1} < j \leq j_r \text{ and } k_{s-1} < k \leq k_s\}$, $h_r = j_r - j_{r-1}$ and $h_s = k_s - k_{s-1}$,

where

$h_{r,s} = h_r \overline{h_s}$, $r, s = 1,2,3, \ldots$, and $I'_{m,n} = \{(j,k) : j'_{m-1} < j' \leq j'_m \text{ and } k'_{n-1} < k' \leq k'_n\}$, $h'_m = j'_m - j'_{m-1}$ and $h'_n = k'_n - k'_n$, where $h'_{m,n} = h'_m h'_n$, $m, n = 1,2,3, \ldots$, if there exists $\delta > 0$, such that

$\frac{f_{j,k}(h'_{m,n})}{f_{j,k}(h_{r,s})} \geq \delta$ for every $I'_{m,n} \subset I_{r,s}$,

Then $x \in S_{\theta_{r,s}}^{f_{j,k}}$ and $x \in S_{\theta'_{r,s}}^{f_{j,k}}$.

Proof: For any $\varepsilon > 0$, and for $I'_{m,n} \subset I_{r,s}$, we can find $I_{r,s}$ such that $I'_{m,n} \subset I_{r,s}$, then we have

$$\begin{aligned}\frac{1}{f_{j,k}(h'_{m,n})} f_{j,k}\left(\left|\{(j,k) \in I'_{m,n}: |x_{jk} - L| \geq \varepsilon\}\right|\right) &\leq \frac{1}{f_{j,k}(h'_{m,n})} f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right) \\ &= \frac{f_{j,k}(h_{r,s})}{f_{j,k}(h'_{m,n})} \frac{1}{f_{j,k}(h_{r,s})} f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right) \\ &\leq \frac{1}{\delta} \frac{1}{f_{j,k}(h_{r,s})} f_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right)\end{aligned}$$

From where it follows that $S_{\theta_{r,s}}^{f_{j,k}} \subset S_{\theta'_{r,s}}^{f_{j,k}}$.

In the next theorem we deal with a more general situation.

Theorem 3.2.3: Let f and g be any two modulus functions of double sequence such that $f(x) \leq g(x)$, for all $x \in [0, \infty)$, and $\theta'_{r,s} = (j'_m, k'_n)$ is a double lacunary refinement of the double sequence $\theta_{r,s} = (j_r, k_s)$. Let $I_{r,s} = \{(j,k): j_{r-1} < j \leq j_r \text{ and } k_{s-1} < k \leq k_s\}$, $h_r = j_r - j_{r-1}$ and $h_s = k_s - k_{s-1}$, where $h_{r,s} = h_r \overline{h_s}$, $r, s = 1,2,3, \ldots$, and $I'_{m,n} = \{(j,k): j'_{m-1} < j' \leq j'_m \text{ and } k'_{n-1} < k' \leq k'_n\}$, $h'_m = j'_m - j'_{m-1}$ and $h'_n = k'_n - k'_n$, where $h'_{m,n} = h'_m h'_n$, $m, n = 1,2,3, \ldots$, if there exists $0 < \delta \leq 1$, such that $\frac{f_{j,k}(h'_{m,n})}{g_{j,k}(h_{r,s})} \geq \delta$ for every $I'_{m,n} \subset I_{r,s}$,

Then $x \in S_{\theta_{r,s}}^{g_{j,k}}$ and $x \in S_{\theta'_{r,s}}^{f_{j,k}}$.

Proof: For any $\varepsilon > 0$, and every $I'_{m,n}$, we can find $I_{r,s}$ such that $I'_{m,n} \subset I_{r,s}$, then we have

$$\begin{aligned}\frac{1}{f_{j,k}(h'_{m,n})} f_{j,k}\left(\left|\{(j,k) \in I'_{m,n}: |x_{jk} - L| \geq \varepsilon\}\right|\right) &\leq \frac{1}{f_{j,k}(h'_{m,n})} g_{j,k}\left(\left|\{(j,k) \in I'_{m,n}: |x_{jk} - L| \geq \varepsilon\}\right|\right) \\ &\leq \frac{1}{f_{j,k}(h'_{m,n})} g_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right) \\ &= \frac{g_{j,k}(h_{r,s})}{f_{j,k}(h'_{m,n})} \frac{1}{g_{j,k}(h_{r,s})} g_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right) \\ &\leq \frac{1}{\delta} \frac{1}{g_{j,k}(h_{r,s})} g_{j,k}\left(\left|\{(j,k) \in I_{r,s}: |x_{jk} - L| \geq \varepsilon\}\right|\right)\end{aligned}$$

From where it follows that $S_{\theta_{r,s}}^{g_{j,k}} \subset S_{\theta'_{r,s}}^{f_{j,k}}$.

In the next theorem we show that the inclusion $S_{\theta_{r,s}}^{f_{j,k}} \subset S_{\theta'_{r,s}}^{f_{j,k}}$ is possible if even if none of $\theta_{r,s}$ and $\theta'_{r,s}$ is refinement of the other.

Theorem 3.2.4: Let f be an unbounded modulus functions such that

$$|f(x) - f(y)| = f(|x - y|), \forall\, x \geq 0, y \geq 0. \tag{7}$$

Suppose $\theta'_{r,s} = (j'_m, k'_n)$ and $\theta_{r,s} = (j_r, k_s)$. Let $I_{r,s} = \{(j,k): j_{r-1} < j \leq j_r \text{ and } k_{s-1} < k \leq k_s\}$, $h_r = j_r - j_{r-1}$ and $h_s = k_s - k_{s-1}$, where $h_{r,s} = h_r \overline{h_s}$, $r, s = 1,2,3, \ldots$, and $I'_{m,n} = \{(j,k): j'_{m-1} < j' \leq j'_m \text{ and } k'_{n-1} < k' \leq k'_n\}$, $h'_m = j'_m - j'_{m-1}$ and $h'_n = k'_n - k'_n$, where $h'_{m,n} = h'_m h'_n$, $m, n = 1,2,3, \ldots$, and $I_{p,q,m,n} = I_{p,q} \cap I_{m,n}$, $p, q, m, n = 1,2,3, \ldots$, if there exists $\delta > 0$ such that $\frac{f_{j,k}(\sigma_{p,q,m,n})}{g_{j,k}(h_{r,s})} \geq \delta$ for every $p, q, m, n = 1,2,3, \ldots$,

provided $\sigma_{p,q,m,n} > 0$, where $\sigma_{p,q,m,n}$ denotes the length of the interval $I_{p,q,m,n}$ then $x \in S_{\theta_{r,s}}^{f_{j,k}}$ implies

$S_{\theta'_{r,s}}^{f\,j,k}$.

Remark 3.2.1: If the condition in theorem 4.4 is replaced by $f(\sigma_{p,q,m,n})/f(h'_{m,n}) \geq \delta$ for every $r,s,m,n = 1,2,3,\ldots$, provided $\sigma_{p,q,m,n} > 0$, where $\sigma_{p,q,m,n}$ denotes the length of the interval $I_{p,q,m,n} = I_{p,q} \cap I_{m,n}$, $p,q,m,n = 1,2,3,\ldots$, it can be seen that $x \in S_{\theta'_{r,s}}^{f\,j,k}$ implies $x \in S_{\theta_{r,s}}^{f\,j,k}$.

Combining remark 4.1 and theorem 4.4, we get the following.

Theorem 3.2.5: Let f be an unbounded modulus functions such that

$$|f(x) - f(y)| = f(|x - y|), \forall\, x \geq 0, y \geq 0. \tag{8}$$

Suppose $\theta'_{r,s} = (j'_m, k'_n)$ and $\theta_{r,s} = (j_r, k_s)$ are two double lacunary sequences. Let $I_{r,s} = \{(j,k): j_{r-1} < j \leq j_r \text{ and } k_{s-1} < k \leq k_s\}$, $h_r = j_r - j_{r-1}$ and $h_s = k_s - k_{s-1}$, where $h_{r,s} = h_r\overline{h_s}$, $r,s = 1,2,3,\ldots, I'_{m,n} = \{(j,k): j'_{m-1} < j' \leq j'_m \text{ and } k'_{n-1} < k' \leq k'_n\}, h'_m = j'_m - j'_{m-1}$ and $h'_n = k'_n - k'_n$, where $h'_{m,n} = h'_m h'_n$, $m,n = 1,2,3,\ldots$, and $I_{p,q,m,n} = I_{p,q} \cap I_{m,n}$, $p,q,m,n = 1,2,3,\ldots$, if there exists $\delta > 0$ such that

$\frac{f_{j,k}(\sigma_{p,q,m,n})}{g_{j,k}(h_{r,s}+h'_{m,n})} \geq \delta$ for every $p,q,m,n = 1,2,3,\ldots$,

provided $\sigma_{p,q,m,n} > 0$, where $\sigma_{p,q,m,n}$ denotes the length of the interval $I_{p,q,m,n}$ then $S_{\theta_{r,s}}^{f\,j,k} = S_{\theta'_{r,s}}^{f\,j,k}$.

4. Discussion

The concept of modulus lacunary statistical convergence of double sequence was introduced via modulus functions where the modulus function is bounded or unbounded. We have also introduced the concept of lacunary strong convergence of double sequence with respect modulus function. We established some inclusion relations between these two lacunary methods and proved some essential analogues results for double sequence. This concept can be further extended in the direction of fuzzy numbers of double sequence.

References

[1] Fast, H. (1951). Sur la convergence statistique, *Colloq. Math.* **2**: 241-244

[2] Schoenberg, I. J. (1959). The integrability of certain functions and related summability methods. *The American mathematical monthly*, *66*(5): 361-775.

[3] Fridy, J. A. and Orhan, C. (1993). Lacunary statistical summability. *Journal of mathematical analysis and applications*, *173*(2): 497-504.

[4] Šalát, T. (1980). On statistically convergent sequences of real numbers. *Mathematica slovaca*, *30*(2): 139-150.

[5] Cakalli, H. (1996). On statistical convergence in topological groups. *Pure and Applied Mathematika Sciences*, *43*: 27-32.

[6] Çakalli, H., and Das, P. (2009). Fuzzy compactness via summability. *Applied Mathematics Letters*, *22*(11): 1665-1669.

[7] Di Maio, G., and Kocinac, L. A. D. (2008). Statistical convergence in topology. *Topology and its Applications*, *156*(1): 28-45.

[8] Maddox, I. J. (1988). Statistical convergence in a locally convex space. In *Mathematical Proceedings of the Cambridge Philosophical Society*, *104*(1): 141-145.

[9] Cheng, L. Lin, G. Lan, Y. and Liu, H. (2008). Measure Theory of Statistical Convergence. *Science in China Series A-Math.*, 51: 2285.

[10] Connor, J. and Swardson, M. A. (1993). Strong integral summability and the Stone-Čech compactification of the half-line. *Pacific Journal of Mathematics*, *157*(2): 201-224.

[11] Miller, H. I. (1995). A measure theoretical subsequence characterization of statistical convergence. *Transactions of the American Mathematical Society*, *347*(5), 1811-1819.

[12] Nuray, F. and Savaş, E. (1995). Statistical convergence of sequences of fuzzy numbers. *Mathematica Slovaca*, *45*(3): 269-273.

[13] Savaş, E. (2001). On statistically convergent sequences of fuzzy numbers. *Information Sciences*, *137*(1-4): 277-282.

[14] Connor, J., Just, W. and Swardson, M. A. (1994). Equivalence of bounded strong integral summability methods. *Math. Japon*, *39*(3): 401-428.

[15] Mursaleen, M., Mohiuddine, S. A., Mursaleen, M. and Mohiuddine, S. A. (2014). Statistical convergence of double sequences. *Convergence Methods for Double Sequences and Applications*, 117-132.

[16] Fridy, J. A. and Orhan, C. (1993). Lacunary statistical convergence. *Pacific Journal of Mathematics*, 160: 43–51.

[17] Savaş, E. and Patterson, R. F. (2005). On some double almost lacunary sequence spaces defined by Orlicz functions. *Filomat*, *19*: 35-44.

[18] Nakano, H. (1953). Concave modulars. *Journal of the Mathematical society of Japan*, *5*(1): 29-49.

[19] Ruckle, W. H. (1973). FK spaces in which the sequence of coordinate vectors is bounded. *Canadian journal of mathematics*, *25*(5): 973-978.

[20] Maddox, I. J. (1986). Sequence spaces defined by a modulus. In *Mathematical Proceedings of the Cambridge Philosophical Society*, 100(1): 161–166.

[21] Aizpuru, A., Listán-García, M. C. and Rambla-Barreno, F. (2014). Density by moduli and statistical convergence. *Quaestiones mathematicae*, *37*(4): 525-530.

[22] Bhardwaj, V. K. and Dhawan, S. (2015). f −Statistical convergence of order α and strong Cesàro summability of order α with respect to a modulus. *Journal of Inequalities and Applications*, *2015*(1): 1-14.

[23] Bhardwaj, V. K., Dhawan, S. and Gupta, S. (2016). Density by moduli and statistical boundedness. In *Abstract and Applied Analysis*, 216: 1-6.

[24] Bhardwaj, V. K. and Dhawan, S. (2016). Density by moduli and lacunary statistical convergence. In *Abstract and Applied Analysis*, 1-11.

RT&A, No 4 (76)
Volume 18, December 2023

ESTIMATION OF STRESS STRENGTH RELIABILITY USING PRANAV DISTRIBUTION

Ankitha Lukose, Chacko V M

•

Department of Statistics
St.Thomas College(Autonomous)
Thrissur, Kerala, 680 001, India
ankithalukose17@gmail.com
chackovm@gmail.com

Abstract

This paper deals with the estimation of stress strength reliability parameter R, which is the probability of Y less than X when X and Y are two independent distribution with different scale parameter and same shape parameter. The maximum likelihood method is used to find an estimator for R. We also obtain the asymptotic distribution of the maximum likelihood estimator of R. Based on this asymptotic distribution, the asymptotic confidence interval can be obtained. We also propose bootstrap confidence interval for the parameter R. Analysis of a simulated data and a real life data have been presented for illustrative purposes.

Keywords: stress strength model, maximum-likelihood estimator, bootstrap confidence intervals, asymptotic distributions and confidence interval

1. Introduction

One of the significant, challenging, but manageable problems in reliability analysis is the calculation of stress-strength reliability using a variety of distributions. Estimating the stress-strength parameter, R, is very helpful in the statistical literature. For instance, if X represents the strength of a system under a stress, Y, then R is a measure of system performance that arises naturally from the mechanical dependability of a system. Only when the applied stress exceeds the system's strength at any point does the system fail. Many lifespan distributions are utilised in reliability analysis. Terms like exponential, Weibull, log-Normal, and their generalisations are commonly used in dependability analysis. The exponential, Lindley, and Weibull distributions are more often used than the gamma and lognormal distributions because their survival functions can both be expressed in closed forms and do not require numerical integration. While sharing a common parameter, the exponential and Lindley distributions differ in that the hazard rate of the exponential is constant whereas the hazard rate of the Lindley is monotonically dropping. Although the Lindley distribution has been used by many academics to model lifetime data and is crucial for understanding stress-strength reliability modelling, there are numerous instances in which modelling actual lifetime data may not be appropriate from a theoretical or practical standpoint. Recently, a number of academics presented many distributions, with the new ones showing a better fit than the currently popular distributions. In order to fix a problem with new models while utilising better fitted models in stress-strength analysis, one may need to carefully examine the estimation process. If the estimation process fails using the available methodologies, one may not be able to do so. Therefore, it is crucial to estimate multiple reliability factors, and researchers must focus more on estimates while utilising better fitted models.

In the literature, it has been debated how to estimate a stress strength model's reliability or survival probability when X and Y have known distributions. A number of authors have examined the survival probabilities of a single component stress-strength (SSS) model for various distributions, including Raqab and Kundu[12], Kundu and Gupta [9], [6], Constantine and Karson[4], and Downtown[5]. The issue of estimating R has been investigated by a number of authors. Church and Harris [3] developed the MLE of R when X and Y are independently and normally distributed. Awad et al study's on the MLE of R under the condition that X and Y have bivariate exponential distributions was published in 1981. In a simulation research, Awad and Gharraf [2] compared three estimations of R when X and Y are independent Burr random variables with different distributions. Estimates for R where X and Y are Burr Type X distributions were reported by Ahmad et al.[1] and Surles and Padgett[16],[17]. Other papers which describes the same idea on different distributions are Akhila and Chacko [18],Kundu and Raqab [11],saraccouglu,Kinaci and Kundu [14] and Shahsanaei, Fatemeh and Daneshkhah, Alireza [15].

In this paper, we consider the problem of estimating the stress strength parameter $R = P[Y < X]$, when X and Y be independent strength and stress random variables having Pranav distribution with parameters θ_1 and θ_2 respectively. Krishna Kumar Shukla [8]introduced Pranav distribution which is a mixture of two distributions, exponential distribution having scale parameter θ and gamma distribution having shape parameter 4 and scale parameter θ, and their mixing proportion of θ, $\frac{\theta^4}{\theta^4+6}$ and $\frac{6}{\theta^4+6}$ respectively. The probability density function (pdf) of Pranav lifetime distribution can be defined as

$$f(x,\theta) = \frac{\theta^4}{\theta^4+6}(\theta + x^3)e^{-\theta x},\ x > 0, \theta > 0 \tag{1}$$

The corresponding cumulative distribution function(c.d.f) is

$$F(x,\theta) = 1 - \left[1 + \frac{\theta x(\theta^2x^2 + 3\theta x + 6}{\theta^4+6}\right] e^{-\theta x},\ x > 0, \theta > 0 \tag{2}$$

The estimation of the stress strength parameter $R = P[Y < X]$, when X and Y are both one-parameter Pranav distributions with parameter θ_1 and θ_2 respectively, is an unsolved problem. Statistical inference on stress-strength parameters are important in reliability analysis. It is observed that the maximum likelihood estimators can be obtained implicitly by solving two nonlinear equations, but they cannot be obtained in closed form. So, MLE' s of parameters are derived numerically. It is not possible to compute the exact distributions of the maximum likelihood estimators, and we used the asymptotic distribution and we constructed approximate confidence intervals of the unknown parameters.

The rest of the paper is organized as follows. In Section 2, the MLE of R is computed. The asymptotic distribution of the MLE' s are provided in Section 3. Bootstrap confidence interval is presented in Section 4. In Section 5, simulation study is given. Theoretical results are verified by analyzing data set in Section 6 and conclusions are given in Section 7.

2. Maximum Likelihood Estimator of R

In this section, the procedure of estimating the reliability of $R = P[Y < X]$ models using Pranav distributions, is considered. It is clear that

$$\begin{aligned} R &= P[Y < X] \\ &= \int_0^\infty f(x,\theta_1)F(x,\theta_2)\,dx \end{aligned} \tag{3}$$

where f (x,y), is the joint probability density function (pdf) of random variables X and Y, having Pranav distributions. If the r.v' s X and Y are independent, then $f(x,y) = f(x)g(y)$, where f (x)

and g(y) are the marginal pdf' s of X and Y, so that

$$R = \int_0^\infty \left(\frac{\theta^4}{\theta^4+6}(\theta+x^3)e^{-\theta x}\right)\left(1-\left[1+\frac{\theta x(\theta^2x^2+3\theta x+6}{\theta^4+6}\right]e^{-\theta x}\right)\,dx \tag{4}$$

On simplification we get

$$R = 1 - \frac{\theta_1\begin{bmatrix}360\theta_1\theta_2^2+1080\theta_2^3+144\theta_2(\theta_1+\theta_2)^2+\\ 6(\theta_1\theta_2^3+\theta_2^4+6)(\theta_1+\theta_2)^3+\\ 6(\theta_1\theta_2)^2(\theta_1+\theta_2)^4+6\theta_1\theta_2(\theta_1+\theta_2)^5+\\ \theta_1(\theta_2^4+6)(\theta_1+\theta_2)^6\end{bmatrix}}{(\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7} \tag{5}$$

If we have two ordered random samples representing strength (X1,X2...Xn) and stress (Y1,Y2,....,Ym) having sizes n and m respectively, following Pranav distribution with parameters θ_1 and θ_2, respectively. A technique for figuring out a model's parameter values is called maximum likelihood estimation. The parameter values were selected to maximise the likelihood that the model's described process resulted in the observed data. The optimal distribution for a collection of data is chosen using the maximum likelihood estimate (MLE). A reliable method for estimating parameters is maximum likelihood. So, a variety of estimation scenarios can apply maximum likelihood estimations. With those parameter values, the likelihood of those observed outcomes is the same as the probability of those observed events. Thus, the likelihood function for the combined random sample can be calculated:

$$L = \prod_{i=1}^{n}\frac{\theta_1^4}{\theta_1^4+6}(\theta_1+x_i^3)\;e^{-\theta_1x_i^3}\prod_{j=1}^{n}\frac{\theta_2^4}{\theta_2^4+6}(\theta_2+y_j^3)e^{-\theta_2y_j^3} \tag{6}$$

the log likelihood is

$$l = logL = 4nlog\theta_1 - nlog(\theta_1^4+6) + \sum_{i=0}^{n}(log(\theta_1+x_i^3)) - \theta_1\sum_{i=0}^{n}x_i + 4mlog\theta_2 - mlog(\theta_2^4+6) + \sum_{j=1}^{m}(log(\theta_2+y_j^3)) - \theta_2\sum_{j=1}^{m}y_j \tag{7}$$

The solution of the following non-linear equations yield the MLE of the parameters parameters θ_1 and θ_2. Differentiating (7) with respect to parameters θ_1 and θ_2, we obtain

$$\frac{\partial l}{\partial\theta_1} = \frac{4n}{\theta_1} - \frac{4n\theta_1^3}{(\theta_1^4)+6} + \sum_{i=0}^{n}\frac{1}{\theta_1+x_i^3} - \sum_{i=0}^{n}x_i \tag{8}$$

$$\frac{\partial l}{\partial\theta_2} = \frac{4m}{\theta_2} - \frac{4m\theta_2^3}{(\theta_2^4)+6} + \sum_{j=0}^{m}\frac{1}{\theta_2+x_j^3} - \sum_{j=0}^{m}y_j \tag{9}$$

The second partial derivatives of (7) with respect to parameters θ_1 and θ_2, are

$$\frac{\partial l^2}{\partial\theta_1^2} = \frac{-4n}{\theta_1^2} - \frac{4n\theta_1^2(18-\theta_1^4)}{(\theta_1^4+6)^2} - \sum_{i=0}^{n}\frac{1}{(\theta_1+x_i^3)^2} \tag{10}$$

$$\frac{\partial l^2}{\partial\theta_2^2} = \frac{-4m}{\theta_2^2} - \frac{4m\theta_2^2(18-\theta_2^4)}{(\theta_2^4+6)^2} - \sum_{j=0}^{m}\frac{1}{(\theta_2+y_j^3)^2} \tag{11}$$

MLE of R is obtained as

$$\hat{R} = 1 - \frac{\hat{\theta_1}\begin{bmatrix}360\hat{\theta_1}(\hat{\theta_2}^2+1080\hat{\theta_2}^3+144\hat{\theta_2}(\hat{\theta_1}+\hat{\theta_2})^2+\\ 6(\hat{\theta_1}\hat{\theta_2}^3+\hat{\theta_2}^4+6)(\hat{\theta_1}+\hat{\theta_2})^3+6(\hat{\theta_1}\hat{\theta_2})^2(\hat{\theta_1}+\hat{\theta_2})^4+\\ 6\hat{\theta_1}\hat{\theta_2}(\hat{\theta_1}+\hat{\theta_2})^5+\hat{\theta_1}(\hat{\theta_2}^4+6)(\hat{\theta_1}+\hat{\theta_2})^6\end{bmatrix}}{(\hat{\theta_1}^4+6)(\hat{\theta_2}^4+6)(\hat{\theta_1}+\hat{\theta_2})^7} \tag{12}$$

Where $\hat{\theta_1}$ and $\hat{\theta_2}$ are the maximum likelihood estimators of θ_1 and θ_2 respectively. This is used in estimation of stress strength for given data.

3. Asymptotic Distribution and Confidence Interval

The asymptotic distribution and confidence interval of the MLE of R are obtained in this section. To find an asymptotic variance of the MLE $\hat{R}^{ML}$, let us denote the Fisher information matrix of $\theta = (\theta_1, \theta_2)$ as $I(\theta) = [I_ij(\theta); i, j = 1, 2]$, i.e.,

$$I(\theta) = E \begin{bmatrix} \frac{-\partial^2 l}{\partial \theta_1^2} & \frac{-\partial^2 l}{\partial \theta_1 \partial \theta_2} \\ \frac{-\partial^2 l}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 l}{\partial \theta_2^2} \end{bmatrix} \tag{13}$$

To establish normality assumption we define

$$d(\theta) = (\frac{\partial R}{\partial \theta_1}, \frac{\partial R}{\partial \theta_2})' \tag{14}$$

$$= (d_1, d_2)' \tag{15}$$

R is in the form

$$R = 1 - \frac{U}{V} P \tag{16}$$

Hence

$$\frac{\partial R}{\partial \theta_i} = -\left[(\frac{U}{V})' P + P' \frac{U}{V} \right], i = 1, 2 \tag{17}$$

$$U = \theta_1 \tag{18}$$

$$V = (\theta_1^4 + 6)(\theta_2^4 + 6)(\theta_1 + \theta_2)^7 \tag{19}$$

$$\begin{aligned} P = {} & 360\hat{\theta_1}(\hat{\theta_2}^2 + 1080\hat{\theta_2}^3 + 144\hat{\theta_2}(\hat{\theta_1} + \hat{\theta_2})^2 + \\ & 6(\hat{\theta_1}\hat{\theta_2}^3 + \hat{\theta_2}^4 + 6)(\hat{\theta_1} + \hat{\theta_2})^3 + 6(\hat{\theta_1}\hat{\theta_2})^2(\hat{\theta_1} + \hat{\theta_2})^4 + \\ & 6\hat{\theta_1}\hat{\theta_2}(\hat{\theta_1} + \hat{\theta_2})^5 + \hat{\theta_1}(\hat{\theta_2}^4 + 6)(\hat{\theta_1} + \hat{\theta_2})^6 \end{aligned} \tag{20}$$

The partial derivatives of $\frac{U}{V}$ and P with respect to θ_1 is

$$\begin{aligned} \frac{\partial P}{\partial \theta_1} = {} & 360\theta_2^2 + 288\theta_2(\theta_1 + \theta_2) + 18\theta_1\theta_2^3 + (\theta_1 + \theta_2)^2 + \\ & 6\theta_2^3(\theta_1 + \theta_2)^3 + 18\theta_2(\theta_1 + \theta_2)^2 + 18(\theta_1 + \theta_2)^2 + \\ & 24(\theta_1\theta_2)^2(\theta_1 + \theta_2)^3 + 12\theta_1\theta_2^2(\theta_1 + \theta_2)^4 + 30\theta_1\theta_2(\theta_1 + \theta_2)^4 + \\ & 6\theta_2(\theta_1 + \theta_2)^5 + 6\theta_1(\theta_2^4 + 6)(\theta_1 + \theta_2)^5 + (\theta_2^4 + 6)(\theta_1 + \theta_2)^6 \end{aligned} \tag{21}$$

$$(\frac{U}{V})' = \frac{\begin{bmatrix} (\theta_1^4 + 6)(\theta_2^4 + 6)(\theta_1 + \theta_2)^7 - \\ 7\theta_1(\theta_1^4 + 6)(\theta_2^4 + 6)(\theta_1 + \theta_2)^6 - \\ 4\theta_1^4(\theta_2^4 + 6)(\theta_1 + \theta_2)^7 \end{bmatrix}}{[(\theta_1^4 + 6)(\theta_2^4 + 6)(\theta_1 + \theta_2)^7]^2} \tag{22}$$

Then partial derivative of R with respect to θ_1 becomes

$$\frac{\partial R}{\partial \theta_1} = -\frac{\begin{bmatrix}-7\theta_1(\theta_2^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^6 - 4\theta_1^4(\theta_2^4+6)(\theta_1+\theta_2)^7 + \\ (\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7\end{bmatrix}}{((\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7)^2}$$
$$\times \begin{bmatrix}360\theta_1\theta_2^2 + 1080\theta_2^3 + 144\theta_2(\theta_1+\theta_2)^2 + \\ 6(\theta_1\theta_2^3+\theta_2^4+6)(\theta_1+\theta_2)^3 + 6(\theta_1\theta_2)^2(\theta_1+\theta_2)^4 + \\ 6(\theta_1\theta_2)^2(\theta_1+\theta_2)^4 + 6\theta_1\theta_2(\theta_1+\theta_2)^5 + \theta_1(\theta_2^4+6)(\theta_1+\theta_2)^6\end{bmatrix}$$
$$+ \begin{bmatrix}360\theta_2^2 + 288\theta_2(\theta_1+\theta_2) + 18\theta_1\theta_2^3(\theta_1+\theta_2)^2 6\theta_2^3(\theta_1+\theta_2)^3 + \\ 18\theta_2(\theta_1+\theta_2)^2 + 18(\theta_1+\theta_2)^2 + 24(\theta_1\theta_2)^2(\theta_1+\theta_2)^3 \\ +12\theta_1\theta_2^2(\theta_1+\theta_2)^4 + 30\theta_1\theta_2(\theta_1+\theta_2)^4 + \\ 6\theta_2(\theta_1+\theta_2)^5 + 6\theta_1(\theta_2^4+6)(\theta_1+\theta_2)^5 + (\theta_2^4+6)(\theta_1+\theta_2)^6\end{bmatrix}$$
$$\times \frac{\theta_1}{(\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7} \tag{23}$$

The partial derivatives of $\frac{U}{V}$ and P with respect to θ_2 is

$$\begin{aligned}\frac{\partial P}{\partial \theta_2} &= 720\theta_1\theta_2 + 3240\theta_2^2 + 288\theta_2(\theta_1+\theta_2) + 144(\theta_1+\theta_2)^2 \\ &+ 18\theta_1\theta_2^3(\theta_1+\theta_2)^2 + 18\theta_1\theta_2^2(\theta_1+\theta_2)^3 + 18\theta_2^4(\theta_1+\theta_2)^2 + 24\theta_2^3(\theta_1+\theta_2)^3 \\ &+ 18(\theta_1+\theta_2)^2 + 24(\theta_1\theta_2)^2(\theta_1+\theta_2)^3 + 12\theta_1^2\theta_2(\theta_1+\theta_2)^4 + 30\theta_1\theta_2(\theta_1+\theta_2)^4 \\ &+ 6\theta_1(\theta_1+\theta_2)^5 + 6\theta_1(\theta_2^4+6)(\theta_1+\theta_2)^5 + 4\theta_1\theta_2^3(\theta_1+\theta_2)^6\end{aligned} \tag{24}$$

$$(\frac{U}{V})' = -\frac{[-7\theta_1(\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^6 - 4\theta_1\theta_2^3(\theta_1^4+6)(\theta_1+\theta_2)^7]}{((\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7)^2} \tag{25}$$

Then the partial derivative of R with respect to θ_2 becomes

$$\frac{\partial R}{\partial \theta_2} = -\frac{[-7\theta_1(\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^6 - 4\theta_1\theta_2^3(\theta_1^4+6)(\theta_1+\theta_2)^7]}{((\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7)^2}$$
$$\times \begin{bmatrix}360\theta_1\theta_2^2 + 1080\theta_2^3 + 144\theta_2(\theta_1+\theta_2)^2 + \\ 6(\theta_1\theta_2^3+\theta_2^4+6)(\theta_1+\theta_2)^3 + 6(\theta_1\theta_2)^2(\theta_1+\theta_2)^4 + \\ 6(\theta_1\theta_2)^2(\theta_1+\theta_2)^4 + 6\theta_1\theta_2(\theta_1+\theta_2)^5 + \theta_1(\theta_2^4+6)(\theta_1+\theta_2)^6\end{bmatrix}$$
$$+ \begin{bmatrix}720\theta_1\theta_2 + 3240\theta_2^2 + 288\theta_2(\theta_1+\theta_2) + 144(\theta_1+\theta_2)^2 + \\ 18\theta_1\theta_2^3(\theta_1+\theta_2)^2 + 18\theta_1\theta_2^2(\theta_1+\theta_2)^3 + 18\theta_2^4(\theta_1+\theta_2)^2 + 24\theta_2^3(\theta_1+\theta_2)^3 + \\ 18(\theta_1+\theta_2)^2 + 24(\theta_1\theta_2)^2(\theta_1+\theta_2)^3 + 12\theta_1^2\theta_2(\theta_1+\theta_2^4) + 30\theta_1\theta_2(\theta_1+\theta_2)^4 + \\ 6\theta_1(\theta_1+\theta_2)^5 + 6\theta_1(\theta_2^4+6)(\theta_1+\theta_2)^5 + 4\theta_1\theta_2^3(\theta_1+\theta_2)^6\end{bmatrix}$$
$$\times \frac{\theta_1}{(\theta_1^4+6)(\theta_2^4+6)(\theta_1+\theta_2)^7} \tag{26}$$

We obtain the asymptotic distribution of $\hat{R}^{ML}$ as

$$\sqrt{n+m}(\hat{R}^{ML} - R) \to N(0, d^{'}(\theta)I^{-1}(\theta)d(\theta)) \tag{27}$$

$$AV(\hat{R}^{ML}) = \frac{1}{n+m}d^{'}(\theta)I^{-1}(\theta)d(\theta) \tag{28}$$

$$i.e, AV(\hat{R}^{ML}) = V(\hat{\theta_1})d_1^2 + V(\hat{\theta_2})d_2^2 + 2d_1d_2Cov(\hat{\theta_1}, \hat{\theta_2}) \tag{29}$$

Asymptotic $100(1-\alpha)$ percentage confidence interval for R

$$\hat{R}^{ML} \pm Z_{\alpha/2}\sqrt{AV(\hat{R}^{ML})} \tag{30}$$

4. Bootstrap Confidence Interval

In this section, we use confidence intervals based on the parametric percentile bootstrap methods (we call it from now on as Boot-p) based, [10]. Bootstrapping is a technique for estimating the variability in a statistic by sampling with replacement from observed data. Permutation tests are a type of re-sampling that is linked to re-sampling. The bootstrap is frequently used to evaluate the accuracy of an estimate based on a sample of data from a larger population. The bootstrap's main advantage is that it allows statisticians to construct confidence intervals on parameters without making irrational assumptions. This was one of the first of many discoveries in computational statistics, which has now become the standard method for practically all work. It creates multiple re-samples (with replacement) from a single set of observations, and computes the effect size of interest on each of these re-samples. To estimate confidence intervals of R in this methods, the following steps are used.

1. Estimate θ, say $\hat{\theta}$, from the sample using MLE method.
2. Generate a bootstrap sample using $\hat{\theta}$. Using these bootstrap sample obtain the bootstrap estimate of θ, say $\hat{\theta}*$ and compute the bootstrap estimate of R.
3. Repeat Step [2] N-BOOT times to get the parametric bootstrap estimates of R
4. Let $\widehat{CDF} = P(\hat{\theta}* \leq x)$, be the cumulative distribution function of $\hat{R}$. Define $\hat{\theta}_{Boot-p}(x) = \widehat{CDF}^{-1}(x)$ for a given x. The approximate $(100-\alpha)\%$ confidence interval for θ is given by

$$(\hat{\theta}_{BOOT-P}(\alpha/2), \hat{\theta}_{BOOT-P}(1-(\alpha/2))) \tag{31}$$

5. Simulation Study

To measure the efficacy of R's estimators, we provide some results based on the inversion method in this section. According to its definition, data simulation is the process of using a significant amount of data to replicate or simulate real-world settings in order to determine the optimal course of action, predict future events, or test a model. The many forms of data simulations are numerous. In simulation statistics, artificially generated data are used to test a hypothesis or statistical method. Every time a new statistical method is created or used, certain assumptions need to be verified. Simulated data is used by statisticians to test their theories. For this purpose, we have generated 1000 samples from independent Pranav(θ_1) and Pranav(θ_2) distributions. We considered sets of parameter values (1) and (1.99) parameter values. The bias and the mean square error (MSE) of the parameter estimates are calculated. In Table 5.1, the average biases, mean squared errors (MSE) and confidence intervals of the estimates of R is given.

Table 1: *Simulation Results*

(n,m)	R(ML)	Bias	MSE	AS(CI)	BT(CI)
(7,7)	0.8982136	0.1017854	0.01036027	0.8783155,0.9181117	0.8947541,0.8999972
(15,15)	0.8982252	0.1017738	0.0103579	0.8782696,0.9181808	0.8966189,0.8994737
(20,20)	0.8982344	0.1017646	0.01035603	0.8783453,0.9181235	0.8965753,0.8999412
(22,22)	0.8985076	0.1014914	0.0103005	0.8787993,0.9182159	0.8969416,0.8993731
(25,25)	0.8985383	0.1014607	0.01029428	0.8789089,0.9181677	0.8973348,0.898984

From the simulation results, it is observed that as the sample size (n, m) increases the biases and the MSE decrease. Thus the consistency properties of all the methods are verified.

6. Data Analysis

In this section we present a data analysis of the strength data reported by Badar and Priest(1982). The data represent the strength data measured in GPA, for single carbon fibers and impregnated 1000-carbon fiber tows. Single fibers were tested under tension at gauge lengths of 1, 10, 20 and 50 mm. Impregnated tows of 1000 fibers were tested at gauge lengths of 20, 50, 150 and 300 mm. It is already observed that the Weibull model does not work well in this case. Surles and Padgett[16], [17] and Raqab and Kundu[12] observed that generalized Rayleigh works quite well for these strength data. For illustrative purposes we are also considering the same transformed data sets as it was considered by Raqab and Kundu [9], the single fibers of 20 mm (Data Set I) and 10 mm (Data Set II) in gauge lengths with sample sizes 69 and 63 respectively. They are presented below:
Data Set I: 0.312 0.314 0.479 0.552 0.700 0.803 0.861 0.865 0.944 0.958 0.966 0.997 1.006 1.021 1.027 1.055 1.063 1.098 1.140 1.179 1.224 1.240 1.253 1.270 1.272 1.274 1.301 1.301 1.359 1.382 1.382 1.426 1.434 1.435 1.478 1.490 1.511 1.514 1.535 1.554 1.566 1.570 1.586 1.629 1.633 1.642 1.648 1.684 1.697 1.726 1.770 1.773 1.800 1.809 1.818 1.821 1.848 1.880 1.954 2.012 2.067 2.084 2.090 2.096 2.128 2.233 2.433 2.585 2.585
Data Set II: 0.101 0.332 0.403 0.428 0.457 0.550 0.561 0.596 0.597 0.645 0.654 0.674 0.718 0.722 0.725 0.732 0.775 0.814 0.816 0.818 0.824 0.859 0.875 0.938 0.940 1.056 1.117 1.128 1.137 1.137 1.177 1.196 1.230 1.325 1.339 1.345 1.420 1.423 1.435 1.443 1.464 1.472 1.494 1.532 1.546 1.577 1.608 1.635 1.693 1.701 1.737 1.754 1.762 1.828 2.052 2.071 2.086 2.171 2.224 2.227 2.425 2.595 3.220
These data were first used by Badar and Priest(1982) and later used by Raqab and Kundu[13] and Hassan and Kumaraswamy [7].

Table 2: *Data Analysis Results*

Length(in mm)	MLE	K-S statistic	P-value
10	1.596362	0.62319	0.7571
20	1.715981	0.79365	0.4375

Table 6.1 gives the result of the goodness of fit test. Maximum likelihood estimates are 1.596362 and 1.715981.The estimated value of the stress strength reliability using these estimates is 0.8850866.The 95% asymptotic confidence interval for R is (0.8530591,0.9171141) and 95% bootstrap confidence interval for R is (0.8748553 ,0.8895613).

7. Conclusion

In this paper, we considered the problem of estimating stress strength reliability using Pranav distribution. The maximum likelihood estimate of stress strength reliability, $\hat{R}$ is obtained. Also, asymptotic $100(1-\alpha)\%$ confidence interval for the reliability parameter is computed. Bootstrap confidence interval for the reliability parameter is also computed. When the sample size is increased, mean square error caused by the estimates comes nearer to zero by extensive simulation. Finally, real data sets are analyzed

References

[1] KE Ahmad, ME Fakhry, and ZF Jaheen. Empirical bayes estimation of $p(y < x)$ and characterizations of burr-type x model. *Journal of Statistical Planning and Inference*, 64(2):297–308, 1997.
[2] Adnan M Awad and Mohamed K Gharraf. Estimation of $p(y < x)$ in the burr case: A comparative study. *Communications in Statistics-Simulation and Computation*, 15(2):389–403, 1986.

[3] James D Church and Bernard Harris. The estimation of reliability from stress-strength relationships. *Technometrics*, 12(1):49–54, 1970.
[4] Kenneth Constantine, Siu-Keung Tse, and Marvin Karson. Estimation of p (y< x) in the gamma case. *Communications in Statistics-Simulation and Computation*, 15(2):365–388, 1986.
[5] F Downton. The estimation of $pr(y < x)$ in the normal case. *Technometrics*, 15(3):551–558, 1973.
[6] Rameshwar D Gupta and Debasis Kundu. Closeness of gamma and generalized exponential distribution. *Communications in statistics-theory and methods*, 32(4):705–721, 2003.
[7] Marwa KH Hassan. The kumaraswamy exponential-weibull distribution and its application in reliability. *International Journal of Reliability and Applications*, 19(2):95–108, 2018.
[8] Shukla KK. Pranav distribution with properties and its applications. *Biom Biostat Int J*, 7(3):244–254, 2018.
[9] Debasis Kundu and Rameshwar D Gupta. Estimation of $p[y < x]$ for generalized exponential distribution. *Metrika*, 61(3):291–308, 2005.
[10] Debasis Kundu, Nandini Kannan, and N Balakrishnan. Analysis of progressively censored competing risks data. *Handbook of statistics*, 23:331–348, 2003.
[11] Debasis Kundu and Mohammad Z Raqab. Estimation of $r = p(y < x)$ for three-parameter weibull distribution. *Statistics & Probability Letters*, 79(17):1839–1846, 2009.
[12] Mohammad Z Raqab and Debasis Kundu. Comparison of different estimators of $p[y < x]$ for a scaled burr type x distribution. *Communications in Statistics"Simulation and Computation®*, 34(2):465–483, 2005.
[13] Mohammad Z Raqab, Mohamed T Madi, and Debasis Kundu. Estimation of $p(y < x)$ for the three-parameter generalized exponential distribution. *Communications in Statistics"Theory and Methods*, 37(18):2854–2864, 2008.
[14] Buğra Saraçoğlu, Ismail Kinaci, and Debasis Kundu. On estimation of $r = p(y < x)$ for exponential distribution under progressive type-ii censoring. *Journal of Statistical Computation and Simulation*, 82(5):729–744, 2012.
[15] Fatemeh Shahsanaei and Alireza Daneshkhah. Estimation of stress-strength model in the generalized linear failure rate distribution. *arXiv preprint arXiv:1312.0401*, 2013.
[16] JG Surles and WJ Padgett. Inference for $p(y < x)$ in the burr type x model. *Journal of Applied Statistical Science*, 7(4):225–238, 1998.
[17] JG Surles and WJ Padgett. Inference for reliability and stress-strength for a scaled burr type x distribution. *Lifetime data analysis*, 7:187–200, 2001.
[18] Akhila K Varghese and VM Chacko. Estimation of stress-strength reliability for akash distribution. *Reliability: Theory & Applications*, 17(3 (69)):52–58, 2022.

FAST AND ROBUST BIVARIATE CONTROL CHARTS FOR INDIVIDUAL OBSERVATIONS

Sajesh T A

Department of Statistics, St. Thomas College (Autonomous), Thrissur, Kerala, India
*sajesh.t.abraham@gmail.com

Abstract

There are various circumstances where it is important to simultaneously monitor or control two or more related quality characteristics. Independently tracking these quality characteristics might be quite deceptive. Hotelling's T2 chart, in which the T2 statistics are generated using the classical estimates of location and scatter, is the most well-known multivariate process monitoring and control approach. It is well known that the existence of outliers in a dataset has a significant impact on classical estimators. Any statistic that is computed using the classical estimates will be distorted by even a single outlier. The non-robustness issue is investigated in this study, which also suggests four robust bivariate control charts based on the robust Gnandesikan-Kettenring estimator. This study employs four highly robust scale estimators, with the best breakdown point, namely the Qn estimator, Sn estimator, MAD estimator, and τ estimator, in order to robustify the Gnandesikan- Kettenring estimator. Through the use of a Monte Carlo simulation and a real-life data, the performance of the suggested control charts is assessed. The four techniques all outperform the traditional method and provide greater computing efficiency.

Keywords: Gnandesikan- Kettenring estimator, Qn estimator, Sn estimator, MAD, τ estimator.

1. Introduction

Bivariate control charts are specifically designed for situations where two variables are observed simultaneously. It enables the detection of patterns or trends that signal a shift or alteration in the process. There are two separate phases, namely Phase I and Phase II when constructing the control chart [1]. Historical data is used in Phase I to determine control limits, estimate the unknown parameters of the in-control process, and evaluate the process' stability. Phase II involves applying the estimated parameters and control limits discovered in Phase I to the data gathered during the actual production process in order to analyse it and find any deviations or out-of-control signals. Phase II's goal is to keep track of and maintain the process' stability in accordance with the defined control limits.

The most frequently used multivariate control chart for monitoring the variability of a multivariate industrial process is the Hotelling's T^2 control chart. Let $\boldsymbol{x}_i$, i = 1, 2, …, n be a two-dimensional vector of measurements made on a process at the time period i, then for the sample $\boldsymbol{x} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$, the Hotelling's T^2 statistic is defined as

$$T^2(\boldsymbol{x}_i) = (\boldsymbol{x}_i - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}_i - \boldsymbol{\theta}). \quad (1)$$

It is assumed that, when the process is in statistical control, $\boldsymbol{x}_i$'s are independent bivariate normal random vectors with mean vector $\boldsymbol{\theta}$ and covariance matrix. If both $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ are known, T_i^2 follows a Chi-square distribution with 2 degrees of freedom. When the population parameters $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ are unknown, the T^2 statistic is constructed using the classical estimators of mean vector ($\bar{\boldsymbol{x}}$) and covariance matrix ($\boldsymbol{S}$) as given below,

$$T^2(\boldsymbol{x}_i) = (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}), \quad (2)$$

where $\bar{\boldsymbol{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}$ and $\boldsymbol{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, S_{12} = Cov(X_1, X_2)$. When constructing control charts using phase I data, the classical estimators, which are unfortunately highly susceptible to the influence of outliers, can produce inaccurate findings (known as the masking problem). Researchers have suggested several methods in the literature to lessen the negative effects of outliers in response to this problem. The control charts can be made more robust and reliable in the presence of outliers by using robust estimators. Through the use of robust estimators, which are more resistant to the presence of outliers, the conventional estimators are swapped out in these alternative methods.

This paper proposes a robust bivariate control chart that can effectively handle spurious outliers. The proposed control chart makes use of the covariance estimator introduced by Gnanadesikan and Kettenring [2], which is based on the identity

$$\mathrm{cov}(X, Y) = \frac{1}{4}(\sigma(X + Y)^2 - \sigma(X - Y)^2), \quad (3)$$

where σ is the standard deviation and X, Y is a pair of random variables. By replacing σ by a robust scale estimator, one can easily robustify the Gnanadesikan and Kettenring (GK) estimator. The robust GK estimator is defined as

$$\mathrm{cov_R}(X, Y) = \frac{1}{4}(s_R(X + Y)^2 - s_R(X - Y)^2), \quad (4)$$

where s_R is a robust scale estimator. This study has considered four robust scale estimators with an optimal breakdown point to robustify the GK estimators. These robust GK estimators were then used for the construction of bivariate control charts. The Lower control limits (LCL) of these control charts are set to zero and the Upper Control limits (UCL) were estimated by fitting the quantiles. The performance of these control charts is examined and compared with that of classical bivariate control chart through Monte Carlo simulation.

2. Robust scale estimators

For a variety of applications, from genuine scale problems to outlier identification, and as auxiliary factors for more complicated analysis, robust estimates of scale are crucial. The broad population of users of statistical methods seems to have a somewhat lower level of acceptance for robust estimation of scale. Previously, the interquartile range, which has a breakdown point of 25%, was the only robust scale estimator to be found in the majority of statistical software packages.

i) Median Absolute Deviation (MAD)

A preliminary or auxiliary estimate of scale is frequently required in robust estimation. A very robust scale estimator is the median absolute deviation about the median (MAD), given by

$$\mathrm{MAD} = b\ \mathrm{med}_i\{|x_i - \mathrm{med}_j(x_j)|\}, \quad (4)$$

where 'med' denotes median. The MAD has the best possible breakdown point (50%, twice as much as the interquartile range), and its influence function is bounded, with the sharpest possible bound among all scale estimators. The MAD was first promoted by Hampel [3], who attributed it to Gauss. The constant b in equation (4) is needed to make the estimator consistent for the parameter of interest. In the case of the usual parameter a at Gaussian distributions, we need to set b = 1.4826. In spite of many advantages, the MAD also has some drawbacks. Its efficiency at Gaussian distributions is very low;

whereas the location median's asymptotic efficiency is still 64%, the MAD is only 37% efficient. Also, it takes a symmetric view on dispersion.

ii) τ Estimator

Yohai and Zamar [4] introduced a new class of robust scale estimates called τ estimates which possesses optimal breakdown value. Let ρ be a real function satisfying the following properties:

1. $\rho(0) = 0$.
2. $\rho(-u) = \rho(u)$.
3. $0 \leq u \leq v$ implies that $\rho(u) \leq \rho(v)$.
4. ρ is continuous.
5. Let $a = \text{Sup } \rho(u)$; $0 < a < \infty$
6. If $\rho(u) < a$ and $0 \leq u < v$, then $\rho(u) < \rho(v)$.
7.

Let ρ_1 an ρ_2 bet two functions satisfying above assumptions. Then for a given sample $\boldsymbol{u} = (u_1, \ldots, u_n)$ the τ estimate for scale is defined as

$$\tau^n(\boldsymbol{u}) = s^2(\boldsymbol{u}) \frac{1}{n} \sum_{i=1}^{n} \rho_2 \left(\frac{u_i}{s(\boldsymbol{u})} \right), \tag{5}$$

where s be a M estimate of scale based on ρ_1. This estimator possesses approximately 80% efficiency when c = 3. Moreover, τ estimate is asymptotically normal and has bounded influence function. Maronna and Zamar [5] used this estimate for introducing a multivariate outlier detection technique in which they have considered s = MAD and $\rho_2(x) = \min(x^2, c^2)$. This study also used the considerations of Maronna and Zamar for s and ρ_2.

iii) S_n and Q_n Estimator

To address the lower efficiency drawback of MAD, Rousseeuw and Croux [6] introduced two robust scale estimators with optimal breakdown value of 50%, namely, S_n estimator and Q_n estimator. S_n estimator is defined as

$$S_n = c \operatorname*{med}_i \{ \operatorname*{med}_j |x_i - x_j| \}. \tag{6}$$

The factor c is for consistency, and its default value is 1.1926. Moreover, the asymptotic efficiency S_n is 58.23% which is much higher than MAD. A drawback of MAD, and S_n, is that their influence functions have discontinuities. The Q_n estimator is solution for this drawback. It is defined as

$$Q_n = d\{|x_i - x_j|; i < j\}_{(k)}, \tag{7}$$

where d is a constant factor and $k = \binom{h}{2} \approx \frac{\binom{n}{2}}{4}$, where $h = \left[\frac{n}{2}\right] + 1$ is roughly half of the observations. The estimator Q_n, shares the attractive properties of S_n simple and explicit formula, a definition that is equally suitable for asymmetric distributions, and a 50% breakdown point. In addition, we will see that its influence function is smooth, and that its efficiency at Gaussian distributions is very high (about 82%). Rousseeuw and Croux [6] showed that, although Q_n is more efficient, S_n is more preferable in most of the applications because of its low gross-error sensitivity.

3. Proposed Bivariate Robust Control Charts

Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be a set of Phase I data follows bivariate normal distribution with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}$. Let $\boldsymbol{y} \notin \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$be a Phase II observation, then it is known that

$$T^2(\boldsymbol{y}) \sim \left[\frac{2(n^2-1)}{n(n-2)}\right] F_{(2,n-2)} \tag{8}$$

where $T^2(\boldsymbol{y})$ is as defined in Equation (1) and $F_{(v_1,v_2)}$ is F distribution with (v_1, v_2) degrees of freedom [7]. Since this statistic is not robust to the presence of outliers, in the proposed robust control chart we use robust Hotelling's T^2 statistic, denoted by T_R^2. It is obtained by replacing the classical estimates used for the computation of Hotelling's T^2 statistic by their robust counterpart. Suppose $\boldsymbol{x}_{med}$ and $\boldsymbol{S}_{GK}$ represent the component wise median vector and robust GK covariance matrix, respectively. We define a robust Hotelling's T^2 for $\boldsymbol{y}$ based on these estimates by

$$T_R^2(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{x}_{med})^T \boldsymbol{S}_{GK}^{-1} (\boldsymbol{y} - \boldsymbol{x}_{med}), \tag{9}$$

where $\boldsymbol{S}_{GK} = \begin{bmatrix} s_R(X_1)^2 & cov_R(X_1,X_2) \\ cov_R(X_1,X_2) & s_R(X_2)^2 \end{bmatrix}$. By applying Slutsky theorem [8], the asymptotic distribution of T_R^2 can be obtained as Chi square distribution with 2 degrees of freedom. As $n \to \infty$

$$(\boldsymbol{y} - \boldsymbol{x}_{med})^T \boldsymbol{S}_{GK}^{-1} (\boldsymbol{y} - \boldsymbol{x}_{med}) \xrightarrow{D} (\boldsymbol{y} - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{\theta}) \sim \boldsymbol{\chi}^2_{(2)}. \tag{10}$$

This asymptotic distribution, though, only holds true for large sample sizes. We employ Monte Carlo simulations to estimate quantiles for different sample sizes in order to determine the control limits for the suggested control charts. The sample size and quantiles of T_R^2 were then fitted with a smooth curve. For modest Phase I sample sizes, these fits can be utilised to determine the proper control limits of the proposed control charts.

3.1. Estimation of Control Limits for the proposed control charts

Upper control limits of the proposed control charts are obtained by modelling the quantiles of T_R^2, for a given Phase I sample size n, computed from N=10000 trials. Phase I samples are generated from a standard bivariate normal distribution $N_2(\boldsymbol{0}, \boldsymbol{I})$ and 99%, 95%, 99.73% and 99.9% quantiles of T_R^2 are computed using S_n, Q_n, MAD and τ estimates. In each trial, for each data set, we also generate a new random observation $\boldsymbol{y}_i$ from $N_2(\boldsymbol{0}, \boldsymbol{I})$ (treated as a Phase II observation) and calculate the corresponding $T_R^2(\boldsymbol{y}_i)$ value using robust GK covariance estimates. By inverting the empirical distribution function of $T_R^2(\boldsymbol{y}_i)$, computed for n = 10, 15, 20, . . . ,500, we obtain Monte Carlo estimates of the 99%, 95%, 99.73% and 99.9% quantiles.

Scatter plots of the empirical quantiles of $T_R^2(\boldsymbol{y}_i)$ versus the sample size n suggest that we could model the quantiles using a family of regression curves of the form $f(n) = a + \frac{b}{n^c}$. Scatter plots of the empirical 99% quantiles of $T_R^2(\boldsymbol{y}_i)$ computed using the four robust scale estimates are shown in figure 1. Since the T_R^2 statistic asymptotically follow $\chi^2_{(2)}$ distribution, following two parameter family of curves is used for both robust GK control charts:

$$f_{1-\alpha}(n) = \boldsymbol{\chi}^2_{(2,1-\alpha)} + \frac{b_{1-\alpha}}{n^{c_{1-\alpha}}}, \tag{11}$$

where $\boldsymbol{\chi}^2_{(2,1-\alpha)}$ is the $1-\alpha$ quantile of the χ^2 distribution with 2 degrees of freedom and $b_{1-\alpha}$ and $c_{1-\alpha}$ are constants with overall false alarm probability α. Fitting this curve to the data will help us to estimate the desired upper control limits of the proposed control charts for any Phase I sample of size n. Note that, as n increases, $f_{1-\alpha}(n)$ approaches $\chi^2_{(2,1-\alpha)}$. Table 1 gives the least-square estimates of the parameters $b_{1-\alpha}$ and $c_{1-\alpha}$. Using Table 1 and Equation (11), we can compute the 99%, 95%, 99.73% and 99.9% quantiles of $T_R^2(\boldsymbol{y}_i)$ for Phase I sample size n. The regression curves given by Equation (9) fit well to all the cases in Table 1, yielding R^2 values of at least 90.7%.

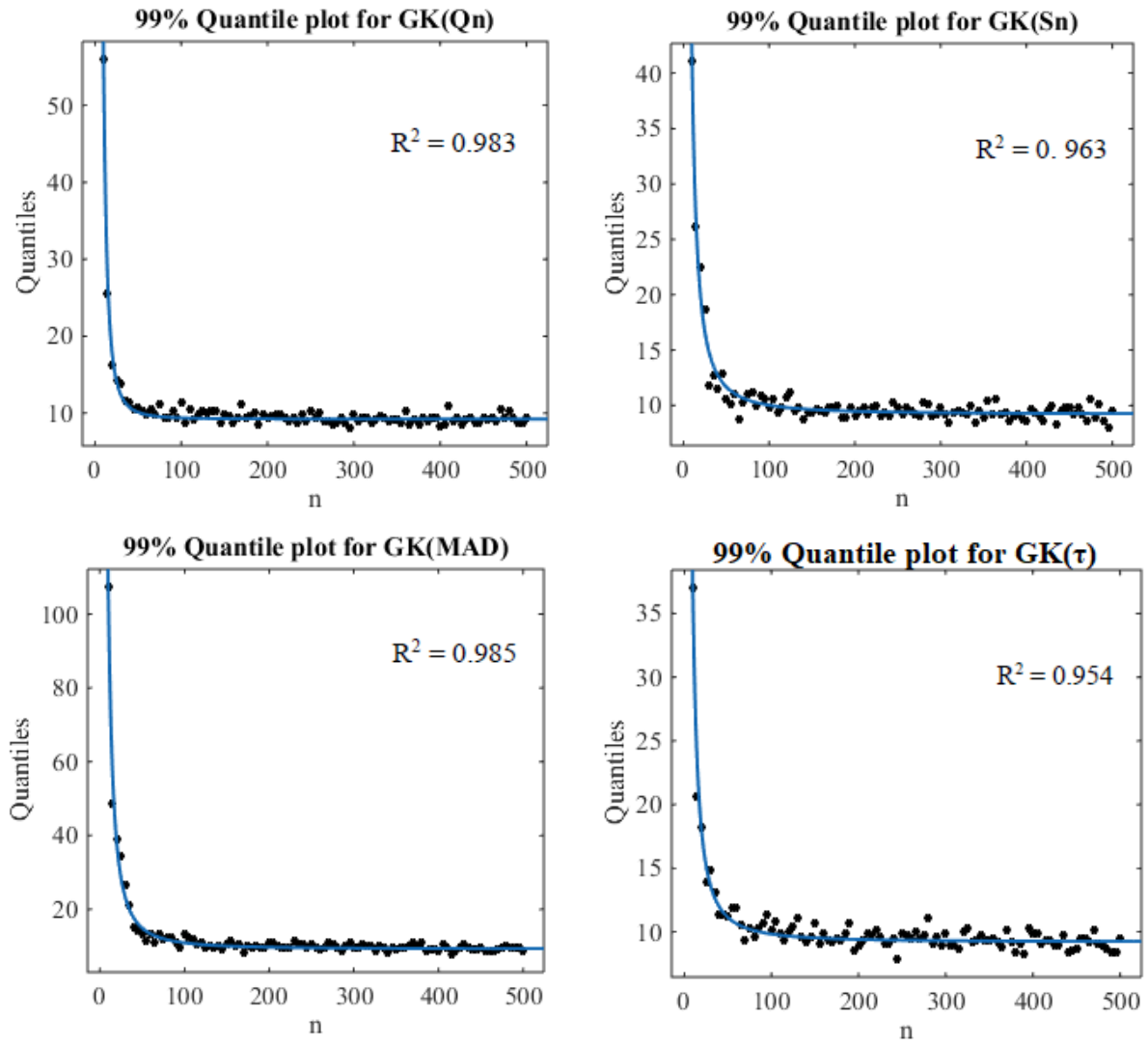


Figure 1: *Simulated Quantiles of T^2_{GK} and the Fitted Curves for $\alpha = 0.01$*

Table 1: *The Least-Squares Estimates of the Regression Parameters $b_{1-\alpha}$ and $c_{1-\alpha}$ for Confidence Levels $1 - \alpha = 0.99, 0.95, 0.9973$ and 0.999.*

α	Parameters	GK(Q_n)	GK(S_n)	GK(MAD)	GK(τ)
0.01	$b_{1-\alpha}$	13120.00	1281.00	5496.00	1145.00
	$c_{1-\alpha}$	2.451	1.598	1.758	1.632
	R^2	0.983	0.963	0.985	0.954
0.05	$b_{1-\alpha}$	228.80	249.50	462.60	324.20
	$c_{1-\alpha}$	1.437	1.399	1.373	1.447
	R^2	0.916	0.930	0.976	0.942
0.0027	$b_{1-\alpha}$	120600.00	11440000.00	721800.00	2610.00
	$c_{1-\alpha}$	2.945	4.66	3.22	1.656
	R^2	0.987	0.990	0.990	0.907
0.001	$b_{1-\alpha}$	556600.00	26090.00	41240.00	3973.00
	$c_{1-\alpha}$	3.526	1.947	1.769	1.826
	R^2	0.968	0.920	0.950	0.930
0.1	$b_{1-\alpha}$	31530.00	248.30	219.4	118.60
	$c_{1-\alpha}$	1.65	1.574	1.31	1.281
	R^2	0.951	0.919	0.967	0.947

3.2. Computational algorithm for obtaining the proposed control chart

A step-by-step approach for constructing proposed robust control chart is given as follows:

Phase I
1. Select the confidence level $1 - \alpha$ and sample size n.
2. Collect the Phase I data $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ at predetermined periodic intervals and compute robust estimates of location and scale parameters using this data.
3. Use Table 1 to select the least square estimates of the parameters $b_{1-\alpha}$ and $c_{1-\alpha}$ for the desired α, then use equation (11) to determine the upper control limit.

Phase II
4. Compute T_R^2 for each of the new observation as per Equation (11) and plot it on a control chart with the control limits determined in Phase I.
5. Analyse the chart and look for any out-of-control points or non-random patterns.

4. Performance of the Proposed Control Charts

We carried out numerous simulations to assess and compare the efficiency of the proposed control charts under normal and contaminated situations. The control chart's efficacy is assessed by analysing its ability to detect changes and the rate of false detections in the process behaviour using different estimators in Phase I. We measure the performance of the control chart by *success rate* (SR)- which is the proportion of statistic values that exceed the control limits across 1000 replications, which provides an estimate of the likelihood of detecting changes and *false alarm rate* (FAR) – which is the false detections in the process behaviour based on the Phase II data. Phase I data are generated using the following contaminated model:

$$(1 - \varepsilon)\, N_2(\boldsymbol{\theta}_0, \boldsymbol{\Sigma}_0) + \varepsilon\, N_2(\boldsymbol{\theta}_1, \boldsymbol{\Sigma}_1),$$

where ε is the proportion of outliers, $\boldsymbol{\theta}_0$ and $\boldsymbol{\Sigma}_0$ are the in-control parameters and $\boldsymbol{\theta}_1$ and $\boldsymbol{\Sigma}_1$ are the out-of-control parameters of location and scatter. Without loss of generality, $\boldsymbol{\theta}_0$ is set to be a zero vector. We have generated the phase I data sets for n = 25, 50, 100 and 1000, ε = 0, 0.1 and 0.2 and α = 0.001, 0.0027, 0.01, 0.05 and 0.1. The probability of detecting a change depends on the values of $\boldsymbol{\theta}_1$, $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ and hence we consider three different contaminated models in which the values of $\boldsymbol{\theta}_1$, $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ vary.

Case A: Independent Variables

In this case, the two variables (Quality Characteristics) $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are assumed to be independent. The contaminated normal model considered is as follows:

$$(1 - \varepsilon)\, N_2(0, \boldsymbol{I}_2) + \varepsilon\, N_2(\boldsymbol{\theta}_1, \boldsymbol{I}_2),$$

where I_2 is the identity matrix of size 2. In this case, we compare the behaviour of different robust alternatives when there are different-sized changes in the average of all the variables if the variables are independent.

Case B: Correlated Variables

In this case, two variables, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are assumed to be correlated. The contaminated normal model considered is as follows:

$$(1 - \varepsilon)\, N_2(\mathbf{0}, \boldsymbol{\Sigma}_0) + \varepsilon\, N_2(\boldsymbol{\theta}_1, \boldsymbol{\Sigma}_0),$$

where $\boldsymbol{\Sigma}_0 = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$. We have used this value of $\boldsymbol{\Sigma}_0$ to analyse whether the correlation level affects the detection probability of each alternative.

Table 2: *SR (FAR) obtained for Case A with different values of* $\boldsymbol{\theta}_1$ *and* ε

n	$\boldsymbol{\theta}_1$	ε	GK(Q_n)	GK(S_n)	GK(MAD)	GK(τ)	Classical
50	(0,0)	0	100 (1.9)	100 (1.0)	99.8 (0.8)	100 (1.3)	100 (1.3)
		10	100 (1.4)	100 (1.0)	99.9 (0.4)	100 (1.1)	100 (1.1)
		20	100 (1.3)	100 (0.9)	99.8 (0.6)	100 (0.6)	100 (0.6)
	(5,5)	0	100 (2.0)	100 (1.6)	99.7 (1.3)	100 (1.2)	100 (1.5)
		10	98.8 (0.3)	98.8 (0.3)	97.7 (0.4)	98.6 (0.1)	22.1 (0.2)
		20	86.2 (0.0)	80 (0.0)	76.8 (0.0)	74.3 (0.1)	2 (0.2)
	(10,10)	0	100 (1.1)	99.9 (0.3)	99.7 (0.5)	100 (1.7)	100 (0.8)
		10	99.4 (0.3)	99.1 (0.1)	97.9 (0.3)	99.1 (0.4)	0.6 (0.4)
		20	88.3 (0.0)	80.8 (0.0)	79.8 (0.1)	75.8 (0.2)	0.4 (0.3)
100	(0,0)	0	100 (1.2)	100 (1.0)	100 (1.0)	100 (1.2)	100 (1.2)
		10	100 (1.5)	100 (1.4)	100 (1.3)	100 (1.3)	100 (1.3)
		20	100 (1.4)	100 (1.3)	100 (1.2)	100 (0.7)	100 (0.7)
	(5,5)	0	100 (0.6)	100 (0.7)	100 (0.9)	100 (1.5)	100 (0.6)
		10	99.7 (0.2)	99.6 (0.1)	99.8 (0.3)	99.8 (0.4)	30.2 (0.4)
		20	92.5 (0.0)	90.4 (0.0)	93.5 (0.0)	83.9 (0.0)	2.5 (0.1)
	(10,10)	0	100 (1.5)	100 (1.2)	99.9 (0.7)	100 (0.7)	100 (1.1)
		10	99.8 (0.2)	99.8 (0.3)	99.8 (0.5)	99.6 (0.4)	0.1 (0.2)
		20	91.1 (0.1)	89.4 (0.1)	92.5 (0.1)	85.8 (0.2)	0.4 (0.4)
500	(0,0)	0	100 (1.9)	100 (1.9)	100 (1.8)	100 (1.8)	100 (1.8)
		10	100 (1.5)	100 (1.4)	100 (1.4)	100 (1.4)	100 (1.4)
		20	100 (0.8)	100 (0.9)	100 (0.9)	100 (0.8)	100 (0.8)
	(5,5)	0	100 (1.4)	100 (1.2)	100 (1.4)	100 (1.2)	100 (1.2)
		10	99.8 (0.0)	99.8 (0.0)	99.9 (0.0)	99.8 (0.1)	34.5 (0.0)
		20	94.8 (0.0)	95.5 (0.0)	98.1 (0.0)	89.3 (0.0)	2.4 (0.3)
	(10,10)	0	100 (0.6)	100 (0.7)	100 (0.8)	100 (1.8)	100 (0.6)
		10	99.8 (0.1)	99.9 (0.1)	99.9 (0.2)	99.9 (0.2)	0.7 (0.1)
		20	94.6 (0.0)	95.1 (0.0)	97.8 (0.1)	88.1 (0.0)	0.4 (0.2)
1000	(0,0)	0	100 (1.0)	100 (0.9)	100 (0.8)	100 (0.9)	100 (0.9)
		10	100 (0.8)	100 (0.8)	100 (0.9)	100 (0.8)	100 (0.8)
		20	100 (0.6)	100 (0.5)	100 (0.5)	100 (0.7)	100 (0.7)
	(5,5)	0	100 (0.8)	100 (0.8)	100 (1.0)	100 (0.7)	100 (0.6)
		10	99.9 (0.2)	99.9 (0.2)	100 (0.3)	100 (0.0)	34.9 (0.3)
		20	94.9 (0.0)	95.8 (0.1)	98.3 (0.1)	89.5 (0.0)	2 (0.4)
	(10,10)	0	100 (1.8)	100 (1.9)	100 (1.7)	100 (0.8)	100 (1.8)
		10	99.8 (0.1)	99.8 (0.1)	100 (0.2)	99.8 (0.1)	0.8 (0.2)
		20	93.8 (0.0)	94.7 (0.0)	98.3 (0.1)	89 (0.0)	0.3 (0.2)

Case C: Correlated Variables and Regression Outliers

Here, the two variables $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are assumed to be correlated and regression outliers are introduced. The contaminated normal model considered is as follows:

$$(1 - \varepsilon)\, N_2(\mathbf{0}, \boldsymbol{\Sigma}_0) + \varepsilon\, N_2(\boldsymbol{\theta}_1, \boldsymbol{\Sigma}_1),$$

where $\Sigma_0 = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$ and $\Sigma_1 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$. In this case, we analyse and compare the proposed robust methods in terms of the so-called good leverage and regression outliers.

In all the three cases, we consider $\boldsymbol{\theta}_1$ as a vector of size 2 where the elements are all 0 (when there is no change), 5 or 10 (which is a good leverage point). This process is repeated 1000 times, and in each trial, a random observation, $\boldsymbol{z}_{1k}$ from $N_2(\mathbf{0}, \boldsymbol{\Sigma}_U)$ and another observation, $\boldsymbol{z}_{2k}$ from $N_2(\boldsymbol{\theta}_C, \boldsymbol{\Sigma}_C)$ are generated, for k = 1, 2, …, 1000. Here, $\boldsymbol{\Sigma}_U$ is the scale estimator used for generating uncontaminated observations in Phase I data, and $\boldsymbol{\theta}_C$ and $\boldsymbol{\Sigma}_C$ are the location and scale estimates used for generating contaminated observations in Phase I model. The success rates are computed as the percentages of $\boldsymbol{z}_{2k}$'s that are successfully detected, and the false alarm rates are computed as the percentages of $\boldsymbol{z}_{1k}$'s that are falsely detected. The results obtained for α = 0.01 is presented here.

Table 3: *SR (FAR) obtained for Case B with different values of* $\boldsymbol{\theta}_1$ *and* ε

n	$\boldsymbol{\theta}_1$	ε	GK(Q_n)	GK(S_n)	GK(MAD)	GK(τ)	Classical
50	(0,0)	0	95.6 (4.1)	90.6 (4.6)	84.6 (2.7)	97 (2.8)	97.4 (0.8)
		10	95.6 (4.3)	91.4 (3.2)	84.6 (2.7)	96.3 (2.7)	96.3 (0.9)
		20	95.5 (3.4)	90.9 (4.8)	82.4 (1.7)	95.7 (2.6)	96.4 (1.4)
	(5,5)	0	94.7 (4.5)	91.2 (3.8)	83.6 (3.0)	96.9 (1.4)	97 (0.9)
		10	84.7 (3.0)	77.1 (2.9)	67.2 (1.7)	79.7 (3.1)	17.9 (0.4)
		20	57 (2.3)	48.2 (1.9)	40.5 (1.7)	40.2 (1.1)	2.3 (0.2)
	(10,10)	0	95.3 (3.4)	91.2 (3.5)	82.4 (2.8)	97.1 (2.5)	98.3 (1.2)
		10	84.4 (3.9)	76.2 (1.9)	67.4 (2.6)	80.3 (2.6)	0.7 (0.2)
		20	57.1 (2.5)	46.3 (1.3)	43.8 (0.8)	41.1 (2.1)	0.8 (0.4)
100	(0,0)	0	98.6 (3.3)	96.3 (4.7)	94.2 (3.9)	98.6 (2.4)	98.4 (1.3)
		10	97.9 (2.0)	95 (3.7)	92.2 (3.4)	97.6 (2.2)	97.4 (1.1)
		20	98.5 (1.8)	95.8 (3.4)	92.4 (3.5)	97.9 (0.9)	97.8 (0.4)
	(5,5)	0	98.5 (2.7)	96 (4.8)	93.2 (3.6)	98.1 (0.9)	98.1 (1.0)
		10	90.4 (1.5)	86.9 (2.6)	86.7 (3.1)	85.8 (2.0)	23.5 (0.9)
		20	63.3 (0.9)	56.8 (2.0)	61.4 (2.7)	49.6 (1.9)	2.8 (0.2)
	(10,10)	0	98.8 (2.1)	96 (3.7)	93 (3.6)	98.5 (1.7)	98.7 (1.1)
		10	91.2 (0.7)	87.4 (2.2)	87.6 (2.6)	88.2 (2.1)	0.8 (0.3)
		20	62.4 (0.7)	57.5 (2.5)	60.1 (2.0)	51.2 (1.6)	0.7 (0.6)
500	(0,0)	0	99.2 (1.0)	99 (2.2)	98.8 (2.8)	99.3 (1.2)	99.2 (1.0)
		10	99.1 (1.1)	98.9 (2.7)	98.1 (3.5)	99 (1.1)	99 (1.1)
		20	98.6 (1.3)	98.6 (2.6)	97.8 (3.5)	98.7 (1.5)	98.6 (1.5)
	(5,5)	0	98.6 (1.1)	98.7 (2.9)	98.3 (3.9)	99 (1.2)	98.4 (1.1)
		10	90.8 (0.7)	91.9 (2.6)	94 (3.2)	90 (0.7)	25.6 (0.6)
		20	62.4 (0.0)	63.8 (0.8)	74.8 (1.6)	53.5 (0.8)	3.1 (0.3)
	(10,10)	0	98.3 (1.7)	98.2 (2.2)	97.6 (3.8)	98.5 (1.4)	98.4 (1.7)
		10	90.9 (0.1)	92.4 (1.5)	93.1 (2.5)	91.1 (0.4)	0.7 (0.2)
		20	64.6 (0.0)	66.6 (0.9)	77.1 (1.4)	55 (0.3)	0.6 (0.4)
1000	(0,0)	0	99 (0.2)	98.9 (0.5)	98.7 (1.3)	99 (0.3)	98.9 (0.2)
		10	98.1 (1.3)	98.1 (1.6)	97.8 (1.7)	98 (1.1)	97.9 (1.1)
		20	99.1 (0.6)	98.9 (1.6)	98.8 (2.1)	99.1 (0.8)	98.9 (0.7)
	(5,5)	0	98.8 (1.3)	98.7 (1.5)	98.7 (2.6)	98.7 (1.2)	98.8 (1.1)
		10	91.3 (0.0)	92.4 (0.5)	93.9 (2.0)	91.8 (0.5)	25.8 (0.2)
		20	64.7 (0.0)	66.7 (0.1)	76.4 (1.1)	53.4 (0.3)	3.2 (0.6)
	(10,10)	0	98.3 (1.6)	98.1 (2.1)	98 (2.7)	98.7 (1.2)	98.3 (1.6)
		10	90.1 (0.1)	91.8 (0.4)	93.8 (1.4)	90.8 (0.4)	0.5 (0.2)
		20	64.8 (0.0)	67.2 (0.0)	77.3 (1.1)	53.1 (0.3)	0.7 (0.1)

Table 2 presents the results obtained from simulations conducted for Case A samples. The findings clearly indicate that when the phase I data contains outliers, the control charts based on robust methods exhibit a high success rate with minimal FAR compared to the classical control chart. Furthermore, the success rates of the proposed control charts improve with increasing sample size. Conversely, when the phase I data is uncontaminated, the robust control charts demonstrate similar performance to their classical counterpart.

Table 4: *SR (FAR) obtained for Case C with different values of* $\boldsymbol{\theta}_1$ *and* ε

n	$\boldsymbol{\theta}_1$	ε	GK(Q_n)	GK(S_n)	GK(MAD)	GK(τ)	Classical
50	(0,0)	0	97.9 (4.7)	96.2 (4.9)	91.6 (2.3)	99.6 (2.4)	100 (1.2)
		10	98.8 (3.8)	97.1 (3.1)	95.3 (2.4)	99.3 (2.7)	100 (1.3)
		20	99 (4.1)	98.8 (4.1)	96.5 (3.9)	99.1 (4.3)	100 (1.8)
	(5,5)	0	98.3 (3.7)	96.8 (3.3)	90 (2.7)	99.2 (1.9)	100 (0.6)
		10	95.1 (2.1)	85.6 (2.8)	75.9 (2.1)	91.5 (3.7)	8.4 (0.2)
		20	78.6 (0.2)	45.9 (1.3)	38 (1.1)	36.7 (2.0)	0.9 (0.3)
	(10,10)	0	98.8 (5.2)	94.9 (4.0)	90.8 (2.9)	99.1 (2.6)	100 (1.0)
		10	96.3 (3.2)	87.8 (2.5)	77.4 (1.7)	92.7 (2.9)	0.8 (0.6)
		20	76.6 (1.7)	47.6 (1.4)	37.1 (1.0)	38 (1.5)	0.6 (0.4)
100	(0,0)	0	100 (2.5)	98.2 (3.3)	96 (3.7)	99.9 (1.6)	100 (1.0)
		10	100 (2.1)	99.2 (2.6)	97.7 (3.5)	99.9 (1.7)	100 (0.6)
		20	99.9 (2.8)	99.7 (3.1)	99.5 (5.2)	99.8 (3.3)	100 (1.9)
	(5,5)	0	99.8 (2.6)	98.4 (2.9)	95.7 (2.5)	100 (1.6)	100 (0.5)
		10	99.7 (0.6)	94.6 (3.5)	93.3 (2.0)	98.4 (2.7)	10.6 (0.1)
		20	91.2 (0.1)	62.4 (2.3)	68.4 (1.6)	47.4 (1.3)	1.2 (0.4)
	(10,10)	0	99.9 (3.5)	98 (4.5)	96.4 (4.3)	100 (1.9)	100 (1.5)
		10	99.9 (0.7)	96.1 (3.4)	94 (3.2)	98 (2.7)	0.6 (0.1)
		20	92.9 (0.2)	65.5 (1.8)	68.9 (1.5)	50.6 (1.5)	0.2 (0.1)
500	(0,0)	0	100 (1.0)	99.9 (1.9)	99.4 (3.4)	100 (1.0)	100 (0.8)
		10	100 (1.5)	100 (1.8)	99.8 (3.5)	100 (1.6)	100 (1.3)
		20	100 (2.1)	100 (2.4)	100 (3.8)	100 (2.1)	100 (1.7)
	(5,5)	0	100 (1.3)	100 (2.5)	99.9 (3.3)	100 (1.7)	100 (1.4)
		10	100 (0.4)	99.6 (1.5)	98.7 (2.3)	100 (0.5)	11 (0.6)
		20	98.9 (0.0)	85.3 (1.0)	95.7 (1.0)	56.4 (1.4)	1.8 (0.0)
	(10,10)	0	100 (1.3)	99.9 (2.5)	99.6 (3.3)	100 (1.5)	100 (1.4)
		10	100 (0.2)	99.8 (0.8)	98.9 (2.6)	100 (0.4)	0.9 (0.4)
		20	99.2 (0.0)	87.9 (1.2)	95.4 (1.9)	58.8 (1.0)	0.4 (0.2)
1000	(0,0)	0	100 (1.1)	100 (1.5)	99.9 (1.9)	100 (1.3)	100 (1.3)
		10	100 (1.7)	100 (2.0)	100 (3.3)	100 (2.0)	100 (1.7)
		20	100 (1.9)	100 (2.5)	100 (3.5)	100 (1.7)	100 (1.7)
	(5,5)	0	100 (0.9)	100 (1.2)	99.9 (1.8)	100 (0.8)	100 (0.8)
		10	100 (0.3)	99.9 (0.8)	100 (1.2)	100 (0.5)	10 (0.3)
		20	99.7 (0.0)	89.9 (0.3)	98.9 (0.9)	57.1 (0.4)	2 (0.3)
	(10,10)	0	100 (1.0)	100 (1.0)	100 (2.2)	100 (1.3)	100 (1.0)
		10	100 (0.1)	100 (0.8)	99.8 (1.6)	100 (0.0)	0.6 (0.2)
		20	99.9 (0.0)	88.2 (0.1)	98.5 (1.1)	59.4 (0.0)	0.2 (0.1)

Results obtained for Case B samples are presented in table 3. It is clear that the proposed control charts perform well with respect to SR and FAR in almost all cases. Even though, their performance is not very satisfactory in cases of small samples with high contamination, their SR and FAR are getting better with increasing sample size.

In all the three cases the proposed robust GK based control charts are performing better than the classical control chart in contaminated situations and they show similar performance as that of classical chart in uncontaminated cases. It is interesting to see that, among the robust control charts, the one based on GK(Q_n) shows better success rate than other robust control charts when $n < 500$ while the control chart based on GK(MAD) outperform others for large samples when the data from Case A and Case B. But, when the data come from Case C, in all the cases GK(Q_n) based control chart shows superior performance.

4.1. Time Complexity

The proposed methods' significantly faster computation times are a key benefit. To compare the computation times of these approaches for various sample sizes, a simulation study was done. The simulation study consists of 10,000 trials, and the average running time is presented in table 5. Among the robust methods, the control chart based on GK(S_n) is multiple times faster than the other three methods when $n \leq 500$ while GK(MAD)-based control chart performs faster than the other compared methods when $n > 500$.

Table 5: *Running time (in seconds)*

n	GK(Q_n)	GK(S_n)	GK(MAD)	GK(τ)
50	0.000128	0.000080	0.000207	0.000361
100	0.000222	0.000093	0.000231	0.000361
500	0.001065	0.000290	0.000304	0.000516
1000	0.002231	0.000526	0.000397	0.000992

4.2. Real life data

A data set given by Quesenberry [9] has been used to evaluate the performance of the proposed methods in real life data. The original data consists of 11 quality variables measured on 30 products from a production process. For our comparison purposes, we consider the third and fourth variables as our bivariate data. Bivariate control charts using the proposed robust methods and the classical methods are developed for this data and presented in figure 3. From charts it is clear that the data is outlier-free, and none of the methods, including classical method, commit any false detection. In order to evaluate performance in a contaminated situation, we artificially created two outlying observations. Observations 7 and 16 are changed from (21.5, 5.08) and (21.5, 15.32), to (22.75, 5.08) and (22.75, 15.32) respectively, by adding a very small shift of 1.25 in the first variable. Control charts of the contaminated data set are given in figure 3, and it is clear that except for the classical control chart, all the proposed control charts detected these outliers.

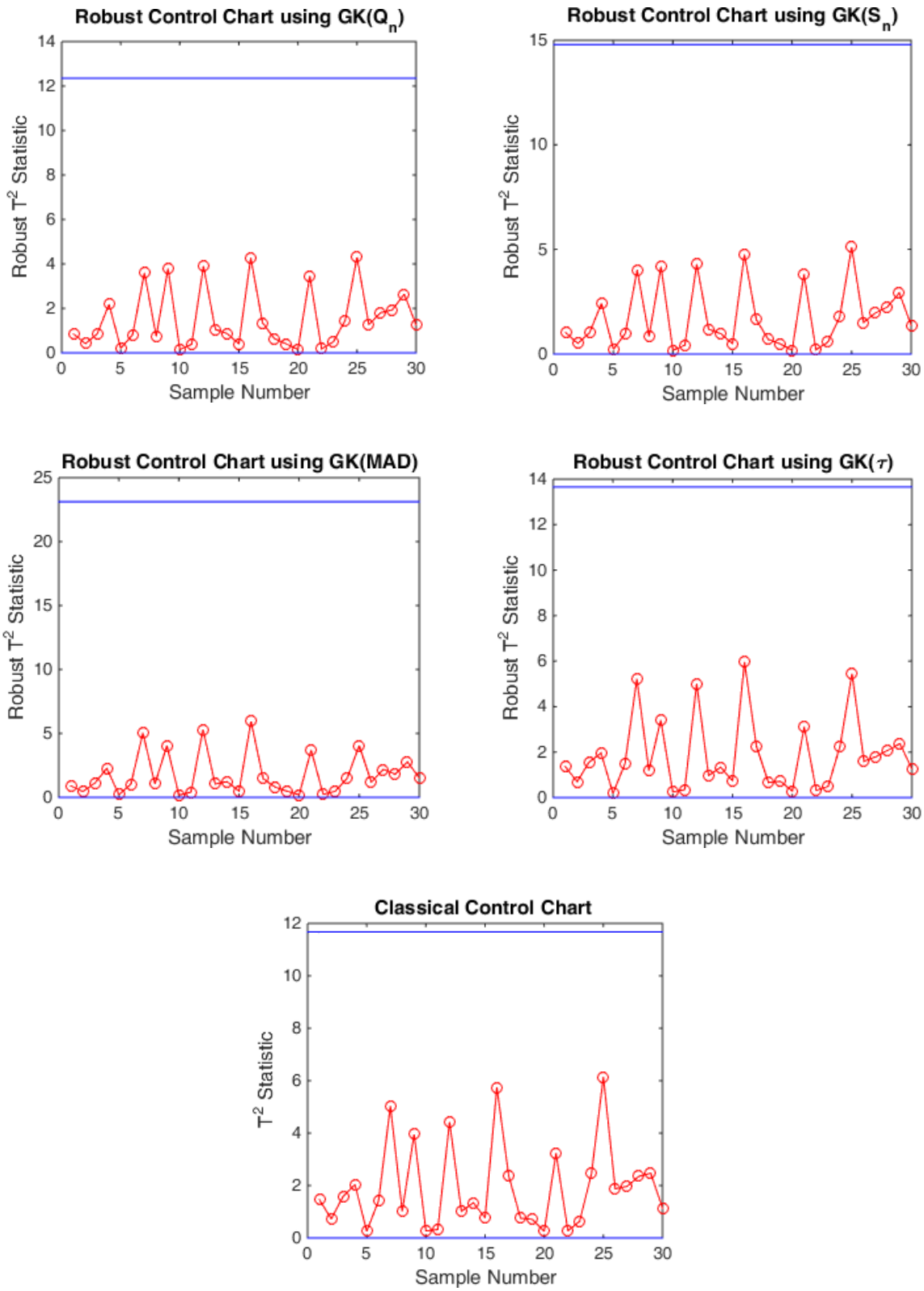


Figure 2: *Control charts for uncontaminated Quesenberry data*

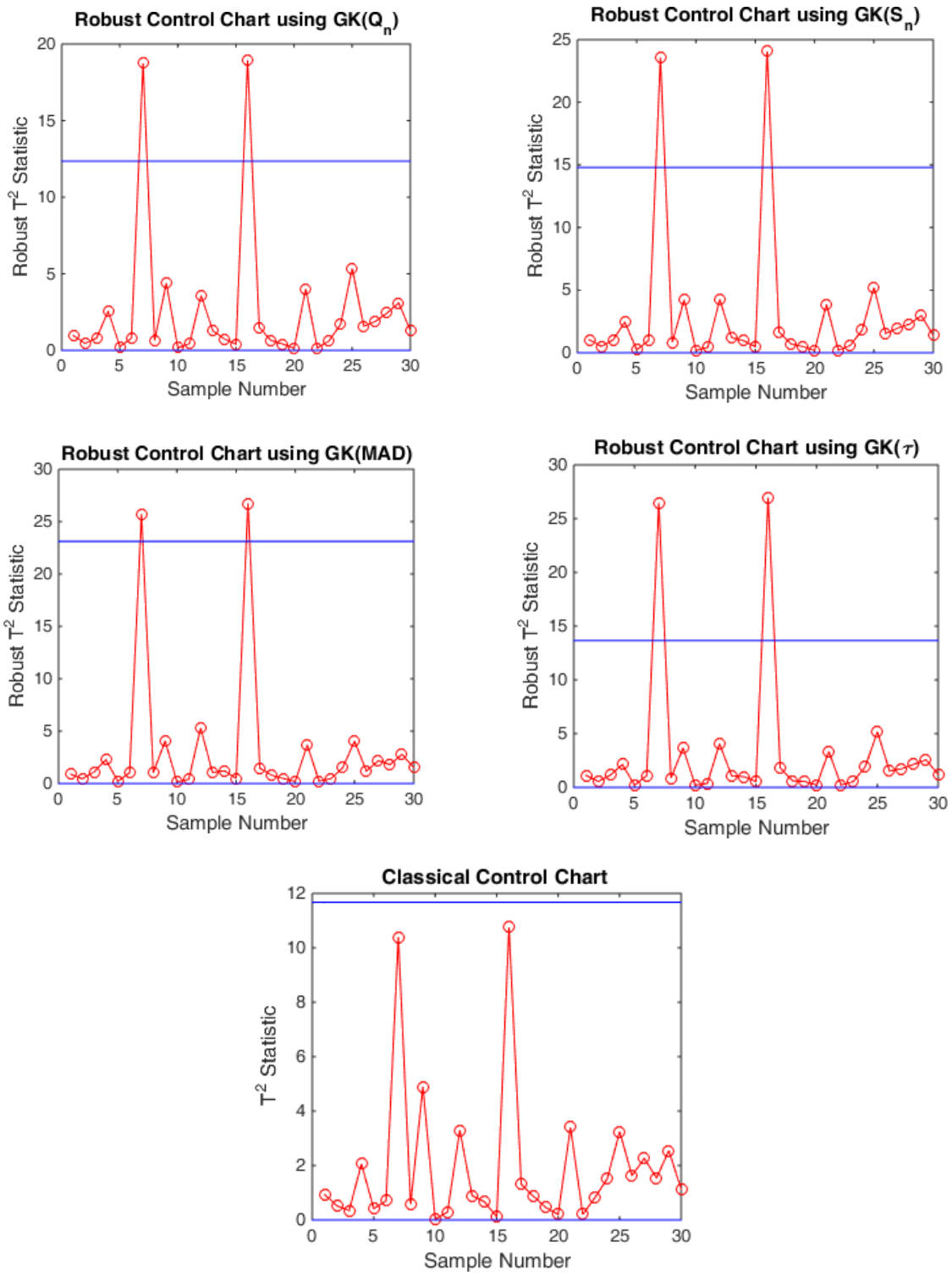


Figure 3: *Control charts for contaminated Quesenberry data*

5. Conclusions

When quality characteristics are interdependent, monitoring them simultaneously is crucial. Using univariate techniques to monitor or analyse these data is frequently ineffective. The Hotelling's T^2 control chart, which is created using the classical estimates of location and scatter, is the most well-known multivariate process monitoring and control method. Unfortunately, the existence of outliers

has a significant impact on the classical estimates, which results in incorrect T^2 statistics computation. In light of this, using Hotelling's T^2 control chart based on classical estimates will be quite deceptive. We address this problem for the bivariate case in this paper and suggest four robust control charts to handle outliers. The proposed control charts make use of the covariance estimator introduced by Gnanadesikan and Kettenring [2]. Four highly robust estimators - the Q_n estimator, S_n estimator, MAD estimator, and τ estimator - were used to robustify the GK estimator. These estimators have a bounded influence function and an ideal break down value.

Four different robust control charts for bivariate quality characteristics were introduced using GK(Q_n), GK(S_n), GK(MAD) and GK(τ) estimators. The upper control limits of these robust control charts are obtained by simulating the empirical quantiles of robust T^2 statistics, while the lower control limits are set to zero. Performance of the proposed methods is evaluated through Monte Carlo simulation. Three cases of contaminated Phase I datasets were considered for various amounts of contaminations and in all the cases the proposed methods outperform the classical chart. The proposed methods were also applied to real-life datasets, and even though the data contained outliers, they were still able to identify the out-of-control observation. Another advantage of the proposed methods is their fast computation which will reduce computational complexity.

References

[1]. Alt, F.B. Multivariate quality control. In S. Kotz and N. Johnson (Eds.), The Encyclopaedia of Statistical Sciences, Volume 6 (110-122), John Wiley & Sons, 1985.

[2]. Gnanadesikan, R. and Kettenring, J.R. (1972), Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics*, 28:81-124.

[3]. Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69:383-393.

[4]. Yohai, V. J. and Zamar, R. H. (1986). High breakdown point estimates of regression by means of the minimization of an efficient scale. *Journal of the American Statistical Association*, 86:403–413.

[5]. Maronna, R.A. and Zamar, R.H. (2002). Robust estimates of location and dispersion for high dimensional datasets, *Technometrics*, 44(4): 307-317.

[6]. Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283.

[7]. Wilks, S. S. Mathematical Statistics. John Wiley & Sons, 1962

[8]. Serfling, R.J. Approximation Theorems of Mathematical Statistics. John Wiley & Sons, 1980.

[9]. Quesenberry, C.P. (2001). The multivariate short-run snapshot Q chart. *Quality Engineering*, 13(4):679-683.

STATISTICAL MODELS FOR FORECASTING EMERGENCY SITUATIONS OF MAN-CAUSED CHARACTER

Valery Akimov, Ekaterina Ivanova, Yuri Shishkov
•
"All-Russian Research Institute for Civil Defense and Emergency Situations of the Ministry of Emergency Situations of Russia" (Federal Center for Science and High Technologies)
akimov@vniigochs.ru

Abstract

The aim of the study is to develop predictive and analytical solutions for technogenic threats for urban areas, the mathematical basis of which is Bayesian classifiers. The result of the work is a formalized description of models for predicting the consequences of a heat supply shutdown; consequences of a power outage; consequences of oil and oil products spills; the consequences of the discharge of liquid technological waste into the hydrosphere; the consequences of the release of hazardous chemicals into the environment.

Keywords: predictive and analytical solutions; Bayes method; man-made emergencies; shutdown of heat supply; power outage; spill of oil and oil products; discharge of liquid technological waste into the hydrosphere; release of hazardous chemicals into the environment.

I. Introduction

The subject of the work is technogenic threats to urban areas with the aim of their formalized description using Bayesian classifiers. The main part of the work contains five sections: a model for predicting the consequences of a heat supply shutdown; a model for predicting the consequences of a power outage; a model for predicting the consequences of an oil spill and oil products; a model for predicting the consequences of the discharge of liquid technological waste into the hydrosphere; model for predicting the consequences of the release of hazardous chemicals into the environment.

II. Methods

The main input data for the formation of the basic training set of the model for predicting the consequences of a heat supply shutdown are the following data: characteristics of heat supply systems and consumers; failure characteristics of heat supply systems; characteristics of reduced (emergency) heat supply to consumers; parameters of the meteorological situation [1].

The main calculated dependencies for determining the reliability indicators of heat supply to a consumer connected to the heat network of the heat supply system are presented in [2].

In the short-term forecasting model, probabilistic assessment using a Bayesian classifier is subject to hypotheses that estimate the deviation of the predicted time until the restoration of heat

supply as a result of the failure of the heat network section from its calculated value.

In the medium-term forecasting model, the hypotheses given in Table 1 are subject to probabilistic assessment using a Bayesian classifier.

Table 1: *List of hypotheses of the medium-range forecasting model*

Hypothesis number	Content of the hypothesis
1.	During the heating period, there were (will be recorded) cases of a decrease in the air temperature in consumer buildings below the limit value
2.	During the heating period, cases of interruptions in the supply of the estimated amount of heat and a decrease in air temperature in the buildings of consumers of the first category below the standard values were recorded (will be recorded)
3.	During the heating period, cases of exceeding the time for eliminating accidents in the TS system by more than 54 hours were recorded (will be recorded) for consumers of the second category

Based on the results of the probabilistic assessment, the level of threat is determined for each consumer (by the worst value).

Statistical model for predicting the consequences of a power outage

The initial data of the model for predicting the consequences of a power outage are: characteristics of electrical energy consumers located on the territory of a municipal district; characterization of massive damage to power grid facilities [3].

In case of accidents on power grids, it is important to timely identify consumers, the consequences of stopping the transmission of electrical energy to which, due to damage to the power grid facilities of grid organizations or equipment of power generation facilities, cause the greatest damage to the life of the population.

To this end, for each de-energized consumer of electrical energy, an index of the priority of restoration of power supply is determined, which takes into account the following factors: classifying the consumer of electrical energy as consumers whose limitation of the mode of consumption of electrical energy may lead to economic, environmental or social consequences; the degree of participation of the consumer of electric energy in ensuring the life of the population; the number of people in the buildings of the consumer of electrical energy; the estimated time at which the consumer can function when the main power supply is turned off if he has an independent power source - an autonomous backup power source; the time of the year and the day of the week in which the consumers of electrical energy were de-energized [4].

The power supply restoration priority index of the j-th transformer substation is determined according to Table 2.

Table 2: *Criteria for determining the index of priority of restoration of electricity supply to consumers*

Index value	Power restoration priority	Index criteria
1	Максимальный	$F_j \leq 50$, in the absence of de-energized consumers with a relative importance coefficient equal to 1
2	Average	$50{,}01 < F_j \leq 80$, in the absence of de-energized consumers with a coefficient of relative importance equal to 1; $F_j \leq 50$, in the presence of de-energized consumers with a coefficient of relative importance equal to 1
3	Minimum	$80{,}01 < F_j \leq 100$, in the absence of de-energized consumers with a coefficient of relative importance equal to 1; $50{,}01 < F_j \leq 100$, in the presence of de-energized consumers with a coefficient of relative importance equal to 1

Statistical Model for Predicting the Consequences of an Oil and Oil Product Spill

In the general case, the process of developing a statistical model for predicting the consequences of an oil and oil spill includes: collecting initial information and forming a basic training set; selection of a Bayesian classifier, preparation of methods for analyzing and interpreting the results of statistical processing.

The main initial data describing the characteristics of an oil and oil product spill (hereinafter referred to as ROP) are: input data characterizing the main parameters of the ROP; input data characterizing the storage tanks for oil and oil products (hereinafter referred to as NN); input data characterizing the meteorological situation; input data characterizing the properties of NN; input data characterizing the site of a possible RUI [5].

The main predictive parameters of the RNN are: the predicted area of the RNN in time corresponding to the forecast step; the predicted mass of the spilled HH after a time corresponding to the step of the forecast; forecasting water pollution, fire, explosion.

The output data of the model for medium-term forecasting are the results of a probabilistic assessment of the possibility of RNR from the storage tanks of HH during the forecasting period (10 days).

The output data of the model for short-term forecasting are [6]:

a) data implemented using the Bayesian classifier: the deviation of the actual value of the mass of the spilled HH onto the site as a result of RNN from its calculated value during the day every 3 hours; deviation of the actual value of the area of the contaminated territory as a result of RNR from its calculated value during the day every 3 hours; probabilistic assessment of the emergence of man-made threats as a result of RNN; probabilistic assessment of pollution of a water body as a result of ROP;

b) initial data used at each step of forecasting.

Statistical model for predicting the consequences of the discharge of liquid technological waste into the hydrosphere

The main input data for the formation of the basic training set of the model for predicting the consequences of the discharge of liquid technological waste into the hydrosphere are the following groups of parameters: parameters of systems (posts) for monitoring discharges of liquid technological waste (hereinafter referred to as LTW) located in close proximity to the sources of discharges of liquid technological waste from industrial facilities; parameters of systems (posts) for monitoring discharges of liquid waste; characteristics of sources of liquid waste disposal; characteristics of sections of a water body (hereinafter - VO); characteristics of the hydrological situation [7].

In this model, hypotheses are subject to probabilistic assessment using a Bayesian classifier: for predicting the concentrations of substances that make up the LTO; to predict the decline in chemical oxygen demand (hereinafter - COD) and biochemical oxygen demand (hereinafter - BOD) [8].

To predict the decline in COD and BOD, the hypotheses indicated in Table 3 are applied.

Table 3*: List of hypotheses of the model for predicting the decline in COD and BOD*

Number hypotheses	Content of the hypothesis
1.	At the observed values of the concentration of the substance that is part of the LTO, as a result of its discharge into the WW, the COD indicator dropped (drops) below 6.0 mg/dm^3
2.	At the observed values of the concentration of the substance that is part of the LTO, as a result of its discharge into the WA, the BOD indicator dropped (drops) below 2.1 mg /dm^3
Note - The content of each hypothesis in the past tense is used at the training stage of the	

SAM-SO, and the content of the hypothesis in the future tense is used when predicting the corresponding events on new values of the observed parameters.

The output data of the model are: a probabilistic assessment of the predicted concentrations of various types of LTL in a given period of time at systems (posts) for monitoring discharges during the spread of the LTL discharge; probabilistic assessment of the decrease in COD and BOD indicators as a result of the discharge of liquid waste; assessment of the level of water pollution based on the overall water quality index SCWQI.

Statistical model for predicting the consequences of the release of hazardous chemicals into the environment

The main input data for the formation of the basic training set of the model for predicting the consequences of the release of hazardous chemicals into the environment are the following groups of parameters: parameters of systems (posts) for monitoring emissions of hazardous substances located along the perimeter of industrial facilities that have sources of hazardous chemical emissions; parameters of systems (posts) for monitoring emissions of OHV; characteristics of the meteorological situation; characteristics of sources and parameters of OHV emissions [1].

In this model, the following parameters are subject to probabilistic assessment using a Bayesian classifier: the release rate and concentration of OHV at the location of the system (post) for monitoring OHV emissions [9].

The output data of the model are: probabilistic assessment of the predicted concentrations of various types of OHV in a given period of time at systems (posts) for monitoring emissions during the propagation of the OHV release; assessment of air pollution levels based on the SCAQI general air quality index.

III. Results

Thus, this article discusses predictive and analytical solutions for technogenic threats for urbanized areas, the mathematical basis of which is Bayesian classifiers. The result of the work is a formalized description of statistical models for predicting the consequences of a heat supply shutdown; consequences of a power outage; consequences of oil and oil products spills; the consequences of the discharge of liquid technological waste into the hydrosphere; the consequences of the release of hazardous chemicals into the environment.

IV. Discussion

The discussion of the verbal and mathematical foundations of predictive modeling of technogenic emergencies is quite active in the scientific literature [10 - 12].

The scientific novelty of the developed models lies in a single scientific approach to their creation, namely, the use of a statistical processing method based on Bayes' theorem.

References

[1] Predictive and analytical solutions for natural, man-made and biological and social threats of a unified system of information and analytical support for the safety of the environment and public order "Safe City" / V. A. Akimov, A. V. Mishurny, O. V. Yakimyuk [and etc.]. - Moscow: All-Russian Research Institute for Civil Defense and Emergency Situations of the Ministry of Emergency Situations of Russia, 2022. - 315 p. – ISBN 978-5-93970-278-2. – EDN MGXNYI.

[2] Ivanova, E. O. Accidents in heat supply systems: a probabilistic assessment of the development of the consequences of failures in a heat network / E. O. Ivanova, A. V. Mishurny // Civil Security Technologies. - 2022. - T. 19, No. 4 (74). - S. 48-50. – EDN ZYZBAW.

[3] Akimov, V. A. Accidents in power supply systems: determining the priority index for the restoration of power supply / V. A. Akimov, A. V. Mishurny // Civil Security Technologies. - 2022. - T. 19, No. 4 (74). - S. 44-47. – EDN RWBXUK.

[4] Akimov, V. A. Determination of the priority index of the restoration of power supply / V. A. Akimov, E. O. Ivanova, Yu. scientific and practical conference, Khimki, March 01, 2023. - Khimki: Academy of Civil Protection of the Ministry of the Russian Federation for Civil Defense, Emergency Situations and Elimination of Consequences of Natural Disasters named after Lieutenant General D.I. Mikhailika, 2023. - S. 14-17. – EDN HXCFJH.

[5] Akimov, V. A. Mathematical model for predicting the consequences of oil spills and oil products / V. A. Akimov, E. O. Ivanova, A. V. Mishurny // Civil Security Technologies. - 2023. - T. 20, No. 1 (75). – S. 68-70. – EDN BGHOAH.

[6] Preliminary national standard of the Russian Federation PNST 775-2022 “Safety in emergency situations. Safe city. Forecasting the consequences of an oil spill and oil products. General requirements".

[7] Akimov, V. A. Mathematical model for predicting the consequences of discharge of liquid technological waste into the hydrosphere / V. A. Akimov, S. V. Koleganov, A. V. Mishurny // Civil Security Technologies. - 2023. - T. 20, No. 1 (75). - S. 71-73. – EDN BPEHQV.

[8] Preliminary national standard of the Russian Federation PNST 770-2022 “Safety in emergency situations. Safe city. Forecasting the Consequences of the Discharge of Liquid Process Wastes into the Hydrosphere. General requirements".

[9] Preliminary national standard of the Russian Federation PNST 771-2022 “Safety in emergency situations. Safe city. Forecasting the consequences of the release of hazardous chemicals into the environment. General requirements".

[10] Akimov, V. A. Mathematical models for forecasting technogenic emergencies / V. A. Akimov, E. O. Ivanova, Yu. A. Shishkov // Civil defense on guard of peace and security: Materials of the VII International Scientific and Practical conference dedicated to the World Civil Defense Day in the year of the 90th anniversary of the formation of the Academy of the State Fire Service of the Ministry of Emergency Situations of Russia. In 5 parts, Moscow, March 01, 2023. Volume Part II. - Moscow: Academy of the State Fire Service of the Ministry of the Russian Federation for Civil Defense, Emergencies and Disaster Relief, 2023. - P. 126-133. – EDN UIZLCO.

[11] Akimov, V. Forecast modeling of man-made emergencies with modern methods/ V. Akimov, M. Bedilo, O. Derendiaeva, E. Ivanova, I. Oltyan // RT&A, Special Issue № 4 (70), Volume 17, November 2022. – S. 318 - 323. – DOI: https://doi.org/10.24412/1932-2321-2022-470-318-323.

[12] Security of Russia. Legal, socio-economic and scientific and technical aspects. Analysis and provision of security from emergency situations / V. A. Akimov, A. A. Antyukhov, E. V. Arefieva [and others]; Security Council of the Russian Federation, Russian Academy of Sciences, EMERCOM of Russia, Rostekhnadzor, Russian Science Foundation, Rostec State Corporation, Rosatom State Corporation, Rosneft Oil Company PJSC, Russian Railways OJSC, Transneft PJSC, Gazprom PJSC. - Moscow: MGOF "Knowledge", 2021. - 500 p. – ISBN 978-5-87633-199-1. – EDN FXIJPZ

CHARACTERIZATION OF CHANDBHAS-P DISTRIBUTION AND ITS APPLICATIONS IN MEDICAL SCIENCE

Praseeja C B[1], Prasanth C B[2*], C Subramanian[3], Unnikrishnan T[4]

[1]Research Scholar, Department of Statistics, Annamalai University, Tamil Nadu, India, prasinikhil@gmail.com
[2]Assistant Professor, Department of Statistics, Sree Keralawarma College, Kerala, prasanthwarriercb@keralavarma.ac.in(*Corresponding author)
[3]Associate Professor, Department of Statistics, Annamalai University, Tamil Nadu, manistat@yahoo.co.in
[4]Assistant Professor, Department of Statistics, Shri C Achuthamenon Govt. College, Kerala, t.unnikrishnan@gmail.com

Abstract

The current research attempts the length biased version of new two-parameter Sujatha distribution, which is referred as ChandBhas-P Distribution (CBPD). Its different structural properties are discussed and the model parameters of this novel distribution are predicted by using the Maximum Likelihood Estimation. The distribution was examined with two real lifetime sets of data. The first set of data is birth weight of new born babies, randomly selected from a hospital at Thrissur, kerala and the second set is the weight of children of age range between three months to four years-collected from a few babysitting centres and play schools across Thrissur, and both are employed in order to discuss the goodness of fit.

Keywords: estimation, Survival analysis, two-parameter Sujatha distribution, Length biased distribution, data.

1. Introduction

Statistical probability distributions are often the foundation of data analysis in the real world. However, data from other domains, including environmental, biomedical, financial, and other areas, may not match the conventional distributions. Thus, it becomes necessary to create new distributions that will effectively capture extreme skewness and kurtosis while improving the goodness-of-fit of empirical distributions.

The most common classical distributions often fail to fit and forecast data in a variety of applied fields, such as engineering, the biological and environmental sciences, the medical and life sciences, finance, and economics. As a result, many generalized families are seen to be an improvement for developing and expanding the typical classical distributions. The recently created families provide more application versatility and have been thoroughly investigated in many fields. Because of their versatility, these extended families have been used to represent data in many applied fields. The statistical distribution theory often involves adding a new parameter to a

distribution family function that already exists. A class of distribution functions may often be made more flexible by adding parameters, which could be extremely helpful for data analysis.

The introduction of an extra parameter into the paradigm, which generates flexibility in nature, allows for the weighted distributions to play a vital part in distribution theory since it offers a new comprehension of the classical distributions that are already in use. In particular, the weighted distributions have been employed to help in the acceptable model's selection for observed data when samples are taken without a suitable frame. The weighted distributions are used to describe heterogeneity, extraneous variance, and clustered sampling in the data set. The weighted distribution arises when data from a sample are noted with uneven probabilities, and it offers a solution for issues when the observations fall into non-replicated, non-experimental, as well as non-random categories. The weighted distributions offer a special method for addressing issues with model formulation and data interpretation. The probability of occurrences as seen and transcribed is modified by using the weighted distributions. The weighted distributions are reduced to length-biased distributions when the function of weight considers just the unit's length of interest.

The notion of weighted distributions was created to record the observation in accordance with certain weight functions. Fisher [8] executed the weighted distribution to analyze how the approach of ascertainment might affect the type of recorded observation's distribution. Rao [13] later developed it into a unified theory w.r.t modeling statistical data when it was discovered that the common practice of utilizing standard distributions was incorrect. Zelen [18] and Cox [4] first proposed the idea of length-biased sampling. In the system of renewal concept, Cox [5] first presented the statistical explanation of length-biased distribution. Length-biased sampling circumstances could arise in survival analysis, reliability theory, clinical trials, as well as population research when an appropriate sample frame is missing.

A significant contribution was conducted by various investigators to create certain important length-biased probability distributions by handling its different lifetime data sets from heterogeneous fields of application. Dar et al. [6] presented Poisson size-biased "Lindley distribution" and its uses. Ganaie and Rajagopalan [9] detailed a "length-biased weighted quasi-gamma distribution" with characterizations and uses. Beghriche and Zeghdoudi [3] explained a size-biased gamma Lindley distribution. Ekhosuehil et al. [7] proposed the Weibull length-biased exponential distribution with statistical properties and uses. Alhyasat et al. [1] expressed power size biased 2-parameter Akash distribution. Sen et al. [15] discussed weighted X-gamma distribution with application and properties. Rasool and Ahmad [12] introduced power "length-biased weighted Lomax distribution" with uses. Mobarak et al. [10] studied the size-biased weighted transmuted Weibull distribution. Sanat [14] proposed the beta-length biased Pareto distribution along with properties. Hussein and Al-Kadim [2] presented Rayleigh distribution and length-biased weighted exponential with the application. Recently, Khan and Mustafa [11] implemented the length-biased powered inverse "Rayleigh distribution" with uses, which demonstrates more flexibility as compared to conventional distribution.

The new two-parameter Sujatha distribution is a recently suggested two-parametric lifespan model by Tesfay and Shanker [16], of which the one-parameter Akash distribution and Sujatha distribution are special cases of the same. Tesfay and Shanker [17] also proposed two parameters of Sujatha distribution which include size-biased Lindley along with Sujatha distribution.

2. ChandBhas-P Distribution (CBPD)

The probability density function(pdf) of the new two-parameter Sujatha distribution is,

$$f(x;\theta,\alpha)=\frac{\theta^3}{\theta^2+\alpha\theta+2}\left(1+\alpha x+x^2\right)e^{-\theta x};\ x>0,\theta>0,\alpha\geq 0 \quad (1)$$

and the cumulative density function (cdf) of the new two-parameter Sujatha distribution is presented by

$$F(x;\theta,\alpha)=1-\left(1+\frac{\theta x(\theta x+\alpha\theta+2)}{\theta^2+\alpha\theta+2}\right)e^{-\theta x};x>0,\theta>0,\alpha>0 \quad (2)$$

Suppose, $f(x)$ is the pdf of X (a non-negative random Variable), then the pdf of weighted X_w is presented as

$$f_w(x)=\frac{w(x)f(x)}{E(w(x))},\quad x>0.$$

where w(x) be the non-negative weight function &

$E(w(x))=\int w(x)f(x)dx<\infty.$

There are several forms of weighted models. Particularly when w(x) = x^c, results a distribution known as a weighted distribution. The length-biased form of the new two -parameter Sujatha distribution is analyzed herewith. The distribution that results from our consideration of the "weight function" at $w(x) = x$ is termed the "length-biased distribution" with pdf,

$$f_1(x)=\frac{x.f(x)}{E((x)} \quad (3)$$

Where $E(x)=\int_0^\infty x\,f(x;\theta,\alpha)dx$

$$E(x)=\frac{\theta^2+2\alpha\theta+6}{\theta(\theta^2+\alpha\theta+2)} \quad (4)$$

We get the pdf of the CBP distribution by putting eqs. (1) & (4) into eq. (3).

$$f_l(x)=\frac{x\theta^4}{\theta^2+2\alpha\theta+6}\left(1+\alpha x+x^2\right)e^{-\theta x} \quad (5)$$

and the cdf of CBP Distribution is,

$$F_l(x)=\int_0^x f_l(x)dx$$

$$=\frac{1}{\theta^2+2\alpha\theta+6}\left(\theta^4\int_0^x x e^{-\theta x}dx+\alpha\theta^4\int_0^x x^2e^{-\theta x}dx+\theta^4\int_0^x x^3e^{-\theta x}dx\right) \quad (6)$$

Put $\theta x=t \Rightarrow \theta dx=dt \Rightarrow dx=\frac{dt}{\theta}$, When $x\to 0, t\to 0$, Also $x=\frac{t}{\theta}$

After the simplification of eq. (6), we get the cdf of CBPD distribution as

$$F_l(x)=\frac{1}{\theta^2+2\alpha\theta+6}\left(\theta^2\gamma(2,\theta x)+\alpha\theta\,\gamma(3,\theta x)+\gamma(4,\theta x)\right) \quad (7)$$

The nature of the pdf and cdt is clear from the Figure 1. & Figure 2. For different α and θ. When α & θ decreases the pdf curve became less skewed.

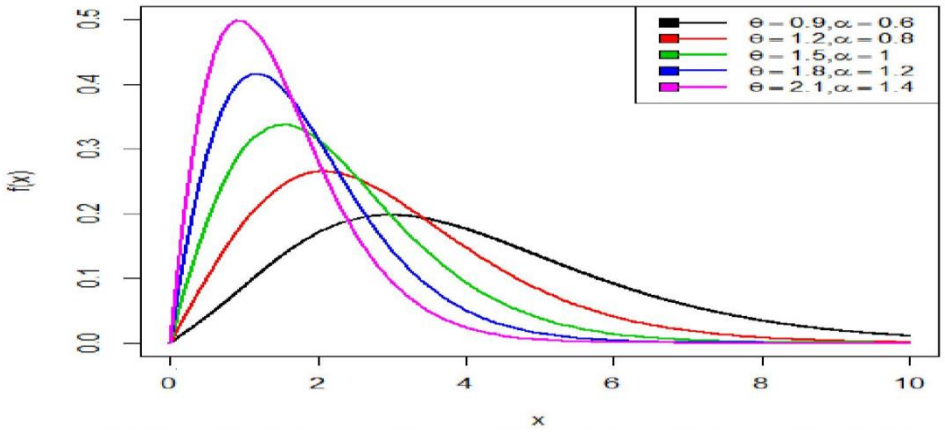


Figure 1. *pdf of CBP Distribution*

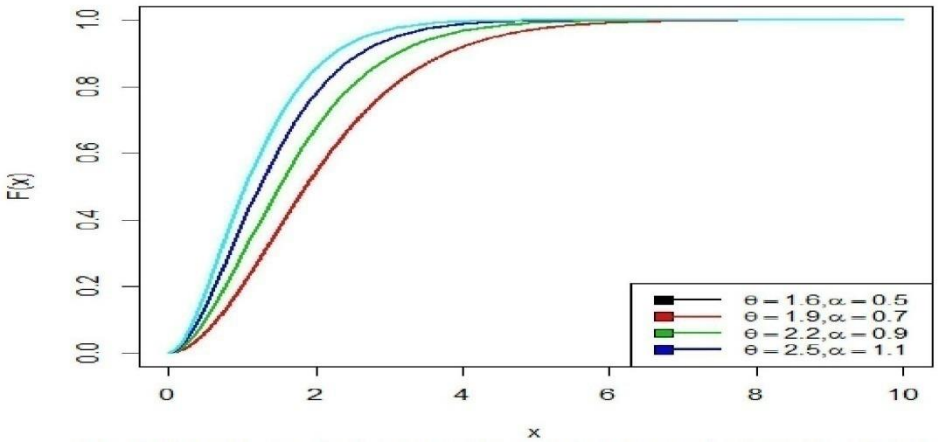


Figure 2. *cdf of CBP Distribution*

3. Characteristics of the CBPD

3.1 Survival function

$$S(x) = 1 - F_l(x) = 1 - \frac{1}{\theta^2 + 2\alpha\theta + 6}\left(\theta^2\gamma(2,\ \theta x) + \alpha\theta\,\gamma(3, \theta x) + \gamma(4,\ \theta x)\right)$$

3.2 Hazard Rate

It is also known as hazard function or failure rate and it is presented as

$$h(x) = \frac{f_l(x)}{1 - F_l(x)} = \frac{x\theta^4(1 + \alpha x + x^2)\ e^{-\theta x}}{(\theta^2 + 2\alpha\theta + 6) - (\theta^2\gamma(2,\ \theta x) + \alpha\theta\,\gamma(3,\ \theta x) + \gamma(4,\ \theta x))}$$

The following figures represent the Survival function and Hazad rate of the new distribution. The decreasing nature of survival functions of CBPD with increase in x for different 'α' & 'θ' is clear from Figure 3. The nature of hazard rate with different α & θ is noticed from Figure 4.

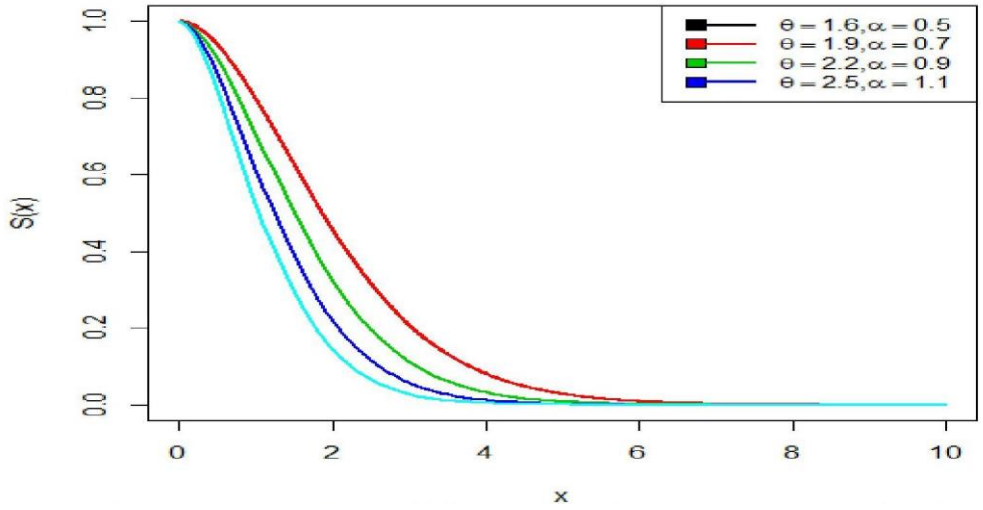


Figure 3. *Survival function of CBP Distribution with different 'α' & 'θ'*

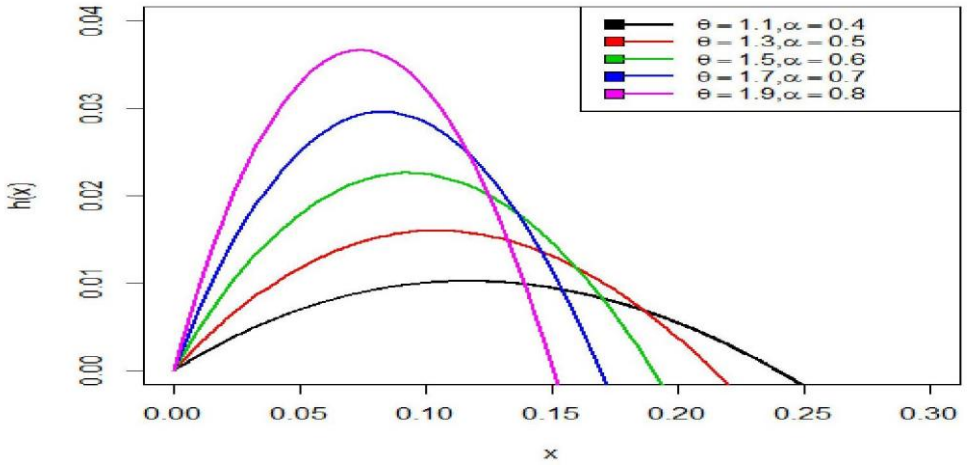


Figure 4. *The nature of hazard rate OF CBPD with different α & θ*

4. Statistical Properties

4.1 Moments

The r^{th} order moment $E(X^r)$ can be calculated by considering the random variable X with the CBPD and parameters θ and α,

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_l(x)dx$$

$$= \int_0^\infty x^r \frac{x\theta^4}{\theta^2 + 2\alpha\theta + 6}\left(1 + \alpha x + x^2\right)e^{-\theta x}dx$$

$$= \int_0^\infty \frac{x^{r+1}\theta^4}{\theta^2 + 2\alpha\theta + 6}\left(1 + \alpha x + x^2\right)e^{-\theta x}dx$$

$$= \frac{\theta^4}{\theta^2 + 2\alpha\theta + 6}\int_0^\infty x^{r+1}\left(1 + \alpha x + x^2\right)e^{-\theta x}dx$$

$$= \frac{\theta^4}{\theta^2 + 2\alpha\theta + 6}\left(\int_0^\infty x^{(r+2)-1} e^{-\theta x} dx + \alpha \int_0^\infty x^{(r+3)-1} e^{-\theta x} dx + \int_0^\infty x^{(r+4)-1} e^{-\theta x} dx\right) \quad (8)$$

By simplifying equation (8), we obtain

$$E(X^r) = \mu_r' = \frac{\theta^2 \Gamma(r+2) + \alpha\theta\Gamma(r+3) + \Gamma(r+4)}{\theta^r(\theta^2 + 2\alpha\theta + 6)} \quad (9)$$

By putting r = 1, 2, 3, & 4 into equation. (9), we get the first four moments of the CBP distribution.

$$E(X) = \mu_1' = \frac{2\theta^2 + 6\alpha\theta + 24}{\theta(\theta^2 + 2\alpha\theta + 6)}, \qquad E(X^2) = \mu_2' = \frac{6\theta^2 + 24\alpha\theta + 120}{\theta^2(\theta^2 + 2\alpha\theta + 6)}$$

$$E(X^3) = \mu_3' = \frac{24\theta^2 + 120\alpha\theta + 720}{\theta^3(\theta^2 + 2\alpha\theta + 6)}, \qquad E(X^4) = \mu_4' = \frac{120\theta^2 + 720\alpha\theta + 5040}{\theta^4(\theta^2 + 2\alpha\theta + 6)}$$

$$\text{Variance} = \frac{6\theta^2 + 24\alpha\theta + 120}{\theta^2(\theta^2 + 2\alpha\theta + 6)} - \left(\frac{2\theta^2 + 6\alpha\theta + 24}{\theta(\theta^2 + 2\alpha\theta + 6)}\right)^2$$

$$S..D(\sigma) = \sqrt{\frac{6\theta^2 + 24\alpha\theta + 120}{\theta^2(\theta^2 + 2\alpha\theta + 6)} - \left(\frac{2\theta^2 + 6\alpha\theta + 24}{\theta(\theta^2 + 2\alpha\theta + 6)}\right)^2}$$

4.2 Harmonic mean

$$H.M = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_l(x) dx = \int_0^\infty \frac{\theta^4}{\theta^2 + 2\alpha\theta + 6}\left(1 + \alpha x + x^2\right) e^{-\theta x} dx$$

$$= \frac{\theta^4}{\theta^2 + 2\alpha\theta + 6}\left(\int_0^\infty x^{(2)-2} e^{-\theta x} dx + \alpha \int_0^\infty x^{(2)-1} e^{-\theta x} dx + \int_0^\infty x^{(3)-1} e^{-\theta x} dx\right) \quad (10)$$

After the simplification of equation (10), we obtain

$$H.M = \frac{\theta(\theta + \alpha\theta + 2)}{\theta^2 + 2\alpha\theta + 6}$$

4.3 Moment generating & Characteristic function

The function of the proposed distribution may be constructed using X as the random variable following length biased new 2 parameters with α and θ distribution.

$$M_X(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx} f_l(x) dx$$

.

We obtain the following with Taylor's series

$$M_X(t) = E(e^{tx}) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + \ldots\right) f_l(x) dx$$

$$= \int_0^\infty \sum_{j=0}^{\infty} \frac{t^j}{j!} x^j f_l(x)dx = \sum_{j=0}^{\infty} \frac{t^j}{j!} \mu_j{}''$$

$$= \sum_{j=0}^{\infty} \frac{t^j}{j!} \frac{[\theta^2\Gamma(j+2) + \alpha.\theta.\Gamma(j+3) + \Gamma(j+4)]}{\theta^j(\theta^2 + 2\alpha\theta + 6)}$$

Hence,

$$M_X(t) = \frac{1}{(\theta^2 + 2\alpha\theta + 6)} \sum_{j=0}^{\infty} \frac{t^j}{j!\theta^j}[\theta^2\Gamma(j+2) + \alpha.\theta.\Gamma(j+3) + \Gamma(j+4)] \quad (11)$$

Similarly the characteristic function of CBP distribution is,

$$\Phi_X(t) = M_X(it) = \frac{1}{(\theta^2 + 2\alpha\theta + 6)} \sum_{j=0}^{\infty} \frac{(it)^j}{j!\theta^j}[\theta^2\Gamma(j+2) + \alpha.\theta.\Gamma(j+3) + \Gamma(j+4)] \quad (12)$$

5. Results and Discussions

5.1 Likelihood Ratio Test.

Consider a random sample "$X_1, X_2, \ldots, X_n$" of size n selected from the CBP distribution. The hypothesis is to be tested for examining its significance.

$H_o: f(x) = f(x;\theta,\alpha)$ vs. $H_1: f(x) = f_l(x;\theta,\alpha)$

To investigate and analyze, whether the random n sample size obtained from the CBP distribution, the given below test statistic rule is applied.

$$\Delta = \frac{L_1}{L_o} = \prod_{i=1}^{n} \frac{f_l(x;\theta,\alpha)}{f(x;\theta,\alpha)}$$

$$= \left(\frac{\theta(\theta^2 + \alpha\theta + 2)}{\theta^2 + 2\alpha\theta + 6}\right)^n \prod_{i=1}^{n} x_i$$

The null hypothesis will not be accepted if

$$\Delta = \left(\frac{\theta(\theta^2 + \alpha\theta + 2)}{\theta^2 + 2\alpha\theta + 6}\right)^n \prod_{i=1}^{n} x_i > k$$

Similarly, we shall reject the "null hypothesis", where

$$\Delta^* = \prod_{i=1}^{n} x_i > k\left(\frac{\theta^2 + 2\alpha\theta + 6}{\theta(\theta^2 + \alpha\theta + 2)}\right)^n \quad \Delta^* = \prod_{i=1}^{n} x_i > k^*, \text{Where } k^* = k\left(\frac{\theta^2 + 2\alpha\theta + 6}{\theta(\theta^2 + \alpha\theta + 2)}\right)^n$$

Whether for the large size $n$sample, $2log\ \Delta$ indicates distributed as "chi-square distribution" with 1degree of freedom as well as this distribution is applied for examining the value of p. Therefore, we reject to maintain the "null hypothesis" when the value of probability is provided as $p(\Delta^* > \beta^*)$,
here$\beta^* = \prod_{i=1}^{n} x_i$ s lower than a particular level of significance and $\prod_{i=1}^{n} x_i$. is the examined value of the statistic Δ^*.

5.2. MLE of CBP distribution

By MLE, estimate the parameters of the CBP distribution. The likelihood function for - *X1, X2,..., and Xn* - a random sample of n size from the CBPD is,

$$L(x)=\prod_{i=1}^{n} f_l(x)=\frac{\theta^{4n}}{(\theta^2+2\alpha\theta+6)^n}\prod_{i=1}^{n}\left(x_i\left(1+\alpha x_i+x_i^{\,2}\right)e^{-\theta x_i}\right)$$

The log-likelihood formula is,

$$InL = 4n\ln\theta - n\ln((\theta^2+2\alpha\theta+6)) + \sum_{i=1}^{n}\ln(x_i) + \sum_{i=1}^{n}\ln(1+\alpha x_i+x_i^{\,2}) - \theta\sum_{i=1}^{n}x_i \qquad (13)$$

By differentiating the log-likelihood eq. (13) w.r.t parameters θ and α. We get,

$$\frac{\partial \log L}{\partial\theta}=\frac{4n}{\theta}-n\left(\frac{(2\theta+2\alpha)}{\theta^2+2\alpha\theta+6}\right)-\sum_{i=1}^{n}x_i=0$$

$$\frac{\partial \log L}{\partial\alpha}=-n\left(\frac{2\theta}{\theta^2+2\alpha\theta+6}\right)+\sum_{i=1}^{n}\left(\frac{x_i}{(1+\alpha x_i+x_i^{\,2})}\right)=0$$

It is vital to note that the preceding system of nonlinear equations has an analytical solution that is too complex to be solved algebraically. So, we predict the suggested distribution's parameters using R and Wolfram mathematics.

The asymptotic normality findings must be used for the purpose of calculating the confidence interval.

We have that if $\hat{\beta}=(\hat{\theta},\ \hat{\alpha})$ represents the MLE of $\beta=(\theta,\ \alpha)$. The results can be executed as,

$$\sqrt{n}(\hat{\beta}-\beta)\rightarrow N_2(0,\ I^{-1}(\beta))$$

Here $I^{-1}(\beta)$indicates the matrix of Fisher's Information.

$$E\left(\frac{\partial^2 \log L}{\partial\theta^2}\right)=-\frac{4n}{\theta^2}-n\left(\frac{2(\theta^2+2\alpha\theta+6)-(2\theta+2\alpha)^2}{(\theta^2+2\alpha\theta+6)^2}\right)$$

$$E\left(\frac{\partial^2 \log L}{\partial\alpha^2}\right)=n\left(\frac{4\theta^2}{(\theta^2+2\alpha\theta+6)^2}\right)-\sum_{i=1}^{n}\left(\frac{(x_i)^2}{(1+\alpha x_i+x_i^{\,2})^2}\right),$$

$$E\left(\frac{\partial^2 \log L}{\partial\theta\,\partial\alpha}\right)=-n\left(\frac{2(\theta^2+2\alpha\theta+6)-2\theta(2\theta+2\alpha)}{(\theta^2+2\alpha\theta+6)^2}\right)$$

Since β being unknown, we predict $I^{-1}(\beta)$ by $I^{-1}(\hat{\beta})$ and this could be utilized to get the confidence intervals for θ&α.

6. Applications of CBP Distribution in real life Data.

We analyzed and examined the two real lifetime sets of Bio-medical data for fitting CBP distribution and the fit was compared with some related distributions. (New Two-Parameter Sujatha, Two-Parameter Sujatha, Sujatha, Exponential, and Lindley distributions).

Data set I: The following set of data represents the birth weight (Kg) of randomly selected new born babies from a hospital at Thrissur, Kerala (Table 1).

Table 1: *Data regarding the birth weight (Kg) of new born babies*

2.910	3.640	2.770	2.190	2.420	1.910
3.1091	3.385	3.145	3.215	2.495	3.515
4.380	3.110	3.280	3.860	4.040	4.170
3.640	3.210	2.870	3.230	4.220	1.520
2.610	3.360	2.840	3.140	2.530	1.580
4.275	3.380	4.870	3.100	2.800	4.280

Data set II: The following data (Table 2) represents the weight (Kg) of 100 randomly selected children- of age range between three months to four years- collected from a few babysitting centres and play schools across Thrissur,

Table 2: *weight (kg) of 100 Children (age between 3months to 4 years) from a babysitting record.*

8.75	12.25	12.50	6.50	16.00	14.50	6.75	14.50	7.50	9.00
7.50	7.00	7.75	7.50	7.25	7.00	6.50	6.75	10.00	5.50
8.75	6.75	12.25	12.50	6.50	14.00	14.50	6.50	15.50	12.50
10.00	8.75	7.25	8.00	7.50	9.00	7.00	8.75	7.00	7.50
12.00	6.75	5.00	7.50	6.50	7.25	5.25	7.50	5.50	8.75
7.25	14.50	6.75	7.50	7.00	6.75	12.25	8.75	7.00	7.50
7.00	7.75	7.50	7.25	7.00	6.50	6.75	10.00	9.00	7.00
6.75	12.25	12.50	6.50	16.00	14.50	6.50	16.00	5.50	6.75
8.75	7.25	8.00	7.50	9.00	7.00	8.75	7.00	12.50	6.50
6.75	5.00	7.50	6.50	7.25	5.25	7.50	5.50	6.50	7.50

Table 3: *Performance & Comparison of Fitted Distributions for a set of Data 1*

"Distributions	MLE	S.E	-2logL	AIC	BIC	AICC
CBP distribution	$\hat{\alpha} = 0.001$ $\hat{\theta} = 1.1139$	$\hat{\alpha} = 0.009$ $\hat{\theta} = 0.0495$	118.94	123.07	126.29	123.49
Two Parameter Sujatha	$\hat{\alpha} = 0.001$ $\theta = 0.819$	$\hat{\alpha} = 0.011$ $\hat{\theta} = 0.0539$	125.79	129.83	133.01	130.20
New Two Parameter Sujatha	$\hat{\alpha} = 0.001$ $\hat{\theta} = 0.698$	$\hat{\alpha} = 0.020$ $\hat{\theta} = 0.2110$	133.96	137.97	141.09	138.32
Sujatha	$\hat{\theta} = 0.691$	$\hat{\theta} = 0.067$	134.19	136.20	137.790	136.29
Exponential	$\hat{\theta} = 0.3047$	$\hat{\theta} = 0.051$	157.59	159.56	161.137	159.68
Lindley	$\hat{\theta} = 0.50694$	$\hat{\theta} = 0.061509$	144.19	146.19	147.775	146.31

Table 4: *Performance &Comparison of Fitted distributions for a set of Data 2*

Distributions	MLE	S.E	-2logL	AIC	BIC	AICC
CBP distribution	$\hat{\alpha} = 0.001$ $\hat{\theta} = 0.469$	$\hat{\alpha} = 0.020$ $\hat{\theta} = 0.025$	223.80	227.81	231.29	228.10
Two Parameter Sujatha	$\hat{\alpha} = 0.001$ $\hat{\theta} = 0.347$	$\hat{\alpha} = 0.0001$ $\hat{\theta} = 0.019$	233.79	237.79	241.36	238.09

New Two Parameter Sujatha	$\hat{\alpha} = 0.001$ $\hat{\theta} = 0.350$	$\hat{\alpha} = 0.010$ $\hat{\theta} = 0.022$	234.20	238.20	241.69	238.49
Sujatha	$\hat{\theta} = 0.336$	$\hat{\theta} = 0.029$	236.58	238.58	240.36	238.67
Exponential	$\hat{\theta} = 0.121$	$\hat{\theta} = 0.0183$	273.58	275.58	277.36	275.67
Lindley	$\hat{\theta} = 0.2208$	$\hat{\theta} = 0.0237$	251.81	253.81	255.59	253.90

The R software is utilized to calculate the model comparison criteria values as well as the estimate of unknown parameters (Table 3 and Table 4). We take into account of standard criteria's like *AICC ("Akaike Information Criterion Corrected"), BIC ("Bayesian Information Criterion"),AIC ("Akaike Information Criterion"),* and -2logL to compare the performance of CBP distribution along with the new 2-parameter Sujatha, two-parameter Sujatha, Sujatha, Lindley and Exponential distributions. The distribution with the lowest values of *AICC, BIC,-2logL, and AIC* is the optimal distribution. From the Table 3 & table 4, the CBP distribution has lower *AIC, AICC,BIC,* as well as *-2logL* values than the New Two-Parameter Sujatha, Sujatha, Two Parameter Sujatha, Lindley and Exponential distributions. Therefore, it is confirmed that the CBP distribution offers a better match than the new Two-Parameter Sujatha, Two-parameter Sujatha, Sujatha, Exponential, and Lindley distributions for fitting of such bio-medical data. Hence the significance of the new distribution is established.

7. Conclusions

Here, a novel distribution is known as the ChandBhas-P distribution was implemented utilizing the length-biased approach in comparison to the conventional distribution. Its various statistical properties were derived and studied. Harmonic mean, moments, the shape of the cdf and pdf, the mean, variance, and standard deviation, as well as survival functions, hazard rate functions, order statistics, reverse hazard rate functions, Bonferroni, & Lorenz curves are observed. The MLE of the parameters of the distribution has been calculated. Two Biomedical data sets have been used to test the new distribution's superiority, and the findings show that the **CBP** distribution offers a far better fit than the new two-parameter Sujatha, Sujatha, two-parameter Sujatha, exponential, as well as Lindley distributions in the case of such biological data.

No Conflict of interest.

We declare there is no conflict of interest.

No funding agencies.

There are no funding agencies for this research article.

References

[1] Alhyasat, K., Kamarulzaman, I., Al-Omari, A. I. and Abu Bakar, M. A. (2020). Power size biased two parameter Akash distribution. *Statistics in Transition new series*, Vol. 21, No. 3, 73-91.

[2] Al-Kadim, K. A. and Hussein, N. A. (2014). New proposed length-biased weighted exponential and Rayleigh distribution with application. *Mathematical Theory and Modeling*, 4(7), 137-152.

[3] Beghriche, A. and Zeghdoudi, H. (2019). A size biased gamma Lindley distribution. *Thailand Statistician*, 17(2), 179-189.

[4]]Cox, D. R. (1969). Some sampling problems in technology, In New Development in Survey Sampling, Johnson, N. L. and Smith, H., Jr .(eds.) *New York Wiley- Interscience*, 506-527.

[5] Cox, D. R. (1962). Renewal theory, *Barnes and Noble, New York*.

[6] Dar, S. A., Hassan, A. and Ahmad, P. B. (2022). Poisson size-biased Lindley distribution and its applications. *International Journal of Modeling, Simulation and Scientific Computing*, Vol. 13, No. 04, 2250031.

[7] Ekhosuehil, N., Kenneth, G. E. and Kevin, U. K. (2020). The Weibull length biased exponential distribution: Statistical properties and applications. *Journal of Statistical and Econometric Methods*, Vol. 9, No. 4, 15-30.

[8] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13-25.

[9] Ganaie, R. A. and Rajagopalan, V. (2020). Length biased weighted quasi gamma distribution with characterizations and applications. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(5), 1110-1117.

[10] Mobarak, M. A., Nofal, Z. and Mahdy, M. (2017). On size-biased weighted transmuted Weibull distribution. *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 7, Issue 3, 317-325.

[11] Mustafa, A. and Khan, M. I. (2022). The length-biased powered inverse Rayleigh distribution with applications. *J. Appl. Math. & Informatics*, 40(1-2), 1-13.

[12] Rasool, S. U. and Ahmad, S. P. (2022). Power length biased weighted Lomax distribution. *RT &A*, 17(4), 543-558.

[13] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; *Pergamum Press and Statistical publishing Society, Calcutta.* 320-332.

[14] Sanat, P. (2016). Beta-length biased Pareto distribution and its properties. *Journal of Emerging Technologies and Innovative Research (JETIR),* Vol. 3, Issue 6, 553-557.

[15] Sen, S., Chandra, N. and Maiti, S. S. (2017). The weighted Xgamma distribution: Properties and application. *Journal of Reliability and Statistical Studies*, 10(1), 43-58.

[16] Tesfay, M. and Shanker, R. (2018). A new-two parameter Sujatha distribution with [17] properties and applications. *Turkiye Klinikleri J Biostat*, 10(2), 96-113.

[17] Tesfay, M. and Shanker, R. (2018). A two-parameter Sujatha distribution. *Biometrics & Biostatistics International Journal*, 7(3), 188-197.

[18] Zelen, M. (1974). Problems in cell kinetic and the early detection of disease, in Reliability and Biometry, F. Proschan & R.J. Sering, eds, *SIAM, Philadelphia*, 701-706.

FORECASTING OF EXTREME RISK USING MARKOV-SWITCHING GARCH MODELS: EVIDENCE FROM GLOBAL ENERGY MARKETS

S. Kavitha[1], G. Mokesh Rayalu*[2], D. Pachiyappan[3], and P. Manigandan[3]

•

[1] Assistant Professor, Department of Statistics,
Periyar University, Salem – 636011, Tamil Nadu, India.
pustatkavitha@gmail.com

[2] Assistant Professor, Department of Mathematics
School of Advanced Sciences, VIT, Vellore-632 014, Tamil Nadu, India.
Corresponding author mail: mokesh.g@vit.ac.in

[3] Department of Statistics,
Periyar University, Salem – 636011, Tamil Nadu, India.
d.pachiyappan321@gmail.com, srimanigandan95@gmail.com

Abstract

This paper investigates the Markov-Switching GARCH and Single-Regime (SR) GARCH models for the extreme-risk prediction of the global energy markets. Using daily data from Jan. 2020 to July. 2022, we find the MS-GARCH-types models are appropriate for both developed and emerging energy markets because they efficiently measure the extreme risk of energy commodities in various cases. Meanwhile, the regime-switching model's capture-dynamic structures in the financial markets and this model is only better than single-regime models in terms of long position risk predicting, rather than short position risk forecasting. That is, on the downside risk predicting, it just outperforms the single regime. Through competitive models, this study examines the risk forecast of energy commodities in different conditions. The findings have strong implications for investors and policymakers in selecting the appropriate model to predict the extreme risk of energy commodities when facing asset allocation, portfolio selection, and risk management.

Keywords: MS- GARCH, extreme risk, energy markets, prediction**.**

I. Introduction

To ensure economic stability and improve national security, energy commodities are among the most important natural resources used by countries as inputs in transportation, industry, and many other economic sectors. Natural gas, oil, coal is the most used major energy sources [1,2]. The oil demand is rose steadily in 2018, with China and India leading its major consumption in the United States. The US, which overtook Saudi Arabia in mid-2020, is currently the world's largest oil exporter and heavy crude oil importer [3]. Natural gas consumption increased by 4.6%, which is almost half of global energy consumption. The global demand for coal energy has continued to rise for two years since 2018. The coal-driven electricity supply is very important in Asia to meet India,

China, Indonesia, South and East Asia. Since natural gas and coal are the main sources of electricity and heating, an increment in the price of energy products is expected to affect household cash flow. In contrast, oil is a fundamental input for industrial production, thus severely affecting inflation rates.

The global pandemic has revealed the vulnerability of the world economy and energy commodities to external shocks. For example, the covid-19 oil demand shock triggered an estimated 10% drop in demand, leading to a more than 60% price drop from Jan. to Apr. 2020. To prevent the shock, OPEC members agreed to reduce oil production by an estimated 9.7 million b/d in Apr.2020. Also, [4] show that volatile oil prices can trigger price fluctuations in other energy products have widespread effects on the international economy. The volatility of oil prices in the 1970s attracted much interest from financial investors, academia, and policymakers as oil-importing and exporting countries became a major factor in various economic sectors. Early pioneers studied the relationship between oil prices and economic activity-demonstrating the significant impact of oil price on macroeconomic activity and the partial responsibility for the post-world war-II US recession from 1973 to 1983 [5]. There are several methods for predicting fluctuations, but the most popular literature is the GARCH models.

Since a key contribution made by [6], who generalized autoregressive conditional heteroscedasticity (GARCH) was introduced by [7], a common in modelling VaR is that GARCH-type models are related to conditional error distribution [8]. This model depends on the suitably estimated volatility. However, conventional GARCH-type specifications belonging to the single regime model are difficult to capture structural changes during economic cycles. Furthermore, by [9], the generalized Markov regime Switching-ARCH model was introduced by [10], a researcher on volatility with regimes in the financial market. A study has found that the MS-GARCH provides far superior to the single regime GARCH on the modelling volatility [11,12].

The contribution of this study is different from two recent extreme risk studies on the two-type of models [13] confirmed that the MS-GARCH-type models had improved predictive accuracy than the standard GARCH-type models on downside risk predictions for global energy commodities. First, they should mainly investigate volatility predictions between three Single Regime-GARCH models and six-types of Regime-Switching models and the just involved two quantiles of the upside Value at Risk content in the appendix. Secondly, according to descriptive analyses of return series, they consider innovations as Normal distribution, but different distributions have a major impact on the model fit, estimation of volatility, and Value at Risk calculation. Based on this point, we consider the most common and effective distributions (Normal (*N*) and student-t (*S*)) for modelling. Third, they only investigated VaR prediction, whereas we also complement ES (expected shortfall) predicting related to extreme risk. To make more reliable conclusions, this operation compares the differences in risk measures between the two types of models in depth.

Fourth, [14] used the same back testing method for evaluating forecasting performance on volatility to assess VaR prediction, whereas we use two prevalent back testing methods used by substantial scientists and researchers [15,16,17,18,19,20] in the field of risk management, to estimate the extreme risk prediction between the two models. When we read them in the global energy commodity risk prediction, we put the RS models and their SR counterparts on the same condition. These procedures ensure that the results are more reliable. As a result, our research differs from [13] and [21] regarding energy commodity risk predicting. The literature related to the risk prediction of energy commodities via the MS-GARCH model is still limited. If the RS model shows better performance in risk measures, it will be recommended to apply for portfolio optimization and risk management. Otherwise, predicting risk values may make capital allocation insufficiently efficient as policymakers and investors set up ineligible assets against market risks.

The rest of this paper is as follows. In section 2, we present the econometric methodology

adopted; section 3 describes data and summaries the descriptive statistics of the global energy commodity return series. In section 4 describes estimation results, and section 5, conclusions of this paper.

2. Methodology

2.1 GJR-GARCH model

The GJR-GARCH model is given by Glosten et al. [22]. GJR-GARCH (1,1) is written as:

$$J^2_{k,t} = \mu_{0,k} + \alpha\lambda^2_{1,k} I\{y_{t-1}\ ^3 0\} - \beta\lambda^2_{2,k} I\{y_{t-1} < 0\} y_{t-1} + \delta J^2_{k,t-1} \quad (1)$$

Where the asymmetric effect is attributed to component $\lambda^2_{2,k} y_{t-1}$ where the parameter $y_{t-1} = 0$ if $\lambda_t > 0$ means shocks on volatility from bad news and $y_{t-1} = 1$ otherwise.

2.2 EGARCH model

The Exponential GARCH (EGARCH) Model of Nelson [23] is given by:

$$\text{In}\left(J_{k,t}\right) = \mu_{0,k} + \alpha_{1,k}\left(\left|\eta_{k,t-1}\right| - E\left[n_{k,t-1}\ |\right]\right) + \alpha_{2,k}\eta_{k,t-1} + \beta_k \text{In}\left(J_{k,t-1}\right) \quad (2)$$

Where $E\left[\left|\eta_{k,t-1}\right|\right]$ is a conditional expectation on regime k, and we have $\theta_k = \left(\mu_{0,k}, \alpha_{1,k}, \alpha_{2,k}, \beta_k\right)^T$ for (k = 1, … , K). This specification model deals with the asymmetric reaction of volatility to previous returns i.e., leverage effect [24,25]. Covariance stationary into each regime is obtained by needful that $\beta_k > 0$.

2.3 TGARCH model

Zakoian 1994[26], introduces the Threshold GARCH (TGARCH) model specification, where the conditional volatility is an explanatory variable instead of the conditional variance. This model is given by:

$$J^{1/2}_{k,t} = \mu_{0,k} + \left(\alpha_{1,k} I\{y_{t-1}\ ^3 0\} - \alpha_{2,k} I\{y_{t-1} < 0\} y_{t-1} + \beta_k J^{1/2}_{k,t-1}\right) \quad (3)$$

We have $\theta = \left(\mu_{0,k}, \alpha_{1,k}, \alpha_{2,k}, \beta_k\right)^T$ for (k = 1 … , K). To ensure positivity, we needful than $\mu_{0,k} > 0,\ \alpha_{1,k} > 0, \alpha_{2,k} > 0$ and $\beta_k \geq 0$. Obtained by requiring that for Co-variance stationarity in each regime $\alpha^2_{1,k} + \beta^2_k - 2\beta_k\left(\alpha_{1,k} + \alpha_{2,k}\right) E\left[\eta_{t,k} I\{\eta_{t,k} < 0\}\right] - \left(\alpha^2_{1,k} - \alpha^2_{2,k}\right) E\left[\eta_{t,k} I\{\eta_{t,k} < 0\}\right] < 1$ Francq and Zakoïan [27]. The quantities of $E\left[\eta_{t,k} I\{\eta_{t,k} < 0\}\right]$ and $E\left[\eta^2_{t,k} I\{\eta_{t,k} < 0\}\right]$ required for the conditions of co-variance stationarity in the TGARCH model [28]. We assume two different probability distributions for D(.) we use student normal (N) and student-t (S) distribution. Then explore the advantages of incorporating the skewness in our analysis by considering the standard skewed versions of N, and S obtained using the mechanism of [29] and [30]. $K \in (1,2,3)$ are the number of regimes: we label our specification SR when K = 1 and MS when K = 3.

2.4 Model Estimation

We estimate all selected models by using maximum likelihood (ML) techniques. This approach requires the estimation of the likelihood function. The first step in this process is to generate the likelihood for MS-GARCH model specifications by collecting the vector of model parameter into $\Omega \equiv (\theta_1, \xi_1, \ldots, \theta_k, \xi_k, P)$. The conditional density of y_t in a state $h_t = k$ given Ω and τ_{t-1} is denoted by $f(y_t \mid h_t = k, \Omega, \tau_{t-1})$ integrating out state h_t we obtain the density of y_t given Ω and τ_{t-1} only. The discrete integration is obtained as

$$f(y_t \mid \Omega, \tau_{t-1}) \equiv \sum_i^k \sum_j^k p_{i,j}\delta_{i,t-1} \times f_D(y_t \mid s_t = j, \Omega, \tau_{t-1}) \quad (4)$$

Where $\delta_{i,t-1} \equiv P[s_{t-1} = i \mid \Omega, \tau_{t-1}]$ represents the filtered probability of state i at time $t-1$ and

where $p_{i,j}$ is the transition probability of moving from i to state j. The filtered probabilities $(\delta_{i,t}; k = 1, \dots, K; 1, \dots, T)$ are obtained via Hamilton filter. Finally, the likelihood function is obtained from Eq. (2) as follows:

$$L(\Omega \mid \tau_T) \equiv \prod_{t=1}^{T} f(y_t \mid \Omega, \tau_{t-1}) \tag{5}$$

The estimator of maximum likelihood $\hat{\Omega}$ is obtained by maximizing the algorithm of Eq (5).

2.6 Risk Measures

Value at risk is the general estimate of the maximal loss when the position declines due to market movements in the financial domain. One step forward conditional probability density with the two regimes is computed:

$$f(y_t \mid \Omega, \tau_t) \equiv \sum_{h=1}^{2} \pi_{h,t} f_D(y_t \mid s_t = k, \Omega, \tau_{t-1}) \tag{6}$$

Since, The PDF is a combination of two-regime distribution, $\pi_{h,t} = \sum_{i=1}^{2} p_{i,h} J_{i,t-1}$ where $\theta_{i,t-1} = P(s_{t-1} = i \mid \Omega, \tau_{t-1})$ is filtered probability, a one step ahead cumulative distribution function (CDF) with regimes is obtained through its conditional probability density distribution:

$$F(y_t \mid \Omega, \tau_{t-1}) \equiv \int_{-¥}^{\varepsilon_t} f(x \mid \Omega, \tau_{t-1}) dx \tag{7}$$

Where the model parameters Ω is estimation by ML in equation (7). Financial regulators utilize VaR to evaluate risks at a particular probability level. The following is how VaR is defined:

$$\begin{aligned} &\Pr[s_t < VaR] = 1 - p \\ &VaR_t^{1-p} = u_t + \hat{\rho} s_{t-1} + z_{1-p} h_t \end{aligned} \tag{8}$$

Where $VaR_{t|t-1}^{1-p}$ is represents the maximal loss of long-position and F_Z is CDF of innovations z_t.

$$\begin{aligned} &F(y_t \mid \Omega, \tau_{t-1}) = F(z_t \mid \Omega, \tau_{t-1}) h_y \\ &VaR_t^{1-p} = u_t + \hat{\rho} s_{t-1} + F^{-1}(1 - p \mid \Omega, \tau_{t-1}) h_t \end{aligned} \tag{9}$$

When calculating the risk of a short position, p substitutes for $1 - p$. Despite its simplicity and ease of implementation, VaR has drawbacks due to its lack of coherence as a risk measure. On the other hand, expected shortfall (ES) as measures of average losses exceeding VaR can overcome this flaw through entailing the magnitude of losses. As a result, we calculate ES to compare the predicting performance of the two types of models. ES is calculated as follows:

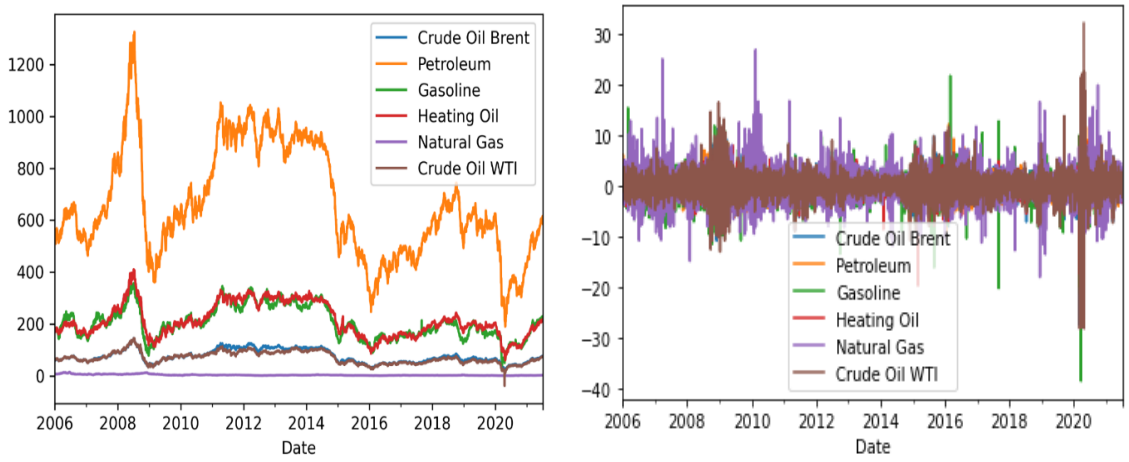
$$\begin{aligned} &ES^{1-p} = E(s_t \mid s_t < VaR^{1-p}) \\ &ES_{t|t-1}^{1-p} = \frac{1}{1-p} \int_{-\infty}^{VaR_t^{1-p}} x f(x \mid \tau_{t-1}) dx \end{aligned} \tag{10}$$

The short and long positions of VaR and ES at two quantiles are investigated in this research.

3. Data and Descriptive Statistics

This study used daily data from January 2, 2019, to July 8, 2022. Each series of datasets includes 4049 observations. The first sub-sample is used for in-sample analysis and parameter estimation, while the second sub-sample, the last 1,500 observations from the entire sample, is used for out-of-sample forecasting. The six types of energy commodities are namely, Crude oil brent, Petroleum,

Gasoline, Heating oil, Natural gas, and Crude oil WTI, which are obtained from the Federal Reserve Economic Database (FRED) (https://fred.stlouisfed.org/). Energy commodity prices used in modelling, are calculated by where is the spot prices of the global energy commodity at time t. Figure 1, shows that the global energy commodity prices and returns are reported in respectively.



Graph 1: *Global energy prices*

Table .1 reports descriptive statistics for the energy price. We obtain that the mean values for Crude Oil Brent, Petroleum, Gasoline, Heating Oil, Natural Gas, and Crude Oil WTI is 0.0054%, 0.0035%, 0.0059%, 0.0046, -0.0187, and 0.0189 respectively. Meanwhile, the values of minimum and maximum have reflected the presence of small extreme returns. The estimation of unconditional volatility is through standard deviation, Natural Gas have the highest volatility.

Table 1: *Descriptive statistics of energy price*

Metrics	Crude Oil Brent	Petroleum	Gasoline	Heating Oil	Natural Gas	Crude Oil WTI
Mean	0.0054	0.0035	0.0059	0.0046	-0.0189	0.0189
Min	-27.976	-19.211	-38.535	-19.995	-18.054	-28.220
Max	19.077	13.723	22.396	10.946	26.771	31.963
Std.dev	2.305	1.955	2.656	2.079	3.203	2.655
Skewness	0.254	0.112	0.998	0.380	0.556	0.212
Kurtosis	13.753	8.473	25.923	9.319	7.729	24.806
PP test	-4328.1*	-4153.4*	-4329.9*	-4286.1*	-4126.3*	-4173.5*
ADF test	-14.092*	-14.391*	-14.265*	-15.412*	-16.037*	-14.219*
Q-(10)	27.973**	24.496**	28.306**	13.445	31.323***	30.555***
Jarque-Bera	19544***	5060.1***	89280***	6832.9***	3979.9***	80217***
ARCH-LM (10)	760***	553 ***	583***	1034***	921***	972***

Table-1 is also showing negative and positive but significant skewness for energy commodities return series which means that energy commodities return series have longer left-tails and fat-tails than the normal distribution. The kurtosis is highly significant for energy commodities return series and Gasoline display larger Kurtosis than the other return series. Values of Jarque-Bera obtained through (Jarque & Bera, 1980) depicting the rejection of normality. The significant values of Phillip Perron (PP) [31] and Augmented Dickey-Fuller (ADF)[32] test indicating that energy commodities return series are stationary. Ljung-Box is showing the (Ljung & Box, 1978) Q-statistics at 10th order for autocorrelation in raw data are extremely significant and rejecting the null hypothesis of no autocorrelation. ARCH-(1) test for restricted heteroscedasticity giving a strong indication of ARCH effect in energy commodities returns series, this evidence suggests that the usefulness and suitability of GARCH-type methods for prediction and modelling their time-varying conditional volatility. These findings usually indicate the high degree of perseverance in

the conditional-volatility procedure of energy commodity prices.

4. Empirical Results

In this section, the estimation result of the MS-GARCH-type models with student normal (*N*), student-t (*S*) distribution is presented in Table 2.

4.1. In-Sample Estimation

The estimated parameters for the MS-GARCH and EGARCH models are given in Table 2. According to our results, the parameters for conditional variance are statistically significant for energy commodity prices. Almost all parameters of the EGARCH and MSGARCH (GJR-GARCH-EGARCH-TGARCH, GARCH-EGARCH, GJR-GARCH-TGARCH, EGARCH-TGARCH, GJR-GARCH, and TGARCH) models are significant, especially β_k is the leverage parameter, which implies the leverage effects of significant volatility. The energy commodity markets reveal strong evidence of asymmetric volatility, while negative news responds with strong shocks to energy commodity fluctuations. Therefore, more useful to capture the volatility of five types of energy commodities based on the two types of models. Meanwhile, Table-2 represents the transition probabilities, which mean three significant regimes volatility in the energy commodities. Therefore, the dynamic structure of energy commodities will change over periods.

Considering practical fitting capabilities, three types of criteria are used to test their appropriate performance, including Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), and log-likelihood (LL). These results reveal significant evidence of fitting efficacy for the energy commodity return series. The MS-GARCH model is successfully for Crude Oil Brent, Petroleum, Gasoline, Heating oil, and Crude Oil WTI, and it is not excellent to the counterpart in the Natural gas due to highest BIC and AIC values. Thus, these results confirm that the MS-GARCH modelling energy commodity prices dataset is appropriate and outperforms the single-regime counterpart in more cases. Nevertheless, good in-sample model fits may not generate accurate and predictions of reliable risk.

Table 2: *Modelling volatilities of energy commodity price by using MS-GARCH*

Global energy markets	**Crude Oil Brent**	**Petroleum**	**Gasoline**	**Heating Oil**	**Natural Gas**	**Crude Oil WTI**
Models (Regime-3)	GJR-GARCH-TGARCH	GARCH-EGARCH	GJR-GARCH - TGARCH	EGARCH-TGARCH	GJR-GARCH	TGARCH
Regime-1						
μ_1	0 .0001 (0.000)***	0.0011 (0.001)**	0.0108 (0.010)**	0.0137 (0.007)**	0.2958(0.188)**	0.1003 (0.041)**
α_1	0.0562 (0.000)***	0.0014 (0.001)**	0.0032 (0.004)**	0.0100 (0.008)**	0.0161(0.016)**	0.0134 (0.013)**
$\alpha_{2\text{-}1}$	0.9206 (0.000)***	0.0083 (0.005)***	0.0092 (0.001)***	0.0462 (0.005)***	0.0436 (0.004)*	0.0311 (0.016)**
β_1	0.0992 (0.000)***	0.9959 (0.001)***	0.9842 (0.012)***	0.9604 (0.010)***	0.9157(0.037)***	0.8966 (0.018)***
$\nu_{\text{-}2}$	-	-	-	-	3.0072(0.002)*	
$\xi_{\text{-}2}$	-	1.0436 (0.0043)**	-	-	-	4.0236 (0.0815)**
Regime-2						
μ_2	0.3544 (0.000)***	0.0164 (0.0052)**	0.0186 (0.0128)**	0.0064 (0.0043)*	0.0825 (0.0378)*	0.0164 (0.0078)**

α_2	0.34790 (0.0000)***	0.0067 (0.0019)**	0.0253 (0.0159)**	0.0361 (0.0412)*	0.0752 (0.0471)**	0.0588 (0.0370)**
$\alpha_{2\text{-}2}$	0.6860 (0.000)***	0.0421 (0.0061)***	0.0112 (0.0036)***	0.0325 (0.0065)***	0.0132 (0.0021)**	0.4301 (0.0531)**
β_2	0.8728 (0.000)***	0.9527 (0.131)***	0.9700 (0.0034)***	0.6424 (0.0017)***	0.9138 (0.0071)***	0.9366 (0.0028)***
ν_{-2}	-	3.8740 (0.8006)**	2.0114 (0.2213)*	-	2.0351 (0.0641)*	-
ξ_{-2}	-	-	-	-	-	-
Regime -3						
μ_3	0.9902 (0.000)***	0.3313 (0.1966)**	0.1103 (0.024)*	11.2219 (2.7450)***	39.3173 (8.7254)***	3.3297 (1.4864)**
α_3	0.0091 (0.0001)***	0.0378 (0.0543)**	4.6579 (2.5728)***	0.9999 (0.000)***	0.0835 (0.0117)***	0.4381 (0.4013)**
$\alpha_{2\text{-}3}$	0.2988 (0.000)***	0.9527 (0.0131)***	0.2188 (1.7872)*	0.0624 (0.0012)***	0.9998 (0.0072)***	0.3463 (7.5013)**
β_3	0.0091 (0.000)***	0.4903 (0.6123)***	0.7803 (0.0079)***	0.0018 (0.000)*	0.000 (0.000)**	0.6535 (0.0058)***
ν_{-3}	-	1.2856 (0.0485)*	-	-	1.2852 (0.085)*	-
ξ_{-3}	-	-	-	-	-	-
Probabilities						
$P_{1,1}$	0.7342	0.9847	0.9425	0.9364	0.7290	0.7287
$P_{2,1}$	0.6555	0.4096	0.2860	0.9592	0.1148	0.9744
$P_{3,1}$	0.8334	0.6531	0.9582	0.5550	0.9172	0.6105
LL	-821.98	-778.19	-874.02	-802.52	-1000.88	-854.2
AIC	1645.96	1559.38	1752.04	1607.05	2003.76	1712.4
BIC	1655.14	1568.38	1761.63	1616.63	2013.34	1721.98

Moreover, risk management sectors and professional are particularly interested in risk predictions. Based on these considerations, we continue to analyze the results of prediction risk at different horizons and significant levels from two types of models.

4.2 Out-of-Sample Risk Forecasting

The previous hypotheses are addressed in this sub-section. To learn more about risk prediction for the two types of models, large-scale comparison studies are conducted. One-ahead and five-ahead forecasts are used in these comparisons between regime-switching and single-regime scenarios. Each energy commodity market's predicting results include the upside and downside risk, as well as the two quantiles of two distributions. The reason for including so many cases in the empirical process is that these adequate experiments reveal the differences between the two types of models. Indeed, this procedure is used to generate much more precise results and subsequently reach robust conclusions. In fact, the VaR and ES methods are combined to measure the extreme risk of energy commodities in a comprehensive way. Downside risk is considered long position, and upside is short position. Table 3 and 4 provide the one-day predict outcomes for six-types of energy commodities, whereas Table 5 and 6 indicates the five-day predict results. Each table has three benchmark models: GARCH, GJR-GARCH, and TGARCH, each of which include the Regime-Switching (RS) model and the Single-Regime (SR) counterpart with two distributions.

Table 3 and 4 shows that the predicting outcomes for Brent, Petroleum, Gasoline, Heating Oil, Natural Gas, and WTI one-day risk predictions, respectively. When the same distribution is employed, the figures in bold in the table denote the better model between the single-regime

model and the regime-switching equivalent in terms of predicting performance. This kind would be better if it has more bold figures than the other under a basic GARCH-type model, such as GARCH, EGARCH, or TGARCH. In this method, the predicting results are assessed using the bolded figures in the table. Apart from the EGARCH case for Natural gas, the RS model is just superior to the corresponding RS in terms of downside risk under the same distribution for the developed energy commodities. Interestingly, the findings on upside risk outcomes of the gasoline have conversed with the downside. In the case of the emerging oil market, the RS model outperforms the SR model in terms of downside risk, apart from the EGARCH example, however, this conclusion is inconsistent with the upside.

The evidence from the one-day ahead findings just shows that the RS model outperforms the SR model only on the downside, not on the upside. This conclusion is appropriate for both developed and emerging markets. Furthermore, the predicting horizon may have an impact on the result, and a five-day-ahead forecast is made based on this consideration.

To summarise, our findings support the two previously proposed predictions about the risk predicting abilities of the two types of conditional variance models. In terms of energy commodities downside risk predicting, the RS model just outperforms the SR model. The RS models are also appropriate for developing energy commodities. More precisely, risk predicting results represent some important findings. First, the MS-GARCH-types models are suitable for both developed and emerging energy commodities, particularly predicting downside risk. If two types of models are used for the five-day prediction, there are some changes in the risk predictions of the upside risk. Second, our findings require policymakers, risk managers, and investors when hedging and investing in energy commodities, as they must carefully control possible extreme risks. The complex model with regime-switching may not always provide far superiority to, all the time, the SR model in the case of risk predictions for both long and short positions. More importantly, some risk management practitioners and scholars may consider the regime-switching model a preferable option for risk predicting. However, the results obtained based on the regime-switching model can lead to massive losses because this model does not always measure the financial tail risk well, especially for the upside risk in this paper.

Table 3: *One-day forward risk predictions of DQ-test in energy commodities*

DQ test	Brent	Petroleum	Gasoline	Heating oil	Natural gas	WTI
			Long Position-0.01			
Single-Regime GARCH-N	0.672	0.4371	0.6745	0.9572	0.3421	0.4351
EGARCH- N	0.8845	0.5076	0.623	0.7643	0.4576	0.2001
TGARCH- N	0.6713	0.5354	0.4032	0.6872	0.432	0.3982
GARCH-S	0.2152	0.3573	0.3561	0.6461	0.04	0.3065
EGARCH- S	0.5302	0.1765	0.431	0.4701	0.5762	0.371
TGARCH- S	0.3965	0.2301	0.319	0.0231	0.3321	0.3361
MS-GARCH	0.0545	0.1298	0.6802	0.1065	0.0365	0.7865
GARCH-EGARCH-TGARCH- N- S						
GARCH- EGARCH- N- S	0.1643	0.5587	0.9171	0.1494	0.0294	0.5431
GARCH -TGARCH N- S	0.0476	0.8385	0.7562	0.2542	0.154	0.8142
EGARCH-TGARCH- N- S	0.1457	0.6814	0.8362	0.1376	0.3751	0.6803
EGARCH- N- S	0.1223	0.5467	0.8062	0.046	0.1361	0.745
TGARCH- N- S	0.1098	0.1452	0.9301	0.0636	0.4316	0.7714

			Short position-0.05			
Single-Regime GARCH-N	0.1764	0.3461	0.6262	0.1265	0.0001	0.1782
EGARCH-N	0.4761	0.4087	0.5342	0.134	0.0089	0.2214
TGARCH- N	0.36	0.5371	0.6942	0.1042	0.1563	0.3672
GARCH-S	0.4753	0.4401	0.5643	0.1243	0.1264	0.4583
EGARCH- S	0.4748	0.415	0.6301	0.1561	0.0463	0.5301
TGARCH- S	0.5851	0.5215	0.7164	0.1164	0.084	0.5751
MS-GARCH	0.084	0.1165	0.823	0.2367	0.1567	0.6632
GARCH-EGARCH-TGARCH- N- S						
GARCH- EGARCH- N- S	0.1315	0.5681	0.8813	0.2224	0.1653	0.6436
GARCH -TGARCH N- S	0.1097	0.8409	0.842	0.3923	0.4303	0.7841
EGARCH-TGARCH- N- S	0.3925	0.8409	0.9217	0.2576	0.0935	0.8024
EGARCH- N- S	0.2517	1	0.8806	0.1604	0.3623	0.6421
TGARCH- N- S	0.32	0.303	0.8325	0.0487	0.138	0.8715

Table 4: *One-day forward risk predictions of CC-test in energy commodities*

CC-test	Brent	Petroleum	Gasoline	Heating oil	Natural gas	WTI
			Long Position-0.01			
Single-Regime GJR-GARCH-N	1783	0.3152	0.1132	0.4214	0.221	0.6571
EGARCH-N	0.4603	0.3105	0.2315	0.3244	0.0431	0.571
TGARCH- N	0.34	0.1004	0.1048	0.3531	0.0935	0.8122
GARCH-S	0.4253	0.419	0.119	0.4413	0.0652	0.6142
EGARCH- S	0.3831	0.3362	0.3362	0.371	0.1043	0.865
TGARCH- S	0.2016	0.3254	0.2431	0.631	0.041	0.5304
MS GARCH EGARCH-TGARCH- N- S	0.6705	0.761	0.1115	0.4294	0.221	0.7853
GARCH- EGARCH- N- S	0.5341	0.7112	0.6704	0.5506	0.1425	0.8664
GARCH -TGARCH N- S	0.8506	0.5044	0.3546	0.8403	0.0972	0.8553
EGARCH-TGARCH- N- S	0.6131	0.7532	0.2437	0.8403	0.1612	0.896
EGARCH- N- S	0.7733	1	0.3356	0.1689	0.1047	0.7364
TGARCH- N- S	1	0.5377	0.2422	0.0559	0.224	0.572
			Short position-0.05			
Single-Regime GJR-GARCH-N	0.5661	0.551	0.7631	0.7451	0.0043	0.587
EGARCH-N	0.6631	0.6142	0.3623	0.4632	0.2305	0.3756
TGARCH- N	0.31	0.4852	0.3621	0.5102	0.065	0.767
GJR-GARCH-S	0.4748	0.3421	0.2306	0.6772	0.0004	0.4603
EGARCH- S	0.5661	0.4212	0.0432	0.0421	0.1267	0.655
TGARCH- S	0.776	0.1502	0.139	0.4682	0.0001	0.467
MS-GARCH GARCH-EGARCH-TGARCH- N- S	0.774	0.3371	0.0079	0.7145	0.1087	0.7541
GARCH- EGARCH- N- S	0.5663	0.3998	0.1446	0.5723	0.0465	0.7567
GARCH -TGARCH N- S	0.4746	0.6073	0.105	0.6147	0.108	1
EGARCH-TGARCH- N- S	0.773	0.5903	0.0105	0.8563	0.1123	0.5366

GJR-GARCH- N- S	0.885	0.3912	0.1845	1	0.1268	0.362
TGARCH- N- S	0.6842	0.2467	0.1343	0.5377	0.0132	0.3789

.

Table 5: *five-day forward risk predictions of Dynamic Quantile (DQ)-test in global energy commodities*

DQ-test	Brent	Petroleum	Gasoline	Heating oil	Natural gas	WTI
		Long Position-0.01				
Single-Regime GJR-GARCH-N	0.773	0.0856	0.6261	0.3685	0.0531	0.614
EGARCH-N	1	0.1506	0.5831	0.371	0.0253	0.4731
TGARCH- N	0.615	0.3132	0.3102	0.4631	0.1564	0.5362
GJR-GARCH-S	0.771	0.331	0.5771	0.74	0.0531	0.4001
EGARCH- S	0.788	0.2310	0.3194	0.5525	0.682	0.3134
TGARCH- S	0.3925	0.3135	0.4849	0.3891	0.5412	0.3134
MS-GARCH	0.4745	0.7253	0.8384	0.4061	0.3048	0.8254
GARCH-EGARCH-TGARCH-N- S						
GARCH- EGARCH- N- S	0.34	0.7468	1	0.311	0.109	0.7856
GARCH -TGARCH N- S	0.6671	0.8734	0.5351	0.2472	0.0362	0.861
EGARCH-TGARCH- N- S	0.5667	0.7887	0.406	0.05	0.0595	0.8662
GJR-GARCH- N- S	0.2361	0.9567	0.6814	0.5632	0.159	0.8988
TGARCH- N- S	0.34	0.8935	0.2943	0.4273	0.5362	0.7411
		Short position-0.05				
Single-Regime GJR-GARCH-N	0.32	0.5661	0.0557	0.3139	0.0045	0.8065
EGARCH-N	0.2041	0.732	0.165	0.7465	0.0288	0.703
TGARCH- N	0.1587	0.661	0.3782	0.3135	0.2766	0.9302
GARCH-S	0.461	0.4741	0.5501	0.3135	0.0119	0.8092
EGARCH- S	0.2351	0.1746	0.1739	0.5131	0.0231	0.7632
TGARCH- S	0.6623	0.5312	0.4428	0.4543	0.0047	0.6134
MS-GARCH	0.3925	0.8314	1	0.2174	0.045	0.706
GARCH-EGARCH-TGARCH-N- S						
GARCH- EGARCH- N- S	0.2573	0.7319	1	0.3411	0.319	0.9302
GARCH -TGARCH N- S	0.885	0.8664	0.6813	0.6339	0.1202	0.819
EGARCH-TGARCH- N- S	0.663	0.945	0.5362	0.5659	0.0288	0.6213
GJR-GARCH- N- S	0.9262	0.9267	0.6813	0.2308	0.9561	0.8975
TGARCH- N- S	0.3926	0.8127	0.2943	0.2945	0.2108	0.6687

Table 6: *five-day forward risk predictions of Conditional Correlation (CC)-test in global energy commodities*

CC-test	Brent	Petroleum	Gasoline	Heating oil	Natural gas	WTI
		Long Position-0.01				
Single-Regime GARCH-N	0.2413	0.4351	0.2413	0.5361	0.0219	0.7472
EGARCH-N	0.7541	0.8484	0.8384	0.1345	0.0192	0.8365
TGARCH- N	0.6748	0.5959	0.5501	0.7696	0.082	0.7541
GARCH-S	0.671	0.406	0.981	0.5377	0.0687	0.81
EGARCH- S	0.8346	0.3742	0.6851	0.7631	0.1216	0.8501

TGARCH- S	0.7812	0.82	0.5351	0.592	0.001	0.6725
MS-GARCH	0.5385	0.5043	0.434	0.5132	0.0425	0.5825
GARCH-EGARCH-TGARCH-N- S						
GARCH- EGARCH- N- S	0.6878	0.4293	0.3132	0.7425	0.0712	0.7113
GARCH -TGARCH N- S	0.7886	0.8403	0.5771	0.7631	0.002	0.7603
EGARCH-TGARCH- N- S	0.941	1	0.3493	1	0.0015	0.7718
EGARCH- N- S	0.5806	0.5484	0.0691	1	0.0245	0.9103
TGARCH- N- S	0	1	0.5342	0.5674	0.0166	0.8826
		Short position-0.05				
Single-Regime GARCH-N	0.7523	0.4134	0.1732	0.4601	0.0661	0.7652
EGARCH-N	0.934	0.1954	0.5506	0.1203	0.3134	0.7164
TGARCH- N	0.6531	0.63	0.3235	0.0261	0.0772	0.5972
GARCH-S	0.778	0.1421	0.1686	0.2876	0.0961	0.7261
EGARCH- S	0.6578	0.1833	0.3136	0.0324	0.591	0.7278
TGARCH- S	0.7751	0.1696	0.1686	0.2621	0.0053	0.5315
MS-GARCH	0.7412	0.1661	0.5384	0.3402	0.0094	0.6811
GARCH-EGARCH-TGARCH-N- S						
GARCH- EGARCH- N- S	0.7415	0.3126	0.6976	0.3914	0.0044	0.1928
GARCH -TGARCH N- S	0.8598	0.0961	0.3135	0.2085	0.0144	0.789
EGARCH-TGARCH- N- S	0.8758	0.5509	0.077	0.3431	0.0054	0.5149
EGARCH- N- S	0.662	0.6813	0.039	0.0187	0.0002	0.6834
TGARCH- N- S	0.7886	0.5362	0.597	0.1543	0.0002	0.5696

These findings on downside risk prediction are consistent with Teterin et al.[32], who demonstrated that the RS model had better prediction accuracy for developed stock markets than the SR model. This conclusion is extended to the commodities market, specifically the developed and emerging energy commodities. However, our conclusions are different from [33,34,35] who indicated that the RS model isn't always better than the SR model. The key reason for this is that our research differs significantly from Zhao et al.[36,37], who did not compare the two types of models under the same distribution**.** Their findings are valid for two specific RS-GARCH models and the three SR-GARCH model instead of the MS-GARCH-type models and their SR counterparts.

5. Conclusion

In this work, we investigate the risk predicting performances between the regime-switching (RS) and single regime (SR) for the global energy commodities. For obtaining robust results, plenty of comprehensive comparisons are implemented, and a related process of comparisons is operated under the same condition. Especially in every energy market, the long and short positions are considered together to see their differences from the downside and upside results. Therefore, empirical results of the in-sample analysis report that the MS model outperforms the SR counterparts in global energy commodity cases. This conclusion that is gained through the risk predictions is suitable for one-day and five-day-ahead cases of energy commodities (Crude oil brent, Petroleum, Gasoline, Heating oil, Natural gas, and Crude oil WTI), some evidence seems interesting in that the upside results are affected by more horizons, but the

findings based on the downside risk are stable.

Investors and policymakers who aim to the specific purpose of economic should be vigilant to use the RS model for risk management for long and short positions. Meanwhile, automotive manufactures and energy-intensive global energy commodity prices are directly and indirectly susceptive. They are also carefully making appropriate productions plans when faced with extreme price changes in the future.

Discloser statement

The authors declare no potential conflict of interest.

References

[1] Ahmed, M. Y. and Sarkodie, S. A. (2021). Counterfactual shock in energy commodities affects stock market dynamics: Evidence from the United States. Resources Policy, 72,102-083.

[2] Aloui, C. and Mabrouk, S. (2010). Value-at-risk estimations of energy commodities via long- memory asymmetry and fat-tailed GARCH models. *Energy Policy*, 38(5), 2326–2339.

[3] ELA, 2020. The U.S Energy Facts Explained U.S Energy Information Administration.

[4] Ardia, D. and Hoogerheide, L. F. (2014). GARCH models for daily stock returns: Impact of estimation frequency on Value-at-Risk and Expected Shortfall forecasts. *Economics Letters*, 123(2), 187–190.

[5] Bauwens, L. and Laurent, S. (2005). A new class of multivariate skew densities with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics*, 23(3), 346–354.

[6] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.

[7] Caporale, G. M. and Zekokh, T. (2019). Modelling volatility of cryptocurrencies using Markov-Switching GARCH models. *Research in International Business and Finance*, 48, 143–155.

[8] di Sanzo, S. (2018). A Markov switching long memory model of crude oil price return volatility. *Energy Economics*, 74, 351–359.

[8] Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74(366), 427–431.

[9] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987-1007.

[10] Ardia, D. Bluteau, K. Boudt, K. & Catania, L. (2018). Forecasting risk with Markov switching GARCH models: A large-scale performance study. *International Journal of Forecasting*, 34(4),733–747.

[11] Fernández, C. and Steel, M. F. J. (1998). On bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441), 359–371.

[12] Francq, C. and Zakoian, J. M. (2019). GARCH models: structure, statistical inference and financial applications. *John Wiley & Sons*.

[13] Giannopoulos, K. and Tunaru, R. (2005). Coherent risk measures under filtered historical simulation. *Journal of Banking & Finance*, 29(4), 979–996.

[14] Gray, S. F. (1996). Modelling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1), 27–62.

[15] Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1–2), 307–333.

[16] Herrera, A. M. Hu, L. and Pastor, D. (2018). Forecasting crude oil price volatility. *International Journal of Forecasting*, 34(4), 622–635.

[17] Jalal, A. and Rockinger, M. (2008). Predicting tail-related risk measures: The consequences

of using GARCH filters for non-GARCH data. *Journal of Empirical Finance*, 15(5), 868–877.

[18] Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity, and serial independence of regression residuals. *Economics Letters*, 6(3), 255–259.

[19] Klaassen, F. (2002). EMPIRICAL ECONOMICS Improving GARCH volatility forecasts with regime switching GARCH. *Empirical Economics*, 27.

[20] Manigandan, P. Alam, M. S. Alagirisamy, K. Pachiyappan, D. Murshed, M. and Mahmood, H. (2023). Realizing the Sustainable Development Goals through technological innovation: Juxtaposing the economic and environmental effects of financial development and energy use. *Environmental Science and Pollution Research*, 30(3), 8239-8256.

[21] Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *In Biometrika*, 68(2).

[22] Manigandan, P. Alam, S. Alharthi, M. Khan, U. Alagirisamy, K. Pachiyappan, D. and Rehman, A. (2021). Forecasting Natural Gas Production and Consumption in United States-Evidence from SARIMA and SARIMAX Models. *Energies*, 14(19), 6021.

[23] Marimoutou, V. Raggad, B. and Trabelsi, A. (2009). Extreme Value Theory and Value at Risk: Application to oil market. *Energy Economics*, 31(4), 519–530.

[24] Muhammad, N. Aviral, T. Sana, M. and Muhammad, S. (2019). Modelling volatility of precious metals markets by using regime switching GARCH models. 64, 1–8.

[25] Nieto, M. R. and Ruiz, E. (2016). Frontiers in VaR forecasting and backtesting. *International Journal of Forecasting*, 32(2), 475–501.

[26] Sobreira, N. and Louro, R. (2020). Evaluation of volatility models for forecasting Value-at-Risk and Expected Shortfall in the Portuguese stock market. *Finance Research Letters*, 32, 101098.

[27] Tan, C. Y. Koh, Y. B. Ng, K. H. and Ng, K. H. (2021). Dynamic volatility modelling of Bitcoin using time-varying transition probability Markov-switching GARCH model. *North American Journal of Economics and Finance*, 56.

[28] Trottier, D.-A. and Ardia, D. (2015). Moments of Standardized Fernandez-Steel Skewed Distributions: Applications to the Estimation of GARCH-Type Models. *SSRN Electronic Journal*, 1–13.

[29] Saltik, O. Degirmen, S. and Ural, M. (2016). Volatility Modelling in Crude Oil and Natural Gas Prices. *Procedia Economics and Finance*, 38, 476–491.

[30] Nelson DB (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*,59(2), 347–370.

[31] Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, 8(2), 153–162.

[32] Teterin, P. Brooks, R. and Enders, W. (2016). Smooth volatility shifts and spillovers in U.S. crude oil and corn futures markets. *Journal of Empirical Finance*, 38, 22–36.

[33] World Bank, (2020). A Shock like no other; the impact of COVID-19 on Commodity Market. http://pubdocs.worldbank.org/en/558261587395154178/CMO-April-202 0-Special-Focus-1.pdf.

[34] Duraisamy, P. Alam, MS. Uddin, MA. Kuppusamy, A. Palanisamy, M. (2022). Modelling volatility of petroleum productions by using markov-switching GARCH models. *International Journal of Agricultural Statistical Science*, 18(1):1657-6.

[35] Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5), 931–955.

[36] Zhao, L. T. Zeng, G. R. Wang, W. J. and Zhang, Z. G. (2019). Forecasting oil price using web-based sentiment analysis. *Energies*, 12(22).

[37] Alam, Md. Shabbir. Muntasir, Murshed, Palanisamy, Manigandan. Duraisamy, Pachiyappan. and Shamansurova, Zilola Abduvaxitovna. (2023). Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: Contextual evidence from India using time series forecasting tools. *Resources Policy*, 81: 103342.

PRICE RISK ANALYSIS USING GARCH FAMILY MODELS: EVIDENCE FROM INDIAN NATIONAL STOCK EXCHANGE FUTURE MARKET

M. Valavan[1], Mohammed Ahmar Uddin[2], and S.Rita*[1]

•

[1] Department of Statistics,
Periyar University, Salem – 636011, Tamil Nadu, India.
valavanmurugesan@gmail.com
*Corresponding Mail: ritasamikannu@gmail.com
[2] Department of Finance and Economics,
College of Commerce and Business Admin, Dhofar University, Salalah, Oman
ahmar@du.edu.om

Abstract

The prediction of time-varying volatility plays an important role in financial data. In the paper, a comprehensive analysis of the mean return and conditional variance of NSE index is performed to use GARCH, EGARCH and TGARCH models with Normal innovation and Student's t innovation. Conducting a bootstrap simulation study which shows the Model Confidence Set (MCS) captures the superior models across a range of significance levels. The experimental results show that, under various loss functions, the GARCH using Student's t innovation model is the best model for volatility predictions of NSE among the six models.

Keywords: time-varying volatility, NSE index, bootstrap simulation, GARCH-type models.

1. Introduction

Forecasting market risk is a widely studied subject that has captured the interest of scholars due to its highly non-linearity and volatility. Thus, several approaches use these data for model testing. Generalized auto regressive conditional heteroscedasticity (GARCH) models are one of the most used models to study volatility. Although response to these models is generally good, they are unable to successfully capture extreme changes in the complete time series. Due to this shortcoming, one of the focuses of research has been to work on alternatives that can better approximate the non-linear part of the series mainly using GARCH, such as Markov Switching GARCH. This algorithm has been used extensively in stock markets and for market risk, because it is able to theoretically approximate any non-linear function with minimal error. In practice, the use of MSGARCH allows us to improve forecasting systems as well as forecasts from econometric models with excellent results.

Multi-period volatility forecasts feature prominently in asset pricing, folio allocation, risk-management, and most other areas of finance where horizon measures of risk are necessary. Such forecasts can be constructed quite different ways. The first approach is to estimate a horizon-

specific of the volatility, such as a weekly or monthly GARCH, that can then form direct predictions of volatility over the next week, month, etc. approach is to estimate a daily model and then iterate forward the daily to obtain weekly or monthly predictions. The forecasting literature refers first approach as "direct" and the second as "iterated". A third method mixed-data sampling (MIDAS) approach introduced by [1]. A MIDAS model uses, for example, daily returns to produce directly multi-period volatility forecasts and can as a middle ground between the direct and the iterated approaches. Volatility literature (see [2]) has mostly focused regressions-based models. It is the purpose of this paper to introduce ideas similar to MIDAS models in GARCH-type models. The advantages of this are that one focuses directly on multi-period forecasts, as in the direct while one preserves the use of high-frequency.

We propose a unifying framework, based on a generic GARCH-that addresses the issue of volatility forecasting involving forecast a different frequency than the information set. Hence, we propose GARCH models that can handle volatility forecasts over the next days and use past daily data, or tomorrow's expected volatility using intra-daily returns. We call the class of models High Frequency Based Projection-Driven GARCH models as the GARCH dynamics by what we call HYBRID processes. HYBRID-GARCH models - nature - relate to many topics discussed in the extensive literature forecasting. These topics include - but are not limited to - iterated forecasting, temporal aggregation, weak versus semi-strong GARCH, and various estimation procedures.

Exchange rates are a relevant topic of study because they serve as indicators of economic competition between nations and also because commercial relationships between countries are regulated by the value of competing currencies. In the past, the value of an exchange rate was set by the economic authorities of each nation based on monetary policy. However, since 1971, the world economy has changed and currently many countries follow a regime where parities are determined based on the supply and demand of the foreign exchange market, making the exchange rate market more volatile and less predictable. Since then, forecasting the variation in exchange rates has been a matter of interest for the decision-making bodies of government entities, banks, insurers, investors or people who trade with parities. Studying these changes poses several challenges such as determining which variables are relevant for a given currency or which method is superior to another for forecasting. In this sense, the use of time series models to model economic variables has been broad such as Autoregressive Moving Average (ARMA) and its derivatives, which include Vector Autoregressive model (VAR), the Vector Error Correction model (VECM), Cointegration model and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model. The remaining paper is organized as follows: In section 2, we present the econometric models. In section 3, we summarize the descriptive statistics of NSE sector. In section 4, we present the estimation procedures. In section 5, we describe the MCS test based on the bootstrap simulation. In section 6, concludes.

ARCH (Autoregressive Conditional Heteroskedasticity) and Generalized ARCH (GARCH) models have emerged as the most proeminent tools for estimating volatility, because they are adequate to capture the random movement of the financial data series. Many researchers have studied over time the performance of GARCH models on explaining volatility of mature stock markets, but only a few have tested GARCH models using daily data from Central and Eastern European stock markets (see, for example, [3],[4],[5],[6],[7],[8]). The focus of our paper is on forecasting stock market volatility in Romania, a market which has not been thoroughly investigated.

Several studies results have confirmed that asymmetric GARCH-models fit better stock markets returns volatility for emerging CEE countries. Lupu [9] found that an EGARCH (Exponential GARCH) model is suitable for the logarithmic returns of the Romanian composite index BET-C covering the period 03/01/2002- 17/11/2005. Furthermore [10] employed different

asymmetric GARCH-family models (EGARCH, PGARCH, and TGARCH) using U.S. and Romanian daily stock return data corresponding to the period 2002-2010. They found that volatility estimates given by the EGARCH model exhibit generally lower forecast error and are therefore more accurate than the estimates given by PGARCH and TGARCH models. [11] examines the presence of volatility at the Karachi Stock Exchange(KSE) through the use of Autoregressive Conditional Heteroskedasticity(ARCH)and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models introduced by Engle (1982), Bollerslev (1986) and Nelson (1991). The empirical result confirms the presence of high volatility at Karachi Stock Exchange throughout the study period. The volatility was found in clustering and stochastic manner. The results of GARCH analysis show a random-walk behavior so market can be termed as very uncertain and very risky for short-term and medium-term investors.

2. Material and Methods

The volatility of a stock price can be used as an indicator of the uncertainty of stock returns. In a financial market, volatility is measured in terms of standard deviation σ or σ2 compute variance from a set of observations as follow:

$$\sigma^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t-\bar{y})^2 \quad (1)$$

here $\bar{y}$ and y_t are the mean return and return, respectively. Return is defined to be the total gain or loss from an investment over a given period of time. In this paper, we compute the daily closing prices are as

$$y_t = 10\log(p_t / p_{t-1}) \quad (2)$$

where p_t is stock closed price at time t. Then prices are converted into logarithmic returns, y_t denotes p_t the continuously compounded daily returns of the underlying assets at time t. We assume that the conditional mean equation of stock return is constructed as the constant term plus residuals error

$$y_t = \mu + \varepsilon_t, \varepsilon_t = \sigma_t z_t \quad (3)$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with zero mean and unit variance, is the conditional variance of derived from mean equation, it is also known as current day's variance or volatility. Larger implies higher volatility and higher risk.

2.1. Parametric models GARCH (1, 1) is written as following is

The standard variance model for financial data is GARCH. GARCH assumes a Gaussian observation model and a linear transition function for the variance: the time-varying variance σ_t^2 is linearly dependent on p previous variance values and q previous squared time series values, so

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q}\alpha_j + \varepsilon_{t-j}^2 + \sum_{i=1}^{p}\beta_i\sigma_{t-i}^2 \quad (4)$$

Hence, GARCH (1,1) is defined as $\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q}\alpha_j + \varepsilon_{t-j}^2 + \sum_{i=1}^{p}\beta_i\sigma_{t-i}^2$ Where $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta < 1$. First, $\omega > 0$ means that volatility cannot have a zero or negative mean. Second, the positive parameters α, β show that the conditional variance forecasts will increase if there is a large

fluctuation in returns, the model thus capturing the stylized feature of volatility clustering. Finally, $\alpha + \beta < 1$ indicates the persistence of shocks to volatility will eventually fade away, which depicts another stylized characteristic of volatility, mean reversion.

2.2. Exponential-GARCH (EGARCH) (1,1) is defined as following is

A more flexible and often cited GARCH extension is Exponential GARCH (EGARCH) (Hamilton, 1994). The default EGARCH (p, q) model in Econometrics is of the form: $\varepsilon_t = \sigma_t z_t$ with Normal innovation or Student's t innovation distributions and

$$\log(\sigma_t^2) = \omega + \beta\log(\sigma_{t-1}^2) + \alpha[\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - E(\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}})] + r(\frac{\varepsilon_{t-1}}{\sigma_{t-1}}) \quad (5)$$

where the parameter α captures the volatility clustering effect and the r measures the leverage effect. The conditional variance is in logarithmic form, which implies that the model has the following features: first, σ_t^2 will always be positive regardless of the sign of the parameters, therefore no constraints of non-negativity are needed. Second, the asymmetrical effect is not quadratic but exponential, if $r < 0$ it indicates a leverage effect. EGARCH model allows good news and bad news to have a different impact on volatility because the level of is included $\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ with a coefficient r.

2.3. MCS test based on bootstrap simulation:

When obtaining the predicted values, we can compare it with the real Proxy variables of a volatility deviation size. However, the loss of function which is used to measure the prediction error is no consensus. The paper uses two loss functions: mean square error and mean absolute deviation to measure the forecasting error. It is not easy to choose the best model which is always the best under all loss function or all data samples. Since, [12] offers some resolution of this quandary, the metric for assessing the forecasts of volatility models is the Bootstrap method of superior predictive ability (SPA) test. But use SPA test, we must need to choose the basic model, it is very vital to choose it which can affect the result. In order to overcome the defects of SPA test, the paper use the MCS test which is a modified version of SPA test.

2.4. MCS test procedure

We define a set of models which are denoted by $M_0 = \{1,\ldots,m\}$, the models are indexed by $i = 1, \ldots, m$, and model is forecasts of σ_t^2 is denoted by $h_{i,t}^2$, We rank models according to their expected loss using one of two loss functions: MSE, $L(h_{i,t}^2, \sigma_t^2) = (h_{i,t}^2 - \sigma_t^2)^2$ and $L(h_{i,t}^2, \sigma_t^2) = |h_{i,t}^2 - \sigma_t^2|$. The loss differential between models *i* and *j*, is given by $d_{ij,t} = L(h_{i,t}^2, \sigma_t^2) - L(h_{j,t}^2 - \sigma_t^2)\, i, j = 1,\ldots,m, t = 1,\ldots,n$ The MCS is determined after sequentially trimming the set of candidate models, M_0. At each step, the hypothesis

$$H_0 : E(d_{ij,t}) = 0, \text{for all } i, j \hat{I} M \grave{I} M_0 \quad (6)$$

The hypothesis, H_0, is a test for (Equal Predictive Ability) EPA over the models in *M* and if H_0 is rejected, the worst performing model is eliminated from *M*. The trimming ends when the first non-rejection occurs. The set of surviving models is the model confidence set $\widehat{M}_\alpha^*$, By holding the significance level,α, fixed at each step of the MCS procedure, we construct a $(1 - \alpha)$-confidence set, $\widehat{M}_\alpha^*$,for the best models in M_0. However, the trimming model which is mentioned in the sequential inspection have a drawback. At each step in the test, we need to test the predictive

power of any two prediction models and calculate a test statistic. To overcome this drawback, our tests for EPA employ the rang statistic, T_R, and the semi-quadratic statistic, T_{sq} given by

$$T_R = \max_{i\,j \in M} \frac{\left|\sum_{i<j} (\bar{d}_{ij})\right|}{\sqrt{\hat{v}(\bar{d}_{ij})}} \qquad T_{sq} = \frac{\sum_{i<j} (\bar{d}_{ij})^2}{\mathrm{var}(\bar{d}_{ij})} \tag{7}$$

Where the sum is taken over the models in M, and $\hat{v}(\bar{d}_{ij})$ is an estimate of $v(\bar{d}_{ij})$.Both of the test statistic value is larger, it means rejecting the EPA hypothesis. In fact, their distribution is very complicated, and the covariance structure depends on the predictive value of each prediction model. So, the paper uses a bootstrap simulation study to find the p-value of the two statistics.

3. Data and Experimental Results

The whole sample consists of 2537 daily data spanning from 4 Jan. 2010 to 16 Mar. 2023, we select subsample of size 2000, dated from 4 Jan. 2010 to 24 Feb. 2023, as the training set for the parameter's estimation for models and the remaining sample of size 537 daily data, from 25 Feb. 2022 to 16 Mar. 2023 is used as the test set or for out of sample forecasting.

Table 1: *Summary statistics of NSE*

		Descriptive statistics	
Sample	2575	Mean	9.844
Std. dev	8.693	Skewness	0.237
Kurtosis	-1.44	JB	243.87

Then we need to calculate logarithmic returns $y_t = 10\log(p_t / p_{t-1})$. Table 1 summarizes the descriptive statistics of NSE index throughout the whole period. Table 1 remarks that these facts suggest a highly competitive and volatile mark. The Skewness is 0.2371853 > 0, the positive skewness indicates that there is a high probability of gain in the market. The value of the Kurtosis is -1.444026 > 3, it suggests that the market is volatile with high probability of extreme events occurrences. The JB.test is 243.87 which shows that the returns deviate from normal distribution significantly and exhibit leptokurtic. Hence the distribution of the index is not the normal distribution, and it has the feature of asymmetric, zero mean and left side.

Table 1: *Unit Root Test of NSE*

Test	Critical value	P.value	alternative hypothesis
ADF test	-13.512	0.01	Stationary
KPSS test	0.065	0.01	Stationary
PP test	-3225.2	0.01	Stationary

The table 2, reports the unit root tests of the NSE. The Augmented Dickey and Fuller-ADF test for the null of non – stationary. Critical value -13.512.KPSS indicates the Kwiatkowski, Phillips, Schmidt and shin test for the null of stationary. Critical value: 0.065915. PP. test indicates the Phillips-Perron test for the null of non- hypothesis. Critical value -3225.2 it means that the series yt is stationary time series.

3.1. Detecting ARCH effects of NSE returns

From the Fig. 1, we can see that the returns appear to fluctuate around a constant level but

exhibit volatility clustering. Large changes in the returns tend to cluster together, and small changes tend to cluster together. So, the preliminary judgment shows that the series exhibits the conditional heteroscedasticity. Now the paper use ARCH-LM to detect whether NSE returns have ARCH effects.

According to the heteroskedasticity test ARCH, the value of F-statistic is 0.004301 and the probability 0.0002< 0.05, R-squared = 9.444 and Adjusted R-squared = 0.9443, the probability is 0.0002 < 0.05, and the number of lags is 1, the test of the residuals for ARCH(1) rejects the null hypothesis of no conditional heteroskedasticity, so it is clear that NSE returns have ARCH effects. Then we can use GARCH-type models to forecast the volatility.

4. Estimation result

Apply the return series to the GARCH and EGARCH models with Normal innovation and Student's innovation, and then we get their parameters. The estimation results and diagnosis are shown in Table 2 and Table 3.

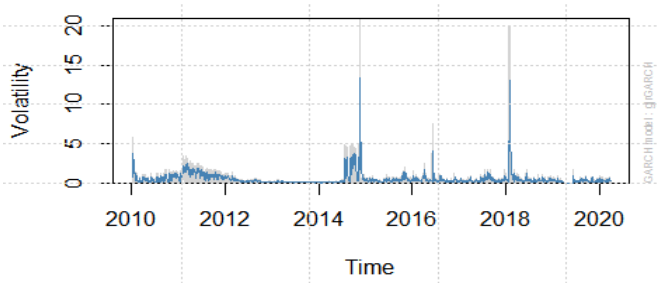


Figure 1. *Daily return and volatility of NSE*

Table 2. *Estimation results by GARCH models*

Statistics	GARCH-T				GARCH-N			
	Parameter	St. Error	t-value	Pr(>|t|)	parameter	St. Error	t-value	Pr(>|t)
mu	-0.002	0.002	-1.104	0.020*	0.023	0.0313	-1.104	0.047*
Omega	0.0008	0.000	4.636	0.000***	0.004	0.025	4.636	0.025*
alpha1	0.114	0.025	4.431	0.000***	0.214	0.356	4.431	0.049*
beta1	0.871	0.24	3.596	0.000***	0.375	0.242	3.039	0.000***
beta2	0.000	0.215	0.6367	0.034*	0.003	0.215	0.580	0.040*
Information Criteria:	Log Likelihood : 476.0085				Log Likelihood : 485.2743			
Akaile	-0.840				-0.740			
Bayes	-0.790				-0.696			

Note. *,**, and *** denotes level of significant at 10%, 5% and 1%, respectively.

Table 3. *Estimation results by EGARCH models*

Statistics	EGARCH-T				EGARCH-N			
	Parameter	St. Error	t-value	Pr(>|t|)	Parameter	St. Error	t-value	Pr(>|t|)
mu	0.036	0.038	0.966	0.333	0.836	0.166	5.027	0.000
ar1	0.982	0.003	2.482	0.000	0.999	0.000	1.512	0.000
ma1	-0.746	0.038	-2 .379	0.000	-0.820	0.019	-4.319	0.000
Omega	-0.027	0.004	-5.884	0.000	-0.012	0.001	-9.364	0.000
alpha1	0.002	0.023	0.101	0.919	0.106	0.021	4.980	0.000
alpha2	0.060	0.031	1.942	0.052	-0.030	0.015	-1.951	0.050
beta1	0.998	0.001	7.618	0.000	0.997	0.000	1.4290	0.000
gamma1	0.493	0.069	7.136	0.000	0.342	0.050	6.799	0.000
gamma2	-0.231	0.074	-3.096	0.001	-0.272	0.049	-5.539	0.000

Information Criteria:	Log Likelihood : 476.0085	Log Likelihood : 485.2743
Akaike	-0.840	-0.740
Bayes	-0.790	-0.696

Note. *,**, and *** denotes level of significant at 10%, 5% and 1%, respectively.

Among the parametric models, with Normal innovation and Student's t innovation, In GARCH-N, the value of LL is 419.724, AIC is -1.989 and BIC -1.063, each parameter is significant. In EGARCH-N, the value of LL is 485.2743, AIC is −0.740 and BIC is −0.863, each parameter is significant. In GARCH-T, the value of LL is 403.819, AIC is -0.929 and BIC is -0.863, each parameter is significant. In EGARCH-T, the value of μ, α1 is not significant. In EGARCH-N, the value of α1 also are not significantly. Hence according to highest value of Log Likelihood (LL) and smallest value of AIC and BIC. Hence the series best fit is EGARCH-N.

4.1. The MCS test results

The Table 4 shows the MCS test results by using bootstrap simulation at 1000 times. Figures in the table represent MCS test p-value. When greater, *p*-value indicates that they more reject the null hypothesis. The paper sets a basis p-value which is p = 0.1. If p-value is less than 0.1, then the volatility forecasting model is poor. So, the model will be removed in the MCS inspection process. Conversely, it survives in MCS.

Table 4. *MCS test results of realized volatility models*

	MSE		MAD	
Model	T_R	T_{sq}	T_R	T_{sq}
GARCH-N	0.161	0.185	0.136	0.042
GARCH-T	0.211	0.231	0.191	0.152
EGARCH-N	0.035	0.052	0.064	0.042
EGARCH-T	0.021	0.035	0.033	0.053

So, the model will be removed in the MCS inspection process. Conversely, it survives in MCS. According to the table, when the loss function is the MSE, the p-values of in the GARCH-N and GARCH-T models are more than 0.1. But other 2 models are less than 0.01. It means that EGARCH-N, EGARCH-T, volatility forecasting models will be removed in the MCS inspection process. Considering the loss function for MAD, we find that only the p-value of GARCH-T model is more than 0.1. Hence using the loss function of MSE and MAD, we find that the value of T_R T_{sq} in the GARCH-T model are more than 0.1. Therefore, GARCH-T model is the best one.

V. Conclusion

The study uses NSE prices to predict daily volatility changes in the stock market. First, we use descriptive statistics to show that the index series has the feature of asymmetric zero mean and left side, it is not the normal distributed. Second, we consider Dickey-Fuller Unit Root Tests to find the series is stationary time series. And then using ARCH-Lagrange multiplier to detect NSE returns have ARCH effects. In this study, NSE index volatility models are estimated with Normal innovation and Student's t innovation distributions to find the effect of distribution selection on forecasting performance of the models. According to highest value of Log likelihood (LL) and smallest value of AIC and BIC, the result suggests that the GARCH model with Student's t innovation enables more accurate forecasting than EGARCH.

The paper use MCS test to find the best model. Under the evaluation criteria of loss functions

MSE and MAD, the empirical results show that GARCH-T model is the best model for forecasting volatility. Although the prediction's results represent that GATCH-T model is not so good, it can be used as an assistant tool in financial applications. The study has also multiple significantly: first, the stock index futures in favor of investors to make rational investment decisions in advance. Second, it helps to improve risk management of institutional and individual investors. Finally, there is conducive to the development of relevant policies and regulatory authorities to improve supervision.

References

[1] Kristjanpoller, W. and Michell, K. (2018). A stock market risk forecasting model through integration of switching regime, ANFIS and GARCH techniques. Applied Soft Computing, 67:106-116.

[2] Proietti, T. and Luati, A. Maximum likelihood estimation of time series models: the Kalman filter and beyond: In Handbook of Research Methods and Applications in Empirical Macroeconomics, Edward Elgar Publishing, (2013).

[3] Deebom, Z. D. and Essi, I. D. (2017). Modeling Price Volatility of Nigerian Crude Oil Markets Using GARCH Model, International Journal of Applied Science and Mathematical Theory, *3*(4):23-49.

[4] Kash-Haroutounian, M. and Price, S. (2001). Volatility in the transition markets of Central Europe, Applied Financial Economics, 11(1):93 105.

[5] Lupu, R. and Lupu, I. (2007). Testing for heteroskedasticity on the Bucharest Stock Exchange, Romanian Economic Journal, 11(23):19-28.

[6] Miron, D. and Tudor, C. (2010). Asymmetric Conditional Volatility Models: Empirical Estimation and Comparison of Forecasting Accuracy, Romanian Journal of Economic Forecasting, 13(3):74-92.

[7] Murinde, V. and Poshakwale, S. (2002). Volatility in the emerging stock markets in Central and Eastern Europe: evidence on Croatia, Czech Republic, Hungary, Poland, Russia and Slovakia, European Research Studies, 4(3): 73-101.

[8] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach, Econometrica, 59:347-370.

[9] Duraisamy, P. Alam, MS. Uddin, MA. Kuppusamy, A. Palanisamy, M. (2022). Modelling volatility of petroleum productions by using markov-switching GARCH models. International Journal of Agricultural Statistical Science, 18(1):1657-67.

[10] Poshakwale, S. and Murinde, V. (2001). Modelling volatility in East European emerging stock markets: evidence on Hungary and Poland. Applied Financial Economics, 11(4):445-456.

[11] Xiao, L. and Aydemir, A. Volatility modelling and forecasting in finance, Forecasting volatility in the financial markets, Oxford: Elsevier Butterworth-Heinemann, (2007).

[12] Kanasro, H. A. Rohra, C. L. and Junejo, M. A. (2009). Measurement of stock market volatility through ARCH and GARCH models: a case study of Karachi stock exchange, Australian Journal of Basic and Applied Sciences, *3*(4):3123-3127.

RT&A, No 4 (76)
Volume 18, December 2023

IDZ DISTRIBUTION: PROPERTIES AND APPLICATION

Idzhar A. Lakibul

•

Department of Mathematics and Statistics
Mindanao State University - Iligan Institute of Technology
Iligan City, Philippines
idzhar.lakibul@g.msuiit.edu.ph

Abstract

This paper introduces a novel two - parameter continuous distribution. This distribution is derived from the mixture of the Exponential, Weibull and Ailamujia distributions. The derived distribution is named as "Idz distribution". The probability density function of the Idz distribution is derived and some of its plots are presented. It can be observed that the Idz distribution can generate right tailed unimodal, non-monotonic decreasing and exponential shapes. Further, survival and hazard functions of the Idz distribution are derived. It reveals that the hazard function of the Idz distribution can accommodate three types of failure rate behaviors, namely, non-monotonic constant, right tailed unimodal and non-monotonic decreasing. Moreover, some properties of Idz distribution such as moments, mean, variance, moment generating function, order statistics and maximum likelihood estimates are derived. In addition, the proposed distribution is applied into a Breast Cancer data and compare with the Exponentiated Generalized Inverse Rayleigh distribution, the Ailamujia Inverted Weibull distribution and the New Extended Exponentiated Weibull distribution. Result shows that the Idz distribution gives better estimates as compared with the said distributions for a given dataset.

Keywords: Weibull distribution, Exponential distribution, Ailamujia distribution.

1. Introduction

Non-negative continuous probability distribution is important in modelling real lifetime data, specifically, in the field of reliability, engineering and biomedical science. There are popular classical limetime distributions such as the exponential, log-normal, log-logistic, Weibull, Rayleigh and the Frechet distributions. But due to the complexity of the lifetime data, the classical distributions need to generalize or extend in order to cater the complex behaviour of the data. One method for extending the classical distribution is by using the generated family of distributions like the Exponentiated - G family of distributions [7], Marshall-Orkin - G family of distributions [11], Beta - G family of distributions [4] and other existing families of distributions.

Another method of facing the complex behaviour of the lifetime data is by using mixture distribution of two or more probability distribution functions. A random variable X is assumed to have a mixture of two or more probability distribution functions $f_1(x)$, $f_2(x)$, $f_3(x)$,...,$f_n(x)$ if its probability density function $m(x) = \sum_{i=1}^{n} a_i f_i(x)$ with $a_i \in [0,1]$ and $\sum_{i=1}^{n} a_i = 1$. Years ago, several distributions have been derived from the mixing distributions, for example, the Aradhana distribution [13] which is a mixtures of the Gamma (2, θ), the Gamma (3, θ) and the Exponential (θ) distributions with corresponding mixing proportions $\frac{2\theta}{\theta^2+2\theta+2}$, $\frac{2}{\theta^2+2\theta+2}$ and $\frac{\theta^2}{\theta^2+2\theta+2}$. Other identified mixture distributions such as the Rama distribution [14], Darna distribution [15], Shanker distribution [12], Gharaibeh distribution [6], Alzoubi distribution [2] and Benrabia

distribution [3].

In this paper, the concept of mixture distribution is used to propose a two - parameter distribution named as Idz distribution which is a mixture of three distributions, namely, the Weibull (λ, β) distribution [17], exponential (λ) distribution [8] and the Ailamujia (λ) distribution [10] with mixing proportions $\frac{\lambda\beta^2}{\lambda\beta^2+\lambda\beta+1}$, $\frac{\lambda\beta}{\lambda\beta^2+\lambda\beta+1}$ and $\frac{1}{\lambda\beta^2+\lambda\beta+1}$, respectively. Other goals of the paper are the following: (i) to derive some properties of Idz distribution such as its moments, moment generating function, mean, variance, order statistics and maximum likelihood estimates of the proposed distribution parameters; and (ii) to apply the proposed distribution into a real dataset and compare with the Exponentiated Generalized Inverse Rayleigh, the Ailamujia Inverted Weibull and the New Extended Exponentiated Weibull distributions.

This paper is arranged as follows: Idz distribution is introduced in section 2. In section 3, some properties of Idz distribution are derived. Order Statistics of the ID distribution is given in section 4 while the maximum likelihood estimates of the ID parameters is presented in section 5. In section 6, the application of Idz distribution is illustrated . Some concluding remarks is presented in section 7.

2. Idz Distribution

This section presents the definition of the Idz distribution and its special cases with the illustration of its pdf.

A random variable X is said to have an Idz distribution (ID) with parameters λ and β if the probability density function of X is given by

$$f(x,\lambda,\beta) = \frac{\lambda^2 e^{-\lambda x}[\beta + 4xe^{-\lambda x} + \beta^3 x^{\beta-1} e^{\lambda(x-x^\beta)}]}{\lambda\beta^2 + \lambda\beta + 1}, \tag{1}$$

where $x \geq 0$, $\lambda > 0$ and $\beta > 0$. The corresponding cumulative distribution function of X is given by

$$F(x,\lambda,\beta) = 1 - \frac{\lambda\beta e^{-\lambda x} + (2\lambda x + 1)\, e^{-2\lambda x} + \lambda\beta^2 e^{-\lambda x^\beta}}{\lambda\beta^2 + \lambda\beta + 1}. \tag{2}$$

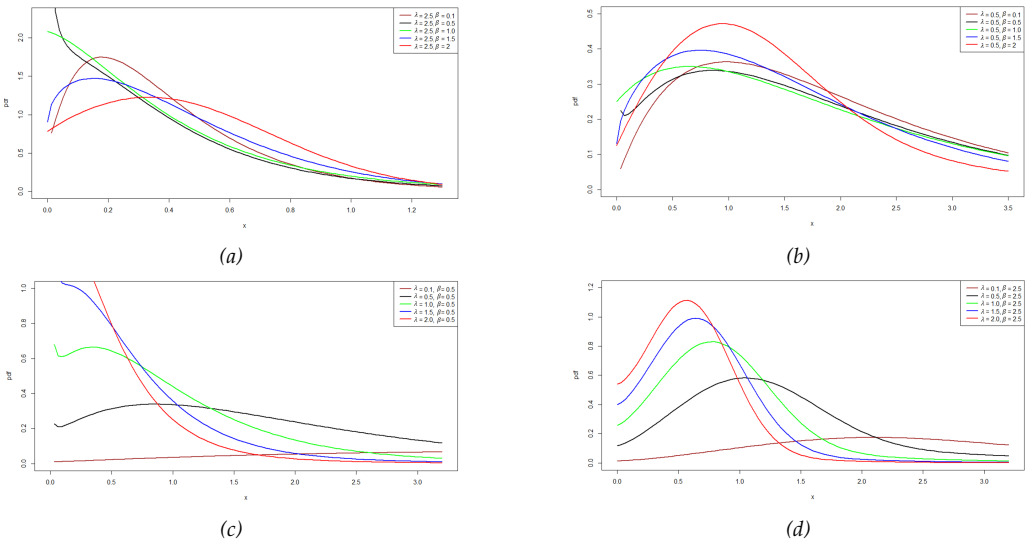


Figure 1: *pdf plots of Idz distribution (ID) for different sets of values of the parameters: (a) $\lambda = 2.5$ and varying values of β; (b) $\lambda = 0.5$ and varying values of β; (c) $\beta = 0.5$ and varying values of λ ; and (d) $\beta = 2.5$ and varying values of λ.*

Figure 1 shows some possible density shapes of the ID distribution and it reveals that the pdf of the ID distribution can generate right tailed unimodal, non-monotonic decreasing and

exponential shapes.

Special Cases of Idz distribution

1. If $\beta = 1$ then ID reduces to

$$f(x,\lambda) = \frac{2\lambda^2 e^{-\lambda x}(1+2xe^{-\lambda x})}{2\lambda+1}. \tag{3}$$

2. If $\lambda = 1$ then ID reduces to

$$f(x,\beta) = \frac{e^{-x}[\beta + 4xe^{-x} + \beta^3 x^{\beta-1}e^{x-x^\beta}]}{\beta^2+\beta+1}. \tag{4}$$

3. If $\beta = 2$ then ID reduces to

$$f(x,\lambda) = \frac{2\lambda^2 e^{-\lambda x}[1+2xe^{-\lambda x} + 4xe^{\lambda(x-x^2)}]}{6\lambda+1}. \tag{5}$$

We name the probability distribution functions (pdf) (3), (4) and (5) as the pdfs of the Edz distribution, Laks distribution and Alds distribution, respectively.

3. Statistical Properties

In this section, we derive some properties of Idz distribution such as its moments, mean, variance, moment generating function, survival fuction and hazard function.

3.1. Moments

Theorem 1. Let X be a random variable that follows an Idz distribution then the rth moment of X denoted by μ'_r is given by

$$\mu'_r = \frac{\lambda^{1-r}}{\lambda\beta^2+\lambda\beta+1}\left[\beta\Gamma(r+1) + \frac{\Gamma(r+2)}{\lambda 2^r} + \frac{\beta^2\Gamma\left(\frac{r}{\beta}+1\right)}{\lambda^{\frac{r}{\beta}-r}}\right], \tag{6}$$

where $r = 1,2,3,...,n$ and $\Gamma(\cdot)$ is a gamma function.

Proof. The rth moment of X is defined by

$$\begin{aligned}
\mu'_r &= \mathbb{E}[X^r] \\
&= \int_{-\infty}^{\infty} x^r f(x)dx \\
&= \int_0^{\infty} x^r \frac{\lambda^2 e^{-\lambda x}[\beta + 4xe^{-\lambda x} + \beta^3 x^{\beta-1}e^{\lambda(x-x^\beta)}]}{\lambda\beta^2+\lambda\beta+1}dx \\
&= \frac{\lambda^2}{\lambda\beta^2+\lambda\beta+1}\left(\beta\int_0^{\infty} x^r e^{-\lambda x}dx + 4\int_0^{\infty} x^{r+1}e^{-2\lambda x}dx + \beta^3\int_0^{\infty} x^r x^{\beta-1}e^{-\lambda x^\beta}dx\right) \\
&= \frac{\lambda^2}{\lambda\beta^2+\lambda\beta+1}\left[\beta\left(\frac{1}{\lambda^{r+1}}\right)\Gamma(r+1) + 4\left(\frac{1}{2\lambda}\right)^{r+2}\Gamma(r+2) + \beta^3\left(\frac{1}{\beta\lambda^{\frac{r}{\beta}+1}}\right)\Gamma\left(\frac{r}{\beta}+1\right)\right] \\
&= \frac{\lambda^{1-r}}{\lambda\beta^2+\lambda\beta+1}\left[\beta\Gamma(r+1) + \frac{\Gamma(r+2)}{\lambda 2^r} + \frac{\beta^2\Gamma\left(\frac{r}{\beta}+1\right)}{\lambda^{\frac{r}{\beta}-r}}\right].
\end{aligned}$$

■

Corollary 1. Let X be a random variable with moment given in equation (6) then the mean μ and variance σ^2 of X are, respectively, given by

$$\mu = \frac{1}{c}\left[\beta\Gamma(2) + \frac{\Gamma(3)}{2\lambda} + \beta^2\lambda^{\frac{\beta-1}{\beta}}\Gamma\left(\frac{1}{\beta}+1\right)\right]$$

and

$$\sigma^2 = \frac{1}{c}\left\{\lambda^{-1}\left[\beta\Gamma(3) + \frac{\Gamma(4)}{4\lambda} + \frac{\beta^2\Gamma\left(\frac{2}{\beta}+1\right)}{\lambda^{\frac{2}{\beta}-2}}\right] - \frac{1}{c}\left[\beta\Gamma(2) + \frac{\Gamma(3)}{2\lambda} + \frac{\beta^2\Gamma\left(\frac{1}{\beta}+1\right)}{\lambda^{\frac{1-\beta}{\beta}}}\right]^2\right\},$$

where $c = \lambda\beta^2 + \lambda\beta + 1$.

Proof. The mean of X is derived when $r = 1$ in (6). Hence,

$$\mu = \frac{1}{c}\left[\beta\Gamma(2) + \frac{\Gamma(3)}{2\lambda} + \beta^2\lambda^{\frac{\beta-1}{\beta}}\Gamma\left(\frac{1}{\beta}+1\right)\right],$$

where $c = \lambda\beta^2 + \lambda\beta + 1$. Next, the variance of X denoted by σ^2 can be computed as

$$\sigma^2 = \mu_2' - (\mu_1')^2.$$

Now, the 2nd raw moment μ_2' of X is obtained by setting $r = 2$ in equation (6). It follows that

$$\mu_2' = \frac{\lambda^{-1}}{c}\left[\beta\Gamma(3) + \frac{\Gamma(4)}{4\lambda} + \frac{\beta^2\Gamma\left(\frac{2}{\beta}+1\right)}{\lambda^{\frac{2}{\beta}-2}}\right].$$

Therefore, the variance σ^2 of X is

$$\begin{aligned}\sigma^2 &= \frac{\lambda^{-1}}{c}\left[\beta\Gamma(3) + \frac{\Gamma(4)}{4\lambda} + \frac{\beta^2\Gamma\left(\frac{2}{\beta}+1\right)}{\lambda^{\frac{2}{\beta}-2}}\right] - \left\{\frac{1}{c}\left[\beta\Gamma(2) + \frac{\Gamma(3)}{2\lambda} + \beta^2\lambda^{\frac{\beta-1}{\beta}}\Gamma\left(\frac{1}{\beta}+1\right)\right]\right\}^2 \\ &= \frac{1}{c}\left\{\lambda^{-1}\left[\beta\Gamma(3) + \frac{\Gamma(4)}{4\lambda} + \frac{\beta^2\Gamma\left(\frac{2}{\beta}+1\right)}{\lambda^{\frac{2}{\beta}-2}}\right] - \frac{1}{c}\left[\beta\Gamma(2) + \frac{\Gamma(3)}{2\lambda} + \frac{\beta^2\Gamma\left(\frac{1}{\beta}+1\right)}{\lambda^{\frac{1-\beta}{\beta}}}\right]^2\right\}.\end{aligned}$$

■

3.2. Moment Generating Function

Theorem 2. Let X be a random variable that follows an Idz distribution then the moment generating function of X is given by

$$M_x(t) = \sum_{r=0}^{\infty}\frac{t^r\lambda^{1-r}}{(\lambda\beta^2+\lambda\beta+1)\,r!}\left[\beta\Gamma(r+1) + \frac{\Gamma(r+2)}{\lambda 2^r} + \frac{\beta^2\Gamma\left(\frac{r}{\beta}+1\right)}{\lambda^{\frac{r}{\beta}-r}}\right],$$

where $t \in \mathbb{R}$.

Proof. The moment generating function of X is defined by

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx.$$

Using equation (1), we have

$$M_X(t) = \int_0^{\infty} e^{tx} f(x,\lambda,\beta)dx.$$

Recall that $e^{tx} = \sum_{r=0}^{\infty} \frac{t^r}{r!} x^r$. Then,

$$M_X(t) = \int_0^{\infty} \sum_{r=0}^{\infty} \frac{t^r}{r!} x^r f(x, \lambda, \beta) dx = \sum_{r=0}^{\infty} \frac{t^r}{r!} \int_0^{\infty} x^r f(x, \lambda, \beta) dx = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu'_r.$$

Using equation (6) and hence,

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r \lambda^{1-r}}{(\lambda\beta^2 + \lambda\beta + 1)\, r!} \left[\beta\Gamma(r+1) + \frac{\Gamma(r+2)}{\lambda 2^r} + \frac{\beta^2 \Gamma\left(\frac{r}{\beta} + 1\right)}{\lambda^{\frac{r}{\beta} - r}} \right],$$

where $t \in \mathbb{R}$. ■

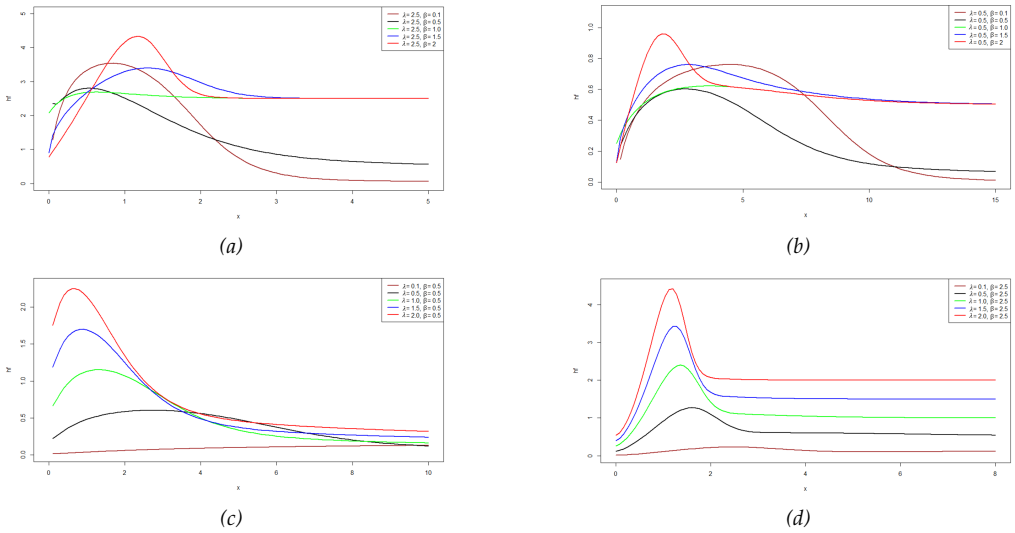
3.3. Reliability Analysis

Let X be a random variable with cdf (2) and pdf (1) then the survival $S(x, \lambda, \beta)$ and hazard $h(x, \lambda, \beta)$ functions of X are respectively, given by

$$S(x, \lambda, \beta) = \frac{\lambda\beta e^{-\lambda x} + (2\lambda x + 1)\, e^{-2\lambda x} + \lambda\beta^2 e^{-\lambda x^{\beta}}}{\lambda\beta^2 + \lambda\beta + 1}, x \geq 0, \lambda > 0, \beta > 0$$

and

$$h(x, \lambda, \beta) = \frac{\lambda^2[\beta + 4xe^{-\lambda x} + \beta^3 x^{\beta-1} e^{\lambda(x - x^{\beta})}]}{\lambda\beta + (2\lambda x + 1)\, e^{-\lambda x} + \lambda\beta^2 e^{-\lambda\left(x - x^{\beta}\right)}}.$$



(a) *(b)* *(c)* *(d)*

Figure 2: *hf plots of Idz distribution (ID) for different sets of values of the parameters: (a) $\lambda = 2.5$ and varying values of β; (b) $\lambda = 0.5$ and varying values of β; (c) $\beta = 0.5$ and varying values of λ ; and (d) $\beta = 2.5$ and varying values of λ.*

Figure 2 presents some possible shapes of the hazard function of the ID distribution and it reveals that the hazard function of the ID distribution can accommodate non-monotonic constant, right tailed unimodal and non-monotonic decreasing behaviors.

4. Order Statistics

Let $X_{(1)}$, $X_{(2)}$,..., $X_{(n)}$ be the order statistics of a random sample X_1, X_2,..., X_n drawn from the continuous population with probability density function (pdf) $f_X(x)$ and cumulative distribution

function $F_X(x)$, then the pdf of rth order statistics $X_{(r)}$ is given by

$$f_{X_{(r)}}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x)[F_X(x)]^{r-1}[1-F_X(x)]^{n-r}. \tag{7}$$

The pdf of rth order statistics $X_{(r)}$ of the ID distribution is derived by inserting (2) and (1) into (7) and is

$$\begin{aligned} f_{X_{(r)}}(x,\lambda,\beta) =& \frac{n!\lambda^2 e^{-\lambda x}[\beta + 4xe^{-\lambda x} + \beta^3 x^{\beta-1} e^{\lambda(x-x^\beta)}]}{(r-1)!(n-r)!\,(\lambda\beta^2+\lambda\beta+1)} \\ & \left[1 - \frac{\lambda\beta e^{-\lambda x} + (2\lambda x+1)\,e^{-2\lambda x} + \lambda\beta^2 e^{-\lambda x^\beta}}{\lambda\beta^2+\lambda\beta+1}\right]^{r-1} \\ & \left[\frac{\lambda\beta e^{-\lambda x} + (2\lambda x+1)\,e^{-2\lambda x} + \lambda\beta^2 e^{-\lambda x^\beta}}{\lambda\beta^2+\lambda\beta+1}\right]^{n-r}. \end{aligned} \tag{8}$$

The pdf of the smallest or 1st order statistics of the ID distribution is obtained by setting $r = 1$ in equation (8) and is

$$\begin{aligned} f_{X_{(1)}}(x,\lambda,\beta) =& \frac{n!\lambda^2 e^{-\lambda x}[\beta + 4xe^{-\lambda x} + \beta^3 x^{\beta-1} e^{\lambda(x-x^\beta)}]}{(n-1)!\,(\lambda\beta^2+\lambda\beta+1)^n} \\ & \left[\lambda\beta e^{-\lambda x} + (2\lambda x+1)\,e^{-2\lambda x} + \lambda\beta^2 e^{-\lambda x^\beta}\right]^{n-1}. \end{aligned}$$

If $r = n$ then the pdf of the nth or largest order statistics of ID distribution is given by

$$\begin{aligned} f_{X_{(n)}}(x,\lambda,\beta) =& \frac{n!\lambda^2 e^{-\lambda x}[\beta + 4xe^{-\lambda x} + \beta^3 x^{\beta-1} e^{\lambda(x-x^\beta)}]}{(n-1)!\,(\lambda\beta^2+\lambda\beta+1)} \\ & \left[1 - \frac{\lambda\beta e^{-\lambda x} + (2\lambda x+1)\,e^{-2\lambda x} + \lambda\beta^2 e^{-\lambda x^\beta}}{\lambda\beta^2+\lambda\beta+1}\right]^{n-1}. \end{aligned}$$

5. Maximum Likelihood Estimation

Let X_1, X_2,..., X_n be a random sample of size n from Idz distribution (ID). Then the likelihood function of ID is given by

$$L = \prod_{i=1}^{n} \frac{\lambda^2 e^{-\lambda x_i}[\beta + 4x_i e^{-\lambda x_i} + \beta^3 x_i^{\beta-1} e^{\lambda(x_i - x_i^\beta)}]}{\lambda\beta^2+\lambda\beta+1}. \tag{9}$$

Then, the log-likelihood function of ID is

$$l = 2n\log(\lambda) - \lambda\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\log[\beta - 4x_i e^{-\lambda x_i} + \beta^3 x_i^{\beta-1} e^{\lambda(x_i - x_i^\beta)}] - n\log(\lambda\beta^2+\lambda\beta+1). \tag{10}$$

The partial derivatives of (10) with respect to parameters β and λ are presented as follow:

$$\frac{\partial l}{\partial\beta} = \sum_{i=1}^{n} \frac{1 + \beta^3 x_i^{\beta-1}\left[\frac{3}{\beta} + \left(1 - \lambda x_i^\beta\right)\log(x_i)\right] e^{\lambda\left(x_i - x_i^\beta\right)}}{\beta - 4x_i e^{-\lambda x_i} + \beta^3 x_i^{\beta-1} e^{\lambda\left(x_i - x_i^\beta\right)}} - \frac{n\lambda(2\beta+1)}{\lambda\beta^2+\lambda\beta+1}; \tag{11}$$

and

$$\frac{\partial l}{\partial\lambda} = \frac{2n}{\lambda} - \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \frac{\beta^3 x_i^{\beta-1}\left(x_i - x_i^\beta\right) e^{\lambda\left(x_i - x_i^\beta\right)} - x_i e^{-\lambda x_i}}{\beta - 4x_i e^{-\lambda x_i} + \beta^3 x_i^{\beta-1} e^{\lambda\left(x_i - x_i^\beta\right)}} - \frac{n\beta(\beta+1)}{\lambda\beta^2+\lambda\beta+1}. \tag{12}$$

The maximum likelihood estimates of the parameters β and λ of Idz distribution can be computed by setting equations (11) and (12) equal to zero. This can be done by using any numerical method like the Newton-Raphson iterative method.

6. Application

This section presents the application of Idz distribution to a medical dataset. In this application, we use breast cancer data from Lee [9]. This dataset is taken from a large hospital in a period from 1929 to 1938 and it represents the survival times of 121 patients with breast cancer. The observations are given as follow: 0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0 and 154.0.

Fatima [5] used the above dataset for their proposed model named as the Exponentiated Generalized Inverse Rayleigh distribution (EGIR) and compared with the Exponentiated Inverse Rayleigh (EIR), the Generalized Inverse Rayleigh (GIR) and the Inverse Rayleigh (IR) distributions. They found that the EGIR had the best fit for the Breast Cancer dataset.

Here, we compare the proposed distribution with the EGIR, the Ailamujia Inverted Weibull distribution (AIW) [16] and the New Extended Exponentiated Weibull distribution (NEEW) [1]. The probability density functions of EGIR, AIW and NEEW are given as follow:

$$f_{EGIR}(x) = 2\frac{\alpha\gamma}{\lambda^2 x^3} e^{-(\lambda x)^{-2}} \left(1 - e^{-(\lambda x)^{-2}}\right)^{\alpha-1} \left[1 - \left(1 - e^{-(\lambda x)^{-2}}\right)^{\alpha}\right]^{\gamma-1}, \; x > 0, \alpha, \lambda, \gamma > 0;$$

$$f_{AIW}(x) = 4\alpha\theta^2 x^{-2\alpha-1} e^{-2\theta x^{-\alpha}}, \; x > 0, \theta, \alpha > 0;$$

and

$$f_{NEEW}(x) = \frac{\alpha\lambda x^{\lambda-1} e^{-\alpha x^{\lambda}} (1 - e^{-\alpha x^{\lambda}}) \left[e^{\theta(1-e^{\alpha x^{\lambda}})}(2 + \theta - \theta e^{-\alpha x^{\lambda}}) + 2\right]}{e^{\theta} + 1}, \; x \geq 0, \alpha, \lambda, \theta > 0.$$

In this application, we use the following diagnostics statistics: (i) Akaike Information Criterion (AIC); (ii) Bayesian Information Criterion (BIC); (iii) Kolmogorov - Smirnov (K-S); (iv) Cramer - von Mises (W^*); and (v) Anderson - Darling (A). Furthermore, a package "fitdistrplus" in R software is also used. In addition, the results are shown in the following tables. Table 1 presents the maximum likelihood estimates of the fitted models for Breast Cancer dataset while Table 2 indicates that Idz distribution gives better estimate for the given dataset since it has a smallest values of some diagnostics statistics as compared with the EGIR, AIW and NEEW distributions. Also, same result is noticed from Figure 3.

Table 1: *ML estimates of the fitted models using different distributions*

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$
ID		1.68335961		0.02058629	
EGIR	0.3331558			5986.6441628	2239.7157204
AIW	0.5159137		4.8684926		
NEEW	0.1099727		1.7984615	0.7582980	

Table 2: *Some diagnostic statistics of the fitted models using different distributions*

Distribution	AIC	BIC	$K-S$	A	W^*
ID	1164.159	1169.751	0.05341806	0.51366194	0.06178559
EGIR	1279.365	1287.753	0.2180867	9.7828973	1.5945699
AIW	1250.729	1256.321	0.1723414	7.0402261	1.1024897
NEEW	1167.178	1175.565	0.07776119	0.45898966	0.06532455

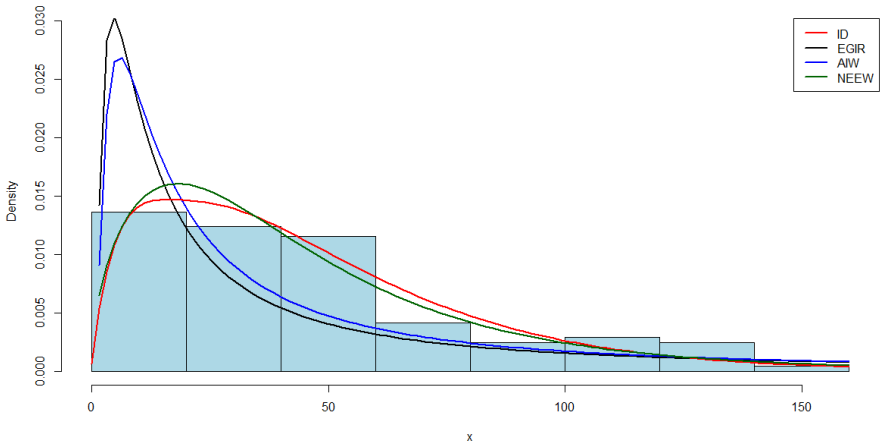


Figure 3: *Estimated pdf of the fitted models for the Breast Cancer dataset.*

7. Concluding Remarks

This paper derives a novel two - parameter continuous distribution called as Idz distribution. Some properties of Idz distribution such as moments, mean, variance, moment generating function, survival function, hazard function and order statistics were derived. Maximum likelihood method was used to estimate the parameters of Idz distribution. The applicability of the proposed distribution was examined by applying into a breast cancer data and compared with the Exponentiated Generalized Inverse Rayleigh (EGIR), the Ailamujia Inverted Weibull (AIW) and the New Extended Exponentiated Weibull (NEEW) distributions. It was found that the Idz distribution provides better fit for the given dataset as compared with the said distributions.

References

[1] Arif, M., Khan, D. M., Aamir, M., Khalil, U., Bantan, R. A. R. and Elgarhy, M. (2022). Modeling COVID-19 Data with a Novel Extended Exponentiated Class of Distributions. *Journal of Mathematics*, Article ID 1908161, https://doi.org/10.1155/2022/1908161.

[2] Benrabia, M. and AlZoubi, L. M. A. (2021). Alzoubi distribution: properties and applications. *Electronic Journal of Statistics Applications and Probability: An International Journal*, 1–16.

[3] Benrabia, M. and AlZoubi, L. M. A. (2022). Benrabia distribution: properties and applications. *Electronic Journal of Applied Statistical Analysis*, 15(2):300–317.

[4] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, 31:497–512.

[5] Fatima, K., Naqash, S. and Ahmad, S. P. (2018). Exponentiated Generalized Inverse Rayleigh Distribution with Application in Medical Sciences. *Pakistan Journal of Statistics*, 34(5):425–439.

[6] Gharaibeh, M. (2021). Gharaibeh Distribution and Its Application. *Journal of Statistics Applications and Probability*, 10(2):1–12.

[7] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communication Statistics Theory and Methods*, 27:887–904.

[8] Kingman, J. (1982). The coalescent. *Stochastic Processes and their Applications*, 13(3):235–248.
[9] Lee, E. T. (1992). Statistical Methods for Survival Data Analysis. John Wiley: Newyork.
[10] Lv, H.Q., Gao, L.H. and Chen, C.L. (2002). Ailamujia distribution and its application in supportability data analysis. *Journal of Academy of Armored Force Engineering*, 16:28–52.
[11] Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84:641–652.
[12] Shanker, R. (2015). Shanker distribution and its Applications. *International Journal of Statistics and Applications*, 5(6): 338–348.
[13] Shanker, R. (2016). Aradhana Distribution and its Applications. *International Journal of Statistics and Applications*, 6(1):23–34.
[14] Shanker, R. (2017). Rama Distribution and its Applications. *International Journal of Statistics and Applications*, 7(1): 26–35.
[15] Shraa, D. and Al-Omari, A. (2019). Darna distribution: properties and application. *Electronic Journal of Applied Statistical Analysis*, 12(2):520–541.
[16] Smadi, M. M. and Ansari, S. I. (2022). Ailamujia Inverted Weibull Distribution with Application to Lifetime Data. *Pakistan Journal of Statistics*, 38:341–358.
[17] Weibull, W. (1951). A statistical distribution function of wide applicability. *J. Appl. Mech.*, 18:293–296.

RT&A, No 4 (76)
Volume 18, December 2023

A LITERATURE REVIEW ON DISCRETE-TIME QUEUEING MODELS

Harini R[1] and Indhira K[1,*]

•

[1] Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore, 632014, Tamil Nadu, India
harini220696@gmail.com kindhira@vit.ac.in

Abstract

In this paper, a quantitative research survey is carried out on discrete-time queueing models. In real-life scenarios, the idea of discrete-time queues has taken on a new meaning. This survey mainly focuses on the unfolding of discrete-time queueing models in recent decades, challenges implied on them and their influence in various fields. The ultimate goal of this paper is to provide enough information to all the researchers and analysts who toil in this field and wish to know more about these models. A few open issues and intriguing future research paths has been discussed.

Keywords: Discrete-time queues, $Geo/G/1$ and $GI/Geo/1$ queues, $Geo/G/1$ retrial queues, Discrete-time retrial queues.

1. Introduction

Waiting in a queue has been a part of our day-to-day lives because, as a process, it has various significant purposes. Queueing theory analyses the entire structure of waiting in a queue. Erlang was a pioneer in the discipline of queueing theory in the early 20^{th} century, and now it has become a well-established research area in the past few decades. Initially, the research was done with the succor of continuous-time (CT) queueing models. Yet in the contemporary period, there has been a recognisable fluctuation from CT to discrete-time (DT) queueing systems (QS) and their counterparts as well. This change is primarily because of the noticeable relevance of DT queues in the fields of communication and computer systems, where the digital information is disseminated in the mould of fixed-length "packets," each of which requires a fixed-length transmission time known as slots. Events are constrained to take place during these slots. A DT queue, for example, might accept and serve at most one packet during a slot. Multiple events may occur during each slot on a network-wide basis. Herwig Bruneel [25] analysed a general single-server DT queueing model and its major principal performance measures were procured with the succor of a unifying analysis. Also, numerous fundamental relationships among the main quantitative measurements of a QS in general were derived, which was mainly helpful in studying the queueing phenomena, especially in the discipline of computerized messaging sectors. Many researchers then referred to books by Bruneel and Kim [24], Woodward [94], Takagi [72] and Miyazawa and Takagi [60], which helped them learn more about DT queueing models.

Due to network complexity and an expanding customer base, consumer behaviour and the retry phenomenon may have had a substantial impact on how well a system performs. Consumers who show up, discover the service completely occupied and hesitant to stay, choose to leave momentarily, but return at a subsequent period. Before retrying to occupy a server, such consumers are supposed to be in orbit, a virtual waiting room. Retrial queues (RQ) are frequently

used to mimic these situations. A basic RQ model is presented in Fig. 1.

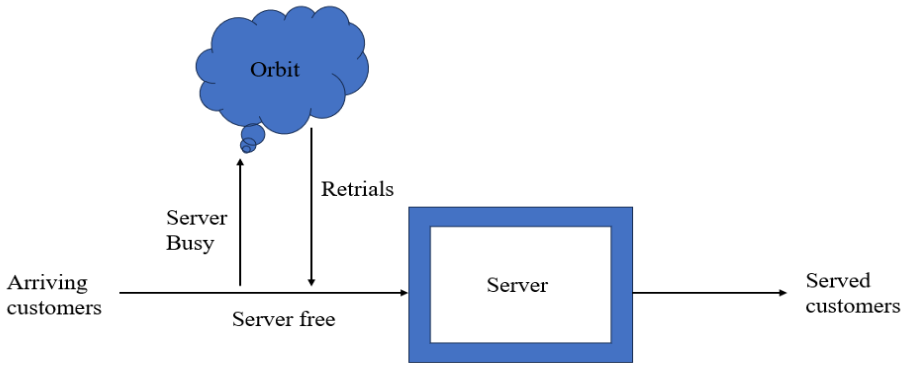


Figure 1: *A basic retrial queueing system*

To illustrate, if a consumer attempts to contact a call centre for assistance while the agents are busy, the consumers are likely to call at an inconvenient hour. This modelling is utilised in assorted fields, including call centres, cellular networks, and LAN. Meanwhile, in the recent two decades, researchers have begun to focus on DT retrial models. Rein Nobel [63] provided a high-level overview of several models for discrete-time late arrival retrial queues and demonstrated the importance of striking a balance between customers' active and passive roles in order to achieve optimal performance. Various DT retrial QS were investigated, and many results were established, which have been used in solving numerous issues in telecommunication networks, telephone switching sectors, computer systems, and networks in the past two decades.

A brief survey of various DT queueing models and DT retrial queueing models have been conducted. Readers who practice in these fields will find it quite useful. The remaining portion of this paper has been organised in the following manner: In Section 2, DT queueing models are reviewed. Section 3 presents the work on DT $Geo/G/1$ and $GI/Geo/1$ queues. DT $Geo/G/1$ retrial queues have been examined in Section 4. Section 5 reviews some of the other DT retrial queueing models. Section 6 details the difficulties in using the DT concepts, while Section 7 describes the methods employed in DT queueing models. Few distinct applications of DT queues are addressed in Section 8. Section 9 presents a few recent advancements in DT queues. Finally, the future scope and some open problems and the concluding remarks of the review are offered in Sections 10 and 11 respectively.

2. Discrete-time Queueing Models

In queueing theory, initially appearing only in scientific literature, DT queueing models in recent decades have been well-proportioned. It has largely been driven by potential applications that utilize slotted time. By analysing the proficiency of DT multi-server queues with priorities, K. Laevens and H. Bruneel [41] derived basic formulae to reckon some key metrics like variances, mean values, and tail probabilities. Dieter Fiems et al. [30] proved that DT single-server queue subjugated to server interruptions was able to reckon the proficiency of low-priority traffic in a two-priority Head-of-Line scheduling discipline. The issues of channel errors on wireless ATM multiplexers were investigated by Mehmet Ali et al. [59] by analysing the functioning of the DT queue with countless buffer. Dieter Fiems et al.[31] again studied a discrete $GI/G/1$ QS subjugated to server interruptions generated by a renewal process and further showed that this system can reckon the proficiency of a multi-class priority scheduling system. M. Mehmet Ali and X. Song [58] examined DT priority queues with correlated arrivals (CA) (i.e.with binary Markov sources) and derived a closed form of expression for the probability generating function (PGF) of queue length and finally extended the results to multiple priority queues. Alfa [4] presented an

informative work about utilizing a matrix-analytic approach to study various queues in DT.

Jinting Wang et al. [84] analysed the DT correlated on-off source QS with negative customers (NC) and demonstrated that this system can simulate loss in the radio interface of HSDPA systems. Also, they came to the conclusion that the impression of NC should be acknowledged in pursuance of the wireless communication networks in an error-prone transmission environment. I. Atencia and A. V. Pechinkin [20] studied a DT QS with optional LCFS priority discipline and examined most of its principal characteristics. An $MMBP/Geo/1$ DT queue has been scrutinized by Jinting Wang et al. [85] with correlated positive and NC arrivals and its mean buffer content and the stationary probabilities have been determined. Ultimately, they concluded that when time slots with persistent size are defined, this queue will be supportive in analysing the radio link layer of HSDPA systems. I. Atencia [11] after analysing a DT QS with the customers arriving where the system follows LCFS discipline or to enroll in the queue where servers may fail and yet be repaired, proposed that the study of the server breakdowns (SB) can be brought into a realistic approach by reckoning with the modifications in the renewal times. With the succor of the Markovian approach, A. Senthil Vadivu et al. [80] scrutinized a DT infinite capacity QS with an on-off source and service time with two states of server and NC. There has been a rise in the popularity of non-deterministic service times in DT models because of the more intricate and erratic service mechanisms in today's communications networks. As a result, Peixia Gao et al. [34] devised a generating-functions method for analyzing a DT multi-server QS with geometrically distributed service times. In the analysis of a finite buffer DT multiple working vacation queue with impatient clients and congestion-dependent service rates, Jyothsna et al. [37] determined the steady-state system length distributions via a matrix technique and give a recursive solution. From the recursive solution of the DT queue, the corresponding CT queue is further calculated.

Recently, F. Verdonck et al. [82] focused on a DT multi-server QS with varying server availability by introducing server interruptions and a CA process. They also assumed the QS to alternate between two distinct dynamic characteristics having a random cluster of servers and determined the allocation of the system content at several observation points. The repercussions of assorted forms of variance on the setup behavior was also investigated and finally, they concluded that this paradigm is suitable for many real-life applications. Jens Baetens et al. [22] investigated a variable-capacity DT two-class batch-service queue in which all clients wait in a single queue and are served in the order in which they arrived and moreover clients of the same class are the only ones the server can service, hence the number of clients in a batch is equal to the number of successive clients of the same class. Using a supplementary variable technique (SVT), the repercussions of disasters on the processing of a DT single-server QS were studied by Mustafa Demircioglu et al. [29] and they concluded that the mean system content reduces with increasing irregularity for the QS with disasters. In order to evoke retransmission in communication packet networks, Kempa [38] has studied the time-dependent queue behavior (i.e., transient state) of a DT queue with a finite buffer and feedback. Larisa and Andrey [45] studied a discrete-time queueing system with a regenerative input flow and heterogeneous servers that may have independent interruptions, where the subsequent moments of breakdowns are specified by a renewal mechanism and are independent of the system state.

Alfa [5] introduced a new way to analyze the $GI/G/1/K$ system via the matrix-analytical method. The novelties stem from the Markov-based depiction of the arrival and service processes, at least one of which is based on the elapsed time approach, and the unusual organization of the state space. The packet loss distribution of a finite buffer single server queue subject to a DT batch Markovian arrival process (DBMAP) and a DT Markovian service process (DMSP)is investigated through a matrix-analytical method by Yung-Chung Wang et al. [90]. Using statistical measures, we will analyze the sporadic behavior of packet loss in terms of both loss period and loss distance. Using a DT $GI^{[X]}/G^{[Y]}/1$ queueing model, Alfa and He [6] investigated an algorithmic strategy based on the matrix-analytic approach. They also introduced and studied the Markov chain that underlies the ageing of service-receiving clients.

3. Discrete-time $Geo/G/1$ and $GI/Geo/1$ Queueing Models

Zhe George Zhang and Naishuo Tian [101] scrutinized the DT $Geo/G/1$ system with collective adjustable vacations and showed that the optimal control issue of vacation policies can be discussed with the succor of this paradigm. Again, a DT $GI/Geo/1$ queue with server vacations was analysed by Naishuo Tian and Zhe George Zhang [74]. The stationary distributions and the waiting times of the customers were procured by using the matrix-geometric method (MGM). Also, this model can be stretched to a vacation with a finite maximum value that is quite widely distributed. Ji-Hong Li and Nai-Shuo Tian [49] examined a DT $GI/Geo/1$ queue with working vacations (WV) and vacation interruption and showed that this QS can be helpful in modelling few realistic obstacles in communication networks and computers. Ji-Hong Li et al.[50] extended a $GI/M/1$ queue with WV to a DT $GI/Geo/1$ queue with umpteen WV, which was further discussed using MGM. For negative binomial distributions, the closed property of conditional probability was also determined and verified.

A DT $Geo^{[X]}/G/1$ queue with an uncertain server and collective adaptive delayed vacation policies was scrutinized by Yinghui Tang et al. [73] where the transient and steady-state distributions of the queue length were obtained and the stochastic decomposition (SD) property of the steady-state queue length was demonstrated. It is feasible to obtain a connection amidst the generating functions of steady-state queue length at departure epoch and an arbitrary epoch.The study of reliability problems became significant since the QS is unreliable. Furthermore, some key metrics were employed to assess the repercussions of diverse specifications on numerous functioning characteristics. A DT $Geo/G/1$ QS with a random threshold policy, namely a (p, N) policy, was analysed by Tsung-Yin Wang and Jau-Chuan Ke [91] and with reference to p and N, they conferred the convex combination of the QS characteristics. A randomised vacation policy was pioneered by Tsung-Yin Wang et al. [92] for a DT $Geo/G/1$ QS that takes at most J vacations. Using the generating function technique, few quantitative measurements of the QS were derived, and this model has great relevance in practical systems as well. The estimation of M/G/1 CT QS with D-policy from a DT $Geo/G/1$ system with D-policy was probed by Se Won Lee et al. [46]. Moreover, the repercussions of threshold D were also studied. Doo Ho Lee and Won Seok Yang [47] were the first ones to look into the conjuction of the disaster phenomenon and the N-policy under DT on $Geo/G/1$ QS and concluded that this could be useful to minimise power consumption in wireless sensor networks (WSNs). At the same time it also help network engineers to design and operate WSNs. Renbin Liu and Zhaohui Deng [53] examined a DT N-policy $Geo/G/1$ QS with repairable server and feedback, and the steady-state system size distribution (SSD) was derived by renewal process theory and probability analysis whose application is applied in a network access proxy system.

Jianjun Li and Liwei Liu [48] scrutinized a DT $Geo/G/1$ QS with vacations in a random environment and obtained the PGF using SVT and showed that it can be estimated to $M/G/1$ QS. A more realistic approach to real life problems was offered by the DT $Geo/G/1/$ QS with changes in vacation times, which was analysed by Ivan Atencia [12]. A state that occurs in a computer network centre with virus infection was investigated and compared with a DT $Geo/G/1$ QS with single vacation and disaster arrivals by S. Jeyakumar and P. Gunasekaran [35] and also the PGF of the queue length and the mean queue length of the QS were derived, which have proven to be valuable in a variety of real-world settings. Shaojun Lan and Yinghui Tang [42] were the initial ones to analyse the DT $Geo/G/1$ QS Bernoulli feedback, modified multiple vacations, and N-policy. The customers waiting time has been reduced, the system's queue length has been controlled, moreover the switching costs of the QS has been largely economised by this QS. Finally, to investigate the cost optimization problem, a cost structure was also established. F. P. Barbhuiya and U. C. Gupta [23] studied a DT batch arrival $GI/Geo/1$ QS with numerous WV. Using SVT and the shift operator method, the closed-form expressions of steady-state system content distribution have been derived, and the impression of assorted specification on the functionality of the QS have also been conferred. A repairable DT $Geo/G/1$ QS with tragedies and working breakdown was analysed by Shan Gao et al. [33] which contributes to a more practical service schedule for

the growth of DT queues with tragedies and breakdowns. Futhermore, a manufacturing system can be feasibly modelled using this. Recently, Vaishnawi [81] looked at the accuracy of a DT $Geo^{[X]}/G/1$ recurrent model with Bernoulli feedback and two distinct forms of vacations.

3.1. Discrete Markovian Queueing models

Bara Kim and Jeongsim Kim [40] investigated a DT batch Markovian arrival process $(D-BMAP)/G/1$ RQ and found a light-tailed asymptotic distribution for the amount of consumers at embedded epochs. Jesus R. Artalejo and Quan-Lin Li [9] introduced an efficient generalisation of the DT Markovian arrival process (D-MAP) by analysing an RQ that follows a new discrete block state-dependent arrival (D-BSDA) distribution. Also, this model aids in creating more sophisticated models. Matrix analysis and the embedded Markov chain technique were used to determine the joint state probabilities at different epochs (arbitrary, pre-arrival, and departure) in Yu and Alfa [99] study of a DT single-server finite-buffer queue with a Markovian arrival process. Nandi and Samanta [86] provided a consolidated view of a system subject to shocks following the rules of a D-MAP, a stochastic model in which the arrival timings of events rely on one another. The results of this research can be used to real-world issues that manufacturing engineers encounter, improving a certain class of stochastic systems in which the shock magnitudes are isolated. By re-blocking the transition probability matrix in its desired $GI/G/1$ structure with rectangular boundary blocks, Das and Samanta [27] was able to determine the queue-length distribution at pre-arrival epochs for both the late arrival system with delayed access and the early arrival system of an $GI^{[X]}/DMSP^{(a,b)}/1$ queue using the RG-factorization method based on censoring technique. Similarly, Samanta and Das [68] examined $DBMAP/DBMSP/1$ queue by re-blocking the transition probability matrix to the desired $M/G/1$ structure via a censoring methodology-based UL-type RG-factorization approach in order to determine the stationary probability vectors at an external observer's epoch for a fixed-size batch service queue via a matrix-analytic approach. Samanta and Nandi [69] examines a single-server, infinite-buffer queueing system in which customers arrive in varying-sized groups in DT. The pre-arrival epoch probabilities are calculated using the UL-type RG-factorization for the Toeplitz type block-structured Markov chain, which is based on the censoring method. A renewal-input N-policy DT $GI/D-MSP/1/\infty$ QS was again studied by Samanta and Nandi [70]. The pre-arrival epoch duration distribution of the system is calculated using a matrix-geometric approach. We use the Markov renewal theory to calculate the random epoch distribution of system lengths. The optimal value of N is found by minimizing the expected cost function, which is assumed to be linear in terms of units of time.

3.2. Discrete-time $Geo/Geo/1$ Queueing Models

The enhancement of queueing theory on NC and disasters to the DT $Geo/Geo/1$ QS was probed by Ivan Atencia and P. Moreno [19] and some of the key metrics of the QS have been found along with its ergodicity condition. A restricted source DT $Geo/Geo/1$ QS with disasters was analysed by F. Jolai et al. [36] that has been used to design an email contact center plus with the succor of disasters, a clearing performance in a real-life source system can also be modelled. R. Sudhesh and R. Sebasthi Priya [71] investigated a DT $Geo/Geo/1$ QS with repair, feedback, and disaster. Busy-period distribution was determined in reference to Catalan numbers, and quantitative measurements like reliability and availability were determined, on top of some key metrics for steady and transient state system-size probabilities. Lee and Ke [52] constructed and studies the steady-state solutions of a $Geo/Geo/1$ queue under the QBD (quasi-birth-death) model, which accounts for server failures and reboots. In this case, the QBD model is level-independent up to a fixed threshold but level-dependent indefinitely above that. Cramer's rule, which goes along with QR factorization, can be used to solve the level-dependent structure of the suggested model. For this reason, they first calculated the states' probabilities using the

finite level-independent technique, and then switch to the algorithm of infinite level-dependent QBDs. Artalejo et al. [8] investigated the performance metrics of a $Geo/Geo/c$ type DT queue with geometric repeated attempts and provided many algorithmic approaches for their efficient computation.

4. Discrete-time $Geo/G/1$ Retrial Queueing Models

Tao Yang and Hui Li [97] were the pioneer to peruse the steady-state distribution of the DT $Geo/G/1$ RQ and showed that the SD law holds for this QS. Also, they proved that this QS can be ultilized to imprecise the CT $M/G/1$ RQ. Not only that, but also, this model has great applications in communication and computer systems where consumer retrials are a conventional phenomenon. A DT $Geo_1, Geo_2/G/1$ RQ with two kind of calls was studied by B. D. Choi and J. W. Kim [26], which seems to have applicability in telephone switching systems and mobile cellular systems. The association between a CT $M/G/1$ RQ with a Bernoulli schedule and its DT counterpart was established by Hui Li and Tao Yang [51] by analysing a DT $Geo/G/1$ RQ with a Bernoulli schedule. I. Atencia and P.Moreno [18] studied a DT $Geo^{[X]}/G_H/1$ RQ with Bernoulli feedback and demonstrated that this QS is sometimes utilized to approximate $M^{[X]}/G_H/1$ RQ with Bernoulli feedback, plus they derived various quantitative measurements of the QS. An estimation of $M/G/1$ RQ with general retrial times was again given by I. Atencia and P.Moreno [15] by examining a DT $Geo/G/1$ RQ with general retrial times. With impatient consumers and a server prone to startup problems, Aboul-Hassan et al. [3] dealt with a $Geo/G/1$ RQ in DT. To simplify the computation of significant distributions, recursive formulas were constructed. To learn how impatience impacts system performance generally, a simulation study has also been performed constructed.

R. Artalejo et al. [7] scrutinized a DT $Geo^{[X]}/G/1$ RQ with batch arrivals where individual customers have admission control, plus the underlying Markov chain was also studied, and finally they concluded by approximating CT $M/G/1$ RQ with batch arrivals and admission control. I. Atencia and P.Moreno [16] investigated a DT $Geo/G/1$ RQ with starting failures once more. They devised two SD laws (SDL) and found the measure of proximity betwixt the SSDs. Also, they proved that this QS can be estimated to M/G/1 RQ with starting failures. I. Atencia and P. Moreno [17] studied a DT $Geo/G/1$ RQ with a breakdown and repairs again, which led to even more findings of DT RQ with active breakdowns. A stochastic distribution law was also derived and perhaps they even pioneered the idea of generalized service time. Jinting Wang and Qing Zhao [89] investigated a DT $Geo/G/1$ RQ contingent on starting fiaso and second optional service. Formulae for the stationary distribution along with two SDL were procured and few statistical paradigm were presented. The amount of consumers within the system's generating functions and in the orbit had been determined along with two SDLs and limits for the distance between the SSDs were given by Jinting Wang and Qing Zhao [88] by analysing a DT $Geo/G/1$ RQ with general retrial time and starting failures. A. Aboul-Hassan et al. [2] scrutinized a DT $Geo/G/1$ RQ with general retrial time and balking customers. They proved that as a fusion of two unrelated random variables, the SSD can be indicated. Also, a set of recursive formulae were obtained, and finally, they concluded that when the QS is considerably charged, the repercussions of balking is more apparent, whereas it has a slight impact on busy probability and the average system magnitude for light loads. Bara Kim and Jeongsim Kim [39] examined a DT $Geo/G/1$ RQ and they found the tail of the queue size distribution and proved that it was asymptotically geometric and also showed that it has been inconsistent with the CT M/G/1 retrial queue. A DT $Geo^{[X]}/G/1$ RQ with general retrial times had been scrutinized by Abdel-Karim Aboul-Hassan et al. [1] where the generating functions of the system state distribution plus the orbit magnitude and SSD were derived. Also, a study was conducted on the consequences of bulk arrivals and general retrial times on the stability region, and they concluded that the upper bound of the stability region rapidly reduces by boosting the average bulk size (i.e)., the whole mean batch size has a significant issue on the QS's functioning.

The repercussions of periodic customers was studied by I. Atencia et al. [13] by analysing

a DT $Geo/G/1$ RQ with recurrent customers, that has various relevance on computers and communication networks. This QS is also affiliated to the vacation queue. Using Discrete Fourier transform inversion, the probability mass function of the QS was also obtained. I. Atencia et al. [14] scrutinized a DT $Geo/G/1$ RQ with Bernoulli feedback, starting failures, and general retrial times in order to obtain various analytical expressions and moreover, CT QS can be roughly estimated by this model. Also, this QS can be asigned to assorted versions of server interruptions. Jinbiao Wu et al. [95] studied a DT $Geo/G/1$ RQ with anticipatory restart and crash, and its ergodicity constraint was derived by analysing the Markov chain. Using the generating function technique, the system state distribution plus the orbit magnitude and SSD was obtained, and furthermore, the SD property was also investigated. Shan Gao et al. [32] examined a repairable $Geo/G/1$ RQ with Bernoulli feedback, recurrent cunsumers, and general retrial times. Using SVT, the Markov chain was studied, and some queueing measures were estimated by solving the Kolmogrov equations. Numerical illustrations were also done to prove the sensitiveness of the QS performance.

A DT $Geo_1, Geo_2^X/G_1, G_2/1$ RQ with two grouping of consumers and feedback was scrutinized by Zaiming Liu and Shan Gao [54] where they studied the Markov chain and procured some quantitative measurements of the QS in the steady-state. Also, the connection between DT system and its continuous counterpart was investigated. A DT $Geo/G/1$ RQ with favoured, intolerant consumers and general retrial times was analysed by Jinbiao Wu et al. [96]. The system state distribution plus the orbit size and SSD were determined. Apart from the SD property, its equivalent CT QS were also investigated. A DT $Geo/G/1$ RQ with general retrial time and Bernoulli vacation was probed by Jinting Wang [83] and he obtained its ergodicity condition along with its two SDL. This paradigm is mainly presented to enhance the view of DT scenarios with vacations from retrial queueing theory.

Yue Dequan and Zhang Feng [100] studied a DT $Geo/G/1$ RQ with general retrial times and a J-vacation policy. The generating abilities of the amount of consumers in the orbit and in the QS, along with the SD outcome for the system size, were procured, and finally, the end result of some frameworks of the QS were also investigated. A DT $Geo/G/1$ RQ with J vacations where the server subjugated to two contrasting kinds of breakdowns was examined by Feng Zhang and Zhifeng Zhu [102] and a few quantitative measurements of the QS were derived. Further numerical analysis was carried out to look into the implications of vacations and collapse of the system. Feng Zhang and Zhifeng Zhu [103] studied DT $Geo/G/1$ RQ with two contrasting kinds of vacations, the non-exhaustive random vacation policy and the exhaustive single vacation policy, and illustrated that the system magnitude has a solely unique SD property. Finally, the repercussions of assorted parameters on few of the main key metrics of the system were presented.

Sheweta Upadhyaya [76] shown how to obtain expressions for system size, orbit size, and other critical metrics using the generating function methodology and other key metrics by analysing a DT $Geo^{[X]}/G/1$ RQ using Bernoulli feedback. The congestion issues endured frequently in various digital systems can be feasibly cleared up with the succor of this paradigm, and it was also proved that this model assit network developers in detaching its system size into two queueing models, one with lacking retrials and the other being the conventional DT RQ. Lately, Shweta Upadhyaya [77] employed the generating function approach and the SVT to analyze a DT $Geo^{[X]}/G/1$ RQ with Bernoulli feedback and a starting failure, and then she deployed the particle swarm optimisation technique to achieve the optimal values for a few critical system parameters, for instance the total system cost, by minimizing it. The initial work on a DT $Geo/G/1$ RQ with non-pre-emptive priority, general retrial times, working vacations, and vacation disruption was probed by Shaojun Lan and Yinghui Tang [43]. The closed-form expressions for the PGF of the stationary distribution of diverse server states, the amount of consumers in the priority queue, in the orbit, and in the QS were found using SVT and the generating function method, and this model can be executed in many real-life congestion scenarios. Again, the pioneer work on a DT $Geo/G/1$ RQ with probabilistic pre-emptive priority, balking consumers, initial defects, and substitutions of repair times was reported by Shaojun Lan and Yinghui Tang [44] where they introduced the concept of replacements. Some quantitative measurements were obtained, and a

cost structure for determining the optimal replacement probability while minimising the system cost was also established.

Pavai Madheswari et al. [55] investigated a general service RQ with recurring clients that operated on a single server in DT and as a special case of the suggested model, several of the current results are generated. Malik and Upadhyaya [57] have recently discussed a DT $Geo/G/1$ RQ, with retrials, preferred and impatient units, single vacations, and state-dependent policies and again they [56] focused on studying a DT $Geo/G/1$ RQ in which, in order to ensure client satisfaction and enhance service quality, the server offers both essential and discretionary services. Shweta Upadhyaya [78] discussed about the periodic customers, high-priority customers, and impatient customers make up the arrival stream in an unreliable DT $Geo/G/1$ RQ with Bernoulli feedback. Recent research conducted by Rajasudha et al. [67] compares the performance of a batch arrival single server DT RQ under three distinct vacation regimes.

4.1. Discrete-time $Geo/Geo/1$ Retrial Queueing Models

An investigation on DT $Geo/Geo/1$ RQ with a server subjugated to failures and a general server lifetime was probed by P. Moreno [61] and diverse quantitative measurements were also estimated. Finally, some numerical analyses were presented. Jin-Ting Wang and Peng Zhang [87] extended the evaluation of retrial queueing theory on NC to DT retrial G QS by analysing a DT $Geo/Geo/1$ retrial G-queue with server breakdowns and repairs. Using G-queues, the reliability modelling was widen to the DT scenario for the maiden time using this model. The marginal functions of the orbit size, besides the ergodicity condition had also been determined. Sheweta Upadhyaya's [75] study on a DT $Geo^{[X]}/Geo/1$ RQ with a functioning vacation plan, aided to design numerous realistic as well as economic gridlock circumstances. Using the MGM, diverse quantitative measurements were obtained. Futhermore, with the succor of a straightforward delve technique gleaned from a heuristic method, joint optimal values were also determined. Also, this model is feasibly regarded as an economical tool. An elementary and basic study on DT $Geo/Geo/1$ retrial G-queues with server breakdown and repair was probed by A.Azhagappan et al. [21] where they acquired the model's ergodicity condition by analysing the Markov chain underlying the QS.

5. **Other Discrete-time Retrial Queueing Models**

R. Artalejo and M. J. Lopez-Herrero [10] conducted a contemplation on DT multi-server RQ with a defined population. In the study conducted by Rein Nobel and Pilar Moreno [64] on a DT single-server QS with retrials, they calculated ergodicity conditions, the vestigial service epoch, and average orbit size, besides some quantitative measurements. A distinct aspect to handle the combinations of various conditions like positive arrival, negative arrival, server breakdowns, and customer retrial was given by Jinting Wang and Peng Zhang [93] by analysing a DT single-server RQ with NC and an unreliable server. A DT inventory system with retrial demands that tends to access in aggrement with a Bernoulli process where the inventory has been refilled in accordance with (s, S) policy was probed by C. Periyasamy [65]. A DT $GI/G/1$ RQ with Bernoulli retrials and time restraint vacation was examined by Jin-ting Wang et al. [86] which further, with the aid of the Markov-based methodology, feasibly be analysed by a level-dependent quasi-birth-and-death (LDQBD) method that enhances the system's computability with a technical perspective. Recently, Upadhyaya et al. [79] analysed a DT bulk-entry recurrent- RQ to examine problems with traffic management and control in ATM networks. First, they used the generating function approach to obtain the required performance indices, and then they utilized an adaptive neuro-fuzzy interface system (ANFIS) to get a rough approximation of all the estimated findings. Finally, they used computational analysis of the model, involving particle swarm optimisation and genetic algorithm techniques, to make the system more cost-effective. Rajasudha and Arumuganathan [66] adopted a hybrid method to simulate and analyze the behavior of a system infected with malware. Malware is filtered out and system performance is enhanced due to a likelihood

estimation. To account for the peculiarities of the damaged system, a DT single-server RQ with two kinds of arrivals is addressed. Xiaoyun Yu et al. [98] introduced an innovative k-out-of-n : G repairable QS that incorporates Bernoulli shocks and retrial into it under the DT assumption as a mechanism of redundancy to considerably increase the system's reliability. Additionally,the DT Markov processes theory is used to assess the system's performance.

6. Challenges on using Discrete Time Queueing models

While many of the current contributions focus on single arrival DT systems, there are various practical scenarios where batch arrival DT systems can exist but have received little attention. An additional difficulty in DT queues is posed by working vacation polices. Moreover, multi-stage DT systems have not been explored due to the focus on single-stage DT systems until this point. Traditional queuing systems are well-known and widely utilized, but their assumptions are often too far-fetched for comfort, therefore fuzzy queues are employed to reflect real scenarios instead. Fuzzy queues are more accurate and realistic than the usual queue method. Fuzzy parameter analysis of DT queues is a growing problem.

7. Techniques employed in Discrete-time models

Many authors have presented new methods for analyzing queueing models that take DT into consideration. They employ strategies such as supplementary variable technique, generating function method, matrix-geometric method (matrix analytic method), quasi-birth-and-death process, quasi Newton method, Markov based approach, UL-type RG-factorization, (s,S) policy, Newton–Raphson's method, method of exclusion, discrete Fourier transform method, partial repeat-after-interruption service strategy, truncation method, transform technique, routine method, direct search method based on heuristic approach to obtain the optimal cost, numerical inversion and maximum entropy techniques, theory of difference equation, renewal process theory, probabilistic analysis method, embedded Markov chain technique, etc. In the study of DT queueing systems, most of the parameters we consider display a stochastic process, therefore the variables we denote arrival rate, retrial rate, service rate, balking rate, reneging rate, breakdown rate, repair rate, vacation rate, feedback rate as random variables.

8. Applications of Discrete-time Queueing models

Several DT queueing models can be extrapolated from practical situations. The creators of these DT queueing models can use the resulting overview to direct their decision-making. This section provides a brief summary of some application-focused research articles that have addressed DT queueing models.

Wireless sensor networks: While digital communication protocols are managed by a " slot,"which is equivalent to a time interval, DT queues are better suited for this analysis. Since data packets in wireless sensor networks are vulnerable to being lost due to external attacks or shocks, DT queues are applied to a power-saving technique. Understanding the appropriate operation of the N-policy for power saving in a wireless sensor networks is realistically vital when dealing with sensor nodes under unstable network connections. Since the N-policy lessens the amount of initial power used to toggle a radio server in sensor nodes between its busy and idle states, it is an effective power-saving method. Thus, Doo and Won [47] analysed a DT queue with N-policy and applied in wireless sensor networks.

Wireless ATM multiplexer: Mehmet Ali et al. [59] looked into how wireless ATM is affected by channel faults. Their efforts will be useful for a cell in a wireless ATM network that includes a base station and a number of users on mobile devices. Since the wireless access point acts as an interface to the wired ATM network, the latter can now serve mobile users. While all the mobile users are assumed to experience the same channel condition at any one time, the uplink

wireless channel can be modeled as a DT queueing system fed by On-Off sources and served by a two-state Markovian server, and then analyzed accordingly.
Wireless local communication: In order to put a number on how much of an impact multimedia services have over a wireless local communications channel with Rayleigh fading, a DBMAP queuing model [90] is used. In this case, a wireless link with sporadic defects and a high data rate is used, with the time axis slotted. Error processes of varying granularity can be modeled using this framework. It can monitor problems at multiple levels, including the bit and packet levels, among others.
Email contact center: The DT model may find use in the increasingly common incoming email contact center. Users from various offices, faculties, and research centers send and receive emails over the network using the office automation system to communicate with potential customers.Everything from requesting a channel to sending and receiving data happens at regular intervals. Thus, with these systems, sending emails, preprocessing, and processing requests are all done in a defined amount of time. While this may be fine for email, it is blatantly unacceptable for incoming phone calls. This is why a DT model of the system is more appropriate here [[36], [92]].
Computer and communication systems: Like bits and bytes, the inter-arrival times of packets and their forward transmission periods are the elementary units of time in computer and communication systems (such as time division multiple access (TDMA)).DT queueing activities can only happen at regularly spaced epochs, which has provided a strong impetus for studying this phenomenon. Furthermore, DT queues can be used to approximate the continuous systems, but not the other way around, making them a better fit for defining the behaviors of data communication and computer networks [42], [74].

9. Advancements in Discrete Time Queueing models

Once a model has been built and is ready to go into operation, the first thing a manufacturer or analyst will want to know is whether or not it is cost effective. This highlights the importance of 'cost analysis' in determining a model's sufficiency. As a result, system designers are able to make more informed choices and reduce potential dangers. In view of this, many authors [[79],[77], [81], [66]] employing methods such as particle swarm optimization, artificial bee colony, genetic algorithm, etc., have attained the cost optimization for their queueing models.
Combining the strengths of neural networks and fuzzy logic systems, an adaptive neuro fuzzy inference system (ANFIS) is a form of artificial intelligence. Like a neural network, ANFIS can learn and make judgments based on data, but it can also deal with fuzzy or incomplete data in the same way as a fuzzy logic system can. Because of this, ANFIS is particularly well-suited for uses where data is volatile or unreliable. Recently, many studies [[79],[81]] have been conducted recently comparing the numerical findings of retry QS to those obtained using neuro-fuzzy methods.

10. Future Scope

The focus of research has shifted in recent years from continuous-time queues to their DT analogues. Applications of DT queues in environments where time is slotted are mostly responsible for this change. However, there have been many studies and publications in the field of DT models, there are still many questions that need to be answered. This section primarily suggests some potential avenues for future study.
Many potential extensions on the DT models with heterogeneous service have not been investigated because previous research mostly focused on single service phase DT models. These unanswered questions may, in part, be attributable to the " curse of dimensionality, "which results from the complexity of the underlying systems. One promising area of study is the formation of approximations to these intricate DT models. Moreover, few works have been done on quorum queues, so investigating DT quorum queues is a promising area for future study.

Additionally, the transient behaviour of DT queues will be an intriguing aspect to investigate as well. Further, complex vacation policies, such as hybrid vacation, coxian vacation, and Bernoulli working vacation, are yet another path to explore while analyzing DT queues.

11. Conclusion

A comprehensive investigation on DT queueing models and DT retrial queueing models has been conducted in this paper. The primary purpose of this review is to access the current level of understanding in the discipline of DT queueing models, identify key authors, research papers, theories and conclusions, and identify the knowledge gaps in this field. The applications of these QS are widely applied in our day-to-day life, precisely in the field of communication and computer systems, where time is slotted. The ideas affiliated to the DT Geo/G/1 QS and retrial QS scrutinized in several research papers have been synthesized. Distinct techniques and various challenges involved in DT queues are discussed. Further, recent advancements in DT queues are also presented. Thus, a diverse collection of literature has been reviewed, with appropriate references mentioned.

Acknowledgements: Not applicable
Authors contribution All the authors made substantial contributions to the conception or design of the work.
Availability of Data: Not applicable.
Funding: Not applicable.
Conflicts of Interest: The authors report no conflict of interest.

References

[1] Aboul-Hassan, A. K., Rabia, S. I. and Taboly, F. A. (2009). Performance evaluation of a discrete-time $Geo^{[X]}/G/1$ retrial queue with general retrial times. *Computers and Mathematics with Applications*, 58(3):548-557. https://doi.org/10.1016/j.camwa.2009.03.101.

[2] Aboul-Hassan, A. K., Rabia, S. I. and Taboly, F. A. (2008) Discrete-time $Geo/G/1$ retrial queue with general retrial times and balking customers. *Journal of the Korean Statistical Society*, 37(4):335-348 .

[3] Aboul-Hassan, A. K., Rabia, S. I. and Al-Mujahid, A. A. (2010). A discrete-time $Geo/G/1$ retrial queue with starting failures and impatient customers. *Transactions on computational science VII*, 22-50. https://doi.org/10.1007/978-3-642-11389-5_2

[4] Alfa, A. S. (2002). Discrete time queues and matrix-analytic methods. *Top*, 10:147-185. https://doi.org/10.1007/BF02579008

[5] Alfa, A. S. (2003). An alternative approach for analyzing finite buffer queues in discrete time. *Performance Evaluation*, 53(2):75-92. https://doi-org.egateway.vit.ac.in/10.1016/S0166-5316(02)00225-0

[6] Alfa, A. S. and He, Q. M. (2008). Algorithmic analysis of the discrete time $GI^{[X]}/G^{[Y]}/1$ queueing system. *Performance Evaluation*, 65(9):623-640. https://doi-org.egateway.vit.ac.in/10.1016/j.peva.2008.02.001

[7] Artalejo, J. R., Atencia, I. and Moreno, P. (2005). A discrete-time $Geo^{[X]}/G/1$ retrial queue with control of admission. *Applied Mathematical Modelling*, 29(11):1100-1120. https://doi.org/10.1016/j.apm.2005.02.005.

[8] Artalejo, J. R., Economou, A. and G?mez-Corral, A. (2008). Algorithmic analysis of the $Geo/Geo/c$ retrial queue. *European Journal of Operational Research*, 189(3):1042-1056. https://doi-org.egateway.vit.ac.in/10.1016/j.ejor.2007.01.060

[9] Artalejo, J. R. and Li, Q. L. (2009). Performance Analysis of a Block-Structured Discrete-Time Retrial Queue with State-Dependent Arrivals. *Discrete Event Dynamic Systems*, 20:325-347. https://doi.org/10. 1007/s10626-009-0075-6

[10] Artalejo, J. R. and Lopez-Herrero, M. J. (2007). A simulation study of a discrete time multi-server retrial queue with finite population. *Journal of Statistical Planning and Inference*, 137(8):2536-2542. https://doi.org/10.1016/j.jspi.2006.04.018.

[11] Atencia, I. (2015). A discrete-time queueing system with server breakdowns and changes in the repair times. *Annals of Operations Research*, 235:37-49. https: //doi.org/10.1007/s10479-015-1940-3.

[12] Atencia, I. (2016). A Discrete-time Queueing Systems with changes in the Vacation Times. *International Journal of Applied Mathematics and Computer Science*, 26(2):379-390. https://doi.org/10.1515/amcs-2016-0027.

[13] Atencia, I., Fortes, I., Nishimura, S. and Sanchez, S. (2010). A discrete-time retrial queueing system with recurrent customers. *Computers and Operations Research*, 37(7):1167-1173. https://doi.org/10.1016/j.cor.2009.03.029.

[14] Atencia, I., Fortes, I. and Sanchez, S. (2009). A discrete-time retrial queueing system with starting failures, Bernoulli feedback and general retrial times. *Computers and Industrial Engineering*, 57(4):1291-1299. https://doi.org/10. 1016/j.cie.2009.06.011.

[15] Atencia, I. and Moreno, P. (2004). A Discrete-Time $Geo/G/1$ Retrial Queue with General Retrial Times. *Queueing Systems*,48:5-21. https://doi.org/10.1023/B:QUES.0000039885.12490.02.

[16] Atencia, I. and Moreno, P. (2006). A Discrete-Time $Geo/G/1$ retrial queue with the server subject to starting failures. *Annals of Operations Research*,141:85-107. https://doi.org/10.1007/s10479-006-5295-7.

[17] Atencia, I. and Moreno, P. (2006). A Discrete-time $Geo/G/1$ Retrial Queue with Server Breakdowns. *Asia-Pacific Journal of Operational Research*, 23(02):247-271. https://doi.org/10.1142/S0217595906000929.

[18] Atencia, I. and Moreno, P. (2004). Discrete-Time $Geo^{[X]}/G_H/1$ Retrial Queue with Bernoulli Feedback. *Computers and Mathematics with Applications*,47(8-9):1273-1294. https://doi.org/10.1016/S0898-1221(04)90122-8.

[19] Atencia, I. and Moreno, P. (2004). The discrete-time $Geo/Geo/1$ queue with negative customers and disasters. *Computers and Operations Research*, 31(9):1537-1548.https://doi.org/10.1016/S0305-0548(03)00107-2.

[20] Atencia, I. and Pechinkin, A.V. (2013). A discrete-time queueing system with optional LCFS discipline. *Annals of Operations Research*, 202:3-17. https: //doi.org/10.1007/s10479-012-1097-2.

[21] Azhagappan, A., Sridevi, A. and Deepa, T. (2018). Discrete Time $Geo/Geo/1$ Retrial G – Queue with Server Break Down and Repair. *International Journal of Engineering Research and Technology*.

[22] Baetens, J., Steyaert, B., Claeys, D. and Bruneel, H. (2016). System occupancy of a two-class batch-service queue with class-dependent variable server capacity. *Proceedings In Analytical and Stochastic Modelling Techniques and Applications: 23rd International Conference, ASMTA 2016, Cardiff, UK, Springer International Publishing*, 32-44. https://doi.org/10.1007/978-3-319-43904-4_3

[23] Barbhuiya, F. P. and Gupta, U. C. (2020). A Discrete-Time $GI^{[X]}/Geo/1$ Queue with Multiple Working Vacations Under Late and Early Arrival System. *Methodology and Computing in Applied Probability*, 22:599-624. https://doi. org/10.1007/s11009-019-09724-6.

[24] Bruneel, H. and Kim, B.G. (2012). Discrete-time models for communication systems including ATM. *Springer Science & Business Media*,205.

[25] Bruneel, H. (1993). Performance of Discrete-Time Queueing Systems. *Computers Operation Research*, 20(3):303-320. https://doi.org/10.1016/ 0305-0548(93)90006-5.

[26] Choi, B.D. and Kim, J.W. (1997). Discrete-Time $Geo_1, Geo_2/G/1$ Retrial Queueing Systems with Two Types of Calls. *Computers and Mathematics with Applications*, 33(10):79-88.

[27] Das, K. and Samanta, S. K. (2023). Computational analysis of GI [X]/D-MSP (a, b)/1 queueing system via RG-factorization. *Mathematical Methods of Operations Research*, 1-39.https://doi.org/10.1007/s00186-023-00816-1

[28] Das, K. and Samanta, S. K. (2021). Modelling and analysis of D-BMAP/D-MSP/1 queue using RG-factorization. *Quality Technology & Quantitative Management*, 18(3):355-381. https://doi.org/10.1080/16843703.2020.1830477

[29] Demircioglu, M., Bruneel, H. and Wittevrongel, S. (2021) Analysis of a Discrete-Time Queueing Model with Disasters. *Mathematics*, 9(24):3283. https://doi.org/10.3390/math9243283.

[30] Fiems, D., Steyaert, B. and Bruneel, H. (2001). Discrete-time queues with general service times and general server interruptions. *In Internet Quality and Performance and Control of Network Systems*, 4211: 93-104. SPIE. https: //doi.org/10.1117/12.417475.

[31] Fiems, D., Steyaert, B. and Bruneel, H. (2004). Discrete-time queues with generally distributed service times and renewal-type server interruption. *Performance Evaluation*, 55(3-4):277-298. https: //doi.org/10.1016/j.peva.2003.08.004.

[32] Gao, S., Liu, Z. and Dong, H. (2012). A repairable discrete-time retrial queue with recurrent customers, Bernoulli feedback and general retrial times. *Operational Research*, 12(3):367-383.https://doi. org/10.1007/s12351-010-0098-7.

[33] Gao, S., Wang, J. and Van Do, T. (2019). Analysis of a discrete-time repairable queue with disasters and working breakdowns. *RAIRO-Operations Research*, 53(4):1197-1216. https://doi.org/10.1051/ro/2018057.

[34] Gao, P., Wittevrongel, S. and Bruneel, H. (2004). Discrete-time multiserver queues with geometric service times. *Computers & Operations Research*, 31(1):81-99. https://doi-org.egateway.vit.ac.in/10.1016/S0305-0548(02)00155-7

[35] Jeyakumar, S. and Gunasekaran, P. (2017). An analysis of Discrete Queue with Disaster and Single Vacation. *International Journal of Pure and Applied Mathematics*, 113:82–90.

[36] Jolai, F., Asadzadeh, S. M. and Taghizadeh, M. R. (2008). Performance estimation of an email contact center by a finite source discrete time Geo/Geo/1 queue with disasters. *Computers & Industrial Engineering*, 55(3):543-556. https: //doi.org/10.1016/j.cie.2008.01.009.

[37] Jyothsna, K., Vijaya Kumar, P. and Gopala Rao, Ch. (2022). Discrete tiem working vacations queue with impatient clients and congestion dependent service rates. *Reliability: Theory & Applications*, 17(3). URL

[38] Kempa, W. M. (2019). Transient study of a discrete-time queueing model with feedback mechanism. *In AIP Conference Proceedings,AIP Publishing*, 2172(1). https://doi.org/10.1063/1.5133595

[39] Kim, B. and Kim, J. (2009). Tail asymptotics for the queue size distribution in a discrete-time $Geo/G/1$ retrial queue. *Queueing Systems*, 61(2-3):243-254. https://doi.org/10.1007/s11134-009-9106-0.

[40] Kim, B. and Kim, J. (2010). Queue size distribution in a discrete-time $D-BMAP/G/1$ retrial queue. *Computers and Operations Research*, 37(7):1220-1227. https://doi.org/10.1016/j.cor.2009.04.016.

[41] Laevens, K. and Bruneel, H. (1998). Discrete-time mult-iserver queues with priorities. *Performance Evaluation*, 33(4):249-275.https:// doi.org/10.1016/S0166-5316(98)00024-8.

[42] Lan, S. and Tang, Y. (2019). An N-policy discrete-time $Geo/G/1$ queue with modified multiple server vacations and Bernoulli feedback. *RAIRO-Operations Research*, 53(2):367-387. https://doi.org/10.1051/ro/2017027.

[43] Lan, S. and Tang, Y. (2019). Performance analysis of a discrete-time Geo/G/1 retrial queue with non-preemptive priority, working vacations and vacation interruption. *Journal of Industrial & Management Optimization*, 15(3). https://doi.org/10.3934/jimo.2018102

[44] Lan, S. and Tang, Y. (2020). An unreliable discrete-time retrial queue with probabilistic pre-emptive priority, balking customers and replacements of repair times. *AIMS Mathematics*, 5(5):4322-4344. https://doi.org/10. 3934/math.2020276.

[45] Larisa G. Afanasyeva. and Andrey Tkachenko. (2018). Stability of Discrete Multi-Server Queueing Systems with Heterogeneous Servers, Interruptions and Regenerative Input Flow. *Reliability: Theory & Applications*, 13(1).

[46] Lee, S. W., Lee, H. W. and Baek, J. W. (2011). Analysis of discrete-time $Geo/G/1$ queue under the D-policy. *In Proceedings of the 6th international conference on queueing theory and network applications*, 107-115. https://doi.org/10.1145/2021216.2021232.

[47] Lee, D. H. and Yang, W. S. (2013). The N-policy of a discrete time $Geo/G/1$ queue with disasters and its application to wireless sensor networks. *Applied Mathematical Modelling*, 37(23):9722-9731. https://doi.org/10.1016/ j.apm.2013.05.012.

[48] Li, J. and Liu, L. (2016). On the Discrete-Time $Geo/G/1$ Queue with Vacations in Random Environment. *Discrete Dynamics in Nature & Society*. http://dx.doi.org/10.1155/2016/4029415.

[49] Li, J. H. and Tian, N. S. (2007). The discrete-time $GI/Geo/1$ queue with working vacations and vacation interruption. *Applied Mathematics and Computation*, 185(1):1-10. https://doi.org/10.1016/j.amc.2006.07.008.

[50] Li, J. H., Tian, N. S. and Liu, W. Y. (2007). Discrete-time $GI/Geo/1$ queue with multiple working vacations. *Queueing Systems*, 56:53-63. https: //doi.org/10.1007/s11134-007-9030-0.

[51] Li, H. and Yang, T. (1998). $Geo/G/1$ discrete time retrial queue with Bernoulli schedule. *European Journal of Operational Research*, 111(3): 629-649. https://doi. org/10.1016/S0377-2217(97)90357-X.

[52] Lin, C. H. and Ke, J. C. (2011). On the discrete-time system with server breakdowns: Computational algorithm and optimization algorithm. *Applied Mathematics and Computation*, 218(7):3624-3634. https://doi-org.egateway.vit.ac.in/10.1016/j.amc.2011.09.003

[53] Liu, R. and Deng, Z. (2014). On the steady-state system size distribution for a discrete-time $Geo/G/1$ repairable queue. *Discrete Dynamics in Nature and Society*. http://dx.doi.org/10.1155/2014/ 924712

[54] Liu, Z. and Gao, S. (2011). Discrete-time $Geo_1, Geo_2^X/G_1, G_2/1$ retrial queue with two classes of customers and feedback. *Mathematical and Computer Modelling*, 53(5-6):1208-1220. https://doi.org/10.1016/j.mcm.2010.11.090

[55] Madheswari, S. P., Suganthi, P. and Vasanthamani, K. (2021). A Study on $Geo/G/1$ Retrial Queueing System with Starting Failure and Customer Feedback. *Mathematical Modeling and Computation of Real-Time Problems*, 79-98. https://doi.org/10.3934/jimo.2018102

[56] Malik, G. and Upadhyaya, S. (2021). Performance Modelling of a Discrete-Time Retrial Queue with Preferred and Impatient Customers, Bernoulli Vacation and Second Optional Service. I*n Recent Trends in Mathematical Modeling and High Performance Computing: M3HPCST-2020, Ghaziabad, India, Springer International Publishing*, 331-345.https://doi.org/10.1007/978-3-030-68281-1_25

[57] Malik, G. and Upadhyaya, S. (2020). Single vacation policy for discrete-time retrial queue with two types of customers. *Strategic System Assurance and Business Analytics*, 335-349. https://doi.org/10.1007/978-981-15-3647-2_25

[58] Mehmet Ali, M. and Song, X. (2004). A performance analysis of a discrete-time priority queueing system with correlated arrivals. *Performance Evaluation*, 57(3):307-339. https://doi.org/10.1016/j.peva. 2004.01.001.

[59] Mehmet Ali, M., Zhang, X. and Hayes, J. F. (2003). A performance analysis of a discrete-time queueing system with server interruption for modelling wireless ATM multiplexer. *Performance Evaluation*, 51(1):1-31. https://doi.org/10.1016/S0166-5316(02)00087-1.

[60] Miyazawa, T. and Takagi, H. (1994) Advances in discrete-time queues. *Queueing Systems*.

[61] Moreno, P. (2006). A Discrete-time retrial queue with Unreliable Server and General Server Lifetime. *Journal of Mathematical Sciences*, 132:643-655. https: //doi.org/10.1007/s10958-006-0009-x.

[62] Nandi, R. and Samanta, S. K. (2018). Performance Analysis of Shock Models Governed by a Discrete-Time Markovian Arrival Process. *Journal of the Indian Society for Probability and Statistics*, 19:455-468. https://doi.org/10.1007/s41096-018-0053-0

[63] Nobel, R. (2016). Retrial queueing models in discrete time: a short survey of some late arrival models. *Annals of Operations Research*, 247:37-63. https://doi.org/10.1007/s10479-015-1904-7

[64] Nobel, R. and Moreno, P. (2008). A discrete-time retrial queueing model with one server. *European Journal of Operational Research*, 189(3):1088-1103. https://doi.org/10.1016/j.ejor.2007.03.051.

[65] Periyasamy, C. (2013). A Discrete Time Inventory System with Retrial Demands. *International Journal of Engineering Research and Technology*, 2(7):165-170 .

[66] Rajasudha, R. and Arumuganathan, R. (2022). Reliability and Optimum Cost Analysis of Malware Infected System by Discrete-Time Negative Arrival Retrial Queue. *IETE Journal of Research*, 1-20. https://doi.org/10.1080/03772063.2022.2125091

[67] Rajasudha, R., Arumuganathan, R. and Dharmaraja, S. (2022). Performance analysis of discrete-time GeoX/G/1 retrial queue with various vacation policies and impatient customers. *RAIRO-Operations Research*, 56(3):1089-1117. https://doi.org/10.1051/ro/2022042

[68] Samanta, S. K. and Das, K. (2023). Detailed Analytical and Computational Studies of $D-BMAP/D-BMSP/1$ Queueing System. *Methodology and Computing in Applied Probability*, 25(1):12. https://doi.org/10.1007/s11009-023-10012-7

[69] Samanta, S. K. and Nandi, R. (2021). Analysis of $GI^{[X]}/D\text{-}MSP/1/\infty$ queue using *RG*-factorization. *Journal of Industrial and Management Optimization*, 17(2):549-573. https://doi.org/10.3934/jimo.2019123

[70] Samanta, S. K. and Nandi, R. (2020). Performance analysis of the GI/D-MSP/1 queue with N-policy and its optimal control. *Quality Technology & Quantitative Management*, 17(4):399-422. https://doi.org/10.1080/16843703.2019.1662217

[71] Sudhesh, R. and Priya, R. S. (2019). An analysis of discrete-time Geo/Geo/1 queue with feedback, repair and disaster. *International Journal of Operational Research*, 35(1):37-53. https://doi.org/10.1504/IJOR.2019.

[72] Takagi, H. (1993). Queueing Analysis: A Foundation of Performance Evaluation in Discrete-time Systems, *Vol.3, North-Holland, Amsterdam*, (1993).

[73] Tang, Y., Yun, X. and Huang, S. (2008). Discrete-time $Geo^{[X]}/G/1$ queue with unreliable server and multiple adaptive delayed vacations. *Journal of Computational and Applied Mathematics*, 220(1-2):439-455. https: //doi.org/10.1016/j.cam.2007.08.019

[74] Tian, N. and Zhang, Z. G. (2002). The discrete-time GI/Geo/1 queue with multiple vacations. *Queueing systems*, 40:283-294. https://doi. org/10.1023/A:1014711529740

[75] Upadhyaya, S. (2015). Working vacation policy for a discrete-time $Geo^{[X]}/Geo/1$ retrial queue. *Opsearch*, 52(4):650-669. https://doi.org/10. 1007/s12597-015-0200-2.

[76] Upadhyaya, S. (2016). Performance prediction of a discrete-time batch arrival retrial queue with Bernoulli feedback. *Applied Mathematics and Computation*, 283:108-119. https://doi.org/10.1016/j.amc.2016.02.026.

[77] Upadhyaya, S. (2020). Cost optimisation of a discrete-time retrial queue with Bernoulli feedback and starting failure. *International Journal of Industrial and Systems Engineering*, 36(2):165-196. https://doi.org/10.1504/IJISE.2020.110245

[78] Upadhyaya, S. (2021). Performance Analysis of a Discrete-Time Retrial Queue with Bernoulli Feedback, Starting Failure and Single Vacation Policy. *In Recent Trends in Mathematical Modeling and High Performance Computing: M3HPCST-2020, Ghaziabad, India, Springer International Publishing*, 365-379. https://doi.org/10.1007/978-3-030-68281-1_27

[79] Upadhyaya, S., Malik, G. and Sharma, R. (2022). Neuro-fuzzy computing and optimisation results for batch discrete time retrial queue. *International Journal of Mathematics in Operational Research*, 23(1):119-146. https://doi.org/10.1504/IJMOR.2022.126049

[80] Vadivu, A. S., Arumuganathan, R. and Kumar, M. S. (2016). Analysis of discrete-time queues with correlated arrivals, negative customers and server interruption. *RAIRO-Operations Research*, 50(1):67-81. https://doi.org/ 10.1051/ro/2015012.

[81] Vaishnawi, M., Upadhyaya, S. and Kulshrestha, R. (2022). Optimal Cost Analysis for Discrete-Time Recurrent Queue with Bernoulli Feedback and Emergency Vacation. *International Journal of Applied and Computational Mathematics*, 8(5):254. https://doi.org/10.1007/s40819-022-01445-8

[82] Verdonck, F., Bruneel, H. and Wittevrongel, S. (2019). Analysis of a 2-state discrete-time queue with stochastic state-period lengths and state-dependent server availability and arrivals. *Performance Evaluation*, 135:102026. https://doi.org/10.1016/j.peva.2019.102026.

[83] Wang, J. (2012). Discrete-time $Geo/G/1$ retrial queues with general retrial time and Bernoulli vacation. *Journal of Systems Science and Complexity*, 25:504-513. https://doi.org/10.1007/s11424-012-0254-7.

[84] Wang, J., Huang, Y. and Dai, Z. (2011). A discrete-time on–off source queueing system with negative customers. *Computers & Industrial Engineering*, 61(4):1226-1232. https://doi.org/10.1016/j.cie.2011.07.013.

[85] Wang, J., Huang, Y. and Van Do, T. (2013). A single-server discrete-time queue with correlated positive and negative customer arrivals. *Applied Mathematical Modelling*, 37(9):6212-6224. https://doi.org/10.1016/j.apm. 2012.12.021.

[86] Wang, J. T., Wang, N. and Alfa, A. S. (2013). Discrete-time $GI/G/1$ retrial queues with time-controlled vacation policies. *Acta Mathematicae Applicatae Sinica, English Series*, 29(4):689-704. https://doi.org/ 10.1007/s10255-013-0244-0.

[87] Wang, J. T. and Zhang, P. (2009). A single-server discrete-time retrial G-queue with server breakdowns and repairs. *Acta Mathematicae Applicatae Sinica, English Series*, 25(4):675-684. https://doi.org/10.1007/ s10255-008-8823-1.

[88] Wang, J. and Zhao, Q. (2007). Discrete-time $Geo/G/1$ retrial queue with general retrial times and starting failures. *Mathematical and Computer Modelling*, 45(7-8):853-863. https://doi.org/10.1016/j.mcm.2006.08.005.

[89] Wang, J. and Zhao, Q. (2007). A discrete-time $Geo/G/1$ retrial queue with starting failures and second optional service. *Computers & Mathematics with Applications*, 53(1):115-127. https://doi.org/10.1016/j.camwa.2006.10.024.

[90] Wang, Y. C., Chou, J. H. and Wang, S. Y. (2011). Loss pattern of $DBMAP/DMSP/1/K$ queue and its application in wireless local communications. *Applied Mathematical Modelling*, 35(4):1782-1797.https://doi-org.egateway.vit.ac.in/10.1016/j.apm.2010.10.009

[91] Wang, T. Y. and Ke, J. C. (2009). The randomized threshold for the discrete-time $Geo/G/1$ queue. A*pplied Mathematical Modelling*, 33(7):3178-3185. https://doi.org/10.1016/j.apm.2008.10.010.

[92] Wang, T. Y., Ke, J. C. and Chang, F. M. (2011). On the discrete-time $Geo/G/1$ queue with randomized vacations and at most J vacations. *Applied Mathematical Modelling*, 35(5):2297-2308. https://doi.org/10.1016/j.apm. 2010.11.021.

[93] Wang, J. and Zhang, P. (2009). A discrete-time retrial queue with negative customers and unreliable server. *Computers & Industrial Engineering*, 56(4):1216-1222. https://doi.org/10.1016/j.cie.2008.07.010.

[94] Woodward, M.E. (1994). Communication and Computer Networks: Modelling with Discrete-Time Queues, *IEEE Computer Society, London*.

[95] Wu, J., Liu, Z. and Peng, Y. (2011). A discrete-time Geo/G/1 retrial queue with pre-emptive resume and collisions. *Applied Mathematical Modelling*, 35(2):837-847. https://doi.org/10.1016/j.apm.2010.07.039.

[96] Wu, J., Wang, J. and Liu, Z. (2013). A discrete-time $Geo/G/1$ retrial queue with preferred and impatient customers. *Applied Mathematical Modelling*, 37(4):2552-2561. https://doi.org/10.1016/j.apm.2012.06.011.

[97] Yang, T. and Li, H. (1995). On the steady-state queue size distribution of the discrete-time $Geo/G/1$ queue with repeated customers. *Queueing systems*, 21:199-215. https://doi.org/10.1007/BF01158581.

[98] Yu, X., Hu, L. and Ma, M. (2023). Reliability measures of discrete time k-out-of-n: G retrial systems based on Bernoulli shocks. *Reliability Engineering & System Safety*, 109491.https://doi.org/10.1016/j.ress.2023.109491

[99] Yu, M. and Alfa, A. S. (2015). Algorithm for computing the queue length distribution at various time epochs in $DMAP/G^{(1,a,b)}/1/N$ queue with batch-size-dependent service time.

European Journal of Operational Research, 244(1):227-239. https://doi.org/10.1016/j.ejor.2015.01.056

[100] Yue, D. and Zhang, F. (2013). A discrete-time $Geo/G/1$ retrial queue with J-vacation policy and general retrial times. *Journal of Systems Science and Complexity*, 26:556-571. https://doi.org/10.1007/s11424-013-1121-x.

[101] Zhang, Z. G. and Tian, N. (2001). Discrete time $Geo/G/1$ queue with multiple adaptive vacations. *Queueing Systems*, 38: 419-429. https://doi. org/10.1023/A:1010947911863.

[102] Zhang, F. and Zhu, Z. (2013) A Discrete-Time $Geo/G/1$ Retrial Queue with J-Vacations and Two Types of Breakdowns. *Journal of Applied Mathematics*. http://dx.doi.org/10.1155/2013/834731.

[103] Zhang, F. and Zhu, Z. (2015). A Discrete-Time $Geo/G/1$ Retrial Queue with Two Different Types of Vacations. *Mathematical Problems in Engineering*, http://dx.doi.org/10.1155/2015/835158.

QUANTILE RESIDUAL ENTROPY FOR SOME LIFE TIME DISTRIBUTIONS

Javid Gani Dar[1], Mohammad Younus Bhat[2], Shahid Tamboli[3], Shaikh Sarfaraj[4], Aafaq A. Rather[5,*], Maryam Mohiuddin[6], Showkat Ahmad Dar[7]

•

[1]Department of Applied Sciences, Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Lavale, Pune, Maharashtra, India
[2]Department of Mathematical Sciences, Islamic University of Science and Technology, Kashmir, India
[3,4]Mechanical Engineering Department, Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Lavale, Pune, Maharashtra, India
[5,*]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[6]Department of Mathematics, National Institute of Technology, Calicut, Kerala, India
[7]Department of Statistics, University of Kashmir, J&K, India
[1]javid.dar@sitpune.edu.in, [2]younis.bat@islamicuniversity.edu.in, [3]shahidt@sitpune.edu.in, [4]sarfarajs@sitpune.edu.in, [5,*]aafaq7741@gmail.com, [6]masmariam7@gmail.com, [7]darshowkat2429@gmail.com

Abstract

This study explores the concept of residual entropy as an alternative approach to traditional entropy measures. The field of information theory, built upon Shannon's entropy, has been instrumental in understanding the dynamics of systems. However, existing literature has recognized the limitations of applying traditional entropy measures to systems that have already been in existence for a certain duration. This study delves into the concept of residual entropy, acknowledging the need for a more suitable approach for such systems. Specifically, we investigate the characteristics of residual entropy using a quantile-based framework. By deriving the quantile residual entropy function for various lifetime models, we gain insights into the reordering and ageing phenomena captured by the quantile version of the residual entropy equation. Our findings contribute to an enhanced understanding of residual entropy and provide a novel perspective on analyzing and interpreting the behavior of established systems.

Keywords: Shannon entropy, Residual entropy, Quantile function, Residual entropy.

1. Introduction

Physicists first created the idea of entropy in the backdrop of the equilibrium thermodynamics, which was then broadened with the creation of statistical mechanics. Claude Shannon devised a mathematical notion called entropy to describe the stochastic nature of missing data in phone-line transmissions [19]. According to the concept of differential entropy for a continuous random variable X having density functions $f_X(x)$ is given by

$$H(X) = -\int_0^\infty f_X(x) \log f_X(x) dx \quad (1)$$

In order to generalize the Shannon entropy evaluation, Mathi and Haubold [12] created a new generalized entropy evaluation linked to a random variable X having pdf $f(x)$, is expressed as

$$M_\alpha(X) = \frac{1}{\alpha-1}\left(\int_{-\infty}^{\infty}[f(x)]^{2-\alpha}dx - 1\right), 0 < \alpha < 2, \alpha \neq 1 \quad (2)$$

When $\alpha \to 1$ (2) goes to the Shannon entropy measure defined in (1).

One feature of (2), as described in [11] and [20], is that combining the maximal entropy concept with normalization and power limitations prompts the widely recognized route model [12]. It should be highlighted that the route model incorporates several well-known statistical distributions as special instances.

The date is frequently abbreviated in reliability as well as life evaluation, thus equation (1) isn't an acceptable metric in such cases. Thus, if there is data about the present age, that may be utilized for determining its degree of unpredictability, Shannon's entropy is inapplicable. Ebrahimi [7] proposes a more realistic method that takes ageing into consideration and is characterized as

$$H(X;t) = -\int_t^{\infty}\frac{f(x)}{\bar{F}(t)}\log\frac{f(x)}{\bar{F}(t)}dx \quad (3)$$

where $\bar{F}(t)$ is the survival function. For $t = 0$, (3) reduces to (1)

The remainder of the Mathai-Haubold (Dar and Al-Zahrani, [5]) entropy functional can be expressed for a positive random parameter X, which represents the lifespan of a unit at time t. It is given by equation (4)

$$M_\alpha(X;t) = \frac{1}{\alpha-1}\left\{\int_t^{\infty}\frac{f^{2-\alpha}(x)}{\bar{F}^{2-\alpha}(t)}dx - 1\right\}, 0 < \alpha < 2, \alpha \neq 1 \quad (4)$$

Any theoretical studies and implementations that utilize the aforementioned metrics depend upon the distribution function, however they may not be appropriate in cases when the distributions aren't analytically tractable. A different technique to investigate is to employ quantile functions, which are specified by

$$Q(u) = F^{-1}(u) = inf\{x/F(x) \geq u\}, 0 \leq u \leq 1 \quad (5)$$

Gilchrist [9] has presented an alternative to the distribution function in statistical data analysis and modeling known as the quantile function (QF). QF is often favored since it provides a straightforward analysis with less influence from extreme observations. To learn more about the properties and usefulness of QF in identifying models, Nair and Sankaran [14], Sunoj et al [17], and related sources offer detailed studies. Researchers have become increasingly interested in using quantile-based entropy measures as an alternative approach to measuring the uncertainty of a random variable. These measures possess unique properties compared to the distribution function approach. Sunoj and Sankaran [17] have recently explored the quantile version of Shannon entropy and its residual form as

$$H = \int_0^1 \log q(p)\, dp \quad (6)$$

and

$$H(u) = H\big(X; Q(u)\big) = \log(1-u) + \frac{1}{1-u}\int_u^1 \log q(p)dp \quad (7)$$

respectively, where the quantile density function $q(u) = \frac{dQ(u)}{du}$ can be defined by $fQ(u) = f(Q(u))$ and Using these definitions, equation (8) is expressed as,

$$q(u)f\big(Q(u)\big) = 1 \quad (8)$$

Further, Nanda et.al. [13] introduced quantile based Renyi's entropy function and study properties and applications of the proposed measure.

In this paper, we introduce a novel quantile-based version of the generalized residual entropy function and examine its key properties. We first demonstrate that the proposed measure provides a unique determination of the quantile distribution functions and derive an entropy function for specific quintile functions that lack an explicit form for distribution functions. Additionally, we define ordering and aging properties based on the quantile residual entropy function and analyze their characteristics. Lastly, we present a characterization of certain lifetime models that are valuable in analyzing lifetime data.

This paper is organized as follows. Section 2 outlines the development of our new quantile-based residual entropy measure and explores its properties. Section 3 delves into the aging and

ordering properties of the quantile-based residual entropy. Finally, Section 4 presents a set of characterization theorems based on the quantile residual entropy measure.

2. Quantile based M-H Residual Entropy

From equation (5), we can see that for a continuous distribution function F, the composite function F(Q(u)) is equal to u, represented as FQ(u). We can define the density quantile function as *fQ(u) = f(Q(u))* and the quantile density function as $q(u) = Q'(u)$ where the prime notation indicates differentiation. Equation (4) and (8) together yield the expression for the quantile version of the M-H residual entropy.

$$M_\alpha(X;Q(u)) = \frac{1}{\alpha-1}\left\{\frac{\int_u^1 (fQ(p))^{2-\alpha} q(p)dp}{(1-u)^{2-\alpha}} - 1\right\}, 0 < \alpha < 2, \alpha \neq 1 \quad (9)$$

$$M_\alpha(X;Q(u)) \quad = \frac{1}{\alpha-1}\left\{\frac{\int_u^1 q^{\alpha-1}(p)dp}{(1-u)^{2-\alpha}} - 1\right\} \quad (10)$$

The expression (10) is known as the quantile M-H residual entropy, which quantifies the average level of uncertainty in the conditional density with respect to predicting an outcome of X up until the 100 (1-u)% point of the distribution. Following theorem shows the uniqueness of quantile version of residual entropy.

Theorem 2.1: Show that the quantile version of residual entropy determines the underlying distribution uniquely.

Proof: Differentiating (10) with respect to u, we get

$$\big(q(u)\big)^{\alpha-1} = (1-u)^{1-\alpha}\begin{bmatrix}(2-\alpha)\{(\alpha-1)M_\alpha(X;Q(u)) + 1\}\\ -(1-u)(\alpha-1)M'_\alpha(X;Q(u))\end{bmatrix} \quad (11)$$

The equation establishes a clear connection between the quantile density function q(u) and $M^\alpha(X;Q(u))$, which shows that quantile version of M-H residual entropy determine the underlying distribution uniquely. The next two theorems gives the bounds of $M_\alpha(X;Q(u))$. The proof follows by using (10).

Theorem 2.2: If $M^\alpha(X;Q(u))$ is increasing u, then

$$q(u) \leq (\geq) \frac{1}{1-u}\big[(\alpha-2)\{(\alpha-1)M_\alpha(X;Q(u)) + 1\}\big]^{\frac{1}{1-\alpha}} for\ 0 < \alpha < 1\ (1 < \alpha < 2) \quad (12)$$

If $M^\alpha(X;Q(u))$ is deccreasing u, then

$$q(u) \geq (\leq) \frac{1}{1-u}\big[(\alpha-2)\{(\alpha-1)M_\alpha(X;Q(u)) + 1\}\big]^{\frac{1}{1-\alpha}} for\ 0 < \alpha < 1 (1 < \alpha < 2) \quad (13)$$

In the following section, we will derive the quantile version of the residual entropy for several lifetime models.

2.1 Govindarajulu's Distribution

Firstly, we will consider Govindarajulu's distribution, for which the quantile function and quantile density function are given by:

$$Q(u) = a\{(b+1)u^b - bu^{b+1}\} \text{ and } q(u) = ab(b+1)(1-u)u^{b-1}, 0 \leq u \leq 1; a, b > 0.$$

Quantile based residual M-H Entropy of r^{th} for Govindarajulu distribution as

$$M^\alpha(X;Q(u)) = \frac{1}{\alpha-1}\left\{\frac{(ab)^{\alpha-1}(b+1)^{\alpha-1}\beta_u((b-1)(\alpha-1)+1;\alpha)}{(1-u)^{\alpha-1}} - 1\right\} \quad (14)$$

where $\beta_u(a;b)$ is an incomplete beta function.

2.2 Uniform Distribution

$$Q(u) = a + (b-a)u \text{ and } q(u) = (b-a), 0 \leq u \leq 1; a < b.$$

$$M^{\alpha}(X;Q(u)) = \frac{1}{\alpha-1}\{(b-a)^{\alpha-1}(1-u)^{\alpha-1}-1\} \tag{15}$$

2.3 Pareto-I Distribution

$$Q(u) = b\left\{(1-u)^{-1/a}\right\} \text{ and } q(u) = \frac{b}{a}\left\{(1-u)^{-\left(1+\frac{1}{a}\right)}\right\}, 0 \le u \le 1; a, b > 0.$$
$$M^{\alpha}(X, Q(u) = \frac{1}{\alpha-1}\left\{\left(\frac{b}{a}\right)^{\alpha-1}\frac{a(1-u)^{\frac{1}{a}(1-\alpha)}}{a(2-\alpha)+(1-\alpha)}-1\right\} \tag{16}$$

2.4 Generalized Pareto Distribution

$$Q(u) = \frac{b}{a}\left\{(1-u)^{-a/a+1}-1\right\} \text{ and } q(u) = \frac{b}{a+1}\left\{(1-u)^{-\left(\frac{a}{a+1}+1\right)}\right\}, 0 \le u \le 1; a, b > 0.$$
$$M^{\alpha}(X, Q(u) = \frac{1}{\alpha-1}\left\{\left(\frac{b}{a+1}\right)^{\alpha-1}\frac{(1+a)(1-u)^{\frac{a}{a+1}(1-\alpha)}}{(1+a)(2-\alpha)+a(1-\alpha)}-1\right\} \tag{17}$$

2.5 Re-Scaled Beta Distribution

$$Q(u) = R\left\{1-(1-u)^{1/c}\right\} \text{ and } q(u) = \frac{R}{c}\left\{(1-u)^{\frac{1}{c}-1}\right\}, 0 \le u \le 1; a, b > 0.$$
$$M^{\alpha}(X, Q(u) = \frac{1}{\alpha-1}\left\{\left(\frac{R}{c}\right)^{\alpha-1}\frac{c(1-u)^{\frac{1}{c}(\alpha-1)}}{c(2-\alpha)+(\alpha-1)}-1\right\} \tag{18}$$

2.6 Exponential distribution

$$Q(u) = -\frac{\log(1-u)}{\lambda} \text{ and } q(u) = \frac{1}{\lambda(1-u)}, 0 \le u \le 1; \lambda > 0.$$
$$M^{\alpha}(X, Q(u) = \frac{1}{\alpha-1}\left\{\frac{1}{(2-\alpha)\lambda^{\alpha-1}}-1\right\} \tag{19}$$

2.7 Power distribution

$$Q(u) = au^{\frac{1}{c}} \text{ and } q(u) = \frac{a}{c}u^{\frac{1}{c}-1}, 0 \le u \le 1; \lambda > 0$$
$$M^{\alpha}(X, Q(u)) = \frac{1}{\alpha-1}\left\{\left(\frac{a}{c}\right)^{\alpha-1}\frac{c}{c(2-\alpha)+(\alpha-1)}\left[\frac{1-u^{(2-\alpha)+\frac{1}{c}(\alpha-1)}}{(1-u)^{2-\alpha}}\right]-1\right\} \tag{20}$$

Following tables and figures gives the computations and comparison of quantile of M-H residual entropy respectively for different parameters for some life time distributions

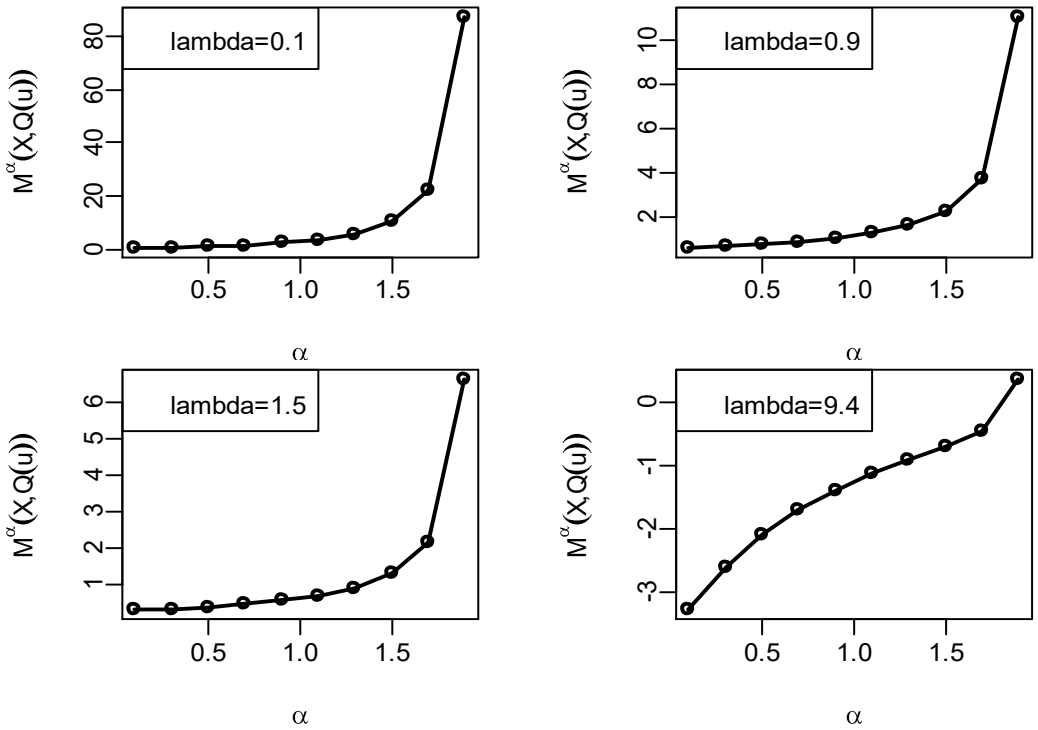
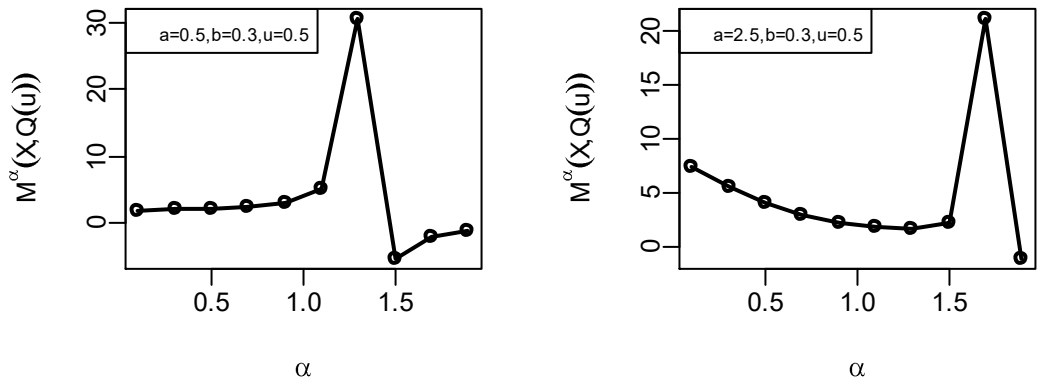


Fig. 1: *Quantile residual entropy plot of exponential distribution*

Table 1: *Qunitile M-H residual entropy values exponential distribution*

Parameter Values		λ					
		0.1	**0.6**	**0.9**	**1.5**	**3.5**	**9.4**
α	**0.1**	1.03749	0.74184	0.57922	0.26877	-0.69467	-3.28247
	0.3	1.26090	0.84087	0.64798	0.31244	-0.59119	-2.60456
	0.5	1.57836	0.96720	0.73509	0.36701	-0.49444	-2.08792
	0.7	2.04824	1.13355	0.84901	0.43757	-0.40051	-1.68863
	0.9	2.77883	1.36182	1.00437	0.53291	-0.30420	-1.37418
	1.1	3.98806	1.69344	1.22880	0.66961	-0.19719	-1.11935
	1.3	6.16792	2.21721	1.58149	0.88318	-0.06324	-0.90201
	1.5	10.64911	3.16398	2.21637	1.26599	0.13809	-0.69534
	1.7	22.43749	5.38030	3.69781	2.15666	0.55265	-0.43639
	1.9	87.14758	16.48519	11.10518	6.60281	2.48718	0.36780



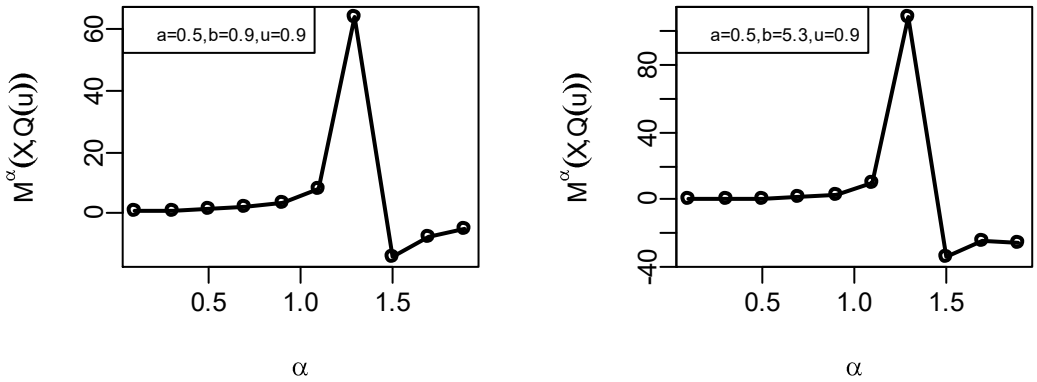


Fig. 2: *Qunitile residual entropy plot for Pareto I Distribution*

Table-2: *Qunitile residual entropy values for Pareto I Distribution*

Parameters		a=0.5, b=0.3, u=0.5	a=1.4, b=0.3, u=0.5	a=2.5, b=0.3, u=0.5	a=0.5, b=0.9, u=0.9	a=0.5, b=2.3, u=0.9	a=5.3, b=0.3, u=0.9
α	**0.1**	1.759	4.442	7.484	0.655	0.281	0.133
	0.3	1.912	3.821	5.670	0.945	0.490	0.273
	0.5	2.066	3.157	4.075	1.431	0.895	0.590
	0.7	2.366	2.695	3.002	2.246	1.695	1.320
	0.9	3.029	2.445	2.322	3.813	3.472	3.194
	1.1	4.955	2.446	1.929	8.073	8.867	9.639
	1.3	30.698	2.973	1.805	63.668	84.366	108.375
	1.5	-5.453	7.239	2.217	-14.245	-22.773	-34.569
	1.7	-2.207	-3.718	21.216	-7.214	-13.912	-24.956
	1.9	-1.275	-0.988	-1.045	-5.188	-12.070	-25.585

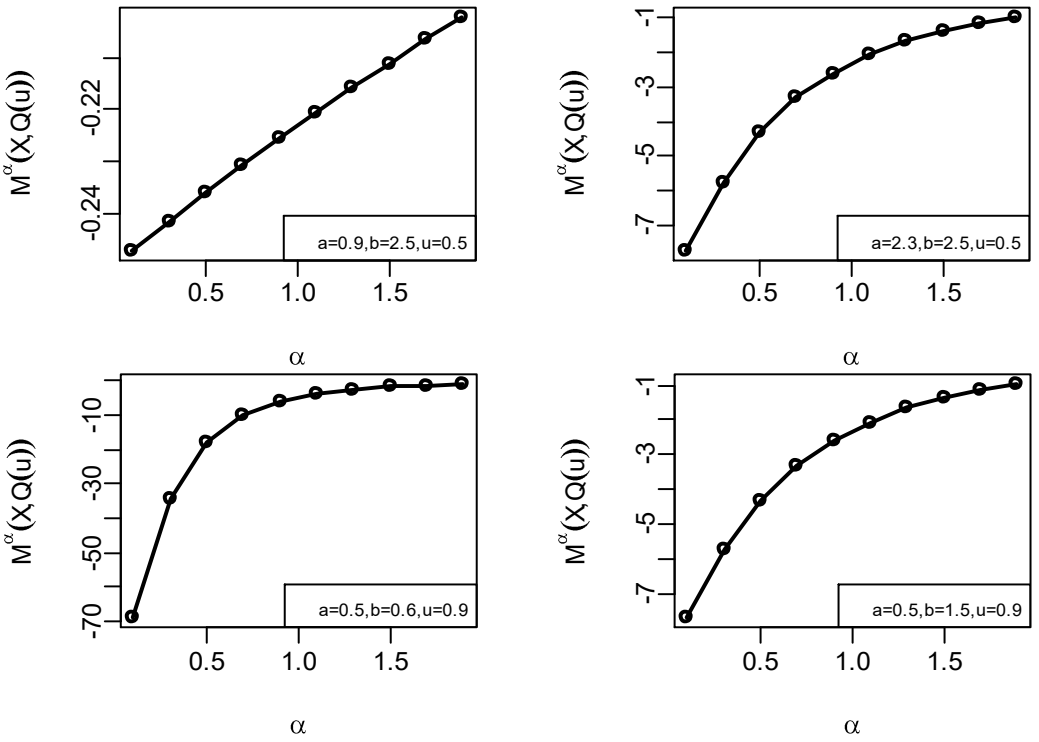


Fig. 3: *Qunitile residual entropy plot for uniform distribution*

Table-3: *Qunitile residual entropy values for uniform distribution*

Parameters		a=0.9, b=2.5, u=0.5	a=1.4, b=2.5, u=0.5	a=2.3, b=2.5, u=0.5	a=0.5, b=0.6, u=0.9	a=0.5, b=0.9, u=0.9	a=0.5, b=1.5, u=0.9
α	**0.1**	-0.247	-0.792	-7.715	-68.995	-19.022	-7.715
	0.3	-0.242	-0.742	-5.731	-34.456	-12.169	-5.731
	0.5	-0.236	-0.697	-4.325	-18.000	-8.000	-4.325
	0.7	-0.231	-0.655	-3.318	-9.937	-5.422	-3.318
	0.9	-0.226	-0.616	-2.589	-5.849	-3.797	-2.589
	1.1	-0.221	-0.580	-2.057	-3.690	-2.752	-2.057
	1.3	-0.216	-0.547	-1.663	-2.496	-2.064	-1.663
	1.5	-0.211	-0.517	-1.368	-1.800	-1.600	-1.368
	1.7	-0.207	-0.489	-1.144	-1.372	-1.278	-1.144
	1.9	-0.202	-0.462	-0.971	-1.094	-1.050	-0.971

The quantile residual entropy plots for the exponential, Pareto-I, and uniform distributions are shown in Figs. 1, 2, and 3. In the case of an exponential distribution, we have an increasing entropy plot for increasing values of the parameters and. For various parameter combinations, the entropy plot under the Pareto-I distribution exhibits both increasing and decreasing behavior. For various parameter values, the entropy plot under a uniform distribution also exhibits an increasing trend. Entropy values for exponential, Pareto-I, and uniform distributions are shown in Tables 1, 2, and 3, which depicts the same behavior as mentioned for graphical displays.

3. Ageing and Ordering Properties of Quantile based M-H Residual Entropy

These nonparametric classes of life distribution are defined using residual M-H quantile entropy. Definition increasing (decreasing) M-H residual quantile entropy (IMHRQE) is claimed to exist for X. (DMHRQE) if $M^{\alpha}(X, Q(u))$ is increasing (decreasing) in $u \geq 0$.

Definition 3.2: If $\left(X \leq_{MHQE} Y\right)$, then X is smaller than Y in the M-H quantile entropy order.

$$M_{\alpha}^{X}(Q_X(u)) \leq M_{\alpha}^{Y}(Q_Y(u)) \text{ for all } u \in [0,1] \tag{21}$$

Definition 3.3: If $\left(X \leq_{MHQFR} Y\right)$, then X is smaller than Y in the M-H quantile failure rate.

$$H_X(u) \geq H_Y \quad \forall\, u \geq 0$$

The following lemma is useful in proving the results in monotonicity of $M^{\alpha}(X, Q(u))$.

Let $f(u,x): R_+^2 \to R_+$ and $g: R_+ \to R_+$ be any two functions, according to Lemma 3.1 Given that the integrals exist, if $_\int_u^{\infty} f(u,x)dx$ is increasing and g(u) is increasing (decreasing) in u, then $\int_u^{\infty} f(u,x)g(x)dx$ is increasing (decreasing) in u.

Theorem 3.1: Consider that X is a continuous random variable that is non-negative and contains the quantile function $Q_X(.)$ and density function $q_X(.)$. Define $Y = \emptyset(X)$, as a nonnegative, where $\emptyset(.)$ increasing, convex(concave) function. (i) Because whenever $M^{\alpha}(X, Q(u))$ is increasing (decreasing) in u, For $1 < \alpha < 2$, $M^{\alpha}(X, Q(u))$ is increasing (decreasing) in u as well. (ii) Because whenever$M^{\alpha}(X, Q(u))$ is increasing (decreasing) in u, For $0 < \alpha < 1$, $M^{\alpha}(X, Q(u))$ is increasing (decreasing) in u as well.

Proof: (i) Using the stated condition, we may infer that $M^{\alpha}(X, Q(u) = \frac{1}{\alpha-1}\left[\frac{\int_u^1 q_X^{\alpha-1}(p)dp}{(1-u)^{2-\alpha}} - 1\right]$ increasing (decreasing) in u and denote $q_X(.)$ as the quantile density function of X.

According to definition 11

$$M_{\alpha}^{Y}(Y,Q_Y(u)) = \frac{1}{\alpha-1}\left[\frac{\int_u^1 q_Y^{\alpha-1}(p)dp}{(1-u)^{2-\alpha}} - 1\right] \quad (22)$$

$$M_{\alpha}^{Y}(Y,Q_Y(u)) = \frac{1}{\alpha-1}\left[\frac{\int_u^1 [q_X(p)\emptyset'(Q_X(p))]^{\alpha-1}}{(1-u)^{2-\alpha}} - 1\right] \quad (23)$$

$$M_{\alpha}^{Y}(Y,Q_Y(u)) = \frac{1}{\alpha-1}\left[\frac{\int_u^1 (q_X(p))^{\alpha-1}(\emptyset'(Q_X(p)))^{\alpha-1}}{(1-u)^{2-\alpha}} - 1\right] \quad (24)$$

$(\emptyset'(Q_X(p))^{\alpha-1}$ is increasing (decreasing) and is non-negative, therefore by Lemma 3.1, (3.1) increasing (decreasing), which proves (i) of the Theorem. This is because $1 < \alpha < 2$ and $\emptyset$ is non-negative, increasing (decreasing), and convex(concave) function.

Similarly, when $0 < \alpha < 1$, $(\emptyset'(Q_X(p))^{\alpha-1}$ increasing (decreasing) in p because is increasing and convex. As a result, theorem (3.1) is increasing (decreasing) in u, demonstrating the second part of the theorem. The preceding Theorem is immediately applied as follows.

Assume $Y = X^{\frac{1}{\alpha}}, \alpha > 0.$ as well as X has an exponential distribution with a failure rate of. The Weibull distribution with $Q(u) = \lambda^{-\frac{1}{\alpha}}(-\log(1-u))^{\frac{1}{\alpha}}.$ becomes then shown in Y. In the occurrence of $1 < \alpha < 2, (0 < \alpha < 1).$, the non-negative increasing function $1 < \alpha < 2, (0 < \alpha < 1).$ is convex (concave). The Weibull distribution is hence increasing (decreasing) M-H quantile residual entropy if EE, according to Theorem 3.1: We now present a lemma that indicates the closure of the MHQE order under increasing convex transformation.

Let $f(u,x)$: $[0,1] \times R_+ \to R_+$ be such that $\int_u^1 f(u,x)dx \geq 0, \forall\, u \in [0,1]$ and g(x) be any non-negative function in x, then prove Lemma 3.2. $\int_u^1 f(u,x)g(x)dx \geq 0, \forall\, u \in [0,1]$ Theorem 3.2 states that if $X \leq_{MHQE} Y.$, consequently let X and Y be two random variables. Following that, for any non-negative increasing convex function $\emptyset$, $\emptyset(X) \leq_{MHQE} \emptyset(Y)$.

Proof: To show that $\emptyset(X) \leq_{MHQE} \emptyset(Y)$, it is enough to show that

$$\frac{1}{\alpha-1}\left[\frac{\int_u^1 (q_X(p))^{\alpha-1}\left(\emptyset'(\emptyset_X(p))\right)^{\alpha-1}dp}{(1-u)^{2-\alpha}} - 1\right] \leq \frac{1}{\alpha-1}\left[\frac{\int_u^1 (q_Y(p))^{\alpha-1}\left(\emptyset'(\emptyset_Y(p))\right)^{\alpha-1}dp}{(1-u)^{2-\alpha}} - 1\right] \quad \forall\, u \in [0,1] \quad (25)$$

Two cases arise:

(i) Consider the case when $1 < \alpha < 2$. Since $X \leq_{MHQE} Y$, we have $\forall\, u \in [0,1]$

$$\frac{1}{\alpha-1}\left[\frac{\int_u^1 (q_X(p))^{\alpha-1}dp}{(1-u)^{2-\alpha}} - 1\right] \leq \frac{1}{\alpha-1}\left[\frac{\int_u^1 (q_Y(p))^{\alpha-1}}{(1-u)^{2-\alpha}} - 1\right] \quad (26)$$

which is equalvient to

$$\int_u^1 (q_X(p))^{\alpha-1}dp \leq \int_u^1 (q_Y(p))^{\alpha-1}dp \quad (27)$$

Thus, $\forall\, u \in [0,1]$ $q_X(u) \leq q_Y(u)$ and consequently $Q_X(u) \leq Q_Y(u)$. have been deduced from (27). Meanwhile, $\emptyset'(\emptyset_X(u)) \leq \emptyset'(\emptyset_Y(u))$ is caused by $\emptyset'(.)$ increasing in u because (.) is convex. Thus inequality (12) follows from (13) and Lemma 3.2.

(ii) Consider the case when $0 < \alpha < 1$.

The proof develops in a way that's similar to the rest of the case (i).

Theorem 3.3: If $X \leq_{MHQFR} Y$, then $X \leq_{MHQE} Y$.

Proof: It is simple and hence omitted.

4. Characterization Theorems

The hazard quantile function, which is a simplification of the well-known hazard function and is valuable in reliability analysis, is a significant quantile measure.

$$H(u) = h\big(Q(u)\big) = \frac{fQ(u)}{(1-u)} = \frac{1}{(1-u)q(u)} \quad (28)$$

Next we state some characterization results of some well-known life time's distribution based on

quantile residual entropy.

Theorem 4.1: Let X represent a random variable with an M-H residual entropy of the following form:

$$M^{\alpha}(X;Q(u)) = \frac{1}{\alpha-1}\left\{C\left(H(u)\right)^{1-\alpha} - 1\right\}, 0 < \alpha < 2, \alpha \neq 1 \quad (29)$$

only if and when

$C = \frac{1}{2-\alpha}$, then X has an exponential distribution.

With a quantile density function, X has a Pareto distribution.

If $C < \frac{1}{2-\alpha}. \; then \; q(u) = \frac{b}{a}\left\{(1-u)^{-\left(1+\frac{1}{a}\right)}\right\}, 0 \leq u \leq 1; a, b > 0$

If $C > \frac{1}{2-\alpha}$., X has a quantile density function of $q(u) = \frac{b}{a}\left\{(1-u)^{\left(\frac{1}{a}-1\right)}\right\}, 0 \leq u \leq 1; \; b > 0, a > 1$ with a finite range distribution. X has uniform distribution if $C = 1$.

Proof: From section 2, the necessary part follows.

For converse part, let (29) is true, then using (9), we have

$$\int_u^1 q^{\alpha-1}(p)dp = C\left(H(u)\right)^{1-\alpha}(1-u)^{2-\alpha} \quad (30)$$

Now using $H(u) = \frac{1}{(1-u)q(u)}$ in (30), subsequently separating it on both sides with regard to u, we obtain

$$\frac{q'(u)}{q(u)} = \left(\frac{c-1}{c\alpha-c}\right)\left(\frac{1}{1-u}\right) \quad (31)$$

This gives

$$q(u) = A(1-u)^{\left(\frac{c-1}{c-c\alpha}\right)} \quad (32)$$

where A remains constant.

As a result, if $C = \frac{1}{2-\alpha}, C < \frac{1}{2-\alpha}, C > \frac{1}{2-\alpha}, C = 1$, X has distributions that are respectively exponential, Pareto, and finite range.

5. Conclusion

This study has shed light on the concept of residual entropy and its relevance in the context of systems that have already been in existence for a specific duration. While Shannon's entropy serves as the foundation of information theory, the notion of residual entropy has emerged due to its inadequacy for such systems. By adopting a quantile-based approach, we have explored the characteristics of residual entropy in detail.

Through the derivation of the quantile residual entropy function for various lifetime models, we have provided a novel perspective on analyzing and understanding the dynamics of established systems. Our investigation has further allowed us to delve into the reordering and ageing aspects inherent in the quantile version of the residual entropy equation.

By extending the application of entropy measures to incorporate quantiles, we have bridged the gap in assessing the behavior of systems with a history of existence. This has opened up new avenues for research in information theory and its practical implications.

Overall, our study emphasizes the importance of considering residual entropy based on quantiles, offering a more comprehensive understanding of system dynamics and enabling more accurate analyses in various domains. Further research can build upon these findings to explore additional applications and refine the quantile-based approach to residual entropy.

Conflicts of Interest

No conflicts of interest exist, according to the authors, with the publishing of this article.

References

[1] Arnold, B. C., Blakrishan, N. and Nagraja, N. H. (1992). A first course in order Statistics. John Wiley and Sons, New York.

[2] Baratpour, S., Ahmadi, J. and Arghami, N. R. (2007). Some characterizations based on entropy of order statistics and record values. Communications in Statistical-Theory and Methods, (36), 47-57.

[3] Behboodian, j. and Tahmasebi, S. (2008). Some properties of entropy for the exponentiated pareto distribution (EPD) based on order statistics. Journal of Mathematical Extension, (3) 43-53.

[4] Crescenzo, A. D and Longobardi , M. (2004). Entropy based measure of uncertainty in past lifetime distribution. Journal of Applied probability, (39) 434-440.

[5] Dar, Javid Gani. and Bandere, A.Z. (2013) On Some Characterization Results of Life Time Distributions using Mathai-Haubold Residual Entropy, International organization of Scientific Research-JM. (5), 56-60

[6] David, H. A. (1981). Order Statistics. Second Edition, John Willey and Sons New York.

[7] Ebrahimi, N (1996). How to measure uncertainty in the residual lifetime distribution, Sankhya Ser. A, (58) 48-56.

[8] Ebrahimi, N. et.,al (2004). Information properties of order statistics and spacing. IEEE Trans. Information Theory: (50), 177-183.

[9] Gilchrist, W. (2000). Statsticl modeling with quantile functions. Chapman and Hall/CRC. Boca Raton, FL

[10] Kullback, S and Leibler R. A. (1951). On information and sufficiency. Ann. Math. Stat. (22), 79-86.

[11] Mathai A (2005) A pathway to matrix-variate gamma and normal densities. Linear Algebra Appl 396:317–328

[12] Mathai A, Haubold H (2007a) On generalized entropy measures and pathways. Phys A Stat Mech Appl 385(2):493–500.

[13] Nanda, A. K, Sankaran, P. G, and Sunoj, S.M. (2014). Renyi's residual entropy: A quantile approach. Statistics and Probability Letters (85) 114-121.

[14] Nair, N. U. and Sankaran, P. G. (2009). Quantile based reliability analysis. Commun. Statist.-Theo.Meth., (38) 222-232.

[15] Park, S. (1995). The entropy of consecutive order statistics. IEEE Trans. Inform. Theory, (41) 2003-2007.

[16] Paul, j. and Yageen, P. (2019). On some properties of Mathai–Haubold entropy of record values. J Indian Soc Probab Stat

[17] Sunoj, S. M. and Sankaran, P. G. (2012). Quantile based entropy function. Statistics and Probability Letters (82) 1049-1053.

[18] Samuel, P. and Thomas, P. Y. (2000). An improved form of a recurrence relation on the product moment of order statistics. Commun. Statist.-Theo.Meth.,(29)1559-1564.

[19] Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal. (27), 379-423 and 623-656.

[20] Sebastian N (2015) Generalized pathway entropy and its applications in difiusion entropy analysis and fractional calculus. Commun Appl Ind Math 6(2):1–20.

[21] Wong, K. M. and Chen, S. (1990). The entropy of ordered sequences and order statistics. IEEE Trans. Information Theory: (36), 276-284.

SAMPLING PLANS BASED ON TRUNCATED LIFE TEST FOR LOGISTIC FAMILY OF DISTRIBUTIONS

Sriramachandran G V
•
Department of Mathematics, Dr. N. G. P. Institute of Technology, Coimbatore, India
gvsriramachandran72@gmail.com

Abstract

In this article, we develops optimal sample size for acceptance number (zero and one) for single and double sampling plans by fixing consumer's risk and test completion time, with the assumption that, the life of the item follows logistic family of distributions (i.e. Logistic Rayleigh distribution/Logistic exponential distribution/Logistic Weibull distribution). The optimal size obtained for single and double sampling plans for logistic family of distributions are compared with baseline distributions and the results are discussed.

Keywords: Consumer risk, single sampling plans; double sampling plans; logistic Rayleigh distribution; logistic exponential distribution; logistic Weibull distribution.

I. Introduction

Based on the quality standards prescribed, quality assurance people uses acceptance sampling plan (ASP) to validate the lot. The main focus of ASP is to reduce consumer's risk (CR) as well as producer's risk (PR). By fixing the CR, the minimum sample sizes (MSS) to be used for ASP are obtained. If time parameter is included with ASP, to determine the MSS to guarantee certain median life of products (MLP), then ASP is called as ASP with truncated life test (TLT). Here it is assumed that MLP follows some probability distributions. ASP with TLT can be performed with a given acceptance number c (ACN) say zero and one to find the MSS that guarantees a given MLP at a given consumer confidence level 1- P* , where P* is called as consumer's risk (CR). Accordingly, the lot is rejected if; actual MLP is less than the specified MLP. If the decision to accept or reject the lot is based on the single sample taken from the lot, then it is called as single acceptance sampling plan (SASP) with TLT, otherwise a second sample is taken and decision is based on both first and second sample, then it is termed as double acceptance sampling plan (DASP) with TLT.

Many researchers have proposed these types of ASP with TLT. In the ASP literature, Epstein [6] was the first to consider TLT with an exponential distribution (ED), SASP and DASP for TLT

based on transmuted Rayleigh distribution (RD) was studied by Mahendra Saha et al.[8], ASP based TLT for the Generalized Weibull Model by Shovan Chowdhury [13], ASP based on TLT for extended ED, by Amer I. Al-Omari et al. [1], TLT based ASP for generalized ED by Muhammad Aslam et al.[10], TLT in ASP based on exponentiated ED by Suresh et al.[11], ASP for Generalized RD based on TLT, by Tzong-ru tsai [14], ASP based on TLT for weighted ED [15] are important to mention.

An approach that helps in defining logistic compounded model was given by Yingjie Lan et.al [16] and they proposed logistic–exponential distribution (L-ED). In this line, by taking baseline distribution as RD, Logistic- Rayleigh distribution (L-RD) was defined by Arun Kumar Chaudhary et al. [2] and stated various properties and applications. By same process, Logistic-Weibull distribution (L-WD) was introduced by taking baseline distribution as WD by Arun Kumar Chaudhary et al. [3] and studied its properties and applications. In this article, we develop SASP and DASP with TLT, assuming that MLP follows L-RD or L-ED or L-WD. For CR fixed at 1% and 2% level, MSS to guarantee certain MLP are obtained for these plans.

The structure of the remaining article is: In Section 2, we described the L-RD/ L-ED /L-WD and studied some of its properties. Section 3 presents the design of SASP with TLT, if MLP follows L-RD, L-ED and L-WD. MSS required to maintain CR at 1% and 2% level are obtained for different test time and the results obtained are tabulated in this section. Section 4 presents the design of DASP with TLT if MLP follows L-RD, L-ED and L-WD and for this ASP, MSS required to maintain CR at 1% and 2% level are obtained for different test time. Section 5 compares the MSS obtained for SASP and DASP with some baseline distributions. Finally, conclusion is placed in Section 6.

II. Logistic Family of Distributions

Yingjie Lan et.al [16] have presented an approach to define the logistic compounded model and introduced the logistic exponential survival distribution. Based on Yingjie Lan et.al [16], the L-RD was defined by Arun Kumar Chaudhary et al. [2] as follows: Let t be a non-negative random variable with a positive shape parameter φ and positive scale parameter ω then CDF of L-RD can be defined as

$$F(\tau;\varphi,\omega)=1-\frac{1}{1+\left(e^{\omega\tau^2/2}-1\right)^{\varphi}}\quad,\quad \tau\geq 0 \text{ and } \varphi,\omega>0\,. \tag{1}$$

The corresponding PDF of L-RD is given by

$$f(\tau;\varphi,\omega)=\frac{\omega\,\varphi\,\tau\,e^{\omega\tau^2/2}\left(e^{\omega\tau^2/2}-1\right)^{\varphi-1}}{\left[1+\left(e^{\omega\tau^2/2}-1\right)^{\varphi}\right]^2},\quad \tau\geq 0 \text{ and } \varphi,\omega>0\,, \tag{2}$$

The median of the L-RD is given by $Md=\left[\frac{2}{\omega}\log 2\right]^{1/2}$.

The plots of CDF and PDF of L-RD are given in Fig 1 and Fig 2 respectively for shape parameters φ=0.25, 1, 2 and the scale parameter $\omega=1$.

Let τ be a non-negative random variable with a positive shape parameter φ and positive scale parameter ω then CDF of L-ED as defined by Sajid Ali et al. [12] as in equation 3.

$$F(\tau;\varphi,\omega)=1-\frac{1}{1+\left(e^{\omega\tau}-1\right)^{\varphi}},\ \tau\geq 0 \text{ and } \varphi,\omega>0 \tag{3}$$

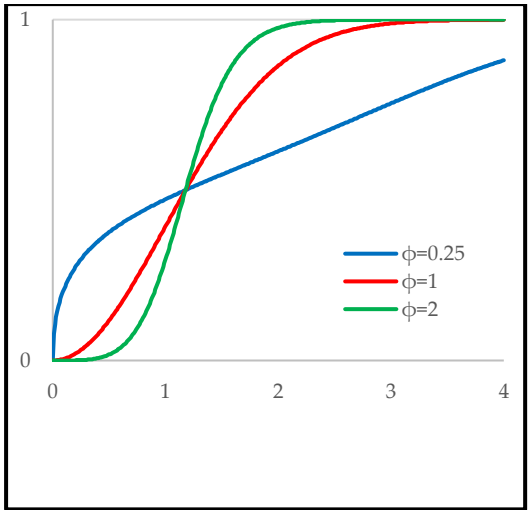


Figure 1 : *cdf plot of L-RD*

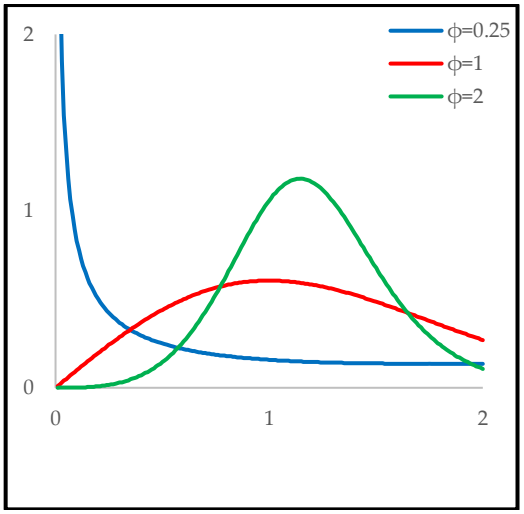


Figure 2 : *pdf plot of L-RD*

Note that the exponential distribution can be obtained as a special case of L-ED when $\phi=1$. The corresponding PDF of L-ED is given in equation 4.

$$f(\tau;\varphi,\omega)=\frac{\omega\varphi e^{\omega\tau}\left(e^{\omega\tau}-1\right)^{\varphi-1}}{\left[1+\left(e^{\omega\tau}-1\right)^{\varphi}\right]^{2}},\ \tau\geq 0 \text{ and } \varphi,\omega>0 \tag{4}$$

The median of the L-ED is given by $Md=\left[\frac{2}{\omega}\log 2\right]$.

The plots of CDF and PDF of L-ED are given in Fig 3 and Fig 4 respectively for shape parameters φ =0.25, 1, 2 and the scale parameter $\omega=1$.

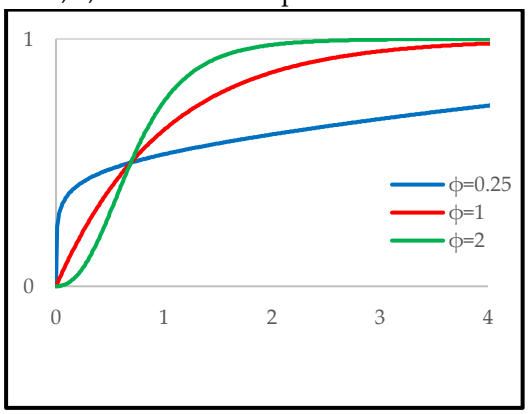


Figure 3 : *cdf plot of L-ED*

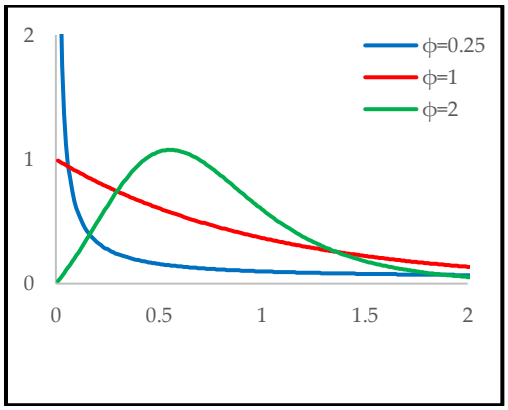


Figure 4 : *pdf plot of L-ED*

Let τ be a non-negative random variable with a positive shape parameter φ and positive scale parameter ω, β then CDF of L-WD as defined by Arun Kumar Chaudhary [3] is given in equation 5.

$$F(\tau;\varphi,\beta,\omega)=1-\frac{1}{1+\left(e^{\omega\tau^{\beta}}-1\right)^{\varphi}},\ \tau\geq 0 \text{ and } \beta,\varphi,\omega>0 \tag{5}$$

The corresponding PDF of L-WD is given by

$$f(\tau;\varphi,\beta,\omega)=\frac{\omega\varphi\beta\tau^{\beta-1}e^{\omega\tau^{\beta}}\left(e^{\omega\tau^{\beta}}-1\right)^{\varphi-1}}{\left[1+\left(e^{\omega\tau^{\beta}}-1\right)^{\varphi}\right]^{2}},\ \tau\geq 0 \text{ and } \beta,\varphi,\omega>0, \tag{6}$$

The median of the L-WD is given by $Md=\left[\frac{2}{\omega}\log 2\right]^{1/\beta}$. The plots of CDF and PDF of L-WD are given in Fig 5 and Fig 6 respectively for shape parameters φ =0.25, 1, 2 and the scale parameter $\omega=1$ and $\beta=0.25$.

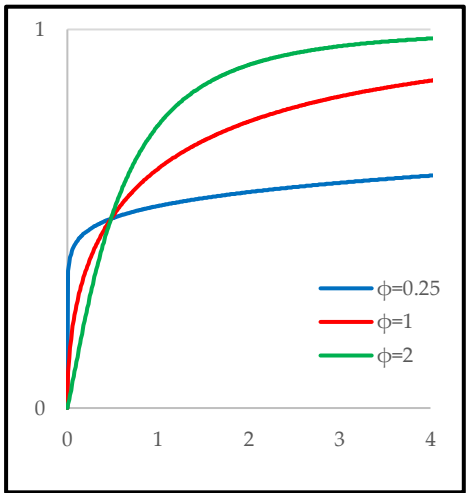


Figure 5 : *cdf plot of L-WD*

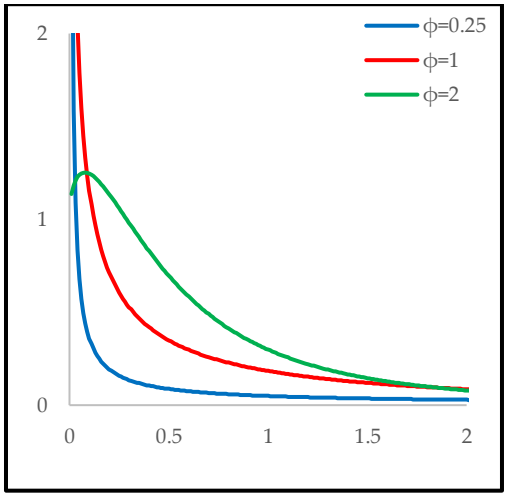


Figure 6 : *pdf plot of L-WD*

III. Design of SASP with TLT Based on Median

The assumption made in designing the SASP with TLT is that the products under study have an MLP of υ. To test the above assumption, we follow the hypothesis test, with null hypothesis H0: $\upsilon\geq\upsilon_0$, the alternative hypothesis H1: $\upsilon<\upsilon_0$ then it is accepted, where υ_0 is a precise MLP of the product. With P^* as CR, the value $1-P^*$ is called the level of significance for the test. Here it may be noted that binomial distribution is applied as the size of the sample is considerably excessive.
As proposed by the ASP, to locate the mss, we have to iterate the inequality

$$\sum_{i=0}^{c}\binom{n}{i}p^{i}(1-p)^{n-i}\leq 1-P^{*} \tag{7}$$

where $F(\tau;\varphi,\omega)$ as in (1) for L-RD , (3) for L-ED and (5) for L-WD. The MSS required for inspection

of the lot for L-RD, L-ED and L-WD for different values of $\frac{\tau}{\upsilon_0}$ with CR = 2% and 1% level with for shape parameters φ=0.25, 1, 2 and the scale parameter $\omega=1$ are calculated and tabulated in the Tables 1 respectively.

Table 1: *MSS for SASP*

CR	Distribution	φ	C	$\frac{\tau}{\upsilon_0}$					
				0.5	0.6	0.7	0.8	0.9	1
1%	L-RD	0.25	0	10	9	8	8	7	7
			1	14	13	12	12	11	11
		1	0	27	19	14	11	9	7
			1	39	28	21	16	13	11
		2	0	131	60	31	17	11	7
			1	-	87	45	25	16	11
	L-ED	0.25	0	8	8	8	8	7	7
			1	12	12	11	11	11	11
		1	0	14	12	10	9	8	7
			1	20	17	15	13	12	11
		2	0	30	20	14	11	9	7
			1	43	29	21	16	13	11
	L-WD	0.25	0	8	8	7	7	7	7
			1	11	11	11	11	11	11
		1	0	10	9	8	8	8	7
			1	15	13	12	12	11	11
		2	0	14	12	10	9	8	7
			1	21	17	15	13	12	11
2%	L-RD	0.25	0	8	8	7	7	6	6
			1	13	12	11	10	10	9
		1	0	23	16	12	9	7	6
			1	35	24	18	14	11	9
		2	0	112	51	26	15	9	6
			1	167	77	40	23	14	9
	L-ED	0.25	0	7	7	7	6	6	6
			1	11	11	10	10	10	9
		1	0	12	10	9	8	7	6
			1	18	15	13	12	10	9
		2	0	25	17	12	9	7	6
			1	38	26	19	14	11	9
	L-WD	0.25	0	7	7	6	6	6	6
			1	10	10	10	10	10	9
		1	0	8	8	7	7	6	6
			1	13	12	11	10	10	9
		2	0	12	10	9	8	7	6
			1	18	15	13	12	10	9

IV. Design of DASP with TLT Based on Median

To give more protection to both consumer as well as producers a two stage ASP called the DASP is preferred. As the name says it provides double protection to producers because we are testing the second sample before taking the final decision on the lot. Hence it provides a total protection to producers and hence minimizes the PR. The parameters of the DASP with TLT are n1 first sample size, it's AN c1, n2 second sample size, it's AN c2 and the testing time t. To accept the lot, it is necessary that the sample supports the hypothesis that sample median to be greater than the median specified. Otherwise, the lot will be rejected. Now, we fix the CR not more than (1-P*).

Then PA of the lot is

$$PA = \sum_{i=0}^{c_1} \binom{n1}{i} p^i (1-p)^{n1-i} + \sum_{x=c_1+i}^{c_2} \binom{n1}{x} p^x (1-p)^{n1-x} \sum_{j=0}^{c_2} \binom{n2}{j} p^j (1-p)^{n2-j} \tag{8}$$

where p is defined in equation (1) for L-RD/ equation (3) for L-ED and equation (5) for L-WD and depends on ratio $\frac{\tau}{\upsilon_0}$. As we are considering only zero-one failure form i.e., c1 =0 and c2 =1,

$$PA = (1-p)^{n1}\left[1 + n1\,p(1-p)^{n_2-1}\right] \tag{9}$$

where p is defined in equation (1) for L-RD/ equation (3) for L-ED and equation (5) for L-WD.

Our aim is to find the MSS for DASP, for this we have to minimize equation 9.

Now, for the given consumer's confidence level P*, the MSS for both the samples n1 and n2, which ensure $\upsilon \geq \upsilon_0$, can be found by the solution of the following optimization problem, given as:

$$\begin{gathered} Min\ \ \mathrm{ASN} = \mathrm{n}_1 + \mathrm{n}_1\mathrm{n}_2\, p(1-p)^{n_1-1} \\ \text{Subject to}: \ (1-p)^{n_1}\left[1 + \mathrm{n}_1\, p(1-p)^{n_2-1}\right] \leq 1 - P^* \\ 1 \leq n_2 \leq n_1 \\ n_1 \text{ and } n_2 \text{ are integers} \end{gathered} \tag{10}$$

While solving the above optimization problem, it provides many solutions for both n1 and n2. We take the solution which minimizes our objective function i.e. our ASN as our best solution. MSS obtained for P* = 0.98 and 0.99 and for different $\frac{\tau}{\upsilon_0}$ are presented in Table 2.

V. Comparative Study

As we are discussing the ASP with TLT based on L-RD, L-ED and L-WD, the CDF and PDF plots of these three distributions are compared at the parameter values at φ=0.25, 1, 2 in Fig 1 and Fig 2. The MSS obtained for L-RD, L-ED and L-WD plans studied here are compared with transmuted RD by Mahendra Saha et al [8], generalized RD by Tzong-ru tsai et al.[14], compound RD by Bhupendra

Singh [4], Wenhao Gui [15], generalized ED by Muhammad Aslam [10], Extended WD by M. S. Hamed et al [7], Generalized WD by Shovan Chowdhury[13], and the results obtained are tabulated in Table 3.

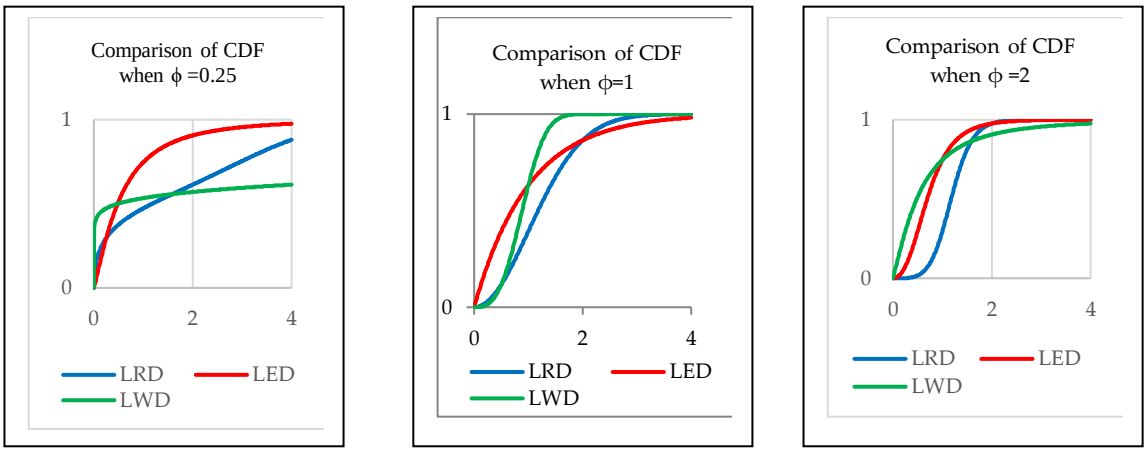


Figure 7: *Comparison of CDF at* φ = *0.25, 1 and 2*

Table 2: *Sample size n1, n2 and ASN for DASP*

CR	Distribution	φ	$\frac{\tau}{\upsilon_0}$							
			0.6		0.8		1		1.2	
			(n1, n2)	ASN	(n1, n2)	ASN	(n1, n2)	ASN	(n1, n2)	ASN
1%	L-RD	0.25	(9,6)	9.28	(8,5)	8.24	(7,5)	7.27	(7,3)	7.12
		1	(19,15)	19.71	(11,7)	11.33	(7,5)	7.27	(5,3)	5.18
		2	(61,49)	63.16	(18,11)	18.47	(7,5)	7.27	(4,2)	4.1
	L-ED	0.25	(8,5)	8.25	(8,4)	8.16	(7,5)	7.27	(7,4)	7.19
		1	(12,7)	12.29	(9,5)	9.23	(7,5)	7.27	(6,4)	6.21
		2	(20,17)	20.81	(11,8)	11.39	(7,5)	7.27	(5,4)	5.24
	L-WD	0.25	(8,4)	8.16	(7,6)	7.36	(7,5)	7.27	(7,4)	7.2
		1	(9,7)	9.36	(8,5)	8.24	(7,5)	7.27	(7,3)	7.12
		2	(12,7)	12.31	(9,5)	9.23	(7,5)	7.27	(6,4)	6.21
2%	L-RD	0.25	(8,5)	8.36	(7,4)	7.31	(6,5)	6.47	(6,3)	6.21
		1	(17,11)	17.76	(10,5)	10.33	(6,5)	6.47	(5,2)	5.12
		2	(52,47)	55.53	(15,13)	16.04	(6,5)	6.47	(4,1)	4.05
	L-ED	0.25	(7,5)	7.4	(7,3)	7.2	(6,5)	6.47	(6,4)	6.33
		1	(10,8)	10.64	(8,4)	8.28	(6,5)	6.47	(5,4)	5.41
		2	(18,11)	18.75	(10,6)	10.41	(6,5)	6.47	(5,2)	5.12
	L-WD	0.25	(7,3)	7.2	(7,3)	7.18	(6,5)	6.47	(6,4)	6.35
		1	(8,5)	8.39	(7,4)	7.31	(6,5)	6.47	(6,3)	6.22
		2	(10,8)	10.68	(8,4)	8.28	(6,5)	6.47	(5,4)	5.41

From Table 3, it is found that the value of MSS for L-WD distribution is low when compared to these distributions. Also the plans studied in this article are compared with respect to MSS at CR=1% and 2%, from table 1, among the three, the L-WD ASP takes MSS for inspection.

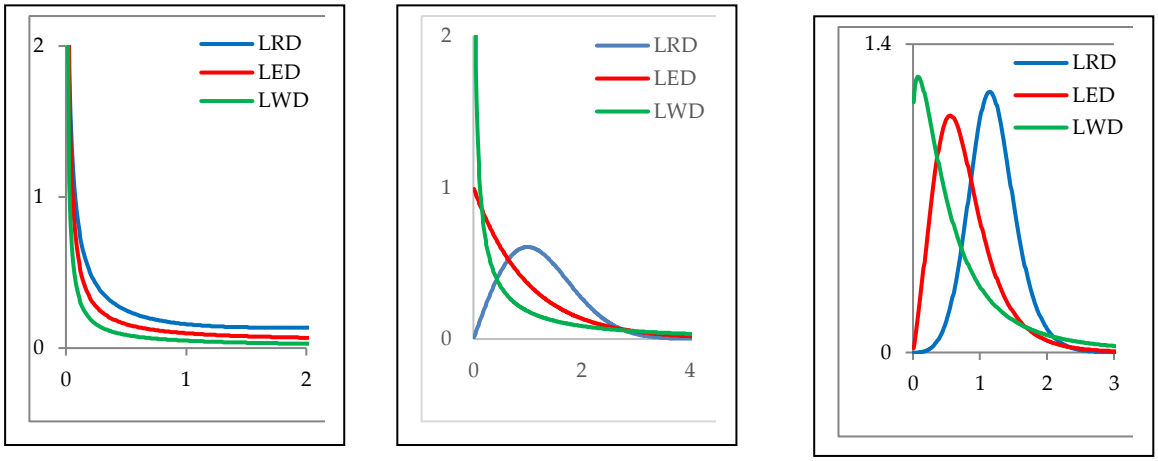


Figure 8 : *Comparison of PDF at* φ = *0.25, 1 and 2*

Table 3: *Comparison of MSS for SASP with baseline distributions for SASP*

Distribution	c=0	c=1
Transmuted RD	16	-
Generalized RD	17	24
Compounded RD	17	24
L-RD at $\varphi=1$	27	39
Generalized ED	14	21
Weighted ED	19	28
L-ED at $\varphi=1$	30	43
Extended WD	6	9
Generalized WD	6	10
L-WD at $\varphi=1$	10	15

VI. Conclusion

In this paper we develop single and double sampling plans by fixing consumer's risk and test completion time, with the assumption that, the life of the item follows logistic family of distributions (i.e. L-RD/L-ED/L-WD). MSS for acceptance number (zero and one) for both SASP and DASP are obtained. The MSS obtained for single and double sampling plans for logistic family of distributions are compared with baseline distributions and the results are discussed.

References

[1] Amer Ibrahim Al-Omari and Al-Hadhrami S.A. (2018). Acceptance sampling plans based on truncated life tests for extended exponential distribution. *Kuwait Journal of Science*, Vol 45, issue 2, pp 30-41.

[2] Arun Kumar Chaudhary and Vijay Kumar. (2020). The logistic-Rayleigh distribution with properties and applications. *International Journal of Statistics and Applied Mathematics*, Vol 5, issue 6, pp 12-19.

[3] Arun Kumar Chaudhary and Vijay Kumar. (2021). The Logistic-Weibull distribution with Properties and Applications. *IOSR Journal of Mathematics*, Vol 17, issue 1, pp 32-41.

[4] Bhupendra Singh. K., Sharma. K. and Dushyant Tyagi. (2013). Acceptance sampling plan based on truncated life tests for compound Rayleigh distribution. *Journal of Reliability and Statistical Studies*, Vol 6 issue 2 pp 1-15.

[5] Canan Hamurkarog, Yigter. A. and Danacıog. N. (2020). Single and Double Acceptance Sampling Plans Based on the Time Truncated Life Tests for the Compound Weibull-Exponential Distribution. *Journal of the Indian Society for Probability and Statistics*, DOI: 10.1007/s41096-020-00087-7

[6] Epstein, B. (1954). Truncated life tests in the exponential case. *Annals of Mathematical Statistics*, DOI: 10.1214/AOMS/1177728723

[7] Hamed. (2020). M. S. Single acceptance sampling plan for truncated life test having the (p-a-l) extended weibull distribution. *Advances and Applications in Statistics*, Vol 61, issue 2, pp 149-159.

[8] Mahendra Saha, Harsh Tripathi and Sanku Dey. (2021). Single and double acceptance sampling plans for truncated life tests based on transmuted Rayleigh distribution. *Journal of Industrial and Production Engineering*, Vol 38, issue 5, pp 356-368.

[9] Mashail M. and Sobhi AL. (2020). The Inverse-Power Logistic-Exponential Distribution: Properties, Estimation Methods, and Application to Insurance Data, *Mathematic*, Vol 8 issue 11, pp 2060.

[10] Muhammad Aslam, Debasis Kundu and Munir Ahmad. (2010). Time truncated acceptance sampling plans for generalized exponential distribution. *Journal of Applied Statistics*, Vol 37, issue 4, pp 555–566..

[11] Suresh. K. K. and Usha. K. (2016). A truncated life test in acceptance sampling plan based on exponentiated exponential distribution. *International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*.

[12] Sajid Ali, Sanku Dey, Muhammad Hussain Tahir and Muhammad Mansoor. (2020). Two-Parameter Logistic-Exponential Distribution: Some New Properties and Estimation Methods. *American Journal of Mathematical and Management Sciences*, DOI: 10.1080/01966324.2020.1728453

[13] Shovan Chowdhury. (2016) Acceptance Sampling Plans Based on Truncated Life Test for the Generalized Weibull Model, *Proceedings of the 2016 IEEE IEEM*.

[14] Tzong-ru tsai and Shuo-jye wu. (2006). Acceptance Sampling Based on Truncated Life Tests for Generalized Rayleigh Distribution. *Journal of Applied Statistics*, Vol 33, issue 6, pp 595–600.

[15] Wenhao Gui and Muhammad Aslam. (2015). Acceptance sampling plans based on truncated life tests for weighted exponential distribution, *Communications in Statistics - Simulation and Computation*, DOI: 10.1080/03610918.2015.1037593 .

[16] Yingjie Lan and Lawrence M. Leemis. (2008). The logistic–exponential survival distribution. *Naval Research Logistics*, Vol 55, issue 3, pp 252-264.

ECONOMIC ORDER QUANTITY MODEL FOR IMPERFECT ITEMS WITH SHORTAGE BACKORDERING

Priyanka Singh[1], A. R. Nigwal[2], U. K. Khedlekar[3]

•

[1]Department of Mathematics, Government Degree College, Piprai, 473440, Ashoknagar, M.P., India
priyanka1singh2013@gmail.com
[2]Department of Mathematics, Ujjain Engineering College, Ujjain, 456010, Madhya Pradesh, India,
arnwnigwal@gmal.com
[3]Department of Mathematics and Statistics, Dr. Harisingh Gour Vishwavidyalaya,
Sagar, 470003, Madhya Pradesh, India, (A Central University)
uvkkcm@yahooo.co.in

Abstract

This study presents the development of an economic order quantity model (EOQ) specifically designed for imperfect quality items. The model takes into consideration three distinct scenarios: (a) Model I trigger a reorder when the inventory level reaches zero; (b) Model II initiates a reorder when the backordered quantity equals the imperfect quantity; (c) Model III initiates a reorder when the shortage persists. To distinguish between perfect and imperfect quality products, a screening process is implemented for each product lot. Upon product delivery from the supplier to the vendor, all received products undergo immediate inspection through the screening process. Following the EOQ ordering policy, the vendor sells imperfect products to customers at a reduced cost at the end of the cycle, rather than returning them to the supplier. To fulfil the remaining demand for high-quality products, the vendor procures such products from a local vendor at a higher price. This study optimizes the duration of positive inventory, selling price, and total profit per unit time. Model I, which exhibits the longest duration of positive inventory, demonstrates greater business stability compared to the other two models. The concavity property is analytically and numerically demonstrated, and a sensitivity analysis is provided to explore the impact of model parameters on outputs.

Keywords: EOQ inventory model, screening process, imperfect items, partial backordering

1. Introduction

In practice production does not lead to perfect items always, there may be a certain portion of imperfect items. Few fractions of imperfect items are usually present in the ordered lot size, shortage takes place due to the presence of imperfect items and also due to lead time. Initially, Rosenblatt and Lee [1] and Porteus[2] were thought to be the creation of things with imperfect quality. An interesting variant has been proposed by Roy et al. [3] who consider an economic order quantity model in which a percentage of imperfect quality lot size follows a uniform distribution function. The associated expected average profit function is generalized for the general distribution function. They discussed the model with partial backlogging and lost sale cases for imperfect quality products. An ideal inventory model for items of imperfect quality, inspection faults, shortfall backorders, and sales returns was proposed by Hsu and Hsu [4]. In their model, a closed-form solution is provided for the optimal order size, the optimal order point and the maximum number of backordering units. A method to estimate the vendor's best

investment in lowering the defect rate in terms of lowering the joint expected yearly total cost was described in Dey and Giri [5].
The impact of the percentage of defective units on the ordering decisions is analysed in Paul et al. [6], who proposed a joint replenishment model with and without price discount for multiple items that indicates the relationship between family cycle length and the integer number of intervals that the replenishment quality of each item will last. The concavity of profit function is shown graphically in each case. According to the Jaber et al. [7], imperfect products make supply chains less sustainable and make the need for change to improve imperfect products worse. They presented two models for the fraction of imperfect quality items received by one shipment. The first model assumes that imperfect items are sent to a repair shop. While the second model assumes that imperfect items are replenished by perfect ones from a local supplier. An EOQ inventory model is developed Taleizadeh et al. ([8]) with partial backordering and studied four cases by taking into account the time when the lot of imperfect products comes back to the buyer's store after the reparation process. Studied a production lot size and backorders level under an EPQ inventory model for imperfect production and considered three cases depending on when the repaired items enter the inventory Taleizadeh et al. [9].

A maintenance policy was implemented by Liao et al. [10] to improve the dependability of a flawed production system. Two states of production operation are executed, namely: the type I state (out-of-control state) and the type II state (in-control state).An EOQ model was suggested by Eroglu and Ozdemir [11] for a production system with backordering for shortages and defective items. The basic assumption in their model is that 100$ screening of each order-lot contains good and defective items both. These defective items are a collection of imperfect quality and scrap items. They observed that the optimal total profit per unit time decreases with an increase in defective and scrap rates individually. An error in an EOQ model with uncertain supply, qualify by a random fraction of imperfect items, and 100$ screening process were fixed by Maddah and Jaber [12] revisiting Salameh and Jaber [14] paper. They proposed a new model that refine this flaw using renewal theory. The EPQ type inventory model with planned backorders was proposed by Cárdenas-Barrón [16] and determines the economic production quantity for a single item, which is manufactured in a single-stage manufacturing system that yield imperfect quality items as well. Also, all the defective items produced are reworked in the same cycle. Two EPQ models were presented by Hsu and Hsu [17], and the results demonstrate that the timing of when to sell the defective goods has a significant impact on the ideal production lot size and the backorder quantity.

Many researchers deliberated about lot-size inventory models with imperfect items. We refer the reader to Rosenblatt and Lee [1] for a detailed review of the defective items with rework and lot size. They considered that the defective quality items could be reworked instantaneously at a certain extra cost and established that the presence of defective items prompts smaller lot-size. According to Salameh and Jaber [14], it was proposed that selling imperfect quality items in a single batch at a reduced price prior to receiving the next shipment leads to an increase in the average percentage of flawed items, affecting the economic lot size. Besides imperfect production, other factors such as breakages and damages during the transportation and handling process may also result in defective items. This result was incorporated in Salameh and Jaber [14]. The problem of shortages of the item may occur when the imperfect quality items are withdrawn from stock. Due to high inventory holding cost, overproduction is not a solution of it assumed by El-Ansary and Robicheaux [18]. The impact of imperfect production processes on the economic lot-sizing policy was examined by Ben-daya and Hariga [19]. An economic production quantity model incorporating the rework of imperfect quality items was suggested by Chiu [20]. He considered that all imperfect items are not repairable, some parts of them are trash and are discarded. According to Chen [21]'s study, imperfect products frequently have an impact on the supply chain's efficiency, inventory holding costs, and production costs. He used a two-echelon supply chain consisting of a manufacturer and a retailer to deal with the imperfect manufacturing system problem. The correct equation and numerical outcomes for the error discovered in the paper by Salameh and Jaber [14] were presented by Cárdenas-Barrón [15]. This error only affects

the optimum value of the order size. For a more recent study on inventory models with imperfect quality items, we mention the readers like Zhou et al. [22], Goyal et al. [23], Patro et al. [24], Keshavarzfard et al. [25], Liao et al. [10], Manna et al. [27], Öztürk [35].

Related recent literature considers that demand is met while screening is being executed, which, in practice, could result in shortages. In view of this, Maddah et al. [13] relaxed this consideration by suggesting an order "overlapping" schemes that provides satisfying demand during the screening process from the "previous order". An economic production quantity model with rework at a single-stage manufacturing system with planned backorders was studied by Chung [33]. He extended Cárdenas-Barrón [16] model and obtained two main results for the annual total relevant cost. Asadkhani et al. [29] developed four EOQ models with different types of imperfect items like salvage, repairable, scrap and reject items. The result show that learning in inspection errors has an important consequence on the profitability. An imperfect production inventory model that includes arbitrary carbon emissions under successive prepayments was developed by Manna et al. [26]. The average profit of the integrated model was maximised by optimising the production rate as well as the defective rate (Manna et al. [28]). This assumption was made that the production system may change from an in-control state to an out-of-control state after a time, which is a random variable. To determine the lower bound of the partial backordering ratio, Yu et al. [36] used an iterative method. By the optimality principle, they show that the shortage is allowed only if the ratio is greater than its lower bound. A periodic delivery policy was put forth for a production-inventory model with vendor-buyer integration by Wee et al. [31]. By combining the assumptions of permissible shortage backordering and the effects of varying backordering cost values, Wee et al. [30] extended Salameh and Jaber [14] model. In order to calculate the expected net profit per unit of time, Chang and Ho [32] revisited the work of Wee et al. [30] and also used the renewal-reward theorem. They used the algebraic methods to obtain the exact closed-form solutions for optimal lot size, backordering quantity and maximum expected profit. The existence of a particular optimal lot size that maximises the anticipated total profit is estimated by Moussawi-Haidar et al. [34]. Also, they studied the effect of deterioration and holding cost on the optimal lot size and backorder level.

It is significant for a manager of any organization to control and handle the inventories of perfect and imperfect quality items. However, one of the weaknesses of current inventory models is the unrealistic assumption that the demand rate is constant. In our study we assume a demand rate is price sensitive, demand varies with respect to selling price \$ p per unit. Each lot received comprises some percentage of imperfect quality, with a known probability density function $f(p)$. The imperfect items are sold in the nearby market by reducing the selling price per unit called salvage price. This type of perfect and imperfect quality item usually occurs in many industries. In this paper, we consider an EOQ inventory model with demand rate as a function of the selling price. At the start of the process, the items received from the supplier, which is far away from the buyer, are inspected. In the meantime, a 100\$ inspection process of the lot is carried out at a rate of α, items of imperfect quality are sorted, kept in one batch and sold at a salvage price of c_s per unit before receiving the next shipment. It is assumed that the imperfect quality items are collected as a single batch and replaced by the perfect items within the replenishment period T. Partial backordering is permissible. The optimum selling price and percentage of duration in which inventory level is positive are obtained by the optimization process to derive an optimal profit.

In traditional inventory models such as the economic order quantity (EOQ) the sole objective is to maximize the total profit, typically to minimize holding cost and ordering cost. These models do not consider the effect of price on the demand rate. Replenishment of imperfect items in an EOQ inventory model with partial backordering proposed model, which dealt with the optimum profit function but considered constant demand rate. However, due to the defective items and inspection process, demand may vary with respect to selling price p. In this direction, this paper develops an EOQ type inventory model with imperfect items and price-sensitive demand rate for determining the optimal unit price, percentage of duration in which inventory level is positive and total profit. The rest of the model explanation is as follows. In section 2, we presented notations

and assumptions used throughout the proposed paper. In section 3, a mathematical model is developed for the three cases when partial backordering is allowed. Section 4 presents numerical examples to exemplify the important aspects of the model. Section 5 show the sensitivity analysis performed on the proposed models. Section 6 presents managerial implications of the models. Finally, a conclusion and future research of the present model are provided in section 7.

2. Notations and Assumptions

I. Notations

The following notations are incorporated to depict the proposed model:

T the cycle length,

t_1 the time duration in which inspection section occur,

α the inspection rate,

x the fraction of imperfect items,

$g(x)$ the probability density function of imperfect items,

y the fraction of backordered quantity,

π the cost of lost sales,

σ the backordered cost,

h_e the holding cost of emergency purchased items,

h the holding cost,

c_o the buyer's ordering cost,

c_p the purchasing cost per unit of an emergency order,

c_u the initial purchasing cost of the unit item,

c_s the salvage price of imperfect items,

c_i the inspection cost,

$D(p)$ the number of units demanded per year where $D(p) = a - bp > 0$, $a, b > 0$,

p unit price,

t the percentage of duration in which inventory level is positive,

q the order quantity,

TP the total profit.

II. Assumptions

The proposed model is constructed based on the subsequent assumptions:

1. The planning horizon is infinite.
2. Shortages are admissible. Partial backordering is allowed.
3. Replenishment is allowed and is equal to the imperfect quality items for one cycle period.

4. Demand is modelled as a price function. The demand rate at selling price p is formulated as $D(p) = a - bp$, where a is representing the maximum demand of the item. b is the coefficient which reflects the choice of price, $a, b > 0$, such that $a - bp > 0$.
5. The proportion of imperfect items x and its probability density function are driven at the end of the cycle length.
6. Purchasing cost c_p from a local supplier is higher than the initial purchasing cost c_u. Moreover, initial purchasing cost c_u is higher than salvage price c_s, $c_s < c_u < c_p$.
7. When the newly purchased items that replace the imperfect items are added to the stock, the inventory level is either zero or negative.
8. Inspection rate is constant and known.
9. The inventory level is I_{max}, at the beginning of the period.
10. In order to distinguish the imperfect items, all products are inspected carefully at the inspection rate α in time duration t_1, where $t_1 = I_{max}/\alpha$.
11. The inspection rate α per unit time is greater than $D(p)$.
12. The reordered items $\rho tTD(p)$ is used to meet the demand during the shortage period partially.

3. Mathematical Formulation

This section extends the work of Taleizadeh *et al.,* (2000) by assuming a price-sensitive demand rate. Three cases are explained with the defective items and partial backordering. Inventory management is a fundamental implementation of price and time control in any modern organization and retail industry. These days, this implementation is important to a supply manager of any supply chain to control and maintain the inventories of perfect and imperfect quality items due to advanced technologies, growing market competition and customer awareness. Consider a vendor who buys products at the beginning of the inventory cycle and satisfy the customer demands during the cycle.

I. *Model I: The reorder is made when the inventory level is zero*

Formulation and solution of the model:
In our model, the initial inventory level at the beginning of the cycle is I_{max} and is equal to $tTD(p)$. This inventory contains perfect and imperfect quality items. Inspection section separates the perfect and imperfect quality items with duration t_1 at an inspection rate α, i.e., $t_1 = \frac{tTD(p)}{\alpha}$. This section identifies $xtTD(p)$ of imperfect quality items and $(1-x)tTD(p)$ is of perfect quality items. The imperfect items are sold at a reduced price c_s so a part of demand of perfect items is still remain to fulfill. To fulfil the demand of buyers, the perfect quality items are purchased from a local supplier at a cost c_p, equal to the imperfect quality items. The reorder is made when the inventory level is zero. At the end of the cycle, the shortage is $(1-t)TD(p)$. y fraction of shortage is backordered at a charge of σ per unit and the rest is lost sales at a charge of π per unit. The total ordered quantity per cycle is $q = tTD(p) + y(1-t)TD(p)$. The logistic representation of the physical scenario is shown in Fig. 1. The inventory holding cost per cycle is calculated from Fig. 1 (summing up the areas) as

$$HC = h\left(\frac{(1-x)^2t^2T(a-bp)}{2} + \frac{xTt^2(a-bp)^2}{\alpha}\right) + \frac{h_e x^2 t^2 T(a-bp)}{2} \tag{1}$$

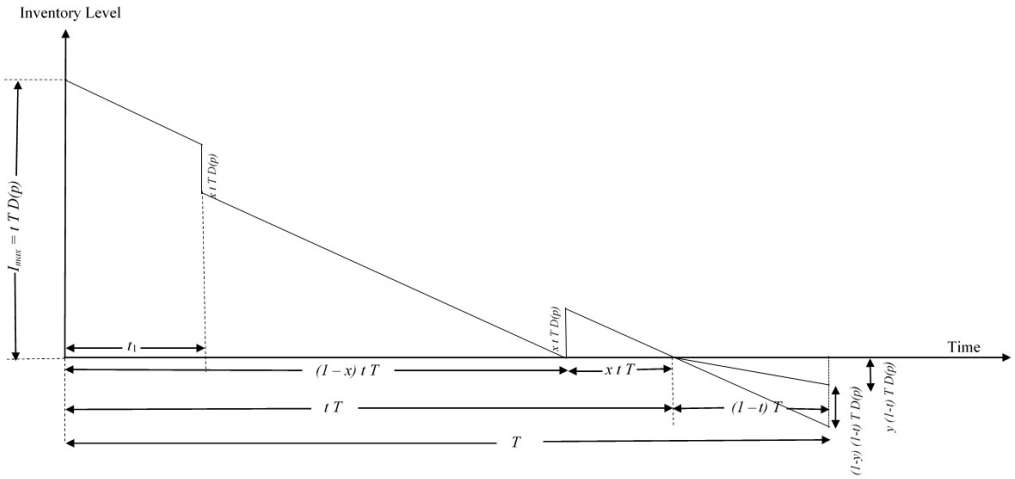


Figure 1: *Behaviour of the inventory level over time*

The shortage cost is given by

$$SC = \frac{(1-t)^2(a-bp)T\sigma y}{2} + \pi(1-y)(1-t)(a-bp) \tag{2}$$

The total revenue per cycle is $p\{tD(p) + y(1-t)D(p)\}$ where p is the unit price. The total profit is the difference of total revenue and the total cost per cycle, and is given as

$$\begin{aligned} TP(t,p) &= p\{t(a-bp) + y(1-t)(a-bp)\} + c_s xt(a-bp) - \frac{c_o}{T} - c_u(t + y(1-t))(a-bp) \\ &- c_p xt(a-bp) - \frac{(1-t)^2(a-bp)T\sigma y}{2} - h\left(\frac{(1-x)^2t^2T(a-bp)}{2} + \frac{xTt^2(a-bp)^2}{\alpha}\right) \\ &- \pi(1-y)(1-t)(a-bp) - c_i t(a-bp) - \frac{h_e x^2 t^2 T(a-bp)}{2} \end{aligned} \tag{3}$$

Our objective is to develop an optimal inventory model to determine the selling price and the percentage of duration in which the inventory level is positive and the total profit. The following cases are considered now

I.1 Case I:

Let the percentage of duration in which inventory level is positive is obtained in terms of p. Then, we have to maximize $TP(p)$, where $t(p)$ ia a function of p, satisfying the constraints $0 < y \leq 1$, $0 < x \leq 1$, $a - bp > 0$, $p > 0$.

In this case, for maximum of TP,

$$dTP(p)/dp = 0 \text{ and } d^2TP(p)/dp^2 < 0,$$

must be satisfied.

Now, differentiate Eq. (3) with respect to t and equating to zero, we obtain

$$t(p) = \frac{\alpha(-c_i + \pi + p - c_u - \pi y - py + \sigma Ty + c_u y - c_p x + c_s x)}{T\left[h\{\alpha(x-1)^2 + 2D(p)x\} + \alpha(\sigma y + h_e x^2)\right]} \tag{4}$$

The expected value of variable x is substituted in Eq. (4)

$$t(p) = \frac{\alpha(-c_i + \pi + p - c_u - \pi y - py + \sigma Ty + c_u y - c_p E(x) + c_s E(x))}{T\left[h\{\alpha(E(x)-1)^2 + 2D(p)E(x)\} + \alpha(\sigma y + h_e E(x)^2)\right]}$$

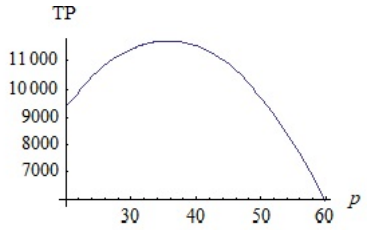


Figure 2: *Graphical representation of the total profit $TP(p)$ with respect to p*

Now Substituting this value in Eq. (3) and differentiating with respect to p leads to

$$\begin{aligned}\frac{dTP(p)}{dp} = {} & \frac{-2bhD(p)h_e\alpha^2x^3(M2)^2}{T(M1)^3} - \frac{2bh^2D(p)\alpha x(M3)(M2)^2}{T(M1)^3} - \frac{D(p)h_e\alpha^2(y-1)x^2(M2)}{T(M1)^2} \\ & - \frac{hD(p)\alpha(y-1)(M3)(M2)}{T(M1)^2} + \frac{bhD(p)\alpha x(M2)^2}{T(M1)^2} + \frac{bh_e\alpha^2x^2(M2)^2}{2T(M1)^2} + \frac{bh\alpha(M3)(M2)^2}{2T(M1)^2} \\ & + \frac{2bc_ihD(p)\alpha x(M2)}{T(M1)^2} + \frac{2bc_phD(p)\alpha x^2(M2)}{T(M1)^2} - \frac{2bhD(p)c_s\alpha x^2(M2)}{T(M1)^2} + \frac{c_iD(p)\alpha(y-1)}{T(M1)} \\ & + \frac{D(p)c_s\alpha(1-y)x}{T(M1)} + \frac{c_pD(p)\alpha(y-1)x}{T(M1)} + \frac{bc_i\alpha x(M4)}{T(M1)} - \frac{bc_s\alpha x(M4)}{T(M1)} \\ & + b\pi(y-1)\left[-1-\frac{x(M2)}{T(M1)}\right] + \frac{1}{2}b\sigma Ty\left[-1-\frac{x(M2)}{T(M1)}\right]^2 \\ & + (D(p)-bp+bc_u)\left[y+\frac{\alpha y(M2)}{T(M1)}-\frac{x(M2)}{T(M1)}\right] - \frac{D(p)\sigma\alpha y(M5)(M4)}{T(M1)^3} \\ & + \frac{(\pi+p-c_u)D(p)\alpha(y-1)(M4)}{T(M1)^2}\end{aligned} \tag{5}$$

where

$$M1 = h(M3) + \alpha(\sigma y + h_e x^2),\ M2 = c_i - p + c_u + \pi(y-1) + py - \sigma Ty - c_u y + c_p x - c_s x,$$
$$M3 = \alpha(x-1)^2 + 2D(p)x,\ M4 = \alpha(y-1)(\sigma y + h_e x^2) + h\left[\alpha(y-1)(x-1)^2 + 2x\{a(y-1)\right.$$
$$\left. + b\left(c_i + c_u\ + \pi(y-1) - \sigma Ty - c_u y + c_p x - c_s x\right)\}\right],\ M5 = c_i\alpha - p\alpha + hT\alpha + c_u\alpha + \pi\alpha(y-1)$$
$$+ p\alpha y - c_u\alpha y + 2D(p)hTx + c_p\alpha x - c_s\alpha x - 2hT\alpha x + (h+h_e)T\alpha x^2$$

whose solution is given by setting the Eq. (5) to zero and solving for p. Clearly, $\frac{d^2TP(p)}{dp^2}$ is negative (see Fig. 2), which implies that $TP(p)$ is concave.

I.2 Case II.

Let p and the percentage of duration in which inventory levels are positive t be decision variables.
Here, we have to maximize $TP(t,p)$
such that $0 < y \leq 1$, $0 < x \leq 1$, $a - bp > 0$, $p > 0$.

The expression for $\frac{\partial TP(t,p)}{\partial t} = 0$ and $\frac{\partial TP(t,p)}{\partial p} = 0$, are complicated and it is possible to obtain the optimal values of p and t analytically by using Mathematica Software and satisfying the conditions $\frac{\partial TP^2(t,p)}{\partial t^2} < 0$, $\frac{\partial TP^2(t,p)}{\partial p^2} < 0$ and $\frac{\partial TP^2(t,p)}{\partial t^2}.\frac{\partial TP^2(t,p)}{\partial p^2} - \left(\frac{\partial TP^2(t,p)}{\partial p \partial t}\right)^2 > 0$.
Therefore, we draw a graph of Eq. (3) for the parameter values interpreted in the numerical example 1. The Fig. 7 demonstrate that the function $TP(p,t)$ is concave having a clear maximum.

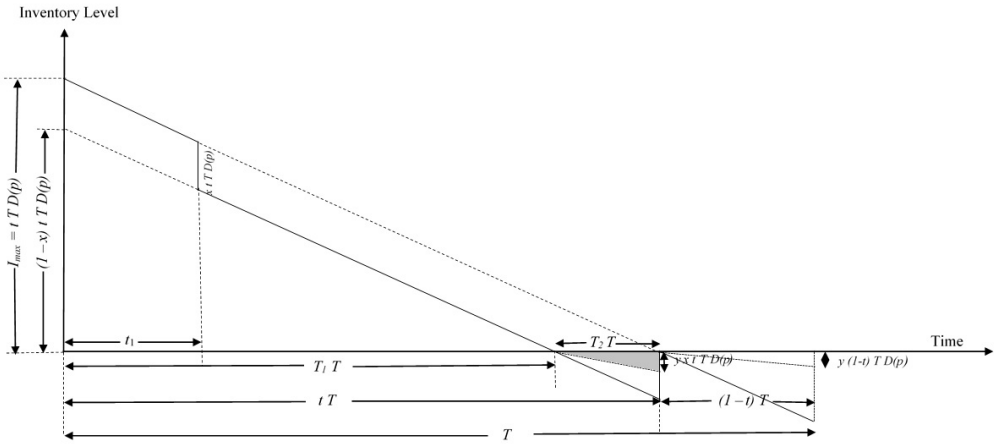


Figure 3: *Behaviour of the inventory level over time*

For this we have,

$$\begin{aligned}&\frac{\partial TP^2(t,p)}{\partial t^2}.\frac{\partial TP^2(t,p)}{\partial p^2}-\left(\frac{\partial TP^2(t,p)}{\partial p\partial t}\right)^2\\&=\frac{1}{\alpha^2}\Big[2bD(p)T\{\alpha y+t\alpha(1-y)+bt^2hTx\}(M1)-\{a(\alpha-\alpha y+4bthTx)\\&+b\left(c\alpha-2p\alpha+thT\alpha+c_u\alpha+\pi\alpha(y-1)+2p\alpha y-\sigma T\alpha y+t\sigma T\alpha y-c_u\alpha y\right.\\&\left.-4bthpTx+c_p\alpha x-c_s\alpha x-2thT\alpha x+thT\alpha x^2+tTh_e\alpha x^2\right)\}^2\Big]\end{aligned}\tag{6}$$

For a particular case it can be proved numerically by considering the data given in section 4, Example 1, $\frac{\partial^2 TP}{\partial p^2}=-19.52<0$, $\frac{\partial^2 TP}{\partial t^2}=-150.48<0$, $\frac{\partial^2 TP}{\partial p^2}.\frac{\partial^2 TP}{\partial t^2}-\left(\frac{\partial^2 TP}{\partial p\partial t}\right)^2=2893.28>0$. This implies that $TP(p,t)$ is concave in nature with respect to p and t.

II. *Model II: The reorder is made when the backordered quantity is equal to the imperfect quality items*

Note that t is the percentage of duration in which the inventory level is positive. In the beginning of cycle, the inventory level begins with the order quantity $I_{max}=tTD(p)$. These products are inspected to distinguish perfect and imperfect items separately. At the end of the inspection process at a rate of α units per unit time during the period t_1, a part of the perfect quality items is used to serve the customer demand at a price of p per unit. At the end of the cycle, the imperfect quality items $xtTD(p)$ are withdrawn and sold at a salvage price c_s. The reorder items are received when the shortage is equal to $xtTD(p)$. After adding these products to the negative inventory lot, the inventory level reaches zero. The shortage is $(1-t)TD(p)$. The backordered percentage of shortage is y and the rest is lost sale. Let x be the fraction of imperfect items. Specifically, we have $t=T_1+T_2$ where $T_1=(1-x)t$, $T_2=xt$. The total ordered quantity per cycle be $q=T_1TD(p)+yT_2TD(p)+y(1-t)TD(p)$. The behaviour of the inventory level over time is illustrated in Fig. 3.

The holding cost per cycle is given as

$$HC=h\left(\frac{(1-x)^2t^2T(a-bp)}{2}+\frac{xTt^2(a-bp)^2}{\alpha}\right)\tag{7}$$

and the shortage cost per cycle is

$$SC = \sigma\frac{yx^2t^2TD(p)}{2} + \sigma\frac{y(1-t)^2TD(p)}{2} + \pi(1-y)T_2D(p) + \pi(1-y)(1-t)D(p)$$

Then the shortage cost can be rewritten as

$$SC = \sigma\frac{yx^2t^2TD(p)}{2} + \sigma\frac{y(1-t)^2TD(p)}{2} + \pi(1-y)(1-T_1)(a-bp) \tag{8}$$

Therefore, the total profit per cycle of length T becomes

$$\begin{aligned} TP(t,p) = {} & p\{(1-x)tD(p) + y(1-(1-x)t)D(p)\} + c_sxtD(p) - \frac{c_o}{T} - c_u(t+y(1-t))D(p) \\ & - c_itD(p) - c_pxtD(p) - h\left(\frac{(1-x)^2t^2TD(p)}{2} + \frac{xt^2TD(p)^2}{\alpha}\right) - \frac{\sigma yx^2t^2TD(p)}{2} \\ & - \frac{\sigma y(1-t)^2TD(p)}{2} - \pi(1-y)(1-(1-x)t)D(p) \end{aligned} \tag{9}$$

The following cases are considered

II.1 Case I:

Let the percentage of duration in which inventory level is positive is obtained in terms of p. Then, we have to maximize $TP(p)$, where $t(p)$ ia a function of p, satisfying the constraints $0 < y \leq 1$, $0 < x \leq 1$, $a - bp > 0$, $p > 0$. In this case, for maximum of TP, $dTP(p)/dp = 0$ and $d^2TP(p)/dp^2 < 0$ must be satisfied. Now, differentiating Eq. (9) with respect to t and equating to zero, we obtain $t(p) = -\frac{\alpha(A6)}{T(A2)}$. Substituting this in Eq. (9) and differentiating with respect to p gives

$$\begin{aligned} \frac{\partial TP}{\partial p} = {} & \frac{-2bhD(p)\sigma\alpha^2yx^3(A1)^2}{T(A2)^3} - \frac{-2bh^2D(p)\alpha x(A5)(A1)^2}{T(A2)^3} + \frac{2bhD(p)c_s\alpha x^2(A1)}{T(A2)^2} + \frac{bh\alpha(A1)^2(A5)}{2T(A2)^2} \\ & - \frac{D(p)\alpha^2\sigma(y-1)(x-1)yx^2(A1)}{T(A2)^2} - \frac{hD(p)\alpha(y-1)(x-1)(A1)(A5)}{T(A2)^2} + \frac{bhD(p)\alpha x(A1)^2}{T(A2)^2} \\ & + \frac{2bc_ihD(p)\alpha x(A6)}{T(A2)^2} - \frac{bc_i\alpha(A6)}{T(A2)} + \frac{2bc_phD(p)\alpha x^2(A6)}{T(A2)^2} + \frac{D(p)c_s\alpha(y-1)(x-1)x}{T(A2)} \\ & + \frac{c_pD(p)\alpha x(A4)}{T(A2)} - \frac{bc_s\alpha x(A1)}{T(A2)} - \frac{bc_p\alpha x(A6)}{T(A2)} + \frac{b\sigma Ty}{2}\left[\frac{\alpha(A1)}{T(A2)} - 1\right]^2 \\ & - b\pi(y-1)\left[1 + \frac{\alpha(x-1)(A1)}{T(A2)}\right] - bp\left[y + \frac{\alpha(y-1)(x-1)(A1)}{T(A2)}\right] + \frac{D(p)\sigma\alpha y}{T(A2)^3}\left[c_i\alpha\right. \\ & -p\alpha + hT\alpha + c_u\alpha + p\alpha y - c_u\alpha y + 2ahTx - 2bhpTx + c_p\alpha x + p\alpha x - c_s\alpha x - 2hT\alpha x \\ & \left. -p\alpha yx + hT\alpha x^2 + \sigma T\alpha yx^2 + \pi\alpha(A4)\right]\left[\sigma\alpha(y-1)y(A3) + h\{\alpha(y-1)(x-1)^3\right. \\ & \left.+2x\left(a(y-1)(x-1)b(A7)\right)\}\right] + \frac{D(p)c_u\alpha(y-1)}{T(A2)^2}\left[\sigma\alpha(y-1)y(A3) + h\{\alpha(y-1)(x-1)^3\right. \\ & \left.+2x\left(a(y-1)(x-1) + b(A7)\right)\}\right] + \\ & \frac{\pi D(p)\alpha(y-1)(x-1)}{T(A2)^2}\left[\sigma\alpha(y-1)y(A3) + h\{\alpha(y-1)(x-1)^3 + 2x\left(a(y-1)(x-1) + b(A7)\right)\}\right] \\ & + \frac{pD(p)\alpha(y-1)(x-1)}{T(A2)^2}\left[\sigma\alpha(y-1)y(A3) + h\{\alpha(y-1)(x-1)^3 + 2x\left(a(y-1)(x-1) + b(A7)\right)\}\right] \\ & + \frac{pD(p)\alpha(y-1)(x-1)}{T(A2)^2}\left[\sigma\alpha(y-1)y(A3) + h\{\alpha(y-1)(x-1)^3 + 2x\left(a(y-1)(x-1) + b(A7)\right)\}\right] \\ & + \frac{c_iD(p)\alpha(A4)}{T(A2)} + bc_u\left[y + \frac{\alpha(A1)}{T(A2)} - \frac{\alpha y(A1)}{T(A2)}\right] + D(p)\left[y + \frac{\alpha(y-1)(x-1)(A1)}{T(A2)}\right] \end{aligned}$$

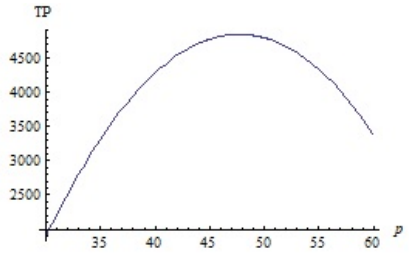


Figure 4: *Graphical representation of the total profit $TP(p)$ with respect to p*

where

$$A1 = (A7) - p(A4),\ A2 = h\{\alpha(x-1)^2 + 2D(p)x\} + \sigma\alpha y(1+x^2),\ A3 = -(1-x)(1+x^2),$$
$$A4 = -(1-x)(1-y),\ A5 = \alpha(x-1)^2 + 2D(p)x,\ A6 = p(A4) + (A7) - Ty(\alpha+\sigma),$$
$$A7 = -c_i - c_u + \sigma Ty + c_u y - \pi(A4) - c_P x + c_s x,$$

$$A8 = \{c_i\alpha - p\alpha + hT\alpha + c_u\alpha + p\alpha y - c_u\alpha y + 2ahTx - 2bhpTx + c_p\alpha x + p\alpha x - c_s\alpha x - 2hT\alpha x - p\alpha yx$$
$$+ hT\alpha x^2 + \sigma T\alpha yx^2 + \pi\alpha(A4)\}\left[\sigma\alpha(y-1)y(A3) + h\{\alpha(y-1)(x-1)^3 + 2x\left(-a(A4) + b(A7)\right)\}\right]$$

Note that as long as $a - bp > 0$, it can be shown by differentiating $TP(p)$ with respect to p twice that $TP(p)$ is concave in p (see Fig. 4).

II.2 Case II.

Let p and the percentage of duration in which inventory levels are positive t be decision variables. Here, we have to maximize $TP(t,p)$ such that $0 < y \leq 1$, $0 < x \leq 1$, $a - bp > 0$, $p > 0$. The expression for $\frac{\partial TP(t,p)}{\partial t} = 0$ and $\frac{\partial TP(t,p)}{\partial p} = 0$, are complicated and it is possible to obtain the optimal values of p and t analytically by using Mathematica Software and satisfying the conditions $\frac{\partial TP^2(t,p)}{\partial t^2} < 0$, $\frac{\partial TP^2(t,p)}{\partial p^2} < 0$ and $\frac{\partial TP^2(t,p)}{\partial t^2}.\frac{\partial TP^2(t,p)}{\partial p^2} - \left(\frac{\partial TP^2(t,p)}{\partial p \partial t}\right)^2 > 0$. Therefore, we draw a graph of Eq. (9) for the parameter values interpreted in the numerical example 1. The Fig. 8 demonstrate that the function $TP(p,t)$ is concave having a clear maximum. For this, we have

$$\frac{\partial TP^2(t,p)}{\partial t^2}.\frac{\partial TP^2(t,p)}{\partial p^2} - \left(\frac{\partial TP^2(t,p)}{\partial p \partial t}\right)^2 = \frac{1}{\alpha^2}\Big[2b(a-bp)T\{bht^2Tx + (t(1-x)(1-y)+y)\,\alpha\}$$
$$\{2hx(a-bp) + \alpha\left(h(x-1)^2 + (1+x^2)y\sigma\right)\} - \{a\,(4bhtTx + (1-x)(1-y)\alpha)$$
$$+b\left(-4bhptTx + \alpha\left(c_i + c_u + c_p x - c_s x + (1-x)^2 htT - c_u y - (2p+\pi)(1-x)(1-y) - Ty\sigma\right.\right.$$
$$\left.\left.+(1+x^2)tTy\sigma\right)\right)\}^2\Big] \tag{10}$$

For a particular case it can be proved numerically by considering the data given in section 4, Example 2, $\frac{\partial^2 TP}{\partial p^2} = -19.48$, $\frac{\partial^2 TP}{\partial t^2} = -150.62$, $\frac{\partial^2 TP}{\partial p^2}.\frac{\partial^2 TP}{\partial t^2} - \left(\frac{\partial^2 TP}{\partial p \partial t}\right)^2 = 2892.35$. Note that we have $\partial^2 TP(p,t)/\partial p^2 < 0$, $\partial^2 TP(p,t)/\partial t^2 < 0$, and $\partial^2 TP(p,t)/\partial p \partial t > 0$, which implies that there exist unique values of p and t that maximize Eq. (9). The optimal solution can be obtained by setting Eqs. (**??**) and (**??**) to zero in Mathematica Software, which leads to $TP(p,t)$ is a concave function.

III. *Model III: The reorder is made when the shortage is yet continued*

We considered the problem of a lot size at the beginning of the cycle $tTD(p)$, which is inspected at an inspection rate α in a time duration t_1 where $t_1 = tTD(p)/\alpha$. We assumed

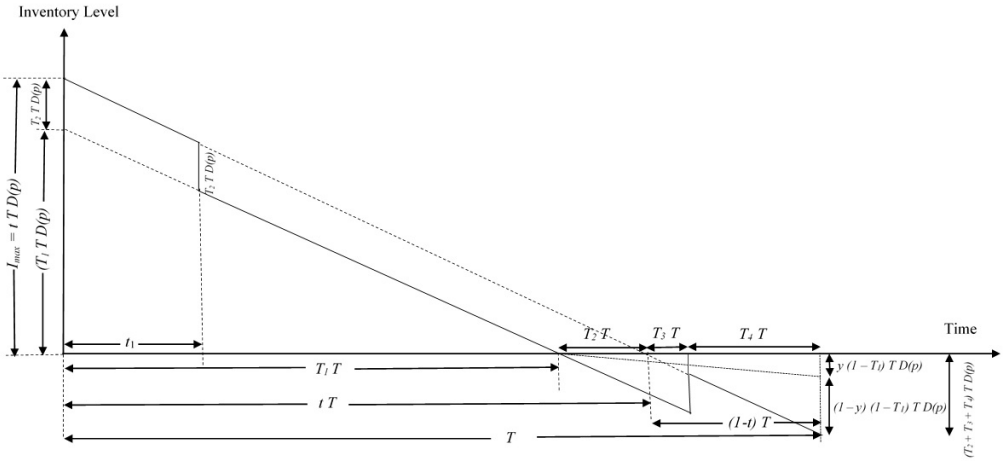


Figure 5: *Behaviour of the inventory level over time*

that the duration length is divided into four parts. The parts are symbolize by T_i, $i = 1, 2, 3, 4$. Let t be the percentage of duration in which inventory level is positive and $t = T_1 + T_2$. Also $1 - t = t_3 + t_4$. Let x be the fraction of imperfect items found and are sold at the end of the cycle at a salvage price c_s. Then $T_1 = (1 - x)t$, $T_2 = xt$. The imperfect items are replaced by perfect items from the local supplier at the cost of c_p. The reorder items are obtained when the shortage level is $(T_2 + T_3)TD(p)$. After adding these items to the inventory system, the inventory level is negative means shortage is still continued. Let y be the fraction of backordered quantity and rest is considered as lost sale. The backordered quantity has backordered cost σ per unit and the lost sale cost per unit is π. At the end of the cycle, the shortage level is $(T_2 + T_3 + T_4)TD(p)$. The ordered quantity per cycle is $q = TD(p) + y(1 - t)TD(p)$. The behaviour of the inventory system is illustrated in Fig. 5. The holding cost per cycle is determined from Fig. 5

$$HC = h\left(\frac{(1-x)^2t^2T(a-bp)}{2} + \frac{xTt^2(a-bp)^2}{\alpha}\right) \tag{11}$$

The shortage cost per cycle is given by

$$SC = \sigma\frac{(T_2+T_3+T_4)(T_3+T_4)yTD(p)}{2} + \pi(1-y)(T_3+T_4)(a-bp) \tag{12}$$

The total profit per cycle of length T becomes

$$\begin{aligned} TP(p,t) = {} & pD(p)\{t + y(1-t)\} + c_s xtD(p) - \frac{c_o}{T} - c_u D(p)\{t + y(1-t)\} - c_i tD(p) \\ & - h\left[\frac{D(p)t^2T(1-x)^2}{2} + \frac{xTt^2D(p)^2}{x}\right] - \frac{\sigma y(1-(1-x)t)(1-t)D(p)T}{2} \\ & - \pi(1-y)(1-t)D(p) - c_p xtD(p) \end{aligned} \tag{13}$$

III.1 Case I:

Let the percentage of duration in which inventory level is positive is obtained in terms of p. Then, we have to maximize $TP(p)$, where $t(p)$ ia a function of p, satisfying the constraints $0 < y \leq 1$, $0 < x \leq 1$, $a - bp > 0$, $p > 0$. In this case, for maximum of TP, $dTP(p)/dp = 0$ and $d^2TP(p)/dp^2 < 0$ must be satisfied.

Differentiating Eq. (13) with respect to t and equating to zero, we obtain $t(p) = -\frac{\alpha(B2)}{2T(B1)}$.

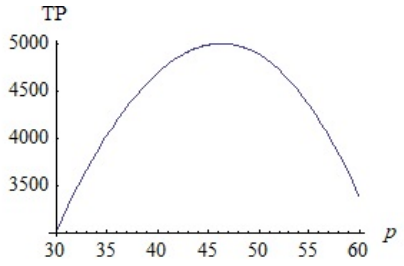


Figure 6: *Graphical representation of the total profit $TP(p)$ with respect to p*

Now, substituting this value in Eq. (13) and differentiating with respect to p, we have

$$\begin{aligned}\frac{\partial TP}{\partial p} &= \frac{2bc_i hD(p)\alpha x(B2)}{T(B1)^2} + \frac{2bc_p hD(p)\alpha x^2(B2)}{T(B1)^2} - \frac{2bhD(p)c_s\alpha x^2(B2)}{T(B1)^2} \\ &+ \frac{ac_i D(p)\alpha(y-1)}{T(B1)} + \frac{2(c_p-c_s)D(p)\alpha(y-1)x}{T(B1)} - \frac{bc_i\alpha(B2)}{T(B1)} - \frac{b\alpha x(c_p-c_s)(B2)}{T(B1)} \\ &+ \frac{b(p-c_u)(B5)}{T(B1)} - \frac{D(p)(B5)}{T(B1)} + 2b\pi(1-y)\left(1+\frac{\alpha(B2)}{2T(B1)}\right) + b\sigma Ty\left(1+\frac{\alpha(B2)}{2T(B1)}\right) \\ &\left(1-\frac{\alpha(x-1)(B2)}{2T(B1)}\right) - \frac{h\alpha(B2)}{4T(B1)^3}\{4bhD(p)\alpha(x-1)^2x(B2) + 8bhD(p)^2x^2(B2) \\ &+ 4D(p)\alpha(y-1)(x-1)^2(B1) + 8D(p)^2(y-1)x(B1) - b\alpha(x-1)^2(B1)(B2) \\ &- 4bD(p)x(B1)(B2)\} + \frac{2D(p)\alpha(\pi+p-c_u)(y-1)(B6)}{T(B1)^2} + \frac{\sigma\alpha yD(p)(x-1)(B6)(B7)}{2T(B1)^3} \\ &- \frac{\sigma\alpha yD(p)(B6)}{(B1)^2}\left(1-\frac{\alpha(x-1)(B2)}{2T(B1)}\right)\end{aligned} \tag{14}$$

where

$$\begin{aligned}B1 &= -\sigma\alpha y(x-1) + h\{\alpha(x-1)^2 + 2D(p)x\} \\ B2 &= 2c_i - 2p + 2c_u + 2\pi(y-1) + 2py - 2\sigma Ty - 2c_u y + 2c_p x - 2c_s x + \sigma Tyx \\ B3 &= -2p\alpha + 2c_u\alpha - 2c_i\alpha(y-1) - 2\pi\alpha(y-1)^2 + 4p\alpha y - 2hT\alpha y - 2\sigma T\alpha y - 4c_u\alpha y \\ &- 2p\alpha y^2 + 2c_u\alpha y^2 + 2c_p\alpha x - 2c_s\alpha x - 4ahTyx + 4bhpTyx - 2c_p\alpha yx + 2c_s\alpha yx \\ &+ 4hT\alpha yx + \sigma T\alpha yx + \sigma T\alpha y^2 x - 2hT\alpha yx^2 \\ B4 &= 2c_i + 2c_u + 2\pi(y-1) - 2\sigma Ty - 2c_u y + 2c_p x - 2c_s x + \sigma Tyx \\ B5 &= -2p\alpha + 2c_u\alpha - 2c_i\alpha(y-1) - 2\pi\alpha(y-1)^2 + 4p\alpha y - 2hT\alpha y - 2\sigma T\alpha y - 4c_u\alpha y \\ &- 2p\alpha y^2 + 2c_u\alpha y^2 + 2c_p\alpha x - 2c_s\alpha x - 4ahTyx + 4chpTyx - 2c_p\alpha yx + 2c_s\alpha yx \\ &+ 4hT\alpha yx + \sigma T\alpha yx + \sigma T\alpha y^2 x - 2hT\alpha yx^2 \\ B6 &= \sigma\alpha y(A4) + h\left[\alpha(y-1)(x-1)^2 + x\{2a(y-1) + b(B4)\}\right] \\ B7 &= 2c_i\alpha - 2p\alpha + 2hT\alpha + 2c_u\alpha + 2\pi\alpha(y-1) + 2p\alpha y - 2c_u\alpha y + 4ahTx - 4bhpTx \\ &+ 2c_p\alpha x - 2c_s\alpha y - 4hT\alpha x - \sigma T\alpha yx + 2hT\alpha x^2\end{aligned}$$

Now, $\frac{dTP(p)}{dp} = 0$ implies a unique value of p using Mathematica software, which is complex in nature. Clearly, $\frac{d^2TP(p)}{dp^2}$ is negative (see Fig. 6), which implies that $TP(p)$ is concave.

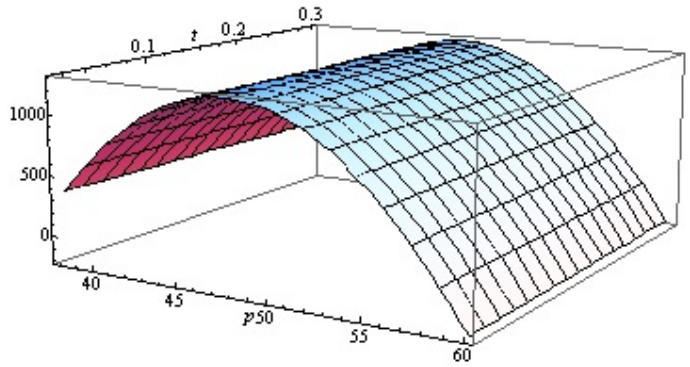


Figure 7: *Graphical representation of the total profit $TP(p,t)$ with respect to p and t*

Table 1: *Optimal value for the decision variables*

	p^*	t^*	$TP^*(p^*,t^*)$
Model I	47.71	21.00%	1278.10
Model II	47.69	14.20%	1276.41
Model III	47.00	16.70%	1272.97

III.2 Case II.

Let p and the percentage of duration in which inventory levels are positive t be decision variables. Here we have maximize $TP(t,p)$ such that $0 < y \leq 1, 0 < x \leq 1, a - bp > 0, p > 0$. The expression for $\frac{\partial TP(t,p)}{\partial t} = 0$ and $\frac{\partial TP(t,p)}{\partial p} = 0$, are complicated and it is possible to obtain the optimal values of p and t analytically by using Mathematica Software and satisfying the conditions $\frac{\partial TP^2(t,p)}{\partial t^2} < 0$, $\frac{\partial TP^2(t,p)}{\partial p^2} < 0$ and $\frac{\partial TP^2(t,p)}{\partial t^2}.\frac{\partial TP^2(t,p)}{\partial p^2} - \left(\frac{\partial TP^2(t,p)}{\partial p \partial t}\right)^2 > 0$, where

$$\left(\frac{\partial^2 TP}{\partial p^2}\right)\left(\frac{\partial^2 TP}{\partial t^2}\right) - \left(\frac{\partial^2 TP}{\partial p \partial t}\right)^2 = -\frac{1}{4\alpha^2}\Big[-8b(a-bp)T\{bht^2Tx + (t+y-ty)\alpha\}\{2hx(a-bp) + (x-1)\alpha(h(x-1)-y\sigma)\} + \{2a(4bhtTx + \alpha - y\alpha) + b(-8bhptTx + \alpha\{2c_i + 2c_u + 2c_p x - 2c_s x + 2htT(1-x)^2 + 4p(y-1) - 2c_u y - 2\pi + 2y\pi - 2Ty\sigma + 2tTy\sigma + Txy\sigma - 2tTxy\sigma\})\}^2\Big] \quad (15)$$

For a particular case it can be proved numerically by considering the data given in section 4, Example 3, $\frac{\partial^2 TP}{\partial p^2} = -19.50 < 0$, $\frac{\partial^2 TP}{\partial t^2} = -151.49 < 0$, $\frac{\partial^2 TP}{\partial p^2}.\frac{\partial^2 TP}{\partial t^2} - \left(\frac{\partial^2 TP}{\partial p \partial t}\right)^2 = 2906.43 > 0$. Our objective is to determine the optimum values of p and t, which maximise the total profit per cycle. We set $\frac{\partial TP}{\partial p} = 0 = \frac{\partial TP}{\partial t}$ in Mathematica Software, to get the optimum values of p and t. Therefore, we draw a graph of Eq. (13) for the parameter values interpreted in the numerical example 3. The Fig. 9 demonstrate that the function $TP(p,t)$ is concave having a clear maximum.

4. Numerical Examples

Example 1. *Model I:* We considered the values of the parameters in suitable units such that cycle length $T = 0.028$ years, market potential $a = 700$, price sensitivity of demand $b = 10$, salvage

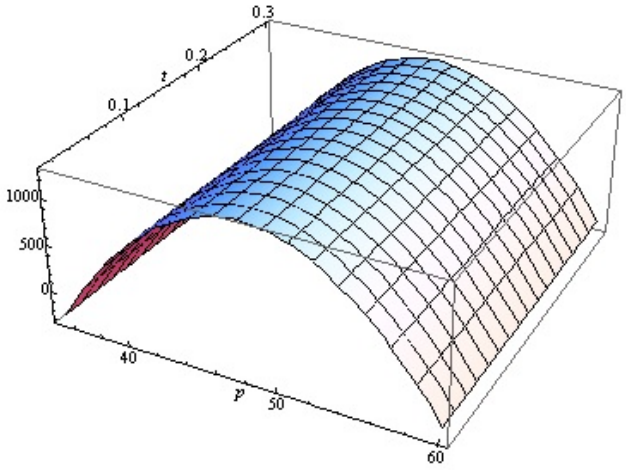


Figure 8: *Graphical representation of the total profit $TP(p,t)$ with respect to p and t*

price $c_s = \$20$ per unit, initial purchasing cost $c_u = \$25$ per unit, inspection cost $c_i = \$0.5$ per unit, purchasing cost of emergency order $c_p = \$40$ per unit, holding cost of emergency order $h_e = \$8$ per unit per year, fraction of backordered quantity $y = 97\%$, fraction of imperfect items $x = 0.03$, ordering cost $c_o = \$100$ per order, inventory holding cost $h = \$5$ per unit per year, inspection rate $\alpha = 175200$ units per year, backordered cost $\sigma = \$20$ per unit per unit time, lost sales cost $\pi = \$0.5$ per unit, $D(p) = a - bp = 222.89$ units per year, which gives optimal unit price $p^* = \$47.71$ per unit, optimal percentage of duration $t^* = 21\%$, optimal total profit $TP^*(p^*, t^*) = 1278.10$ (see Fig. 7).

We assume that the fraction of imperfect quality items follow a probability density function with $x = U\ (0, 0.04)$, $g(x) = 1/(0.04 - 0)$. **Example** 2. *Model II:* We considered a situation with the following parameters: $T = 0.028$, $a = 700$, $b = 10$, $c_s = 20$, $c_u = 25$, $c_i = 0.5$, $c_p = 40$, $h_e = 8$, $y = 97\%$, $x = 0.03$, $c_o = 100$, $h = 5$, $\alpha = 175200$, $\sigma = 20$, $\pi = 0.5$. Then, the optimal unit price and optimal percentage of duration in which inventory level is positive are found by maximizing the total profit function. We obtain the following optimal solution $p^* = \$47.7$, $t^* = 14.2\%$ and $TP^*(p^*, t^*) = \$1276.41$ (see Fig. 8). **Example** 3**.** *Model III:* We assumed a situation with the following parameters: $T = 0.028$, $a = 700$, $b = 10$, $c_s = 20$, $c_u = 25$, $c_i = 0.5$, $c_p = 40$, $h_e = 8$, $\sigma = 20$, $y = 97\%$, $x = 0.03$, $c_o = 100$, $h = 5$, $\alpha = 175200$, $\pi = 0.5$, which gives $p^* = \$47$, $t^* = 16.7\%$ and $TP^*(p^*, t^*) = \$1272.97$ (see Fig. 9).

5. Comments on Numerical Examples

In this section, we prepared numerical analysis for the models. First, we find the optimal profit for a given set of parameters. Table 1, indicates that Model I is the best choice as it gives maximum profit. But if we think in reference of duration, then Model II is the good choice as it gives more profit in less time. Comparative studies in the example show that the total profit per year using the Model I is usually highest than the Model II and Model III. Also, the total profit of the Model II is higher than the Model III.

We then execute sensitivity analysis by varying some model parameters and calculating the optimal price p, optimal percentage of duration in which inventory level is positive t and optimal total profit $TP(p,t)$ for each parameter set. We performed sensitivity analysis of the optimal solutions with respect to the length of cycle, T, and price sensitivity of demand, b, the other parameters are kept fixed at their pre-assumed values. The results are presented in Table 2 and Table 3. In the first column of Table 2, we vary the cycle length, which assumes the values -20%, -10%, $+50\%$, $+60\%$, $+70\%$, $+80\%$. The results indicate that as the cycle length increases, the

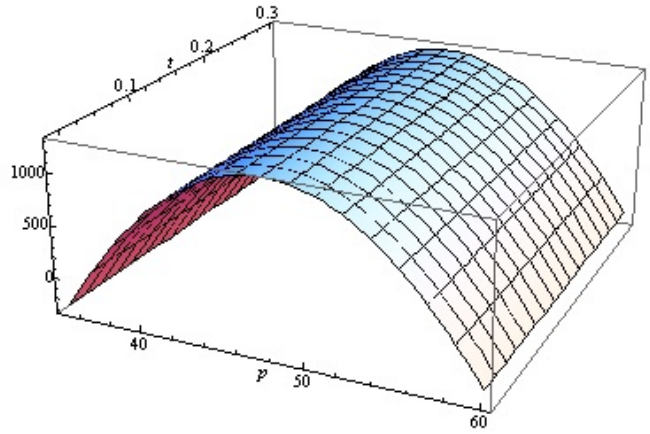


Figure 9: *Graphical representation of the total profit $TP(p,t)$ with respect to p and t*

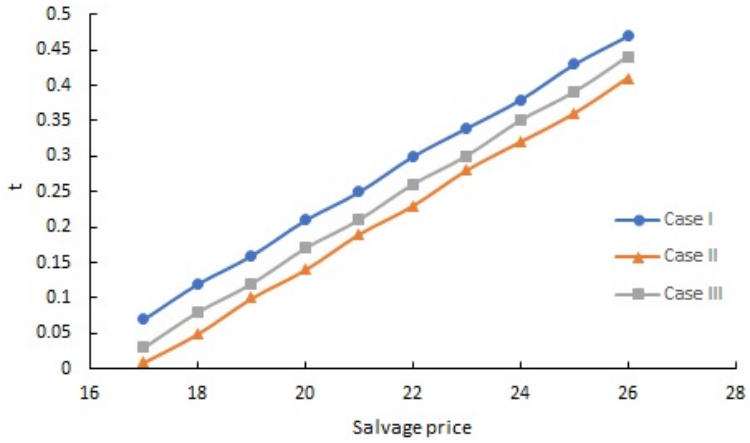


Figure 10: *Behaviour of Percentage of duration t for the three cases against Salvage price c_s*

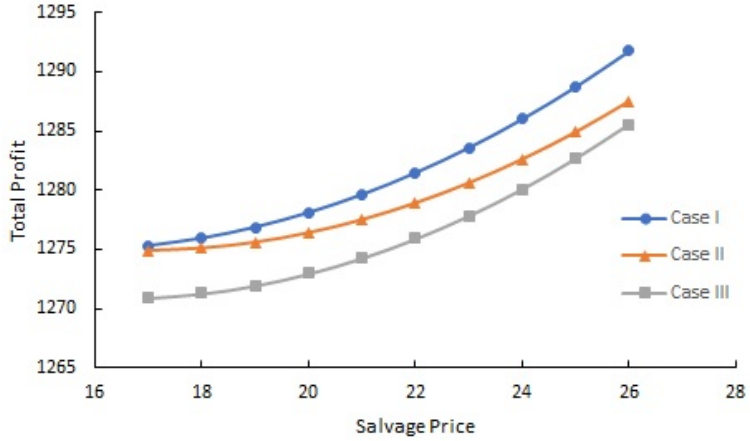


Figure 11: *Behaviour of Total Profit TP for the three models against salvage price c_s*

Table 2: *Model I: Optimal solutions for different cycle time, data used here are* $a = 700$, $b = 10$, $c_s = 20$, $c_u = 25$, $c_i = 0.5$, $c_p = 40$, $h_e = 8$, $\sigma = 20$, $y = 97\%$, $x = 0.03$, $c_o = 100$, $h = 5$, $\alpha = 175200$, $\pi = 0.5$.

T (years)		p^*	t^*	$TP(p^*, t^*)$
$-$ 20%	0.022	47.63	4%	314.00
$-$ 10%	0.025	47.68	13%	854.03
$+$ 50%	0.042	47.81	41%	2453.80
$+$ 60%	0.045	47.83	44%	2610.10
$+$ 70%	0.048	47.84	46%	2746.70
$+$ 80%	0.050	47.85	47%	2828.58

Table 3: *Optimal solutions for different values of parameter* b, *data used here are* $T = 0.028$, $a = 700$, $c_s = 20$, $c_u = 25$, $c_i = 0.5$, $c_p = 40$, $h_e = 8$, $\sigma = 20$, $y = 97\%$, $x = 0.03$, $c_o = 100$, $h = 5$, $\alpha = 175200$, $\pi = 0.5$.

b	Model I			Model II			Model III		
	p^*	t^*	$TP(p^*, t^*)$	p^*	t^*	$TP(p^*, t^*)$	p^*	t^*	$TP(p^*, t^*)$
7	63.02	89%	5969.72	62.98	80.0%	5957.21	63.02	89.7%	5969.54
8	56.62	60%	3965.11	56.59	52.5%	3957.94	56.62	60.5%	3964.64
9	51.67	38%	2451.49	51.64	31.2%	2447.66	51.67	38.0%	2451.04
10	47.71	21%	1278.10	47.69	14.2%	1276.41	47.00	16.7%	1272.97
11	44.48	6%	350.14	44.47	0.3%	349.86	44.48	5.2%	350.05

optimal solutions p, t and $TP(p.t)$ increase, for fixed a, b, c_s, c_u, c_i, c_p, h_e, σ, y, x, c_o, h, α and π. In Table 3, the b takes the values 7, 8, 9, 10 and 11. The results indicate that as the value of b increases, the optimal solutions decrease, for fixed a, T, c_s, c_u, c_i, c_p, h_e, σ, y, x, c_o, h, α and π.

Figure 10, represents the variation of variable t with respect to salvage price c_s for the three cases for fixed p^*, a, T, c_u, c_i, c_p, h_e, σ, y, x, c_o, h, α and π.
For the fixed p^* and the above-mentioned values of parameters, Figure 11, represents the variation of total profit function TP for the three models with respect to salvage price c_s.
The main results from the sensitivity analysis are as follows:

(1) As depicted in Fig. 10, when the salvage value of the imperfect quality items increases, the percentage of the duration in which inventory level is positive, t(Model I), t(Model II) and t(Model III) also increase. The value of t(Model III) is higher than that of t(Model II). But t(Model I) is highest always. This reveals the percentage of the duration of positive inventory is high which shows more stability of the business. Also, t is high that is the shortages time is less, which also shows a better strategy. This shows that Model I with partial backordering performs better than the other two models.

(2) As depicted in Fig. 11, when the salvage price of the imperfect quality items increases, the optimal total profit per year TP for Models I, II and III also increase. The value of TP for Model I is always higher than that of Model II and Model III. So, Model I is the more perfect choice for the proposed objective.
It is noted that TP(Model I) $>$ TP(Model II) $>$ TP(Model III).

6. Managerial Implications

Replenishment of perfect items equal to imperfect quality items is well established and studies in the literature. In this paper, we integrate the EOQ model with replenishment of perfect quality items equal to imperfect quality items after going through the screening process. We find the optimal unit price, optimal percentage of duration in which inventory level is positive and optimal profit for the models with imperfect quality items, and compare the results of the models.

We examine the effect of changing various model parameters, such as cycle time, salvage price and holding cost of emergency purchased items. Also, we incorporated partial backordering into our model. We notice that as the salvage price increases, it is optimal to have more percentage of duration (t) and total profit (TP). In all three cases, the optimal unit price increases as the cycle time (T) increases. The effect of price-sensitive parameter (b) significantly impacts the unit price (p), percentage of duration (t) and total profit (TP). Several studies discuss the techniques to reduce the effect of imperfect quality items at the vendor level. The cost of reducing the effect of imperfect quality items and partial backordering by employing appropriate techniques would help to increase the total profit.

7. Conclusion and Future Research

In this paper, we have developed three different models for the same problem and then selected the most suitable model that is the optimal inventory model with perfect and imperfect quantity items, the inspection process, partial backorders and reorders, which gives better results than the other two. We developed three models: The first model assumes that reordered items are received when the inventory level is zero. Our analytical and numerical results show that there exists a unique optimal sales price, lot size and percentage of duration in which inventory level is positive that maximizes the total profit. The second model assumes that reordered items are received when the backordered quantity is equal to the imperfect quantity items. The third model assumes that reordered items are received when backordered is yet continued. We provided numerical examples and sensitivity analysis to illustrate the outcomes of the above models. Sensitivity analysis performs to study the effect of different parameters of the system like salvage price, cycle time, holding cost of emergency purchased items and price-sensitive parameter on the optimal solutions. Numerical computations show that when the price-sensitive parameter (b) increases, the optimal unit price (p), the optimal percentage of the duration in which inventory level is positive (t) and the optimal total profit (TP) decrease for all three models. The percentage of the duration of positive inventory t is high which shows more stability of the business. Also, t is high that is the shortages time is less, which also show a better strategy. Model I is the more perfect choice for the proposed objective as it gives the maximum value of total profit.

Several extensions of the current model are possible. Future studies can be carried out by considering complete backordering. Another interesting extension is to offering price discounts on the stock-out items to the EOQ models with imperfect quality items and the screening process. One way to extend the model is to consider stochastic demand.

References

[1] Rosenblatt, M. J. and Lee, H. L. (1986). Economic production cycles with imperfect production processes. *IIE Transactions*, 18(1):48–55.

[2] Porteus, E. L. (1986). Optimal lot sizing, process quality improvement and setup cost reduction. *Operations Research*, 34:137–144.

[3] Roy, M. D., Sana S. S. and Chaudhuri, K. (2011). An economic order quantity model of imperfect quality items with partial backlogging. *International Journal of Systems Science*, 42(8):1409–1419.

[4] Hsu, J. T. and Hsu, L. F. (2013b). An EOQ model with imperfect quality items, inspection errors, shortage backordering, and sales returns. *International Journal of Production Economics*, 143(1):162–170.

[5] Dey, O. and Giri, B. C. (2014). Optimal vendor investment for reducing defect rate in a vendorâĂŞbuyer integrated system with imperfect production process. *International Journal of Production Economics*, 155:222–228.

[6] Paul, S., Wahab, M. I. M. and Ongkunaruk, P. (2014). Joint replenishment with imperfect items and price discount. *Computers and Industrial Engineering*, 74:179–185.

[7] Jaber, M. Y., Zanoni, S. and Zavanella, L. E. (2014). Economic order quantity models for imperfect items with buy and repair options. *International Journal of Production Economics*, 155:126–131.

[8] Taleizadeh, A. A., Khanbaglo, M. P. S. and Cárdenas-Barrón, L. E. (2016). An EOQ inventory model with partial backordering and reparation of imperfect products. *International Journal of Production Economics*, 182(2016):418–434.

[9] Taleizadeh, A. A., Khanbaglo, M. P. S. and Cárdenas-Barrón, L. E. (2017). Outsourcing rework of imperfect items in the economic production quantity (EPQ) inventory model with backordered demand. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(12):2688–2699.

[10] Liao, J. J., Huang, K. N., Chung, K. J., Lin, S. D., Chuang, S. T. and Srivastava, H. M. (2020). Optimal ordering policy in an economic order quantity (EOQ) model for non-instantaneous deteriorating items with defective quality and permissible delay in payments. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, 114(1):41, 1–26.

[11] Eroglu, A. and Ozdemir, G. (2007). An economic order quantity model with defective items and shortages. *International Journal of Production Economics*, 106(2):544–549.

[12] Maddah, B. and Jaber, M. Y. (2008). Economic order quantity for items with imperfect quality: Revisited. *International Journal of Production Economics*, 112(2):808–815.

[13] Maddah, B., Salameh, M. K. and Moussawi-Haidar, L. (2010). Order overlapping: A practical approach for preventing shortages during screening. *Computers and Industrial Engineering*, 58(4):691–695.

[14] Salameh, M. K. and Jaber M. Y. (2000). Economic production quantity model for items with imperfect quality. *International Journal of Production Economics*, 64(1–3):59–64.

[15] Cárdenas-Barrón, L. E. (2000). Observation on: "Economic production quantity model for items with imperfect quality". *International Journal of Production Economics*, 67(2):201.

[16] Cárdenas-Barrón, L. E. (2009). Economic production quantity with rework process at a single-stage manufacturing system with planned backorders. *Computers and Industrial Engineering*, 57(3):1105–1113.

[17] Hsu, J. T. and Hsu, L. F. (2013a). Two EPQ models with imperfect production processes, inspection errors, planned backorders, and sales returns. *Computers and Industrial Engineering*, 64(1):389–402.

[18] El-Ansary, A. I. and Robicheaux, R. A. (1974). A theory of channel control: revisited. *Journal of Marketing*, 30:2–7.

[19] Ben-Daya, M. and Hariga, M. (2000). Economic lot scheduling problem with imperfect production processes. *Journal of the Operational Research Society*, 51:875–881.

[20] Chiu, Y. P. (2003). Determining the optimal lot size for the finite production model with random defective rate, the rework process, and backlogging. *Engineering Optimization*, 35(4):427–437.

[21] Chen, T. H. (2017). Optimizing pricing, replenishment and rework decision for imperfect and deteriorating items in a manufacturer-retailer channel. *International Journal of Production Economics*, 183:539–550.

[22] Zhou, Y. W., Chen, J., Wu, Y. and Zhou, W. (2015). EPQ models for items with imperfect quality and one-timeonly discount. *Applied Mathematical Modelling*, 39(3–4):1000–1018.

[23] Goyal, S. K., Singh, S. R. and Yadav, D. (2017). Economic order quantity model for imperfect lot with partial backordering under the efect of learning and advertisement dependent imprecise demand. *International Journal of Operational Research*, 29(2):197–218.

[24] Patro, R., Acharya, M., Nayak, M. M. and Patnaik, S. (2018). A fuzzy EOQ model for deteriorating items with imperfect quality using proportionate discount under learning effects. *International Journal of Management and Decision Making*, 17(2):171–198.

[25] Keshavarzfard, R., Makui, A., Tavakkoli-Moghaddam, R. and Taleizadeh, A. A. (2019). Optimization of imperfect economic manufacturing models with a power demand rate dependent production rate. *SÄĄdhanÄĄ*, 44(9):206.

[26] Manna, A. K., Das, B. and Tiwari, S. (2020(a)). Impact of carbon emission on imperfect production inventory system with advance payment base free transportation. *RAIRO - Operations Research*, 54(4):1103–1117.

[27] Manna, A. K., Dey, J. K. and Mondal, S. K. (2020(b)). Effect of inspection errors on imperfect production inventory model with warranty and price discount dependent demand rate. *RAIRO Operations Research*, 54(4):1189–1213.

[28] Manna, A. K., Dey, J. K. and Mondal, S. K. (2017). Two layers supply chain in an imperfect production inventory model with two storage facilities under reliability consideration. *Journal of Industrial and Production Engineering*, 35(2):57–73.

[29] Asadkhani, J., Mokhtari, H. and Tahmasebpoor, S. (2021). Optimal lot-sizing under learning effect in inspection errors with different types of imperfect quality items. *Operational Research* 22:2631–2665.

[30] Wee, H., Yu, J. and Chen, M. C. (2007). Optimal inventory model for items with imperfect quality and shortage backordering. *Omega*, 35(1):7-11.

[31] Wee, H. M., Yu, J. C. P. and Wang, K. J. (2006). An integrated production-inventory model for deteriorating items with imperfect quality and shortage backordering considerations. *Lecture Notes in Computer Science*, 885–897.

[32] Chang, H. C. and Ho, C. H. (2010). Exact closed-form solutions for "optimal inventory model for items with imperfect quality and shortage backordering". *Omega*, 38(3–4):233–237.

[33] Chung, K. J. (2011). The economic production quantity with rework process in supply chain management. *Computers and Mathematics with Applications*, 62(6):2547–2550.

[34] Moussawi-Haidar, L., Salameh, M. and Nasr, W. (2014). Effect of deterioration on the instantaneous replenishment model with imperfect quality items. *Applied Mathematical Modelling*, 38(24):5956–5966.

[35] Öztürk, H. (2020). Economic order quantity models for the shipment containing defective items with inspection errors and a sub-lot inspection policy. *European Journal of Industrial Engineering*, 14(1):85–126.

[36] Yu, J. C. P., Wee, H. M. and Chen, J. M. (2005). Optimal ordering policy for a deteriorating item with imperfect quality and partial backordering. *Journal of the Chinese Institute of Industrial Engineers*, 22(6):509–520.

DESIGNING A HYBRID SINGLE SAMPLING PLAN FOR LIFE-TIME ASSESSMENT USING THE EXPONENTIAL-RAYLEIGH DISTRIBUTION

Radhika A*
•
Assistant Professor, Department of Statistics, Periyar University, Salem-11
radhisaran2004@gmail.com
Nandhini M
•
Research Scholar, Department of Statistics, Periyar University, Salem-11
nandhinichithra9@gmail.com
Jeslin J
•
Research Scholar, Department of Statistics, Periyar University, Salem-11
jeslin.statistics@gmail.com

Abstract

The approach of statistical quality control known as "product control" deals with the steps involved in making judgments on one or more batches of completed goods produced by production processes. One of the main categories of product control is sampling inspection by variables, which includes processes for selecting numerous individual units based on sample measurements for a quality characteristic under investigation. These approaches are predicated on the knowledge of the functional form of the probability distribution and the presumption that the quality feature is measured on a continuous scale. The literature on product control contains inspection techniques that were created with the implicit presumption that the quality characteristic is distributed normally with the associated attributes. In this study, a single variable sampling plan is developed and assessed under the assumption that the quality characteristic will be distributed using an Exponential-Rayleigh distribution. This article discusses the development of reliability sampling plans for intermittent test batches using type-I and type-II censoring data. To build a sampling strategy using the Exponential-Rayleigh distribution, this work offers a two-parameter continuous probability distribution. One of the main categories of acceptance sampling is sampling inspection by variables, which involves processes for making decisions regarding the disposition of numerous individual units based on sample measurements of those units for a quality feature under investigation. Assume that the sample inspection's number of defective items follows the Poisson distribution. The suggested SSP's ideal parameters are determined using a multi-objective genetic algorithm, which is concerned with concurrently minimizing the average number of samples and inspection costs a maximizing the likelihood of the acceptance sampling plan. The Rayleigh distribution is an appropriate model for life-testing studies, and the Exponential Rayleigh Distribution is studied as a model for a lifetime random variable. The paper also analyses the effectiveness of reliable single sampling plans designed using the median lifetime of products. The efficiency of these sampling plans is evaluated in terms of sample size and sampling risks. Poisson probabilities are used to determine the parameters of the sampling plans, to protect both producers and consumers from risks. For manufacturing enterprises to analyze the viability of the sample plan, necessary tables and procedures are constructed with acceptable examples.

Keywords: Reliability Sampling, Exponential Rayleigh Distribution, Median life-time, Reliable Single Sampling, Single Sampling Plan.

I. Introduction

Statistical quality control (SQC) is a valuable technique that can help improve a company's production process. One important aspect of SQC is sampling for acceptance or rejection of a lot. Acceptance single sampling is a commonly used method for determining whether a lot should be accepted or rejected based on classical attribute quality characteristics. Acceptance sampling is a widely used technique in Quality Control. The primary objective of the plan is to determine optimal plan specifications, including the sample size and acceptance number, to save time and cost during the experiment. The study focuses on a group acceptance sampling plan for items with MOKw-E distribution. Key design parameters are extracted and operating characteristic function is determined for different quality levels. The results will guide future research on Nano quality-level topics with different probability distributions [1]. This is particularly important if the product's quality is defined by its lifetime.

Many authors have created Reliability Sampling Plans that rely on the sample size and acceptance constant but do not provide a guarantee of product reliability for intermittent testing lots or batches. A Modified group chain sampling plan is developed for a truncated life test when the lifetime of an item follows Rayleigh distribution. An Optimal number of groups and operating characteristic values are obtained by obeying the specified consumer risk, test termination time, and mean ratio [2]. The duration of a successful product’s operation is measured in lifetime data, which is gathered during life tests, in hours, miles, cycles before failure, or any other pertinent parameter. This is referred to as 'Life Data'. These reliability sampling plans are designed to be more reliable in terms of accepting products or batches. A new acceptance sampling plan based on truncated life tests and Quasi Shanker distribution was developed for quality control. The suggested plan provides smaller sample sizes and a substantial sampling economy compared to other competitors. It can be used in industry and for further research [3]. Procedures for designing and simulating necessary tables are provided to simplify product selection and testing.

The Exponential and Rayleigh distribution are two of the most important distributions in the field of life testing and reliability theory. They possess significant structural properties and offer great mathematical flexibility. The concept of truncated single acceptance sampling plan at a pre-assigned time. Different acceptance numbers and values for the ratio of the specified test duration to the specified mean life are achieved with the given probability levels. It is also discovered that the minimal sample sizes guaranteed the specified mean lifetime [4]. The problem of acceptance sampling when the life test is truncated at a pre-assigned time is discussed with known shape and scale parameters. For different acceptance ratios, different levels of confidence, and different proportions of the fixed experimental period to the given median lifetime [5]. However, it's important to note that all of these works were carried out assuming that the life testing was done under a hybrid censoring scheme. They also only consider the consumer's risk while ignoring the risk for the producer in rejecting lots of good products. The paper proposes a fuzzy Poisson-based single sampling plan with varying OC curve widths and compares it to the binomial-based plan [6].

This paper aims to establish the dependability of sampling plans based on exponential Rayleigh distribution while considering the levels of producer and consumer risk. Briefly describe the theoretical view of Rayleigh distribution in section II. The OC function of the reliability single sampling plan is derived in section III, and the procedures for determining and operating the sampling plans are explained in section IV. Section V covers the construction of tables for optimal sampling plans in specific cases, and an example is provided to illustrate the selection of the sampling plan. Finally, the results are summarized in section VI. Two acceptance sampling plans based on Weibull Exponential and Weibull Lomax distributions, using the maximum likelihood method to estimate model parameters. The proposed plans are compared with existing plans based on inspected items [7].

II. Theoretical View of Rayleigh Distribution

The Rayleigh distribution is a continuous probability distribution within the domain of probability theory and statistics, applicable to random variables featuring non-negative values. It exhibits a correlation with the chi distribution, specifically when endowed with two degrees of freedom, although this connection involves rescaling. The eponym for this distribution stems from Lord Rayleigh [18]. When the overall magnitude of a vector in the plane is correlated with its directional components, a Rayleigh distribution is frequently seen. The Rayleigh distribution might naturally appear, for instance, when the two-dimensional analysis of wind velocity is performed. The overall wind speed (vector magnitude) will be represented by a Rayleigh distribution if each component has zero mean, equal variance, and is normally distributed. The situation of random complex numbers with real and imaginary components that are independently and identically distributed Gaussian with equal variance and zero mean provides a second illustration of the distribution. In that situation, the complex number's absolute value has a Rayleigh distribution.

Magnetic Resonance Imaging (MRI) is a field in which the estimation is applied. The background data is Rayleigh distributed because MRI pictures are typically interpreted as magnitude images even though they are recorded as complex images. As a result, the noise variance in an MRI image can be calculated from background data using the technique above [19] [20]. The application of the Rayleigh distribution extended to the field of nutrition, where it was employed to establish a link between dietary nutrient levels and the physiological responses of both humans and animals. This technique represents one approach to compute the nutritional response relationship through utilization of the parameter [21].

III. Operating Characteristic Function of RSSPs under the conditions of Exponential Rayleigh distribution

A reliability single sampling plan is a process used to make decisions about submitted lots by conducting a life test on randomly selected items. A single sampling plan by variables assumes a non-normal Inverse Gaussian distribution for the quality characteristic and develops a procedure for determining plan parameters based on specified quality levels [8]. It is characterized by four parameters (N, n, c, t): lot size (N), sample size (n), acceptance number (c), and test termination time (t). A new sampling plan based on the Generalized Poisson Distribution is proposed and studied for its performance measures in lot acceptance [9]. The plan can be implemented by selecting items from the lot according to these parameters. Economic Reliability Test Plan (ERTP) is developed considering that the lifetime of the submitted items follows a generalized exponential distribution. Test termination time is calculated for a given group size, defined acceptance number, and producer's risk [10].

(1) Choose a random selection of n products from the submitted lot of size N.
(2) Conduct the life test for the selected items considering t as the test terminated time. Observe the number of failed items X=x.
(3) Terminate the life test, if either at time of t or X>c before reaching time t, whichever is earlier.
(4) Accept the lot, if x ≤ c at time t; reject the lot if x>c either at time t or earlier.

Let T be the lifetime of the product, which is distributed according to an exponential Rayleigh distribution having the probability density function (PDF)

$$f(x) = \lambda\beta x e^{\frac{\beta}{2}x^2} . e^{-\lambda(e^{\frac{\beta}{2}x^2}-1)} \qquad x\varepsilon R;\ \lambda,\beta > 0 \tag{1}$$

Here λ and β are the shape and scale parameters respectively. The cumulative distribution function of the exponential Rayleigh distribution is given by,

$$F(x) = 1 - \lambda e^{-\lambda(e^{\frac{\beta}{2}x^2}-1)} \qquad x\varepsilon R;\ \lambda, \beta > 0 \tag{2}$$

Estimate to a parameter β respectively

$$\beta = \frac{2}{m^2} log\ (\frac{1 - log\ (\frac{1}{2})}{\lambda})$$

The lot fraction nonconforming, n, p, can be calculated corresponding to each value of 1/m from

$$F(X) = F\left(\frac{1}{m}\right) = p$$

The performance of a sampling plan may be analyzed using their OC functions. The OC function of a sampling plan is given by

$$P_a = P\ (x \leq c) = \sum_{x=0}^{c} P\ (X = x)$$

The probability distribution of X can be assumed appropriately as hyper-geometric distribution. When n/N ≤0.10, n is large and p is small such that np<5, the sampling distribution of X can be approximated by the Poisson (np) distribution [11]. In light of these facts, it is suggested here

$$P_a(p) = \sum_{x=0}^{c} \frac{e^{-np}(np)^x}{x!}$$

IV. Determination of Plan Parameters under the conditions of Exponential Rayleigh Distribution

The most reliable single sampling plans are established for ER (λ, θ) distribution by using the Binomial probability distribution's OC function. A Special Type of Double Sampling (STDS) plan has been proposed to emphasize the importance of acceptance sampling plans in ensuring product quality. This plan uses the Generalized Poisson Distribution to achieve the same level of acceptance with fewer samples than a single sampling plan [12]. A sampling plan is designed to protect both the producer and consumer simultaneously. To ensure protection, two points are specified on the OC curve: (p_1, 1-α) and (p_2, β). Here, p_1 represents acceptable quality, α represents the producer's risk, p_2 represents limiting quality, and β represents the consumer's risk. Derivation of the Operating Characteristic (OC) function of the sampling plan, which describes its performance in terms of the probability of accepting or rejecting a batch based on the observed number of defects. The plan parameters are determined for specific sets of values for (p_1, α, p_2, β), which are parameters of the ZIP distribution [13]. An optimal RSSP can be found based on the points that meet the following requirements.

$$P_a(p_1) \geq 1 - \alpha$$

and

$$P_a(p_2) \leq \beta$$

These conditions may be written as

$$\sum_{x=0}^{c} \frac{e^{-np_1}(np_1)^x}{x!} \geq 1-\alpha \quad (3)$$

and

$$\sum_{x=0}^{c} \frac{e^{-np_1}(np_1)^x}{x!} \leq \beta \quad (4)$$

There are various ways to figure out the best values for n and c while adhering to conditions (3) and (4). Economic Reliability Test Plans (ERTP) is proposed considering that the lifetime of the submitted items follows the Pareto distribution of the second kind [14]. To determine the plan parameters, an iterative process is used as outlined below. In summary, when given λ, t, m_1, m_2, α, and β, the most effective values for n and c can be found using the following steps.

(1) For specified values of m_1 and m_2 with $m_1 > m_2$, calculate $\beta_1 = \frac{2}{{m_1}^2} \log\left(\frac{1-\log\left(\frac{1}{2}\right)}{\lambda}\right)$ and $\beta_2 = \frac{2}{{m_2}^2} \log\left(\frac{1-\log\left(\frac{1}{2}\right)}{\lambda}\right)$

(2) Corresponding to t, β_1 and β_2, determine $p_1 = F_T(1/m_1)$ and $p_2 = F_T(1/m_2)$

(3) Set c=0

(4) Find the largest n, say n_L, such that $P_a(p_1) \geq 1-\alpha$

(5) Find the smallest n, say n_S, such that $P_a(p_2) \leq \beta$

(6) If $n_S \leq n_L$, then the optimum plan is (n_S, c); otherwise increase c by 1.

(7) Till the optimum values of n and c are attained, repeat steps 4 through 6.

A submitted lot may undergo the sampling examination after n and c have been established using the hybrid censoring procedures outlined in section 2.

V. Construction of Tables

The values of n and c of the optimum reliability sampling plans are determined using Poisson probabilities for the combination of λ, t, m_1, m_2, α, and β. The producer's risk and consumer's risk are considered at two different levels such as α=0.05, 0.05 and β=0.05, 0.10. The producer's estimated range of mean product lifetime is taken as m_1=4000, 4500, 5000, 5500, 6000, 6500, 7000, and 7500 hours respectively. Two different levels of test termination time and one value for shape parameter λ as assumed as t=200, 350, 500, and 650 hours and λ=1 respectively. The consumer's projected mean product lifespan is taken as m_2 = 750, 1000, 1250, 1500, 1750, 2000, 2250 hours respectively. The optimum reliability sampling plans' n and c values are shown in Tables 1 through Table 4. Each cell entry (n, c) in every table reflects the ideal value of the pair (n, c) that corresponds to the given values of λ, t, m_1, m_2, α, and β. The Selection of plans from these for the given requirements is described in the following illustration.

Illustration

Let the lifetime of the products submitted for inspection be distributed according to ER (1, β). The mean lifetime of the products meeting the expectation of the producer and consumer are respectively m_1=4000 hours and m_2=2250 hours. Suppose that the quality inspector prescribes to censor the life test at t=500 hours. Then, the values of acceptable quality level and limiting quality level can be computed as p_1=0.0082 and p_2=0.0260. If the producer's risk and the consumer's risk as α=0.05 and β=0.05, then the plan parameters may be obtained using Poisson probabilities from Table 3 as n=556 and c=8.

Now, the life-test-based lot-by-lot sampling inspection can be carried out as follows: A

sample of 556 products may be selected randomly from the submitted lot. Life tests may be conducted on all the sampled products. The life test may be stopped after 500 hours if there have been no more than 8 failures. The lot might be approved. On the other hand, if the ninth failure occurs before t=500 hours, terminate the life test. The lot may be rejected.

Table 1: *Parameters of RSSPs under the conditions of ER (β, λ=1) Distribution with α=0.05, λ=1 and t=200 hours.*

		m1	4000	4500	5000	5500	6000	6500	7000	7500
t=200, λ=1		t/m1	0.05	0.0444	0.04	0.0363	0.0333	0.0307	0.0285	0.0266
		P1	0.0013	0.0010	0.0008	0.0006	0.0005	0.0004	0.0004	0.0003
m2	t/m2	P2								
750	0.2666	0.0374	(104,1)	(104,1)	(104,1)	(62,0)	(62,0)	(62,0)	(62,0)	(62,0)
			(127,1)	(127,1)	(127,1)	(127,1)	(81,0)	(81,0)	(81,0)	(81,0)
1000	0.2	0.0210	(185,1)	(185,1)	(185,1)	(185,1)	(185,1)	(185,1)	(110,0)	(110,0)
			(226,1)	(226,1)	(226,1)	(226,1)	(226,1)	(226,1)	(226,1)	(226,1)
1250	0.16	0.0134	(395,2)	(289,1)	(289,1)	(289,1)	(289,1)	(289,1)	(289,1)	(289,1)
			(468,2)	(468,2)	(352,1)	(352,1)	(352,1)	(352,1)	(352,1)	(352,1)
1500	0.1333	0.0093	(569,2)	(569,2)	(416,1)	(416,1)	(416,1)	(416,1)	(416,1)	(416,1)
			(829,3)	(673,2)	(673,2)	(507,1)	(507,1)	(507,1)	(507,1)	(507,1)
1750	0.1142	0.0068	(972,3)	(774,2)	(740,2)	(774,2)	(566,1)	(566,1)	(566,1)	(566,1)
			(1331,4)	(1128,3)	(916,2)	(916,2)	(916,2)	(690,1)	(690,1)	(690,1)
2000	0.1	0.0052	(1762,5)	(1269,3)	(1269,3)	(1011,2)	1011,2)	(1011,2)	(739,1)	(739,1)
			(2249,6)	(1739,4)	(1473,3)	(1473,3)	1196,2)	(1196,2)	(1196,2)	(901,1)
2250	0.0888	0.0041	(2830,7)	(2230,5)	(1606,3)	(1606,3)	1280,2)	(1280,2)	(1280,2)	(935,1)
			(3470,8)	(2847,6)	(2200,4)	(1864,3)	1864,3)	(1514,2)	(1514,2)	(1514,2)

In each cell, the first pair is the value of (n, c) corresponding to (α=0.05, β=0.10) and the Second pair corresponding to (α=0.05, β=0.05).

Table 2: *Parameters of RSSPs under the conditions of ER (β, λ=1) Distribution with α=0.05, λ=1 and t=350 hours.*

		m1	4000	4500	5000	5500	6000	6500	7000	7500
t=350, λ=1		t/m1	0.0875	0.0777	0.07	0.0636	0.0538	0.0538	0.05	0.0466
		P1	0.0040	0.0031	0.0025	0.0021	0.0017	0.0015	0.0013	0.0011
m2	t/m2	P2								
750	0.4666	0.1144	(34,1)	(34,1)	(34,1)	(21,0)	(21,0)	(21,0)	(21,0)	(21,0)
			(42,1)	(42,1)	(42,1)	(42,1)	(27,0)	(27,0)	(27,0)	(27,0)
1000	0.35	0.0644	(61,1)	(61,1)	(61,1)	(61,1)	(61,1)	(61,1)	(36,0)	(36,0)
			(74,1)	(74,1)	(74,1)	(74,1)	(74,1)	(74,1)	(74,1)	(74,1)
1250	0.28	0.0412	(129,2)	(95,1)	(95,1)	(95,1)	(95,1)	(95,1)	(95,1)	(95,1)
			(153,2)	(153,2)	(115,1)	(115,1)	(115,1)	(115,1)	(115,1)	(115,1)
1500	0.2333	0.0286	(186,2)	(186,2)	(136,1)	(136,1)	(136,1)	(136,1)	(136,1)	(136,1)
			(271,3)	(220,2)	(220,2)	(166,1)	(166,1)	(166,1)	(166,1)	(166,1)
1750	0.2	0.0210	(318,3)	(253,2)	(253,2)	(253,2)	(185,1)	(185,1)	(185,1)	(185,1)
			(435,4)	(369,3)	(299,2)	(299,2)	(299,2)	(226,1)	(226,1)	(226,1)
2000	0.17	0.0161	(576,5)	(415,3)	(415,3)	(331,2)	(331,2)	(331,2)	(242,1)	(242,1)
	5		(735,6)	(568,4)	(481,3)	(481,3)	(391,2)	(391,2)	(391,2)	(295,1)
2250	0.1555	0.0127	(924,7)	(728,5)	(525,3)	(525,3)	(418,2)	(418,2)	(418,2)	(306,1)
			(1133,8)	(930,6)	(719,4)	(609,3)	(609,3)	(495,2)	(495,2)	(495,2)

In each cell, the first pair is the value of (n, c) corresponding to (α=0.05, β=0.10) and the Second pair corresponding to (α=0.05, β=0.05).

Table3: *Parameters of RSSPs under the conditions of ER (β, λ=1) Distribution with α=0.05, λ=1 and t=500 hours.*

		m1	4000	4500	5000	5500	6000	6500	7000	7500
t=500, λ=1		t/m1	0.125	0.1111	0.1	0.0909	0.0833	0.0769	0.0714	0.0666
		P1	0.0082	0.0065	0.0052	0.0043	0.0036	0.0031	0.0026	0.0023
m2	t/m2	P2								
750	0.6666	0.2317	(17,1)	(17,1)	(17,1)	(10,0)	(10,0)	(10,0)	(10,0)	(10,0)
			(21,1)	(21,1)	(21,1)	(21,1)	(13,0)	(13,0)	(13,0)	(13,0)
1000	0.5	0.1312	(30,1)	(30,1)	(30,1)	(30,1)	(30,1)	(30,1)	(18,0)	(18,0)
			(37,1)	(37,1)	(37,1)	(37,1)	(37,1)	(37,1)	(37,1)	(37,1)
1250	0.4	0.0841	(64,2)	(47,1)	(47,1)	(47,1)	(47,1)	(47,1)	(47,1)	(47,1)
			(75,2)	(75,2)	(57,1)	(57,1)	(57,1)	(57,1)	(57,1)	(57,1)
1500	0.3333	0.0584	(92,2)	(92,2)	(67,1)	(67,1)	(67,1)	(67,1)	(67,1)	(67,1)
			(133,3)	(108,2)	(108,2)	(108,2)	(82,1)	(82,1)	(82,1)	(82,1)
1750	0.2857	0.0429	(156,3)	(124,2)	(124,2)	(124,2)	(91,1)	(91,1)	(91,1)	(91,1)
			(214,4)	(181,3)	(147,2)	(147,2)	(147,2)	(111,1)	(111,1)	(111,1)
2000	0.25	0.0329	(282,5)	(204,3)	(204,3)	(162,2)	(162,2)	(162,2)	(119,1)	(119,1)
			(360,6)	(279,4)	(236,3)	(236,3)	(192,2)	(192,2)	(192,2)	(145,1)
2250	0.2222	0.0260	(453,7)	(357,5)	(257,3)	(257,3)	(205,2)	(205,2)	(205,2)	(150,1)
			(556,8)	(456,6)	(353,4)	(299,3)	(299,3)	(243,2)	(243,2)	(243,2)

In each cell, the first pair is the value of (n, c) corresponding to (α=0.05, β=0.10) and the second pair corresponding to (α=0.05, β=0.05).

Table 4: *Parameters of RSSPs under the conditions of ER (β, λ=1) Distribution with α=0.05, λ=1 and t=650 hours.*

		m1	4000	4500	5000	5500	6000	6500	7000	7500
t=650, λ=1		t/m1	0.1625	0.1444	0.13	0.1181	0.1083	0.1	0.0928	0.0866
		P1								
m2	t/m2	P2	0.0139	0.0109	0.0088	0.0073	0.0061	0.0052	0.0045	0.0039
750	0.8666	0.3844	(11,1)	(11,1)	(11,1)	(6,0)	(6,0)	(6,0)	(6,0)	(6,0)
			(13,1)	(13,1)	(13,1)	(13,1)	(8,0)	(8,0)	(8,0)	(8,0)
1000	0.65	0.2205	(18,1)	(18,1)	(18,1)	(18,1)	(18,1)	(18,1)	(11,0)	(11,0)
			(22,1)	(22,1)	(22,1)	(22,1)	(22,1)	(22,1)	(22,1)	(22,1)
1250	0.52	0.1418	(38,2)	(28,1)	(28,1)	(28,1)	(28,1)	(28,1)	(28,1)	(28,1)
			(45,2)	(45,2)	(34,1)	(34,1)	(34,1)	(34,1)	(34,1)	(34,1)
1500	0.4333	0.0987	(54,2)	(54,2)	(54,2)	(40,1)	(40,1)	(40,1)	(40,1)	(40,1)
			(79,3)	(64,2)	(64,2)	(64,2)	(49,1)	(49,1)	(49,1)	(49,1)
1750	0.3714	0.0725	(93,3)	(74,2)	(74,2)	(74,2)	(54,1)	(54,1)	(54,1)	(54,1)
			(127,4)	(107,3)	(87,2)	(87,2)	(87,2)	(66,1)	(66,1)	(66,1)
2000	0.325	0.0555	(167,5)	(121,3)	(121,3)	(96,2)	(96,2)	(96,2)	(70,1)	(70,1)
			(214,6)	(165,4)	(140,3)	(140,3)	(114,2)	(114,2)	(114,2)	(86,1)
2250	0.2888	0.0439	(268,7)	(212,5)	(153,3)	(153,3)	(122,2)	(122,2)	(122,2)	(89,1)
			(329,8)	(270,6)	(209,4)	(177,3)	(177,3)	(144,2)	(144,2)	(144,2)

In each cell, the first pair is the value of (n, c) corresponding to (α=0.05, β=0.10) and the second pair corresponding to (α=0.05, β=0.05).

VI. Conclusion

Under the assumption that the lifespan quality feature is modeled by an Exponential-Rayleigh distribution, a method for determining single sampling plans for life tests is derived. The paper introduces a sampling plan for a truncated life test, specifically for the Exponential Rayleigh distribution with parameters λ and β. The number of groups and acceptance number are provided for the special case where λ=1 and the consumer's and producer's risk plan parameters are specified. Tables are also included to aid in selecting the optimal plan parameters for products with Exponential Rayleigh distribution, ultimately reducing test time and cost.

References

[1] Abdullah M Almarashi, Khushnoor Khan and et al., (2021). Group Acceptance Sampling Plan Using Marshall-Olkin Kumaraswamy Exponential (MOKw-E) Distribution. *Process*, 9: 1066.

[2] Aiman Fikri Jamaludin, Zakiyah Zain and et al., (2016). A Modified Group Chain Sampling Plans for Lifetimes Following a Rayleigh Distribution. *Global Journal of Pure and Applied Mathematics*, 12: 3941-3947.

[3] Amer Ibrahim, Al-Omari and et al., (2021). Acceptance Sampling Plans under two-parameter Quasi Shanker distribution assuring mean life with an application to manufacturing data. *Science Progress*, 104(2): 1-17.

[4] Kaviyarasu V and Sivasankari S (2021). Contribution to time truncating reliability single sampling plan for weighted Rayleigh distribution. *International Journal of Statistics and Applied Mathematics*, 6:2456-1452.

[5] Ramkumar T.B and Sanjana O.K (2009-2011). Reliability Test Plans Based on Burr Distribution from Truncated Life Tests. *International Journal of Mathematics and Computer Applications Research*, 1:28-40.

[6] Baloul Jamkhaneh E, Sadeghpour-Glideh and et.al, (2008). Acceptance Single Sampling Plan by Using of Poisson Distribution. *The Journal of Mathematics and Computer Science*, 1:6-13.

[7] Farooq M.M and Bashir A (2023). Acceptance sampling plan using new truncated Weibull-X family based on run lengths of conforming items. *Scientia Iranica E*, 30(1):273-284.

[8] Geetha S (2016). A Procedure for Selection of Single Sampling Plans by Variables Based on Inverse Gaussian distribution. *International Journal of Scientific Research*, 5:2277-8179.

[9] Kaviyarasu Velappan and Devika Veerasangili (2018). Designing Single Sampling Plan Using Generalized Poisson distribution. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 5:226-232.

[10] Jaffer Hussian, Abdur Razzaque Mughal, et al., (2011). Economic Reliability Group Acceptance Sampling Plans for Lifetimes Following a Generalized Exponential Distribution. *Electronic Journal of Applied Statistical Analysis*, 4: 124-130.

[11] Azarudheen S and Pradeepa Veerakumari (2019). Selection of Tightened-Normal-Tightened sampling scheme under the implications of intervened Poisson distribution. *Pakistan Journal of Statistics and Operation Research*, 1:129-140.

[12] Kaviyarasu V and Devika V (2017). Designing of Special Type of Double Sampling Plan for Compliance Testing through Generalized Poisson distribution. *International Journal of Pure and Applied Mathematics*, 117: 7-17.

[13] Loganathan A and Shalini K (2014). Determination of Single Sampling Plans by Attributes Under the Conditions of Zero-Inflated Poisson distribution. *Communication in Statistics-Simulation and Computation.* 43:538-548.

[14] Muhammad Aslam, Abdur Razzaque Mughal, et al., (2010). Economic Reliability Group Acceptance Sampling Based on Truncated Life Tests Using Pareto Distribution of the Second Kind. *Communications of the Korean Statistical Society*, 17: 725-731.

[15] Muhammad Shoaib, Muhammad Aslam, et al., (2011). Acceptance Decision Rule for Multiple Items under the Truncated Life Test for Birnbaum-Saunders Distribution Percentiles. *World Applied Sciences Journal,* 12: 1745-1753.

[16] Navjeet Singh, Gurcharan Singh and et al., (2022). Repetitive Acceptance Sampling Plan for Lifetimes Following a Skew-Generalized Inverse Weibull Distribution. *Pesquisa Operacional,* 42: 1-13.

[17] Srinivasa Rao Boyapati and Sricharani Poranki (2018). Limited Failure Censored Life Test Sampling Plan in Dagum Distribution. *American Journal of Applied Mathematics and Statistics,* 6: 181-185.

[18] Encyclopedic Britannica (1888). The Problem of the Random Walk. The Wave Theory of Light. *Nature,* 72:318.

[19] Sijbers J, den Dekker, A J, Raman E, Van Dyck D (1999). Parameter estimation from magnitude MR images. *International Journal of Imaging Systems and Technology.* 10(2): 109-114.

[20] den Dekker A J, Sijbers J (2014). Data distributions in magnetic resonance image a review. *Physica Medica*. 30(7):725-741.

[21] Ahmadi, Hamed (2017). A mathematical function for the description of nutrient-response curve. *PLOS ONE,* 12(11).

REDESCENDING M-ESTIMATOR BASED LASSO FOR FEATURE SELECTION

R. MUTHUKRISHNAN[1] AND C. K. JAMES[2]

•

[1,2]Department of Statistics
Bharathiar University, Tamil Nadu
Coimbatore-641046, India
[1]muthukrishnan1970@gmail.com, [2]jamesck74@gmail.com

Abstract

Aim: Regression analysis is one of the statistical methods which helps to model the data and helps in prediction, a large data set with higher number of variables will often create problem due to its dimensionality and hence create difficulties to gather important information from the data, so it is a need of a method which can simultaneously choose important variables which contains most of the information and hence helps to fit the model. Least absolute shrinkage and selection operator (LASSO) is a popular choice for shrinkage estimation and variable selection. But LASSO uses the conventional least squares technique for feature selection which is very sensitive to outliers. As a result, when the data set is contaminated with bad observations (Outliers), the LASSO technique gives unreliable results, so in this paper the focus is to create a method which can resist to outliers in the data and helps in giving a meaningful result. Method: proposed a new procedure, a LASSO method by adding weights which uses the concept of redescending M-estimator, which can resist outliers in both dependent and independent variables. The observation with greater importance receives a higher weight and less weight to the least important observation. Findings: The efficiency of the proposed method has been studied in the real and simulation environment and compared with other existing procedures with measures like Median Absolute Error (MDAE), False Positive Rate (FPR), False Negative Rate (FNR), Mean Absolute Percentage Error (MAPE). The proposed method with the redescending M-estimator shows a higher resistance to outliers compared to conventional LASSO and other robust existing procedures. Conclusion: The study reveals that the proposed method outperforms other existing procedures in terms of MDAE, FPR, FNR and MAPE, indicating its superior performance in variables selection under outlier contaminated datasets.

Keywords: Feature Selection, LASSO, MAPE

I. Introduction

One of the most frequent problems we run into in real-time applications and other scientific fields is data including outliers. The existence of outliers, according to Chatterjee and Hadi [4], may leads to influence the parameter estimation and inaccurate predictions for traditional approaches. Both dependent variables and covariates (predictor variables) may contain outliers. As a result, it's crucial to deal with outliers in regression analysis. For outlier detection problems, numerous robust regression algorithms have been created, such as S- estimator [7], the least median of squares estimator [11], the MM- estimator [18], the τ- estimator [19], and so on. It is well known that there are some M- estimator-based regression methods such as Huber regression which does not delete large residuals, and Tukey regression is not Robust against the outliers in the leverage points. [12], Redescending M-estimators are more resilient than M-estimators since they totally reject extreme outliers. Alamgir et al. [1] proposed an efficient Redescending M-estimator for robust estimation.

In practice, a large number of variables are often incorporated at the beginning of modelling. The

interpretation of models that contain all of the variables is extremely difficult, even when irrelevant factors may increase variance. Therefore, one of the most significant issues in data analysis is the choice of an important variable. Popular methods for variable selection are penalized regression methods such as Least Absolute Shrinkage and Selection Operator (LASSO) [13], Smoothly Clipped Absolute Deviation (SCAD) [6] and adaptive LASSO [20], and so on. Most of the methods mentioned above are closely related to Ordinary Least Squares (OLS) technique. OLS-based methods are not resistant to outliers therefore the outliers can cause problems in a variable selection based on OLS. So, the robust variable selection approach has to be studied. There are numerous effective variable selection techniques in the literature, including the Least Absolute Deviation (LAD)-LASSO [15], which deals with heavy-tailed error, WLAD-LASSO [14], the weighted Wilcoxon-type SCAD method [13], the Huber's criterion and adaptive lasso penalty [9], the quantile regression for analyzing heterogeneity ultra-high dimension [16], the variable selection in the semiparametric varying- coefficient partially linear model via a penalized composite quantile loss [8], the Composite Quantile Regression (CQR) [21], the variable selection with the exponential squared loss [17], Penalized least trimmed square (LTS) [2], Maximum Tangent Likelihood Estimator (MTE) [10], and so on. In this paper, a new robust feature selection method has been introduced. The improved version of the LASSO method uses a weight from a Redescending M-Estimator which can tolerate outliers in the X-Y space. The study based on simulation and real data indicates that the proposed robust feature selection procedure performs better than other existing methods.

The paper is organized as follows. In section II provides brief introduction to LASSO method. A new technique, Alarm Weight LASSO (AW-LASSO) and its corresponding algorithm is described in Section III. In section IV, real data analysis and a simulation study and is carried out to comprehend how well the suggested method works. Finally, section V gives summary and conclusion.

II. LASSO Methods

Regressions models are commonly used in statistical analysis. A popular use is to model the predicted risk of a likely outcome. Unfortunately, using standard regression techniques to create a model from a set of candidate variables often results in overfitting, which increases the number of variables that are eventually included in the model and overestimates how well the included variables explain the observed variability (an effect known as optimism bias). Extreme (extremely low or very high) risk observations are particularly difficult for the model to forecast. A shrinkage and variable selection strategy for regression models is LASSO regression. In order to create a model that minimizes the prediction error, LASSO regression seeks to discover the variables and corresponding regression coefficients. This is accomplished by placing a restriction on the model parameters that causes the regression coefficients to decrease towards zero, or more specifically, by requiring that the total absolute value of the regression coefficients be smaller than a predetermined value (λ). The equation of the LASSO is given below

$$\beta_{LASSO} = \min_{\beta} \left(\sum_{i=1}^{n} \left(y_i - \sum_j \beta_j X_{ij} \right)^2 + \lambda \sum_j |\beta_j| \right) \quad (1)$$

Since λ controls the amount of regularization the choice of λ is often made by using an automated k-fold cross validation approach. If $\lambda = 0$ the LASSO is same as OLS. As λ increases, the number of non-zero components of β decreases, at $\lambda = \infty$, the LASSO gives the null model. The above LASSO method is based on OLS loss function which is not resistant to outliers, therefore, to address this issue, we modified the LASSO by adding a new weight to form Alarm Weight LASSO which is Elaborately discussed in the section III.

III. Alarm Weight LASSO

Consider the linear regression model

$$y_i = \beta_o + x_i^T \beta + \varepsilon_{i,} \qquad i = 1,2,3,...,n \tag{2}$$

where Y_i is the response variable $X_i = (x_{i1}, x_{i2}, ..., x_{ip})^T$ is the p-dimensional covariate vector, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^T$ are the regression parameters, and ε_i are the iid random errors. We assume that $\beta_o = 0$. This can be achieved by centering the covariates and response variable. That is, from now on we will consider the model

$$y_i = x_i^T \beta + \varepsilon_{i,} \qquad i = 1,2,3,...,n \tag{3}$$

To estimate β is to minimize the ordinary least square (OLS) criterion

$$\sum_{i=1}^{n} (y_i - x_i^T \beta)^2 \tag{4}$$

The OLS estimates β by minimizing the error sum of squares, i.e.,

$$\hat{\beta}_{OLS} = \min_{\beta} \left\{ (Y - X\beta)^T (Y - X\beta) \right\} \tag{5}$$

The OLS approach to estimate the regression parameter is very sensitive to the outliers. One of the alternatives to OLS is to use weighted OLS. Weighted regression is its robustness against outliers. Weighted regression can assign less weight to outliers and hence reduce their impact on the estimate of the coefficients. Which is obtained by minimizing the OLS criterion

$$\sum_{i=1}^{n} w_i (y_i - x_i^T \beta)^2 \tag{6}$$

where w_i, for i = 1,2,3,…,n is the weights which is determined by a redescending M-Estimator which can resist outliers in both X and Y space. The influence function describes the sensitivity of the overall estimate of the outlying data and is defined as

$$\psi(r) = \begin{cases} \dfrac{16re^{-2(r/c)^2}}{(1+e^{-(r/c)^2})^4} & , |r| \le c \\ 0 & , |r| > c \end{cases} \tag{7}$$

The functional relationship between ψ and ϱ is given by

$$\psi(r) = \frac{d}{dr}\rho(r) \tag{8}$$

Integrating out the ψ - function under the initial condition, we get the corresponding $\rho(r)$, given by

$$\rho(r) = \begin{cases} \dfrac{2c^2}{3}\left[1 - \dfrac{2(1+3e^{(r/c)^2})}{(1+e^{(r/c)^2})^3}\right] & , |r| \le c \\ \dfrac{2c^2}{3}\left[1 - \dfrac{2(1+3e)}{(1+e)^3}\right] & , |r| > c \end{cases} \tag{9}$$

The weight function w(r)= ψ (r)/r, is as follows

$$w(r)=\begin{cases}\dfrac{(4e^{-(r/c)^2})^2}{(1+e^{-(r/c)^2})^4} & ,|r|\leq c\\ 0 & ,|r|>c\end{cases} \tag{10}$$

where "*r*" denotes residual and "*c*" is tuning constant.

Efficiency and robustness are two characteristic of a robust procedures that are inversely connected. As a result, choose an estimator with the highest resistance and the lowest efficiency loss. Nobody can afford to select a highly robust estimator that is resistant to outliers at the expense of decreasing efficiency. These two properties should be balanced in some way. The weight function ensures that the residuals with the highest weight (close to 1) correspond to the majority of good observations. Figure 1, represents the weight function w(r) of redescending M-estimator

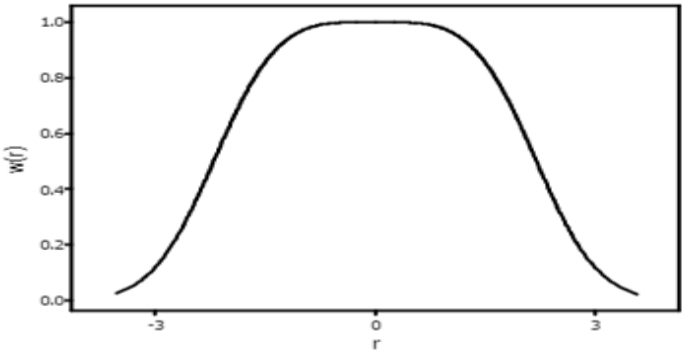


Fig.1: Alarm Weight function of redescending M estimator

From Figure 1, As can be seen, only severely outlying observations are given 0 weights, ensuring that good observations are used to their full potential and that extreme outliers are not overly relied upon. The equation given in (1) thus modified by adding the weight of the redescending M estimator to form Alarm Weight LASSO and the equation is given below.

$$\beta_{AW-LASSO}=\min_{\beta}\left(\sum_{i=1}^{n}w_i\left(y_i-\sum_j\beta_jX_{ij}\right)^2+\lambda\sum_j|\beta_j|\right) \tag{11}$$

w_i represents the weight function and λ is the tuning parameter which is chosen by cross validation method.

I. Computational Algorithm

Consider the LASSO model given in (1). We use Iteratively Least Square Method (IRWLS) algorithm for the computation of the AW-LASSO.

Step1: Find the initial estimates of β by using the ridge regression model.

Step2: obtain the corresponding residuals from our initial estimates

Step3: Compute the corresponding weights based on the proposed weight function

Step4: Calculate the new estimate of AW-LASSO coefficients using the IRWLS algorithm

Step5: Repeat step 2 to 4 until convergence.

IV. Experimental Results

In this section various LASSO-type feature selection techniques are compared to the proposed method ology in a real-world setting. Outliers were present in the real data, which were eliminated using the cook distance [5], and the analysis was done using the R programming language. The obtained results such as Median Absolute Error (MDAE), Mean Absolute Percentage Error (MAPE), number of variabl es, for with and without outliers are also discussed.

I. Real data examples

Here we considered two data sets, namely Boston housing data and diabetes data, detailed descriptions are available in the standard packages. The Boston data set has 506 observations of 15 independent and a dependent variable. The diabetes data set has 442 observations of 9 independent and a dependent variable. The feature selection procedures have been performed after standardizing the variables, with and without outliers and the results are summarised in the table 1.

Table 1: Error values, under with and without outliers

Methods	MDAE	MAPE	No. of Variables Selected	Selected Variables
			Boston Housing Data	
LASSO	0.303 (0.123)	4.27 (0.783)	12(12)	Tract, ion, lat, Crim, Zn, nox, rm, dis, tax, ptratio, b, istat (ion, crim, zn, indus, nox, rm, age, dis, tax, ptratio, b, istat)
LAD LASSO	0.267 (0.177)	4.07 (1.20)	12(15)	Tract, ion, crim, zn, nox, rm, age, dis, tax, ptratio, b, istat (Tract, ion, lat, crim, zn, indus, nox, rm, age, dis, rad, tax, ptratio, b, istat)
Huber LASSO	0.291 (0.189)	3.97 (1.17)	11(12)	Tract, ion, crim, nox, rm, age, dis, tax, ptratio, b, istat (Tract, ion, crim, zn, indus, rm, age, dis, tax, ptratio, b, istat)
MTE LASSO	0.306 (0.245)	4.24 (1.32)	10(11)	Tract, ion, crim, nox, rm, dis, tax, ptratio, b, istat (ion, crim, zn, indus, rm, age, dis, tax, ptratio, b, istat)
AW-LASSO	0.301 (0.124)	3.59 (0.783)	6(11)	ion, rm, tax, ptratio, b, istat (ion, crim, zn, indus, rm, age, dis, tax, ptratio, b, istat)
			Diabetes Data	
LASSO	0.516 (0.511)	1.15 (1.48)	4(4)	BMI, BP, S3, S5, (BMI, BP, S3, S5)
LAD LASSO	0.490 (0.527)	1.29 (1.57)	4(4)	BMI, BP, S3, S5, (BMI, BP, S3, S5)
Huber LASSO	0.501 (0.510)	1.23 (1.53)	4(4)	BMI, BP, S3, S5, (BMI, BP, S3, S5)
MTE LASSO	0.517 (0.504)	1.15 (1.46)	4(4)	Bmi, bp, S3, S5, (BMI, BP, S3, S5)
AW-LASSO	0.485 (0.510)	1.07 (1.49)	4(4)	bmi, BP, S3, S5, (BMI, BP, S3, S5)

(.) without outlier

From the above table it is observed that under with and without outliers the proposed procedure, AW-LASSO produces the error values are minimum and also select the significant variables when compared with the other procedures.

II. Simulation Study

Simulation studies are carried out to check the efficacy of various methods. In our simulation study, the covariates are generated from a multivariate normal distribution with Mean $\mu = [o]_{p*1}$ and variance

$\Sigma = [\sigma ij] = \rho^{|i-j|}$ for various levels of correlation, ρ = 0.01, 0.5, 0.9 and number of variables, p=10, 15, 25. The model consists of six significant variables and the rest is considered noise variables. The performance of the proposed method is compared with various other robust methods along with the classical LASSO method. Various levels of contamination (0%, 5%, 10%, 20%) are studied for sample size n = 100, 200, 1000.

Table 2: False Negative and False Positive rate of each method under various levels of contamination (0% and 5%)

			η = 0									η = 5								
			p = 10			p = 15			p = 25			p = 10			p = 15			p = 25		
			P			ρ			ρ			ρ			ρ			ρ		
Method	N	Error	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90
LASSO	100	FPR	0.36	1.02	0.92	0.64	0.75	0.69	0.43	0.55	0.54	0.93	0.99	0.92	0.69	0.73	0.66	0.47	0.56	0.47
		FNR	0.00	0.00	0.03	0.00	0.00	0.05	0.00	0.00	0.08	0.00	0.00	0.07	0.00	0.00	0.13	0.00	0.01	0.20
	200	FPR	0.96	0.99	1.03	0.68	0.72	0.74	0.46	0.57	0.57	0.88	1.00	1.00	0.66	0.76	0.75	0.45	0.57	0.52
		FNR	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.00	0.03	0.00	0.00	0.07
	1000	FPR	0.90	1.02	0.99	0.67	0.74	0.78	0.44	0.53	0.58	0.96	1.01	1.03	0.62	0.72	0.77	0.42	0.54	0.58
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LAD LASSO	100	FPR	0.59	0.73	0.77	0.27	0.42	0.43	0.10	0.18	0.23	0.49	0.67	0.66	0.22	0.31	0.28	0.06	0.09	0.15
		FNR	0.00	0.01	0.16	0.03	0.06	0.26	0.16	0.18	0.41	0.04	0.05	0.26	0.13	0.14	0.46	0.00	0.40	0.62
	200	FPR	0.67	0.78	0.87	0.31	0.38	0.45	0.09	0.16	0.27	0.49	0.70	0.79	0.20	0.30	0.40	0.06	0.12	0.18
		FNR	0.00	0.00	0.04	0.00	0.00	0.15	0.00	0.04	0.22	0.01	0.00	0.12	0.02	0.04	0.22	0.16	0.14	0.42
	1000	FPR	0.58	0.78	0.81	0.28	0.42	0.52	0.09	0.17	0.29	0.55	0.69	0.78	0.17	0.30	0.43	0.07	0.11	0.22
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.08
Huber LASSO	100	FPR	0.56	0.71	0.74	0.24	0.37	0.40	0.09	0.15	0.20	0.47	0.60	0.63	0.21	0.27	0.25	0.05	0.07	0.12
		FNR	0.00	0.01	0.15	0.03	0.06	0.25	0.17	0.19	0.40	0.04	0.05	0.25	0.13	0.15	0.46	0.47	0.42	0.62
	200	FPR	0.64	0.73	0.82	0.28	0.33	0.39	0.08	0.14	0.22	0.48	0.67	0.74	0.18	0.27	0.36	0.06	0.10	0.15
		FNR	0.00	0.00	0.04	0.00	0.00	0.12	0.00	0.04	0.20	0.01	0.00	0.11	0.02	0.04	0.21	0.15	0.14	0.43
	1000	FPR	0.55	0.74	0.77	0.26	0.39	0.46	0.09	0.15	0.23	0.55	0.64	0.75	0.16	0.27	0.39	0.06	0.09	0.18
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.04
MTE LASSO	100	FPR	0.51	0.69	0.73	0.15	0.29	0.39	0.05	0.08	0.18	0.33	0.47	0.54	0.11	0.15	0.19	0.02	0.03	0.10
		FNR	0.35	0.16	0.22	0.69	0.40	0.33	0.75	0.66	0.54	0.55	0.38	0.40	0.79	0.66	0.64	0.91	0.87	0.73
	200	FPR	0.94	0.92	0.97	0.60	0.60	0.64	0.15	0.26	0.39	0.61	0.78	0.84	0.25	0.32	0.43	0.03	0.07	0.16
		FNR	0.00	0.00	0.00	0.00	0.00	0.02	0.31	0.04	0.09	0.20	0.09	0.08	0.44	0.24	0.17	0.61	0.47	0.36
	1000	FPR	0.90	0.97	0.93	0.67	0.70	0.74	0.44	0.48	0.53	0.96	0.97	0.96	0.61	0.66	0.71	0.42	0.48	0.53
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AW-LASSO	100	FPR	0.37	0.51	0.51	0.22	0.34	0.35	0.13	0.21	0.22	0.27	0.39	0.37	0.13	0.22	0.17	0.07	0.13	0.12
		FNR	0.00	0.00	0.26	0.00	0.00	0.27	0.00	0.01	0.30	0.00	0.01	0.46	0.01	0.05	0.57	0.00	0.10	0.60
	200	FPR	0.32	0.47	0.54	0.15	0.26	0.32	0.11	0.17	0.23	0.25	0.35	0.46	0.12	0.19	0.27	0.06	0.11	0.17
		FNR	0.00	0.00	0.13	0.00	0.00	0.16	0.00	0.00	0.15	0.00	0.01	0.30	0.00	0.01	0.31	0.00	0.02	0.34
	1000	FPR	0.26	0.33	0.49	0.11	0.18	0.30	0.05	0.09	0.19	0.25	0.27	0.46	0.11	0.13	0.25	0.05	0.06	0.15
		FNR	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.09	0.00	0.00	0.10	0.00	0.00	0.11

The performance of each model is measured by using MDAE, False Negative Rate (FNR), and False Positive Rate (FPR). FNR is defined as the proportion of zero coefficient estimates whose corresponding true coefficients are nonzero and FPR is defined as the proportion of nonzero coefficient estimates whose corresponding true coefficients are zero. The obtained results are summarized in Table 2-5. Also, for effective understanding of the performance of various methods, the pictorial representations of error measures are given in Figure 2 and 3 respectively.

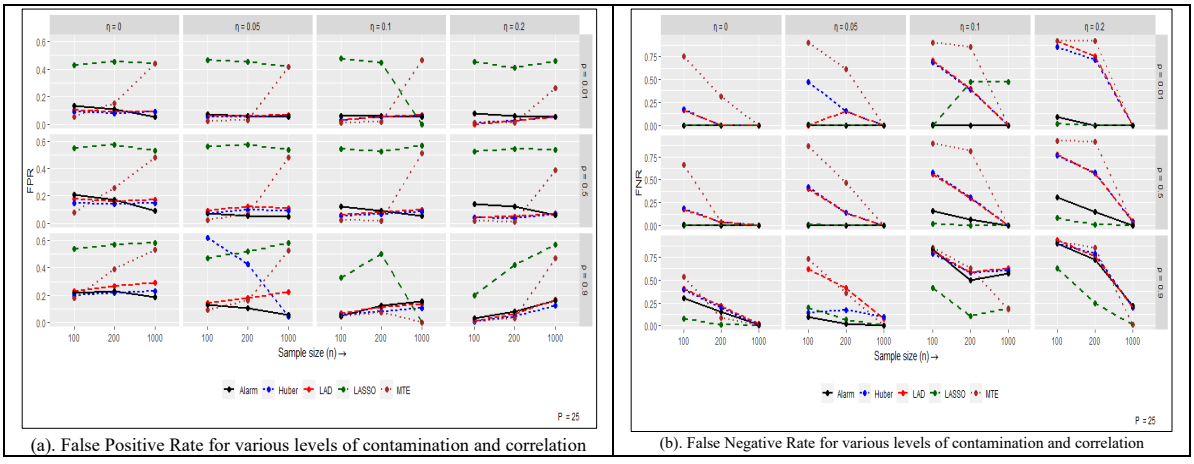
Table 3: False Negative and False Positive Rate of each method under various levels of contamination (10% and 20%)

			η = 10									η = 20								
			p = 10			p = 15			p = 25			p = 10			p = 15			p = 25		
			ρ			ρ			ρ			ρ			ρ			ρ		
Method	N	Error	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90
LASSO	100	FPR	0.93	1.00	0.83	0.65	0.73	0.53	0.48	0.54	0.33	0.90	0.93	0.65	0.63	0.71	0.38	0.45	0.52	0.20
		FNR	0.00	0.00	0.18	0.00	0.01	0.27	0.00	0.02	0.42	0.00	0.04	0.31	0.01	0.04	0.48	0.01	0.08	0.63
	200	FPR	0.92	1.02	0.97	0.65	0.73	0.68	0.45	0.52	0.50	0.93	1.01	0.87	0.65	0.75	0.65	0.41	0.54	0.42
		FNR	0.00	0.00	0.05	0.00	0.00	0.08	0.47	0.00	0.11	0.00	0.00	0.12	0.00	0.00	0.11	0.00	0.01	0.24

	1000	FPR	0.89	0.97	1.01	0.67	0.74	0.76	0.00	0.57	0.00	0.92	1.02	0.99	0.68	0.75	0.75	0.46	0.53	0.57
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.47	0.00	0.20	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02
LAD LASSO	100	FPR	0.44	0.65	0.51	0.16	0.23	0.21	0.03	0.06	0.07	0.39	0.49	0.26	0.09	0.18	0.09	0.00	0.04	0.02
		FNR	0.14	0.09	0.44	0.35	0.32	0.64	0.71	0.57	0.82	0.32	0.29	0.66	0.63	0.49	0.83	0.92	0.77	0.94
	200	FPR	0.52	0.64	0.69	0.21	0.24	0.28	0.05	0.08	0.11	0.42	0.55	0.52	0.15	0.21	0.21	0.02	0.05	0.06
		FNR	0.01	0.02	0.19	0.13	0.10	0.38	0.40	0.30	0.58	0.06	0.07	0.39	0.29	0.23	0.51	0.75	0.57	0.76
	1000	FPR	0.46	0.63	0.79	0.20	0.27	0.39	0.07	0.10	0.14	0.44	0.60	0.78	0.16	0.23	0.38	0.05	0.07	0.17
		FNR	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.63	0.00	0.00	0.03	0.00	0.00	0.11	0.00	0.04	0.20
Huber LASSO	100	FPR	0.43	0.58	0.50	0.16	0.19	0.19	0.03	0.05	0.06	0.38	0.46	0.26	0.09	0.15	0.08	0.01	0.04	0.01
		FNR	0.14	0.09	0.44	0.35	0.31	0.60	0.69	0.58	0.79	0.31	0.29	0.64	0.61	0.50	0.81	0.86	0.77	0.91
	200	FPR	0.50	0.62	0.63	0.20	0.22	0.24	0.05	0.07	0.08	0.41	0.54	0.48	0.14	0.19	0.18	0.02	0.04	0.05
		FNR	0.01	0.02	0.19	0.14	0.10	0.38	0.39	0.31	0.59	0.05	0.07	0.40	0.28	0.24	0.52	0.72	0.58	0.79
	1000	FPR	0.44	0.60	0.76	0.18	0.23	0.34	0.06	0.08	0.11	0.43	0.57	0.71	0.16	0.22	0.33	0.05	0.07	0.13
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.00	0.00	0.02	0.00	0.00	0.08	0.00	0.04	0.20
MTE LASSO	100	FPR	0.25	0.40	0.41	0.07	0.09	0.15	0.01	0.03	0.04	0.24	0.30	0.21	0.04	0.07	0.06	0.01	0.02	0.01
		FNR	0.65	0.51	0.55	0.85	0.79	0.72	0.90	0.90	0.85	0.72	0.64	0.73	0.89	0.85	0.88	0.92	0.92	0.92
	200	FPR	0.37	0.50	0.59	0.12	0.12	0.21	0.01	0.02	0.08	0.22	0.30	0.34	0.06	0.07	0.10	0.01	0.01	0.03
		FNR	0.46	0.26	0.25	0.71	0.58	0.50	0.86	0.82	0.63	0.67	0.53	0.56	0.84	0.80	0.69	0.92	0.92	0.85
	1000	FPR	0.87	0.88	0.94	0.66	0.69	0.69	0.47	0.51	0.00	0.78	0.87	0.90	0.53	0.63	0.65	0.26	0.39	0.47
		FNR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.02	0.02
AW-LASSO	100	FPR	0.29	0.38	0.23	0.12	0.18	0.10	0.06	0.12	0.05	0.27	0.37	0.14	0.17	0.21	0.06	0.08	0.14	0.03
		FNR	0.02	0.11	0.68	0.03	0.16	0.73	0.01	0.16	0.84	0.07	0.25	0.79	0.06	0.27	0.85	0.09	0.31	0.90
	200	FPR	0.25	0.31	0.36	0.11	0.17	0.20	0.06	0.09	0.13	0.26	0.37	0.27	0.12	0.21	0.13	0.06	0.12	0.08
		FNR	0.00	0.04	0.43	0.00	0.04	0.47	0.00	0.07	0.50	0.01	0.10	0.59	0.00	0.11	0.67	0.00	0.15	0.73
	1000	FPR	0.25	0.27	0.46	0.11	0.12	0.24	0.05	0.06	0.16	0.25	0.27	0.46	0.11	0.14	0.26	0.05	0.06	0.16
		FNR	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.58	0.00	0.00	0.18	0.00	0.00	0.18	0.00	0.00	0.22

False Positive Rate (Figure-2a), the selection of insignificant variables by the method. In these circumstances, the conventional LASSO has a high False Positive Rate relative to other approach es, it is because the LASSO tends to select a greater number of coefficients, while the MTE method's False Positive Rate rises as the sample size increases. In almost all situations, the AW-LASSO approac h has a very low False Positive Rate.

False Negative Rate (Figure-2b), or the number of significant variables that the technique failed to choose. The AW-LASSO False Negative Rate is usually always zero at all levels, but when the correlat ion level rises, the approach produces inconsequential results. The MTE technique shows a high rate o f False Negative, however as the sample size grows, the approach tends to converge to zero.



(a). False Positive Rate for various levels of contamination and correlation

(b). False Negative Rate for various levels of contamination and correlation

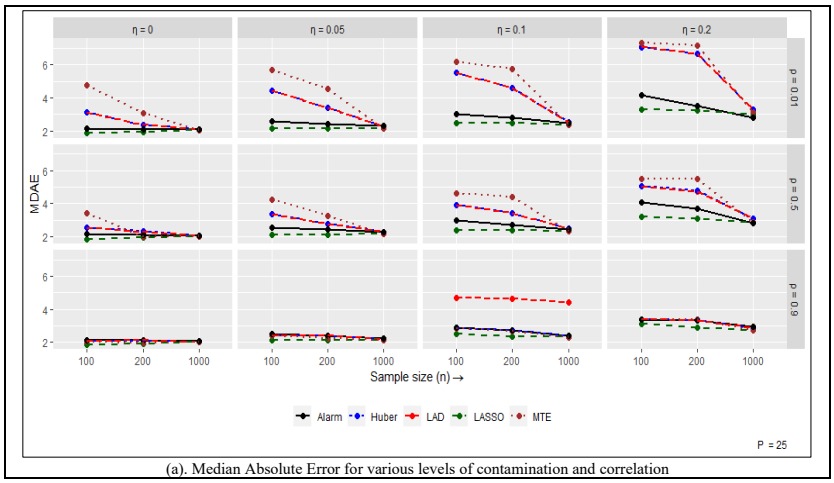
Figure 2: FPR and FNR under various levels of contamination and correlation

Table 4: Median Absolute Error of each method under various levels of contamination (0% and 5%)

Method	N	η = 0									η = 5								
		p = 10			p = 15			p = 25			p = 10			p = 15			p = 25		
		ρ			ρ			ρ			ρ			ρ			ρ		
		0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90
LASSO	100	1.90	1.91	1.91	1.90	1.87	1.90	1.87	1.84	1.85	2.80	2.49	2.49	2.55	2.70	2.50	2.57	2.56	2.50
	200	1.96	1.96	1.96	1.97	1.96	1.96	1.94	1.93	1.94	2.85	2.37	2.34	2.39	2.75	2.37	2.44	2.41	2.39
	1000	2.02	2.03	2.02	2.02	2.02	2.02	2.01	2.01	2.01	2.80	2.35	2.22	2.30	2.80	2.24	2.35	2.30	2.23
LAD LASSO	100	1.99	2.01	2.02	2.37	2.10	2.05	3.10	2.57	2.09	2.51	2.24	2.23	2.40	2.57	2.36	2.44	2.35	2.43
	200	1.98	1.97	1.99	2.11	2.06	2.02	2.36	2.30	2.11	2.29	2.15	2.13	2.20	2.39	2.25	2.37	2.34	2.35
	1000	2.05	2.06	2.08	2.03	2.03	2.02	2.08	2.06	2.04	2.13	2.13	2.14	2.21	2.17	2.17	2.28	2.25	2.21
Huber LASSO	100	1.97	2.00	1.99	2.37	2.10	2.02	3.10	2.54	2.09	2.48	2.22	2.21	2.41	2.64	2.34	2.43	2.35	2.42
	200	1.98	1.97	2.00	2.11	2.06	2.01	2.37	2.31	2.09	2.28	2.15	2.12	2.22	2.37	2.23	2.39	2.30	2.37
	1000	2.06	2.07	2.08	2.03	2.03	2.02	2.08	2.07	2.04	2.13	2.13	2.14	2.24	2.18	2.16	2.28	2.25	2.20
MTE LASSO	100	3.03	2.15	1.99	4.44	2.75	2.06	4.76	3.44	2.10	2.30	2.20	2.16	2.43	2.63	2.36	2.35	2.26	2.41
	200	1.96	1.96	1.99	2.12	2.01	2.07	3.09	2.10	2.16	2.45	2.30	2.08	2.22	2.57	2.14	2.36	2.28	2.24
	1000	2.06	2.06	2.07	2.02	2.02	2.08	2.20	2.15	2.12	2.14	2.15	2.14	2.21	2.20	2.14	2.16	2.15	2.14
AW-LASSO	100	1.95	1.95	1.96	2.12	2.09	1.97	2.14	2.15	2.07	2.14	2.13	2.10	2.14	2.16	2.14	2.15	2.12	2.12
	200	1.97	1.97	1.98	2.10	2.04	1.98	2.11	2.13	2.09	2.16	2.12	2.12	2.15	2.17	2.14	2.14	2.12	2.12
	1000	2.03	2.02	2.02	2.00	1.99	2.00	2.05	2.00	2.02	2.14	2.15	2.13	2.16	2.26	2.17	2.16	2.15	2.15

Table 5: Median Absolute Error of each method under various levels of contamination (10% and 20%)

Method	N	η = 10									η = 20								
		p = 10			p = 15			p = 25			p = 10			p = 15			p = 25		
		ρ			ρ			ρ			ρ			ρ			ρ		
		0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90	0.01	0.50	0.90
LASSO	100	2.90	2.81	2.80	2.96	2.97	2.92	2.96	2.91	2.92	3.88	3.82	3.80	4.12	4.06	4.03	4.20	4.21	4.10
	200	2.89	2.64	2.65	2.74	2.91	2.67	2.96	2.90	2.96	3.97	3.86	3.85	4.06	3.95	3.79	4.21	4.09	4.12
	1000	2.92	2.42	2.55	2.45	2.93	2.41	2.98	2.95	2.93	3.91	3.80	3.81	4.93	3.83	3.73	4.18	4.17	4.14
LAD LASSO	100	2.50	2.49	2.47	2.51	2.46	2.63	2.50	2.41	2.85	2.81	2.96	2.96	3.01	2.89	2.70	3.08	3.07	3.39
	200	2.56	2.35	2.42	2.37	2.49	2.57	2.60	2.40	2.71	2.89	2.99	2.99	3.00	2.85	2.81	3.68	3.75	3.33
	1000	2.53	2.28	2.30	2.36	2.35	2.35	2.51	2.47	2.41	2.77	2.63	2.65	2.84	2.78	2.77	3.29	3.11	2.89
Huber LASSO	100	2.45	2.50	2.46	2.35	2.44	2.61	2.53	2.42	2.83	2.60	2.96	2.97	2.99	2.81	2.80	3.04	3.03	3.39
	200	2.50	2.35	2.39	2.39	2.39	2.55	2.62	2.41	2.71	2.88	3.00	3.03	2.98	2.69	2.80	3.64	3.77	3.33
	1000	2.52	2.28	2.30	2.36	2.35	2.34	2.51	2.48	2.41	2.67	2.63	2.65	2.85	2.78	2.76	3.29	3.11	2.89
MTE LASSO	100	2.60	3.46	2.42	5.72	2.34	2.61	2.70	2.64	2.83	2.82	3.44	3.43	2.92	5.16	3.34	3.29	3.47	3.38
	200	2.62	2.90	2.36	5.23	2.81	2.58	2.75	2.48	2.66	2.84	3.47	3.48	2.96	5.01	3.22	3.17	3.49	3.37
	1000	2.58	2.32	2.31	2.36	2.33	2.31	2.72	2.35	2.42	2.94	2.78	2.79	2.93	2.84	2.72	3.12	2.89	2.75
AW-LASSO	100	2.40	2.32	2.36	2.44	2.37	2.32	2.41	2.48	2.52	2.68	2.62	2.63	2.61	2.68	2.65	2.66	2.69	2.65
	200	2.42	2.34	2.35	2.34	2.41	2.37	2.39	2.40	2.36	2.67	2.64	2.66	2.66	2.65	2.66	2.64	2.70	2.70
	1000	2.44	2.32	2.30	2.34	2.33	2.30	2.36	2.33	2.33	2.65	2.76	2.72	2.69	2.77	2.69	2.70	2.77	2.63



(a). Median Absolute Error for various levels of contamination and correlation

Figure 3: MDAE under various levels of contamination and correlation

The ability of the approaches to predict is seen in Figure 3. In comparison to other approaches AW-LASSO has lower MDAE, but if the data contains no outlier the LASSO exhibits a higher predicti on capacity than other methods and also the MTE method, shows a very high prediction error. Howev er, the AW-LASSO method always has very low prediction error.

IV. Conclusion

Feature selection is a technique that aids in extracting the important variables from a larger range of variables. It is becoming increasingly vital in statistics and is crucial to statistical analysis. In this paper, it is proposed a new feature selection method, which uses a weight function from a redescending M-estimator to modify the ordinary LASSO to form a new feature selection method namely, AW-LASSO. The proposed technique performs well in both with and without outlier condition by examining under the real datasets namely, the Boston Housing and Diabetic data sets. Further, the simulation studies also showed that the superiority of AW-LASSO method over the other methods. The study concluded that the proposed method can be used in the field of statistical learning, specifically in prediction models.

References

[1] Alamgir, Amjad Ali, Sajjad Ahmad Khan, Dost Muhammad Khan, and Umair Khalil, (2013). A New Efficient Redescending M- Estimator: Alamgir Redescending M- estimator. Research Journal of Recent Sciences, 2: 79-91.

[2] Alfons, C. Croux and Gelper, S. (2013). Sparse least trimmed squares regression for analyzing high-dimensional large data sets. The Annals of Applied Statistics, 7: 226-248.

[3] Arslan, O. (2012). Weighted lad-lasso method for robust parameter estimation and variable selection in regression. Computational Statistics & Data Analysis, 56: 1952–1965.

[4] Chatterjee S. and Hadi A. S. Sensitivity Analysis in Linear Regression, John Wiley & Sons, New York, 2009.

[5] Cook, R. D. (2000). Detection of influential observation in linear regression. Technometrics, 42: 65–68.

[6] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of American Statistical Association, 96: 1348–1360.

[7] Iglewicz, B. and Martinez, J. (1982). Outlier detection using robust measures of scale. Journal of Statistical Computation and Simulation, 15: 285–293.

[8] Kai, B., Li, R., and Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. Annals of Statistics, 39: 305.

[9] S. Lambert-Lacroix and Zwald, L. (2011). Robust regression through the Huber's criterion and adaptive lasso penalty. Electronic Journal of Statistics, 5: 1015.

[10] Qin, Y. , Li, S. , Li, Y. and Yu Y. (2017). Penalized maximum tangent likelihood estimation and robust variable selection. ArXiv Preprint ArXiv: 1708.05439.

[11] Wang, H. Li, G. and Jiang, G. (2007). Robust regression shrinkage and consistent variable selection through the lad-lasso. Journal of Business & Economic Statistics 25: 347–355.

[12] Stella Ebele Anekwe, and Sidney Iheanyi Onyeagu, (2021). A Redescending M-Estimator for Detection and Deletion of Outliers in Regression Analysis. Pakistan Journal of Statistics and Operation Research, 17: 997-1014.

[13] Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological), 58: 267-288.

[14] Wang, L. and Li, R. (2009). Weighted Wilcoxon-type smoothly clipped absolute deviation method. Biometrics, 65: 564–571.

[15] Wang, H. Li, G. and Jiang, G. (2007). Robust regression shrinkage and consistent variable

selection through the lad-lasso. Journal of Business & Economic Statistics 25: 347–355.

[16] Wang, L. Wu, Y. and Li, R. (2012). Quantile regression for analysing heterogeneity in ultra-high dimension. Journal of American Statistical Association, 107: 214–222.

[17] Wang, X., Jiang, Y., Huang, M. and Zhang, H. (2013). Robust variable selection with exponential squared loss. Journal of the American Statistical Association, 108: 632-643.

[18] Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. Annals of Statistics, 15: 642–656.

[19] Yohai, V. J. and Zamar, R. H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale. Journal of American Statistical Association, 83: 406–413.

[20] Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of American Statistical Association, 101: 1418–1429.

[21] Zou, H and Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. Annals of Statistics, 36: 1108-1126.

PARAMETER ESTIMATION OF SCALE MUTH DISTRIBUTION(SMD) UNDER TYPE-1 CENSORING USING CLASSICAL AND BAYESIAN APPROACHES

Agni Saroj[1], Prashant K. Sonker[2], Shalini Kumari[3],Rakesh Ranjan[4] & Mukesh Kumar*[5]

•

[1,2,3]Department of Statistics, Banaras Hindu University, Varanasi, 221005, India.

[4] DST Centre for Interdispinary Mathematical Sciences, Banaras Hindu University, Varanasi, 221005, India.

[5]Department of Statistics, MMV, Banaras Hindu University, Varanasi 221005, India.

e-mail: [1]agni.saroj4@bhu.ac.in, [2]prashant.s4@bhu.ac.in, [3]shalini2612@bhu.ac.in, [4]rakesh.ranjan@bhu.ac.in, *[5]mukesh.mmv@bhu.ac.in

*Corresponding Author

Abstract

The lifetime distributions are used to understand and explain the real life circumstances in various fields (medical, engineering, etc.). Many times it is very tough task to complete an experiment with complete data due to lack of time, money or some other factors and get the data in incomplete form. to draw the information from such type of data (incomplete data), we use some censoring techniques. In the field of statistics, there are several censoring techniques available where type-I censoring is most commonly used due to its simplicity. In this article, the scale Muth distribution (SMD) is considered as a lifetime distribution under type-I censoring scheme. The parameter estimation has been done by classical as well as Bayesian approach. Under the classical paradigm, two most popular methods were used maximum likelihood estimation (MLE) and the maximum product of spacing estimation (MPSE). And under the Bayesian paradigm, we used the informative priors for each parameter and obtained the estimates by considering the squared error loss function using an approximation method, Metropolis Hasting (MH) algorithm. The performance of each estimator is evaluated by their mean squared error or simulated risk. Also, a real data set is used to illustrate the real phenomena and to estimate the parameter using above-mentioned techniques under type-I censoring scheme.

Keywords: Scale Muth distribution, type-I censoring, classical approach, Bayesian approach, real data analysis.

1. Introduction

A new popularized lifetime distribution, Muth distribution was introduced by J. E. Muth [1] in the field of reliability theory. Muth distribution has less probability mass on right tail in

comparison of some other well known distribution like Weibull, Gamma and log-normal. The cumulative distribution function (cdf) and probability density function (pdf) of this distribution are defined respectively:

$$F(y;\alpha) = 1 - e^{\left\{\alpha y - \frac{1}{\alpha}(e^{\alpha y}-1)\right\}} \qquad y > 0, \quad \alpha \in (0,1] \tag{1}$$

$$f(y;\alpha) = \begin{cases} (e^{\alpha y} - \alpha)e^{\left\{\alpha y - \frac{1}{\alpha}(e^{\alpha y}-1)\right\}} & y > 0, \quad \alpha \in (0,1] \\ 0 & otherwise \end{cases} \tag{2}$$

A very few properties of this distribution was given in [1] and mainly focuses on its strict positive memory and mean residual lifetime. After this Rinne [2] used this distribution as lifetime model with German data set of car prices for fitting. In 2008, Leemis and McQueston [3] given a well defined figure and showed the relationship with exponential distribution. When the shape parameter $\alpha \to 0$ it converges to standered exponential distribution having parameter 1, and given its name as Muth distribution. In 2015, Jodra et.al. [4] again highlighted this distribution and derived some other statistical properties such as, moment generating function (mgf), Quantile function, r^{th} moments, moments of order statistics , mode, median and given a scale transformation form of this distribution named as SMD with scale parameter $\beta > 0$ and shape parameter $\alpha \in (0,1]$.

Estimation of parameters of SMD with classical methods has been done in [4]. After that many of authors used this distribution with different transformation and derived their properties and applications such as Jodra et. al. [5] define power Muth distribution, Chesneau and Agiwal [6] given inverse power Muth distribution and observed that there is no restriction to select the values of α in eq.[1] as $\alpha < 0$ and $\alpha \in (0,1]$ but $\alpha \neq 0$. In 2022, Saroj et.al. [7] given a new lifetime distribution with up-side down bathtub (UBT) shape hazard, named as inverse Muth distribution(IMD) by taking inverse transformation of Muth distribution, and derived some statistical properties of IMD where it was found that the moment of any order for IMD does not exist for $\alpha \in (0,1]$. Also, the scale transformation was considered from the IMD named as scale inverse Muth distribution with real data applicability [7]. Recently in 2022 Sonker et. al.[8] estimated the parameter under simple stress-strength reliability and multi-component stress strength. In 2018, Almarashi and Elgarhy [9] given a Muth generated (M-G) class of distribution with the help of Muth distribution as a generator and T-X generated family [10]. Then Al-Babtain et. al. introduced a Transmuted M-G class of distribution [11] and Almarashi et. al. introduced a new truncated M-G family of distribution[12]. Some censoring schemes are also used by some authors for different form of Muth distribution such as power Muth distribution under progressive censoring scheme [13] and SMD under type-II censoring scheme [14].

Generally lifetime distributions are used under the life testing experiments, where we get the data in ordered form by failure (or death) of testing units. Many of the times, it is very tough task to record failure time of each event due to lack of time cost effectiveness and many more. In all these cases we need a technique that helps in drawing inferences about population after observing a part of testing units. For this, there are several censoring techniques are available

for different cases and situations. There are many censoring methods available in statistics one of them is type-I censoring scheme which is a simplest censoring technique and the form of right censoring. Under the type-I censoring scheme, we select a time point T_o to terminate the experiment and record the failure time of the event which fails before T_o, and after T_o it assumed that all the remaining items are censored. A very impressive experimental example for real life situation under type-I censoring was given in [15]. For the more details about type-I censoring see [16] & [17].

In this article, classical as well as Bayesian method of estimation are used to estimate the parameter of SMD under type-I censoring scheme. In classical method two most popular method MLE and MPSE are used and Bayesian estimation is performed under squared error loss function (SELF) and an approximation method is used to find the estimates because our posterior is not in simplest form.

2. Classical methods of estimation

2.1. Maximum Likelihood Estimation

In the classical inference, one of the most popular estimation method was proposed by R. A. Fisher in 1920 based on likelihood function called MLE. The main principle of this method is to search for the value of parameter which maximizes the likelihood function. In this article we have to estimate the parameter of SMD under type-I censoring scheme. In this case the likelihood function was defined by Cohen(1965)[18] as:

$$L(\theta|x) = C\prod_{i=1}^{n}[f(x_i)]^{\delta_i}[1-F(T_o)]^{(1-\delta_i)} \tag{3}$$

where δ_i is an indicator function which indicates that an experimental unit is observed or censored. If T_o is the pre-assigned time point to terminate the experiment then δ_i is defined for the $i^{th}; i = 1, 2, 3, ..., n$ unit as:

$$\delta_i = \begin{cases} 1 & if \quad x_i < T_o \\ 0 & if \quad x_i > T_o \end{cases}$$

Let us suppose that the failure time of m units are observed from n before time point T_o, then the constant C will be defined as $\frac{n!}{(n-m)!}$, and the Likelihood function is:

$$L(\theta|x) = \prod_{i=1}^{m}[f(x_i)][1-F(T_o)]^{(n-m)} \tag{4}$$

Let $x \sim SMD(\alpha, \beta)$ having pdf and cdf as:

$$F(x;\alpha,\beta) = 1 - e^{\left\{\frac{\alpha}{\beta}x - \frac{1}{\alpha}\left(e^{\frac{\alpha}{\beta}x}-1\right)\right\}} \qquad x > 0, \quad \alpha \in (0,1] \tag{5}$$

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta}(e^{\frac{\alpha}{\beta}x} - \alpha)e^{\left\{\frac{\alpha}{\beta}x - \frac{1}{\alpha}(e^{\frac{\alpha}{\beta}x}-1)\right\}} & x > 0, \quad \alpha \in (0,1] \\ 0 & otherwise \end{cases} \tag{6}$$

then the Likelihood function is expressed as:

$$L(\alpha,\beta|x) = \frac{n!}{(n-m)!}\prod_{i=1}^{m}\left[\frac{1}{\beta}(e^{(\frac{\alpha}{\beta}x_i)}-\alpha)e^{\left\{\frac{\alpha}{\beta}x_i-\frac{1}{\alpha}(e^{(\frac{\alpha}{\beta}x_i)}-1)\right\}}\right]\left[1-\left\{1-e^{\left\{\frac{\alpha}{\beta}x_i-\frac{1}{\alpha}(e^{(\frac{\alpha}{\beta}x_i)}-1)\right\}}\right\}\right]^{n-m}$$

$$L(\alpha,\beta/x) = C\frac{1}{\beta^m}\prod_{i=1}^{m}\left[(e^{(\frac{\alpha}{\beta}x_i)}-\alpha)e^{\left\{\frac{\alpha}{\beta}x_i-\frac{1}{\alpha}(e^{(\frac{\alpha}{\beta}x_i)}-1)\right\}}\right]\left[e^{\left\{\frac{\alpha}{\beta}x_i-\frac{1}{\alpha}(e^{(\frac{\alpha}{\beta}x_i)}-1)\right\}}\right]^{n-m} \tag{7}$$

where $C = \frac{n!}{(n-m)!}$. after taking the Log of both side of eq.[7] the log-likelihood defined as

$$\begin{aligned} log(L(\alpha,\beta/x)) =& log(C) - mlog(\beta) + \sum_{i=1}^{m} log\left(e^{(\frac{\alpha}{\beta}x_i)}-\alpha\right) + \sum_{i=1}^{m}\left(\frac{\alpha}{\beta}x_i - \frac{1}{\alpha}(e^{(\frac{\alpha}{\beta}x_i)}-1)\right) \\ &+ (n-m)\left(\frac{\alpha}{\beta}T_o - \frac{1}{\alpha}(e^{(\frac{\alpha}{\beta}T_o)}-1)\right) \end{aligned} \tag{8}$$

To obtain the MLE for α and β we have to search the value of α and β which maximizes the eq.[8]. now, differentiate the eq.[8] w.r.t. α and β and equating to 0 we get,

$$\frac{\partial}{\partial\alpha}log(L(\alpha,\beta/x)) = 0$$

$$\sum_{i=1}^{m}\frac{x_i e^{\frac{\alpha}{\beta}x_i}-\beta}{\beta(e^{\frac{\alpha}{\beta}x_i}-\alpha)} + \sum_{i=1}^{m}\left\{\frac{x_i}{\beta} + \frac{e^{\frac{\alpha}{\beta}x_i}-1)}{\alpha^2} - \frac{x_i\, e^{\frac{\alpha}{\beta}x_i}}{\alpha\beta}\right\} + (n-m)\sum_{i=1}^{m}\left\{\frac{T_o}{\beta} + \frac{e^{\frac{\alpha}{\beta}T_o}-1)}{\alpha^2} - \frac{x_i\, e^{\frac{\alpha}{\beta}T_o}}{\alpha\beta}\right\} = 0 \tag{9}$$

and,
$$\frac{\partial}{\partial\beta}log(L(\alpha,\beta/x)) = 0$$

$$-\frac{m}{\beta} - \sum_{i=1}^{m}\frac{\left(\alpha x_i e^{\frac{\alpha}{\beta}x_i}\right)}{\left(e^{\frac{\alpha}{\beta}x_i}-\alpha\right)} + \frac{1}{\beta^2}\sum_{i=1}^{m}x_i\left(e^{\frac{\alpha}{\beta}x_i}-\alpha\right) + \frac{T_o(n-m)}{\beta^2}\sum_{i=1}^{m}\left(e^{\frac{\alpha}{\beta}T_o}-\alpha\right) = 0 \tag{10}$$

now solve these two non-linear eq.[9] and eq.[10] to get the value of $\alpha$and β with satisfying the condition

$$\left(\frac{\partial\, log(L(\alpha,\beta|x))}{\partial\alpha\partial\beta}\right)_{\alpha_{ml},\,\beta_{ml}} < 0 \tag{11}$$

As we can see that eq.[9] and eq.[10] are the complex function for α and β i.e. we can not solve both the equations analytically. To obtain the value of α and β we used Newton Raphson iteration method and get the the value of α and β as $\hat{\alpha}_{ml}$ and $\hat{\beta}_{ml}$ which satisfying the condition eq.[11] and maximize the eq.[8].

2.2. Maximum Product of Spacing Estimation

This (MPSE) method is an another method of estimation alternative to MLE which is originally proposed by Cheng and Amin in 1983 [19] and Ranneby in 1984 [20]. It provides the consistent and asymptotically unbiased estimate. This method cover the drawback of MLE where MLE fails to estimate three or more parameter distribution like gamma, weibull and log-normal [21] and also the major limitation of MLE is that it does not work satisfactorily for Heavy tailed continuous distribution with unknown location and scale parameter [22]. The estimators obtained by MPSE

method possess most of the large sample properties like sufficiency, consistency and asymptotic efficiency, possessed by MLE under more general condition [[23] & [21]] and MPSE holds the invariance property same as the MLE, shown by Coolen and Newby [24]. The consistency of MPS estimator have been discussed in detailed in [25]. MPS method is mostly similar as MLE where in case of MLE we use the likelihood function and search the value of parameter which maximises the likelihood function but in the case of MPSE method we tried to find the estimated value of parameter which maximises the product of spacing function based on cdf of the target density function.

Let $x_1, x_2, ..., x_n$, be an ordered random sample of size n from the univariate distribution. For the MPSE, the general condition is that the density function is considered to be strictly positive in an interval (a,b) and zero elsewhere $(a = -\infty \quad and \quad b = \infty)$. In this article we have the cdf F(x) and pdf f(x) defined in eq.[5] and [6] respectively with supporting value $x > 0; 0 < \alpha \leq 1$ and $\beta > 0$. By this $F(x) = 0$ and $f(x) = 0$ for $x < 0$, and $F(x) = 1$ and $f(x) = 0$ for $x > \infty$. The i^{th} spacing function is defined as $D_i = F(x_i) - F(x_{i-1})$ and the product of spacing function is $S = \prod_{i=1}^{n+1} D_i$. The average of this is measured by geometric mean of spacing denoted by 'G' and defined as $G = S^{\frac{1}{n+1}}$

If the sample is very likely or most probable, it is assumed that they gives the magnitude of spacing more or less equal. Hence, MPSE will be that value of parameter which maximizes the S or G. This study considered under type-I censoring so, the sample values belong in (0,∞) can be divided in no. of parts of interval like $[0, x_1]$, $(x_1, x_2]$, $(x_2, x_3]...(x_{k-1}, x_k]$, $(x_k, T_o]$, (T_o, ∞). MPSE method described in detail for the type-I censoring case in [14]. According to [14] we may write the product of spacing function for chosen distribution as :

$$S(\alpha, \beta|x) \propto \begin{cases} f(T_o) \cdot \prod_{i=1}^{m} D_i \cdot \left\{ \prod_{j=1}^{n+1} D_{m+j} \right\}; & if \quad |T_o - x_m| < \epsilon \\ D_{\xi} \cdot \{ \prod_{i=1}^{m} D_i \} \cdot \left\{ \prod_{j=1}^{n+1} D_{m+j} \right\}; & if \quad otherwise \end{cases} \tag{12}$$

where, $D_i = F(x_i) - F(x_{i-1})$, $D_{m+j} = (1 - F(T_o))/(n - m)$, $D_{\xi} = F(T_o) - F(x_m)$ and $F(x)$, $f(x)$ given in eq.(5), (6) respectively. Thus the MPSE can be obtained by maximizing $S(\alpha, \beta|x) = log(S(\alpha, \beta|x))$. Now differentiating log of eq.(12) w.r.t. α and β we get,

$$\frac{\partial}{\partial \alpha} S(\alpha, \beta|x) \propto \begin{cases} \frac{f'(T_o)}{f(T_o)} + \sum_{i=1}^{m} \frac{F'(x_i) - F'(x_{i-1})}{F(x_i) - F(x_{i-1})} - \frac{(n-m)F'(T_o)}{1 - F(T_o)} & if \quad |T_o - x_m| < \epsilon \\ \frac{F'(T_o) - F'(x_m)}{F(T_o) - F(x_m)} + \sum_{i=1}^{m} \frac{F'(x_i) - F'(x_{i-1})}{F(x_i) - F(x_{i-1})} - \frac{(n-m)F'(T_o)}{1 - F(T_o)} & otherwise \end{cases} \tag{13}$$

where,

$$\frac{\partial}{\partial \alpha} f'(x) = \frac{A_1}{\alpha^2 \beta} \left[(\alpha - e^{\frac{\alpha}{\beta} x}) \left\{ \alpha x (e^{\frac{\alpha}{\beta} x} - \alpha) - \beta e^{\frac{\alpha}{\beta} x} + \beta \right\} + \alpha^2 (x e^{\frac{\alpha}{\beta} x} - \beta) \right]$$

$$\frac{\partial}{\partial \alpha} F'(x) = \frac{A_1}{\alpha^2 \beta} \left[\left\{ \alpha x (e^{\frac{\alpha}{\beta} x} - \alpha) - \beta e^{\frac{\alpha}{\beta} x} + \beta \right\} \right]$$

and $$A_1 = e^{(\frac{\alpha}{\beta} x - \frac{1}{\alpha}(e^{\frac{\alpha}{\beta} x} - 1))}$$

And the partial derivative of log of eq.[12] w.r.t. β will be similar as eq.[13]

$$\frac{\partial}{\partial \beta} S(\alpha, \beta|x) \propto \begin{cases} \frac{f'(T_o)}{f(T_o)} + \sum_{i=1}^{m} \frac{F'(x_i) - F'(x_{i-1})}{F(x_i) - F(x_{i-1})} - \frac{(n-m)F'(T_o)}{1 - F(T_o)} & if \quad |T_o - x_m| < \epsilon \\ \frac{F'(T_o) - F'(x_m)}{F(T_o) - F(x_m)} + \sum_{i=1}^{m} \frac{F'(x_i) - F'(x_{i-1})}{F(x_i) - F(x_{i-1})} - \frac{(n-m)F'(T_o)}{1 - F(T_o)} & o.w. \end{cases} \tag{14}$$

where,

$$\frac{\partial}{\partial\beta}f'(x) = -\frac{A_1}{\beta^2}\left[\beta(e^{\frac{\alpha}{\beta}x} - \alpha) + \alpha x e^{\frac{\alpha}{\beta}x} + x(e^{\frac{\alpha}{\beta}x} - \alpha)^2\right]$$

$$\frac{\partial}{\partial\beta}F'(x) = -\frac{A_1}{\beta^2}\left(e^{\frac{\alpha}{\beta}x} - \alpha\right)$$

Now equating eq.[13], [14] to zero and solve it for α and β, we get the estimated value of α and β which maximizes the eq.[12]. But the analytical solution is not possible because those equations are not in closed form. Hence we use Newton Raphson method to solve it numerically and get estimated value as $\hat{\alpha}_{mp}$ and $\hat{\beta}_{mp}$.

3. Asymptotic Confidence Interval

In this section, we define the interval estimation for $\alpha$and β . Since the exact distribution of MLE and MPSE is not easy to find because of the maximum likelihood and product of spacing function is not in the explicit form. Several authors show that the MPSE method is asymptotically equivalent to the MLE see [[19], [26], [27], [28]]. So, by using the large sample theory we may propose the Asymptotic Confidence Interval based on MLE and MPSE function and we may write the asymptotic distribution of both kind of estimator is:

$$(\hat{\theta} - \theta) \equiv N(0, I^{-1}(\hat{\theta})); \tag{15}$$

where,

$\hat{\theta}$ is the estimate of parameter

θ is the true value of parameter

$I^{-1}(\hat{\theta})$ is the inverse of Fisher information matrix

The $100(1-\alpha^*)$%confidence interval is defined as:

$$\hat{\theta} \pm Z_{\alpha^*_{/2}}\sqrt{Var(\hat{\theta})}; \tag{16}$$

where $Var(\hat{\theta}) = -E\left(\frac{\partial^2(L)}{\partial x^2}\right)^{-1}_{\theta=\hat{\theta}}$ is the digonal element of the inverse of Information matrix $I^{-1}(\hat{\theta})$ In our case the information matrix is defined as:

$$I(\hat{\alpha};\hat{\beta}) = \begin{bmatrix} I_{1,1} & I_{1,2} \\ I_{2,1} & I_{2,2} \end{bmatrix}$$

Here $I_{1,1} = \frac{\partial^2}{\partial\alpha^2}$, $I_{2,2} = \frac{\partial^2}{\partial\beta^2}$ *and* $I_{1,2} = I_{2,1} = \frac{\partial^2}{\partial\alpha\partial\beta}$ are the second order derivative of log-likelihood function eq.[8] and the log of the product spacing function eq.[12] . So, for α the $100(1-\alpha^*)$% confidence interval is defined as:

$$\hat{\alpha} \pm Z_{\alpha^*_{/2}}\sqrt{Var(\hat{\alpha})}; \qquad \textit{where } \hat{\alpha} \textit{ is } \hat{\alpha}_{ml} \textit{ and } \hat{\alpha}_{mp} \tag{17}$$

and for β

$$\hat{\beta} \pm Z_{\alpha^*_{/2}}\sqrt{Var(\hat{\beta})}; \qquad \textit{where } \hat{\beta} \textit{ is } \hat{\beta}_{ml} \textit{ and } \hat{\beta}_{mp} \tag{18}$$

4. Bayes estimates of α and β under squared error loss function

In this section we considered Bayesian approach to estimate the parameter α and β under type-I censoring scheme. In the Bayesian method it is assumed that our parameter are random in nature instead of a fixed quantity. As the random nature of the parameter, we can associate it with distribution function such as called prior distribution. Let us consider that gamma prior for scale parameter β and beta prior for shape parameter α, as informative prior, therefore

$$\pi(\alpha) \propto \alpha^{a-1} \cdot (1-\alpha)^{b-1} \tag{19}$$

and

$$\pi(\beta) \propto \beta^{c-1} \cdot e^{-d\beta} \tag{20}$$

where, a,b and c,d are hyper parameters for beta prior and gamma prior respectively. Thus the joint prior of α and β is

$$\pi(\alpha,\beta) \propto \alpha^{a-1} \cdot (1-\alpha)^{b-1} \cdot \beta^{c-1} \cdot e^{-d\beta} \tag{21}$$

Now, according to the Bayes rule the formula for joint posterior of α and β is

$$\Pi(\alpha,\beta|x) = \frac{\pi(\alpha,\beta)L(\alpha,\beta|x)}{\int \pi(\alpha,\beta)L(\alpha,\beta|x)d\alpha\ d\beta} \tag{22}$$

where, $L(\alpha,\beta|x)$ is likelihood given in eq.[7]. So, combining the eq.[7] and [21] the posterior for α and β can be written as:

$$\Pi(\alpha,\beta|x) \propto C_1^{-1} \cdot \beta^{(c-m-1)} \cdot \alpha^{(a-1)} \cdot (1-\alpha)^{(b-1)} \cdot e^{(n-m)(\frac{\alpha}{\beta}T_o - \frac{1}{\alpha}(e^{\frac{\alpha}{\beta}T_o}-1))-d\beta} \cdot e^{\sum(\frac{\alpha}{\beta}x_i - \frac{1}{\alpha}(e^{\frac{\alpha}{\beta}x_i}-1))} \tag{23}$$

where, $C_1 = \int\ \beta^{(c-m-1)} \cdot \alpha^{(a-1)} \cdot (1-\alpha)^{(b-1)} \cdot e^{(n-m)(\frac{\alpha}{\beta}T_o - \frac{1}{\alpha}(e^{\frac{\alpha}{\beta}T_o}-1))-d\beta} \cdot e^{\sum(\frac{\alpha}{\beta}x_i - \frac{1}{\alpha}(e^{\frac{\alpha}{\beta}x_i}-1))}\ d\alpha\ d\beta$
In the case of Bayesian techniques, loss function play very important role. In this article, to evaluate the Bayes estimator of α and β we use a symmetric loss function, squared error loss function (SELF) which gives the equal weight for under and over estimation problem. The general form of SELF can be expressed as:

$$L(\hat{\theta},\theta) = \lambda(\hat{\theta}-\theta)^2; \qquad \lambda > 0 \tag{24}$$

where $\hat{\theta}$ is the estimate of parameter θ. the Bayes estimator of parameter under SELF is posterior mean i.e. posterior expectation $E_\theta[\theta|x]$. In our case it is found that the posterior mean for both parameter is the ratio of two mathematically non-tractable integrals; so we use the Markov Chain Monte Carlo (MCMC) with Metropolis-Hasting (MH) algorithm to obtain Bayes estimate. MH algorithm is generalization form of basic algorithm originated by Metropolis and Ulman(1949) [29], then Metropolis et. al. (1953) used first time in statistics study the equation of a state of a two-dimensional rigid sphere system. In 1970, W. K. Hasting gives the original method by using arbitrary proposal distribution and it popularized by name MH algorithm. For more detail see[[30], [31] & [32]]. By using this technique we can generate the sample from posterior eq.[23] and evaluate the Bayes estimate. The following iterative steps are involved in MH algorithm:
1. Set an initial guess value θ^0.

2. By using previous point Generate the next value of θ as θ^j from the proposal density $p^*(\theta^j|\theta^{j-1})$. Here we take normal distribution as the proposal density for parameter.

3. Then we generate u $U(0,1)$and calculate the ratio:

$$r = \frac{p(\theta^j|\underset{\sim}{x}) \cdot p^*(\theta^{j-1}|\theta^j)}{p(\theta^{j-1}|\underset{\sim}{x}) \cdot p^*(\theta^j|\theta^{j-1})}$$

4. Set

$$\theta^j = \begin{cases} \theta^j & if \;\; u \leq \frac{p(\theta^j|\underset{\sim}{x}) \cdot p^*(\theta^{j-1}|\theta^j)}{p(\theta^{j-1}|\underset{\sim}{x}) \cdot p^*(\theta^j|\theta^{j-1})} \\ \theta^{j-1} & otherwise \end{cases}$$

5. Repeat step (2-4) for j=1,2,3,...,M and obtain θ_1,θ_2,...,θ_M

6. To find Bayes estimate under SELF, we take

$$E_p\left(\theta|\underset{\sim}{x}\right) = \left(\frac{1}{M - M_0} \sum_{M_0+1}^{M} \theta\right)$$

where M_0 is burn-in period of Markove chain.

5. Highest Posterior density interval

Nowadays, in the field of Bayesian inference is considered that the parameter of interest behaves like a random variable, then what is probability that the parameter (say θ) lies within a specified interval. Edwards et al. (1963) [33] suggested a method to calculate this interval and named it credible interval. By using this many of the statistician summarizes the marginal posterior density by $100(1-\alpha^*)\%$ posterior credible intervals for the parameters. These intervals can be easily calculated by analytically or MCMC simulation technique. The shortest interval among all of the Bayesian credible intervals is called highest posterior density (HPD) interval. The HPD intervals for parameter can be obtained by using method explained by Chen and Shao (1999)[34] on the basis of MCMC sample from posterior density in Eq.[23]. There are also mentioned that when a marginal distribution is not symmetric, a HPD interval is more desirable. Box and Tiao (1973)[35] was already discused that the HPD interval has two main properties as: (i) the density for the points which are lie within the interval is more than that for the points outside the interval, and (ii) for a given probability say, (1 -α^*), the interval is of the shortest length. If we have the posterior as $p(\theta|x)$ then the credible interval for θ is defined as:

$$(\theta_{[(\alpha/2)M]}, \theta_{[1-(\alpha/2)M]})$$

and the HPD interval is defined as:

$$(\theta_{j^*}, \theta_{j^*+[1-(\alpha/2)]})$$

where, j^* is chosen by which length $l_{j^*} = min(\theta_{j+[1-(\alpha/2)]} - \theta_j)$ for $1 \leq j \leq M - [(1-\alpha)M]$

6. Simulation Study

Under this section, we study about the behavior of the estimated value of parameters using the simulation study and compare the performances of MLE, MPSE and Bayes estimator on the basis

of their mean squared error (MSE) values. To compute the value of MSE, we use the formula as; $MSE(\theta)=\frac{1}{N}\sum_1^N(\theta-\hat{\theta}_i)^2$, where N is the number of simulated random sample (we choose N=10000), θ is the true value of parameter and $\hat{\theta}_i$ is the estimated value of the parameter for i^{th} simulated random sample. all the computation was done by R-software like sample generation from the considered distribution and finding the estimated value. To generate the random sample, we considered the following steps:

Step 1. Provide some selected values of the parameters, α and β, and sample size n.

Step 2. Then generate a random sample from a standard uniform distribution U(0, 1), which also be size n denoted as $u_i; \quad i = 1, 2, 3..., n$

Step 3. Now by using the quantile function of SMD given in [4], obtain u^{th} quantile (random sample) from the SMD x_i.

Step 4. Repeat steps 2 and 3, for each i = 1, 2, ... , n we get a random sample of size n from the SMD.

In step 3 the Lambert-W function $W_{-1}(.)$ can be computed using R-command lambertWm1() from 'lamW ' package, see Adler [2]. In the simulation study, we choose the different combination of parameter values as $(\alpha = 0.3, \quad \beta = 0.5)$, $(\alpha = 0.5, \quad \beta = 0.5)$, $(\alpha = 0.7, \quad \beta = 0.5)$, $(\alpha = 0.3, \quad \beta = 2.5)$, $(\alpha = 0.5, \quad \beta = 2.5)$, $(\alpha = 0.7, \quad \beta = 2.5)$, and observed the behavior of MSE by varying sample size. For this we choose the different sample sizes as $n = (30,\ 50,\ 70)$. In case of Bayesian estimate we need the value of hyper parameters for which we take prior means equal to true value of parameter and the variances can be chosen as, for small value of parameter (α=(0.3,0.5) & β=0.5) we take small variance for gamma and beta prior and for the large value of parameter (α=0.7 & β=2.5) we take some large variance for the both prior. An another issue in this simulation study is the choice of T_o. We take $T_o = (0.5, 0.9)$ for β=0.5 with different value of α, and $T_o = (3, 5)$ for β=2.5 with different value of α. On observing the simulation table [1-6] we found that:

In general, as sample size increases the MSE of all estimators decreases. This shows that the estimators are consistent. And this decreasing nature of MSE is also found with increasing T_o.

The average length of asymptotic confidence interval as well as HPD interval decreases as sample size increases.

MSE of Bayes estimates is least in comparison of MLE and MPSE both, and the MSE of MPSE is less that MLE.

Width of the HPD interval is smaller in comparison of width of asymptotic confidence interval and the width of asymptotic confidence interval for MPSE is less than the width for asymptotic confidence interval MLE.

Table 1: *Estimates of the parameters, MSE, CI and HPD interval for α=0.3, β=0.5*

			MLE			MPS			Bayes		
	n	To	Est.	MSE	CI Length	Est.	MSE	CI Length	Est.	MSE	HPDI Length
$\alpha = 0.3$	30	0.5	0.34704	0.07774	0.72461, 0.06792 0.65669	0.33273	0.06535	0.72461, 0.06792 0.65669	0.49504	0.04156	0.73352, 0.25224 0.48128
		0.9	0.32891	0.04485	0.76512, 0.09032 0.67480	0.31103	0.03862	0.75007, 0.08840 0.66167	0.46616	0.03210	0.68793, 0.24349 0.44444
	50	0.5	0.32582	0.05735	0.72646, 0.08661 0.63985	0.31781	0.05174	0.71290, 0.09112 0.62179	0.47571	0.03562	0.69921, 0.24710 0.45211
		0.9	0.31201	0.03038	0.72890, 0.10531 0.62359	0.30214	0.02796	0.71555, 0.10427 0.61129	0.43824	0.02370	0.64006, 0.23556 0.40450
	70	0.5	0.31557	0.04712	0.71985, 0.09472 0.62513	0.31017	0.04380	0.70609, 0.09875 0.60734	0.46208	0.03090	0.67481, 0.24393 0.43088
		0.9	0.29997	0.02274	0.68941, 0.11159 0.57782	0.29338	0.02154	0.67894, 0.11067 0.56827	0.41580	0.01768	0.60281, 0.22852 0.37429
$\beta = 0.5$	30	0.5	0.51621	0.01738	0.80992, 0.36543 0.44448	0.50410	0.01401	0.76774, 0.36775 0.39998	0.50820	0.01338	0.69941, 0.35175 0.34765
		0.9	0.50952	0.00725	0.67859, 0.38679 0.29180	0.50958	0.00640	0.67060, 0.39217 0.27843	0.52538	0.00699	0.67103, 0.39446 0.27657
	50	0.5	0.51008	0.01114	0.73543, 0.37591 0.35951	0.50452	0.00943	0.93305, 0.38101 0.55204	0.49198	0.00676	0.63733, 0.36875 0.26858
		0.9	0.50562	0.00415	0.63644, 0.40359 0.23284	0.50584	0.00366	0.63370, 0.40661 0.22708	0.51580	0.00388	0.62735, 0.41243 0.21492
	70	0.5	0.50932	0.00790	0.70286, 0.38420 0.31866	0.50536	0.00674	0.72334, 0.38939 0.33395	0.48843	0.00462	0.61332, 0.38087 0.23245
		0.9	0.50449	0.00286	0.61525, 0.41471 0.20054	0.50428	0.00271	0.61537, 0.41601 0.19936	0.51197	0.00259	0.60662, 0.42325 0.18337

Table 2: *Estimates of the parameters, MSE, CI and HPD interval for α=0.5, β=0.5*

			MLE			MPS			Bayes		
	n	To	Est.	MSE	CI Length	Est.	MSE	CI Length	Est.	MSE	HPDI Length
$\alpha = 0.5$	30	0.5	0.48074	0.07860	0.83626, 0.10341 0.73285	0.45601	0.06665	0.82580, 0.11771 0.70808	0.52302	0.00395	0.76440, 0.27504 0.48936
		0.9	0.50786	0.04226	0.85294, 0.17571 0.67723	0.47783	0.03675	0.83123, 0.17776 0.65346	0.52332	0.00517	0.74687, 0.29615 0.45072
	50	0.5	0.49154	0.05688	0.84875, 0.15769 0.69106	0.47462	0.05084	0.83208, 0.16627 0.66582	0.52207	0.00485	0.75113, 0.28546 0.46566
		0.9	0.50121	0.02726	0.79694, 0.23084 0.56610	0.48342	0.02516	0.77794, 0.22849 0.54944	0.51618	0.00582	0.71989, 0.30895 0.41094
	70	0.9	0.48544	0.04791	0.83548, 0.18795 0.64754	0.47320	0.04426	0.81647, 0.19507 0.62140	0.51835	0.00532	0.73660, 0.29214 0.44445
		0.9	0.50248	0.01965	0.75472, 0.26570 0.48902	0.48975	0.01847	0.73761, 0.26445 0.47316	0.51345	0.00587	0.70190, 0.32155 0.38035
$\beta = 0.5$	30	0.5	0.53008	0.01885	0.87626, 0.36571 0.51055	0.51941	0.01454	0.88980, 0.36510 0.52470	0.53692	0.01431	0.74564, 0.36762 0.37801
		0.9	0.50709	0.00447	0.65383, 0.39495 0.25888	0.50638	0.00415	0.65251, 0.39464 0.25787	0.51898	0.00463	0.65208, 0.39763 0.25445
	50	0.5	0.52053	0.01047	0.75021, 0.38034 0.36987	0.51528	0.00892	0.73297, 0.38130 0.35167	0.52322	0.00794	0.67951, 0.39025 0.28926
		0.9	0.50307	0.00248	0.61212, 0.41396 0.19816	0.50265	0.00238	0.71388, 0.41340 0.30048	0.51130	0.00250	0.61151, 0.41741 0.19410
	70	0.5	0.51661	0.00723	0.70762, 0.39148 0.31614	0.51257	0.00643	1.18937, 0.39147 0.79790	0.51562	0.00486	0.64536, 0.40269 0.24267
		0.9	0.50260	0.00179	0.59260, 0.42653 0.16608	0.50227	0.00173	0.59251, 0.42607 0.16643	0.50890	0.00179	0.59275, 0.42927 0.16348

Table 3: *Estimate of the parameters, MSE, CI and HPD interval for α=0.7, β=0.5*

			MLE			MPS			Bayes		
	n	To	Est.	MSE	CI Length	Est.	MSE	CI Length	Est.	MSE	CI Length
$\alpha = 0.7$	30	0.5	0.63980	0.06248	0.92610, 0.14204 0.78406	0.59994	0.05700	0.91185, 0.17602 0.73583	0.55456	0.02370	0.79820, 0.30191 0.49629
		0.9	0.50786	0.04226	0.91664, 0.26414 0.65250	0.64827	0.02526	0.89101, 0.28691 0.60410	0.58975	0.01594	0.80713, 0.36497 0.44216
	50	0.5	0.68032	0.03730	0.92746, 0.24556 0.68190	0.65050	0.03410	0.90945, 0.27204 0.63741	0.57610	0.01872	0.80553, 0.33467 0.47086
		0.9	0.70544	0.01720	0.89096, 0.37177 0.51919	0.67673	0.01564	0.87044, 0.38060 0.48984	0.61656	0.01144	0.80940, 0.41673 0.39268
	70	0.9	0.69356	0.02819	0.91520, 0.31706 0.59813	0.67052	0.02584	0.89762, 0.33618 0.56143	0.59264	0.01592	0.80882, 0.36402 0.44481
		0.9	0.70790	0.01276	0.87322, 0.43324 0.43999	0.68693	0.01162	0.85563, 0.43526 0.42037	0.63296	0.00900	0.80789, 0.45184 0.35605
$\beta = 0.5$	30	0.5	0.53196	0.01545	0.93145, 0.37106 0.56039	0.52256	0.01226	0.80758, 0.36741 0.44017	0.56326	0.01757	0.78578, 0.38201 0.40377
		0.9	0.50076	0.00278	0.61889, 0.40570 0.21319	0.49952	0.00257	0.61928, 0.40339 0.21588	0.50111	0.00277	0.61721, 0.39407 0.22314
	50	0.5	0.51477	0.00624	0.69256, 0.39453 0.29802	0.51045	0.00563	0.79455, 0.39222 0.40234	0.54068	0.00797	0.70099, 0.40394 0.29705
		0.9	0.50108	0.00166	0.58866, 0.42670 0.16196	0.49997	0.00158	0.60593, 0.42482 0.18110	0.49808	0.00166	0.58369, 0.41675 0.16695
	70	0.5	0.51100	0.00391	0.64484, 0.41075 0.23409	0.50826	0.00365	0.64154, 0.40867 0.23287	0.53239	0.00491	0.66268, 0.41861 0.24407
		0.9	0.50143	0.00115	0.57417, 0.43798 0.13619	0.50052	0.00110	0.57392, 0.43666 0.13726	0.49782	0.00109	0.56874, 0.42956 0.13918

Table 4: *Estimates of the parameters, MSE, CI and HPD interval for α=0.3, β=2.5*

			MLE			MPS			Bayes		
	n	To	Est.	MSE	CI Length	Est.	MSE	CI Length	Est.	MSE	HPDI Length
$\alpha = 0.3$	30	3	0.33908	0.06482	0.74655, 0.07368 0.67287	0.32510	0.05521	0.73790, 0.07829 0.65961	0.48950	0.04125	0.79820, 0.30191 0.49629
		5	0.33122	0.04056	0.76794, 0.09554 0.67240	0.31172	0.03484	0.75338, 0.09133 0.66206	0.46244	0.03133	0.67978, 0.24478 0.43500
	50	3	0.32277	0.04611	0.74687, 0.09224 0.65463	0.31534	0.04189	0.73531, 0.09535 0.63996	0.46766	0.03350	0.68422, 0.24752 0.43670
		5	0.31184	0.02651	0.71418, 0.10936 0.60482	0.30053	0.02438	0.70142, 0.10662 0.59480	0.43199	0.02197	0.62896, 0.23485 0.39411
	70	3	0.31052	0.03601	0.73439, 0.09710 0.63729	0.30563	0.03362	0.72240, 0.10037 0.62202	0.45009	0.02728	0.65550, 0.24180 0.41370
		5	0.30539	0.01948	0.66915, 0.11946 0.54969	0.29735	0.01840	0.65943, 0.11694 0.54248	0.41117	0.01666	0.59270, 0.22909 0.36360
$\beta = 2.5$	30	3	2.56213	0.33418	3.79115, 1.82942 1.96174	2.53662	0.25619	3.64655, 1.85827 1.78828	2.48647	0.13565	3.18897, 1.86317 1.32580
		5	2.52823	0.18019	3.34280, 1.92556 1.41723	2.53977	0.14785	3.33731, 1.94855 1.38876	2.58703	0.12178	3.19483, 2.02135 1.17349
	50	3	2.53769	0.21201	3.47349, 1.90637 1.56711	2.52595	0.16525	4.80067, 1.93116 2.86951	2.47033	0.10858	3.04908, 1.94862 1.10045
		5	2.52399	0.09420	3.15481, 2.02575 1.12907	2.52977	0.08328	3.15171, 2.03905 1.11266	2.57688	0.07407	3.07648, 2.10567 0.97081
	70	3	2.54392	0.14758	3.34860, 1.96835 1.38025	2.53228	0.12293	3.41558, 1.98083 1.43475	2.47440	0.07359	2.98561, 2.01011 0.97550
		5	2.51170	0.07450	3.04064, 2.07754 0.96310	2.51796	0.06407	3.04172, 2.08817 0.95355	2.55707	0.06053	2.99013, 2.14544 0.84470

Table 5: *Estimates of the parameters, MSE, CI and HPD interval for α=0.5, β=2.5*

			MLE			MPS			Bayes		
	n	To	Est.	MSE	CI Length	Est.	MSE	CI Length	Est.	MSE	HPDI Length
$\alpha = 0.5$	30	3	0.49063	0.06249	0.86146, 0.12443 0.73703	0.46567	0.05350	0.84526, 0.13707 0.70819	0.52890	0.00457	0.76369, 0.28798 0.47572
		5	0.51364	0.03845	0.84534, 0.18898 0.65635	0.48145	0.03331	0.82155, 0.18757 0.63398	0.52551	0.00554	0.74427, 0.30314 0.44113
	50	3	0.49569	0.04508	0.84606 , 0.18054 0.66552	0.47960	0.04058	0.82739, 0.18762 0.63978	0.52462	0.00526	0.74473, 0.29795 0.44678
		5	0.50622	0.02396	0.78083, 0.24494 0.53589	0.48643	0.02201	0.75979, 0.24123 0.51856	0.51811	0.00584	0.71663, 0.31660 0.40003
	70	3	0.49200	0.03487	0.82882, 0.21163 0.61719	0.48054	0.03235	0.81042, 0.21597 0.59444	0.51908	0.00542	0.72744, 0.30437 0.42307
		5	0.50179	0.01767	0.73912, 0.27530 0.46382	0.48773	0.01671	0.72061, 0.27285 0.44775	0.51249	0.00587	0.69522, 0.32677 0.36845
$\beta = 2.5$	30	3	2.60200	0.27134	3.90702, 1.86206 2.04496	2.56793	0.22632	4.41924, 1.86555 2.55369	2.56169	0.11156	3.26219, 1.93257 1.32962
		5	2.52288	0.10553	3.22612, 1.97815 1.24797	2.52547	0.09680	3.23023, 1.97960 1.25063	2.55712	0.07207	3.11970, 2.02915 1.09055
	50	3	2.56577	0.14969	3.42124, 1.96766 1.45359	2.54752	0.13165	3.35205, 1.97103 1.38102	2.54999	0.07483	3.12237, 2.03016 1.09221
		5	2.51464	0.05957	3.03958, 2.08200 0.95758	2.51666	0.05730	3.04707, 2.08138 0.96569	2.54589	0.04696	2.99657, 2.11663 0.87995
	70	3	2.55300	0.09713	3.25062, 2.02987 1.22075	2.53941	0.08819	5.83190, 2.02799 3.80391	2.54610	0.05616	3.04397, 2.08901 0.95496
		5	2.51102	0.04241	2.94791, 2.13973 0.80818	2.51263	0.04122	2.95148, 2.14001 0.81147	2.53840	0.03553	2.92507, 2.16716 0.75791

Table 6: *Estimates of the parameters, MSE, CI and HPD interval for α=0.7, β=2.5*

			MLE			MPS			Bayes		
	n	To	Est.	MSE	CI Length	Est.	MSE	CI Length	Est.	MSE	HPDI Length
$\alpha = 0.7$	30	3	0.66784	0.04386	0.93602, 0.18502 0.75101	0.62763	0.04023	0.91603, 0.21852 0.69750	0.57242	0.01912	0.80563, 0.32935 0.47628
		5	0.69938	0.02430	0.91249, 0.27819 0.63430	0.65117	0.02269	0.88475, 0.29866 0.58609	0.59734	0.01431	0.80928, 0.37808 0.43120
	50	3	0.68838	0.02839	0.91716, 0.29125 0.62591	0.66076	0.02607	0.89869, 0.31178 0.58691	0.59256	0.01550	0.80859, 0.36665 0.44193
		5	0.70956	0.01542	0.88856, 0.38549 0.50307	0.67790	0.01394	0.86624, 0.39026 0.47598	0.62341	0.01003	0.81088, 0.42952 0.38137
	70	3	0.69760	0.01970	0.89736, 0.36228 0.53508	0.67665	0.01817	0.88071, 0.37265 0.50806	0.60782	0.01306	0.80874, 0.39669 0.41204
		5	0.71161	0.01148	0.87161, 0.44849 0.42313	0.68855	0.01042	0.85369, 0.44483 0.40886	0.63988	0.00781	0.80919, 0.46494 0.34425
$\beta = 2.5$	30	3	2.55373	0.14866	3.46516, 1.92668 1.53848	2.52432	0.13224	3.39144, 1.90963 1.48181	2.57792	0.08864	3.25633, 1.96686 1.28947
		5	2.49284	0.06864	3.06494, 2.02924 1.03570	2.49091	0.06396	3.07687, 2.01978 1.05709	2.47366	0.04880	2.97784, 1.99879 0.97905
	50	3	2.53912	0.07923	3.14974, 2.06092 1.08882	2.52252	0.07473	3.13833, 2.04526 1.09306	2.57224	0.06356	3.11481, 2.07595 1.03886
		5	2.50142	0.04142	2.92997, 2.13616 0.79381	2.49971	0.03929	2.93508, 2.12985 0.80523	2.47471	0.03182	2.86599, 2.09898 0.76701
	70	3	2.52649	0.05159	3.00072, 2.13151 0.86921	2.51528	0.04996	2.99912, 2.11911 0.88001	2.56087	0.04968	3.02008, 2.13653 0.88355
		5	2.50640	0.02923	2.86355, 2.19413 0.66942	2.50501	0.02801	2.86692, 2.18957 0.67735	2.48004	0.02321	2.81090, 2.15919 0.65172

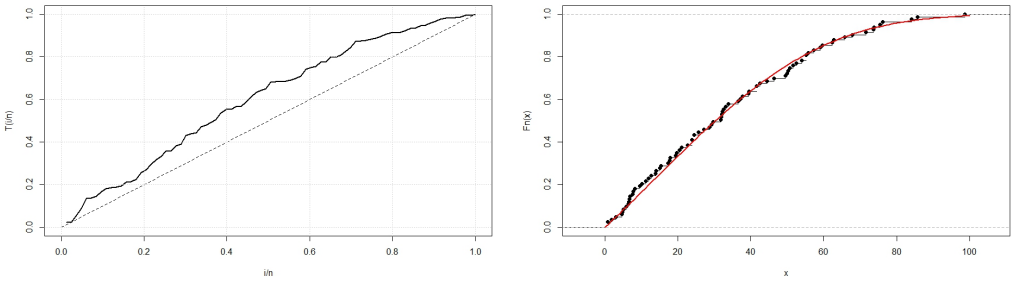
7. Real Data Analysis

In this section, we use a real data set to illustrate the applicability of the proposed work in real life situation. The data-set recently considered in [4], which is taken from the website of the Bureau of Meteorology of the Australian Government (www.bom.gove.au). It represent the monthely rainfall (in mm) between the period of January 2000 to February 2007 in the rain guege station of Carrol, located in the State of New South Wales on the east coast of Australia. The data are given as:

Table 7: *Carrol Data-set (n=83)*

12	22.7	75.5	28.6	65.8	39.4	33.1	84	41.6	62.3	52.5	13.9	15.4	31.9
32.5	37.7	9.5	49.9	31.8	32.2	50.2	55.8	20.4	5.9	10.1	44.5	19.7	6.4
29.2	42.5	19.4	23.8	55.2	7.7	0.8	6.7	4.8	73.8	5.1	7.6	25.7	50.7
59.7	57.2	29.7	32	24.5	71.6	15	17.7	8.2	23.8	46.3	36.5	55.2	37.2
33.9	53.9	51.6	17.3	85.7	6.6	4.7	1.8	98.7	62.8	59	76.1	67.9	73.7
27.2	39.5	6.9	14	3	41.6	49.5	11.2	17.9	12.7	0.8	21.1	24.5	

To ensure that this data-set is appropriate for the illustration of proposed work, first we draw the Total Time on Test (TTT) plot Fig.[1(a)], which shows that the Hazard rate of considered data-set same as the considered distribution (increasing Hazard rate). The Kolmogorov-Smirnov (K-S) test has been used to verify that this data fitted or not on the considered distribution, and found that the the value of K-S statistic is 0.057005 and p-value is 0.9502 which were shows that the considered data fitted on SMD. Graphical representation of K-S test is shown in Fig[1(b)].



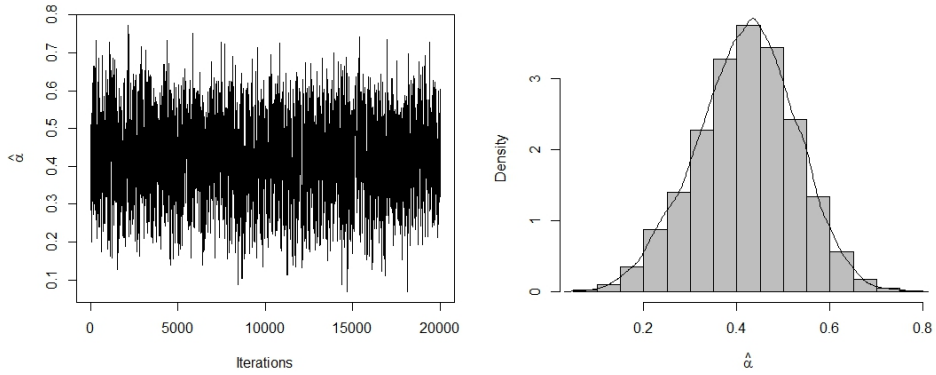
(a) *TTT Plot Carrol Data-set* **(b)** *empirical cdf and fitted cdf plot*

Figure 1: *(a) plot shows the increasing hazard rate and (b) show fitted cdf of real data on SMD*

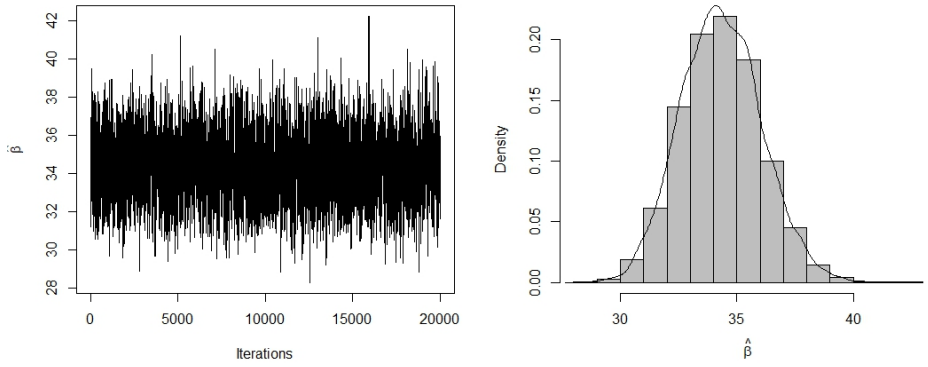
In [4] it is already shows that SMD gives best fit for this data-set in comparison of some other lifetime distribution (Weibull, Gamma, log-Normal and exponential distribution), on basis of their AIC and BIC values shown in table[8]. The MLE, MPSE and Bayes estimate of the parameter for this data-set in case of type-1 censoring given in table[9].

Table 8: *MLE, AIC and BIC values for the Carrol Data-set*

Model	MLE	AIC	BIC
SMD (α, β)	$\hat{\alpha} = 0.4608, \hat{\beta} = 33.9049$	740.3600	741.5220
Weibull (a, b)	$\hat{a} = 1.3665, \hat{b} = 36.9120$	744.4891	745.6511
Gamma(a, b)	$\hat{a} = 1.5160, \hat{b} = 22.3838$	747.3087	748.4706
Exponential (λ)	$\hat{\lambda} = 0.0295$	753.0520	753.6330
Log-Normal (μ, σ)	$\hat{\mu} = 3.1597, \hat{\sigma} = 1.0253$	767.1983	768.3602



(a) *traceplot and marginal density plot for α*



(b) *traceplot and marginal density plot for β*

Figure 2: *Shows that MCMC traceplot and marginal density plot with histogram*

Table 9: *Estimate of parameter for Carrol Data-set*

To	m		MLE est	MLE CI	MPSE est	MPSE CI	Bayes est	Bayes HPD
25	36	α	0.4079	0.8835, 0.0589 0.8246	0.3912	0.8807, 0.0529 0.8278	0.4328	0.7031, 0.1414 0.5617
		β	33.4852	51.7903, 21.6500 30.1403	33.2816	51.4880, 21.5130 29.9749	33.4929	37.3285, 29.2367 8.0918
50	60	α	0.3060	0.7129, 0.0726 0.6403	0.2985	0.7081, 0.0694 0.6387	0.3565	0.6010, 0.1126 0.4884
		β	36.3427	45.1023, 29.2843 15.8180	36.2114	44.9145, 29.1947 15.7197	36.4217	40.0513, 32.6227 7.4286
75	78	α	0.4266	0.6486 , 0.2307 0.4179	0.4152	0.6382 , 0.2223 0.4160	0.4254	0.6131, 0.2169 0.3962
		β	34.1870	39.8435, 29.3336 10.5099	34.2052	39.8796, 29.3382 10.5415	34.3796	37.6232, 31.0938 6.5295

8. Conclusion

In this article, the considered model SMD is another form of Muth distribution by adding a scale parameter. To estimate the parameters of this distribution, we used classical and Bayesian approach under type-1 censoring scheme. In classical approach MLE and MPSE are obtain for time censored data. For the Bayes estimate of the parameters, beta and gamma prior are considered for shape parameter α and scale parameter β respectively. The simulation study was also done by using Monte Carlo method of simulation. From the simulation study it is found that the Bayes estimate perform better than classical method (MLE and MPSE) on the basis of their MSE and length of interval estimation.

References

[1] Muth, E. J. (1977). Reliability models with positive memory derived from the mean residual life function. Theory and applications of reliability, 2, 401-436.

[2] Rinne, H. (1981). Estimating the lifetime distribution of private motor-cars using prices of used cars: The Teissier model. Statistiks Zwischen Theorie und Praxis, 172-184.

[3] Leemis, L. M., & McQueston, J. T. (2008). Univariate distribution relationships. The American Statistician, 62(1), 45-53.

[4] Jodra, P., Jimenez-Gamero, M. D., & Alba-Fernandez, M. V. (2015). On the Muth distribution. Mathematical Modelling and Analysis, 20(3), 291-310.

[5] Jodra, P., Gomez, H. W., Jimenez-Gamero, M. D., & Alba-Fernandez, M. V. (2017). The power Muth distribution. Mathematical Modelling and Analysis, 22(2), 186-201.

[6] Chesneau, C., & Agiwal, V. (2021). Statistical theory and practice of the inverse power Muth distribution. Journal of Computational Mathematics and Data Science, 1, 100004.

[7] Saroj, A., Sonker, P. K., & Kumar, M. (2022). Statistical properties and application of a transformed lifetime distribution: inverse muth distribution. Reliability: Theory & Applications, 17(1 (67)), 178-193.

[8] Sonker, P. K., Kumar, M., & Saroj, A. (2023). Stress–strength reliability models on power-Muth distribution. International Journal of System Assurance Engineering and Management, 14(Suppl 1), 173-195.

[9] Almarashi, A. M., & Elgarhy, M. (2018). A new muth generated family of distributions with applications. J. Nonlinear Sci. Appl, 11, 1171-1184.

[10] Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. Metron, 71(1), 63-79.

[11] Al-Babtain, Abdulhakim A., Ibrahim Elbatal, Christophe Chesneau, and Farrukh Jamal. "The transmuted Muth generated class of distributions with applications." Symmetry 12, no. 10 (2020): 1677.

[12] Almarashi, A. M., Jamal, F., Chesneau, C., & Elgarhy, M. (2021). A new truncated muth generated family of distributions with applications. Complexity, 2021, 1-14.

[13] Kohansal, A., & Bakouch, H. S. (2022). Estimation and prediction on power Muth distribution with progressive censored data: a Bayesian approach. Communications in Statistics-Simulation and Computation, 1-23.

[14] BICER, H. D., & *ÖZTÜ*RKER, B. (2021). Estimation procedures on Type-II censored data from a scaled Muth distribution. Sigma Journal of Engineering and Natural Sciences, 39(2), 148-158.

[15] Basu, S., Singh, S. K., & Singh, U. (2017). Parameter estimation of inverse Lindley distribution for Type-I censored data. Computational Statistics, 32, 367-385.

[16] Epstein, B. (1954). Truncated life tests in the exponential case. The Annals of Mathematical Statistics, 555-564.

[17] Sinha S (1986) Reliability and life testing. Wiley Eastern Limited. https://books.google.co.in/books? id=esWaAAAACAAJ

[18] Cohen, A. C. (1965). Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. Technometrics, 7(4), 579-588.

[19] Cheng, R. C. H., & Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. Journal of the Royal Statistical Society: Series B (Methodological), 45(3), 394-403.

[20] Ranneby, B. (1984). The maximum spacing method. An estimation method related to the maximum likelihood method. Scandinavian Journal of Statistics, 93-112.

[21] Cheng, R. C. H., & Traylor, L. (1995). Non?regular maximum likelihood problems. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 3-24.

[22] Pitman EJ (1979) Some basic theory for statistical inference, vol 7. Chapman and Hall, London

[23] Cheng, R. C. H., & Iles, T. C. (1987). Corrected maximum likelihood in non?regular problems. Journal of the Royal Statistical Society: Series B (Methodological), 49(1), 95-101.

[24] Coolen, F. P. A., & Newby, M. J. (1990). A note on the use of the product of spacings in Bayesian inference. Department of Mathematics and Computing Science, University of Technology.

[25] Shao, Y., & Hahn, M. G. (1999). Maximum product of spacings method: a unified formulation with illustration of strong consistency. Illinois Journal of Mathematics, 43(3), 489-499.

[26] Ghosh, K., & Jammalamadaka, S. R. (2001). A general estimation method using spacings. Journal of Statistical Planning and Inference, 93(1-2), 71-82.

[27] Anatolyev, S., & Kosenok, G. (2005). An alternative to maximum likelihood based on spacings. Econometric Theory, 21(2), 472-476.

[28] Singh, U., Singh, S. K., & Singh, R. K. (2014). A comparative study of traditional estimation methods and maximum product spacings method in generalized inverted exponential distribution. Journal of Statistics Applications & Probability, 3(2), 153.

[29] Metropolis, N., & Ulam, S. (1949). The monte carlo method. Journal of the American statistical association, 44(247), 335-341.

[30] Roberts, G. O., & Smith, A. F. (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. Stochastic processes and their applications, 49(2), 207-216.

[31] Chib, S., & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. The american statistician, 49(4), 327-335.

[32] J. O. Berger, Statistical Decision Theory and Bayesian Analysis, Springer Series in Statistics, Springer, New York, NY, USA, 2nd edition, (1985).

[33] Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. Psychological review, 70(3), 193.

[34] Chen, M. H., & Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. Journal of computational and Graphical Statistics, 8(1), 69-92.

[35] Box, G. and Tiao, G. (1973). Bayesian Inference in Statistical Analysis. Addison-Wesley, Reading, Massachusetts.

ROBUST MAHALANOBIS DEPTH BASED ON MINIMUM REGULARIZED COVARIANCE DETERMINANT ESTIMATOR FOR HIGH-DIMENSIONAL DATA

R Muthukrishnan and Surabhi S Nair
•
Department of Statistics, Bharathiar University, Coimbatore 641046, Tamil Nadu, India
muthukrishnan1970@gmail.com, surabhinair93@gmail.com

Abstract

Handling of high-dimensional data is an important issue in robust literature. For analyzing data, location measure plays a vital role in almost all statistical methods. The location parameter of a distribution is used to find the central value. Many computational methods are used to find the measure of location for analyzing data. The data depth procedure is one approach to finding the true representative of the entire data and it is one of the key concepts in multivariate data analysis. Data depth is a term used to describe how deep a particular point is inside the broad multivariate data cloud. Instead of the typical smallest to biggest rank, the sample points can be ordered from the center outward. Mahalanobis depth is one of the popular depth procedures. The traditional approach used to find Mahalanobis depth is based on Mahalanobis distance, it is based on the classical sample mean vector and covariance matrix. So the conventional Mahalanobis depth is sensitive to outliers and may fail when the data is contaminated. To solve this problem, the Minimum Covariance Determinant (MCD) estimators are used instead of classical estimators. However, the MCD estimators cannot be calculated in high dimensional data sets, in which the variable number is higher than the subset size. To calculate Mahalanobis depth values in high dimensional data, propose a new depth function namely the Robust Regularized Mahalanobis Depth (RRMD), which can be calculated in high dimensional data sets. The proposed procedure is based on Minimum Regularized Covariance Determinants (MRCD) estimators, this study shows that the proposed depth function is successful in finding the deepest point in high dimensional data sets with real and simulation studies up to a certain level of contamination.

Keywords: mahalanobis depth, outliers, robust distance, minimum covariance determinant estimator, minimum regularized covariance determinant estimators

I. Introduction

Data depth is a key concept in the nonparametric method of multidimensional analysis of data. The idea of data depth was proposed by Tukey [13] as a graphical tool for displaying two-dimensional data sets, and it has now been expanded to include multivariate data sets, Donoho and Gasko [3]. Data depth is a statistical technique that describes data distribution in accordance with center-outward ordering instead of density or linear ordering, Liu [6], Modarres [10]. According to Liu et al. [8], the idea of data depth is being used for statistical analysis with multiple variables since it offers a nonparametric method. Several researchers have established various notions of depth preliminaries in the literature.

Researchers are looking for solutions in robust depth procedures to address the sensitivity issue in high-dimensional data analysis. Also can build robust depth procedures to handle the

presence of outliers by replacing the classical estimators with robust estimators such as M-estimators, Minimum Covariance Determinant (MCD) estimators, Mia Hubert and Michiel Debruyne [4], Minimum Volume Ellipsoid (MVE) estimators, Stefan Van Aelst and Peter Rousseeuw [12], and Minimum Regularized Covariance Determinants (MRCD), Kris Boudt, Peter Rousseeuw [1].

A very reliable estimator of multivariate location and scatter is the Minimum Covariance Determinant (MCD) approach. Given an $n \times p$ data matrix $Y = (Y_2, \dots, Y_p)'$ with $y_i = (y_{i1}, y_{i2}, \dots, Y_{ip})'$, its objective is to find h observations whose sample covariance matrix has the lowest possible determinant. Here $h < n$ is fixed. The average of these h points serves as the MCD estimate of location, while the scatter estimate represents a multiple of the covariance matrix. The maximum possible breakdown value is found in the MCD, which also possesses a bounded influence function (i.e. 50%) when $h = [(n + p + 1)/2]$. This covariance matrix of any h-subset must be non-singular for the dimension p to meet $p < h$, which is a significant restriction of the MCD technique. In fact, taking n > 5p is frequently advised for the estimator's accuracy. This restriction leaves a hole in the selection of high-breakdown techniques. In order to reduce this gap, Boudt et al. (2020) suggested modifying the MCD so that it can be used for high dimensions. A regularized covariance estimate, which is a weighted mean of the sample covariance of the h-subset and a preset positive definite target matrix, is intended to replace the subset-based covariance. The regularized covariance based on the h-subset that results in the least overall determinant is then the Minimum Regularized Covariance Determinant (MRCD) estimate. In addition to supporting high dimensions, the MRCD estimator's key characteristics include maintaining the MCD estimator's excellent breakdown qualities and being highly conditioned by construction.

This study proposes a new depth procedure, namely Robust Regularized Mahalanobis Depth (RRMD), which can be used to find the location measure in high dimensional datasets by comparing the existing method. In the suggested algorithm, Mahalanobis depth is obtained based on MRCD estimators instead of classical mean and covariance matrix. Using various kinds of simulation tests and two real datasets, it is examined if the recommended algorithm for location estimation in high dimensional data produces accurate results even if the data is contaminated.

The rest of this paper is structured as follows. Section 2 describes traditional Mahalanobis depth, the robust estimator used, and the proposed method. The results and discussions based on the real data and simulation study will be given in section 3. The conclusion will be provided in the last section.

II. Methods

Today, in numerous domains, huge amounts of data are produced and tainted by noise. It is important to establish training methods that are resilient to data instabilities and disruptions. In this section, the foundations of controlled learning are discussed, including the traditional Mahalanobis Depth, the estimator used in this study – Minimum Regularized Covariance Determinant Estimator, and the proposed depth procedure - Robust Regularized Mahalanobis Depth.

I. Mahalanobis Depth

Mahalanobis depth (MD) was first described by Liu et al. [7] from Mahalanobis distance. Mahalanobis [9] established the statistical idea of generalized distance which is calculated by using a classical mean vector and covariance matrix. For determining the Mahalanobis depth of an observation, the Mahalanobis distance is used. The positive inverse of Mahalanobis distance is termed as Mahalanobis depth. For an observation $y \in S \subset R^d$ about a d- dimensional data,
The squared Mahalanobis distance (D) and Mahalanobis depth (MD) are given by

$$D(Y, \bar{Y}, S) = (Y - \bar{Y})' S^{-1} (Y - \bar{Y})$$

$$MD(Y, \bar{Y}, S) = [1 + D(Y, \bar{Y}, S)]^{-1}$$

where $\bar{Y}$ and S are the mean vector and sample covariance matrix. Since it is reliant on non-robust parameters like the mean and dispersion matrix, this algorithm lacks to be reliable.

To get a reliable result MD is calculated using a robust location vector and covariance matrix using MCD estimator instead of classical mean vector and covariance matrix. Generally, the Minimum Covariance Determinants (MCD) estimators are used for this aim. Due to the failure of MCD estimators to be generated for high-dimensional datasets, this approach is not applicable in these cases. So, the Mahalanobis depth using MCD estimator can't be applicable when want to analyze a high dimensional data set.

II. Minimum Regularized Covariance Determinant Estimator (MRCD)

The MRCD estimator is a modified version of NCD estimators for high-dimensional data and was proposed by Boudt et. al. [1]. To guarantee that the MRCD scatter estimator is scale equivariant and location unvarying, as is common in the literature, first, standardize the variables. The use of a trustworthy univariate location and scale estimate is required for standardization. For this, the median of each subset is calculated and placed in a location vector called m_x. Additionally, each variable's scale using the Qn estimator of Rousseeuw and Croux (1993) is calculated, then insert these scales into d_x, the diagonal matrix.

Let $X = (x_1, x_2, \dots, x_n)'$ be an $n \times p$ matrix with $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$; its goal is to identify the h observations with the sample covariance matrix with the lowest determinant. The term H refers to a set of h variables denoting the data contained in the subset, while the term $\mathcal{H}_h$ refers to the compilation of all of these sets. The corresponding $h \times p$ submatrix of X is denoted by X_H for a particular H in $\mathcal{H}_h$. The mean and sample scatter matrix of X_H are given by

$$M_X(H) = h^{-1} X_H' I_h \tag{1}$$

$$S_X(H) = (h-1)^{-1}(X_H - M_X(H))'(X_H - M_X(H)) \tag{2}$$

After that, the MCD method seeks to minimize the determinant of $S_X(H)$ for all $H \in \mathcal{H}_h$.

$$H_{MCD} = \underset{H \in \mathcal{H}_h}{argmin} \left(det(S_X(H))^{1/p} \right) \tag{3}$$

For statistical considerations, eqn (3) takes the p^{th} root of the determinant. The geometric average of its eigenvalues is the p^{th} root of the determinant of the scatter matrix. It is referred to as the standardized generalized variance.

The mean of the h-subset is used to define the MCD estimate of location M_{MCD}, while the MCD scatter estimate is expressed as a multiple of the sample scatter matrix, and is given by

$$M_{MCD} = M_X(H_{MCD}) \tag{4}$$

$$S_{MCD} = C_\alpha S_X(H_{MCD}) \tag{5}$$

where C_α is a consistency factor that is based on the trimming percentage = $(n-h)/n$ and is similar to the one provided by Rousseeuw and Croux [11].

The standardized observation is given by

$$Z_i = d_x^{-1}(x_i - m_x) \tag{6}$$

The regularized scatter matrix of the standardized observation is

$\mathcal{S}(H) = \rho T + (1-\rho)C_\alpha S_Z(H)$
where $S_Z(H)$ is defined in (2), however, in the case of Z, c is the same consistency parameter as in (5).

Let A be the diagonal matrix containing eigenvalues of T, and the orthogonal matrix Q contains the relevant eigenvectors. Utilizing the spectral decomposition $T = QAQ'$ will be practical.

Now,
$$\mathcal{S}(H) = QA^{1/2}[\rho I + (1-\rho)C_\alpha S_W(H)]AA^{1/2}Q' \quad (7)$$

where W is the $n \times p$ matrix consisting of the transformed standardized observations $w_i = A^{-1/2}Q'Z_i$, and $S_W(H) = A^{-1/2}Q'S_ZQA^{-1/2}$

The MRCD subset is given by

$$H_{MRCD} = \underset{H\in\mathcal{H}_h}{argmin}\left(det(\rho I + (1-\rho)C_\alpha S_W(H))^{1/p}\right) \quad (8)$$

The MRCD location and scatter estimations of the initial data matrix X are defined as follows

$M_{MRCD} = m_X + d_x M_Z(H_{MRCD})$

$S_{MRCD} = C_\alpha S_X(H_{MRCD})$.

III. Robust Regularized Mahalanobis Depth

The proposed depth procedure namely Robust Regularized Mahalanobis Depth (RRMD), it is based on the Minimum Regularized Covariance Determinant estimator, which can be calculated in high dimensional data sets. ie., Mahalanobis depth can be calculated using robust location and covariance matrix calculated, the MRCD estimator instead of classical mean vector and covariance matrix. The robust MRCD estimator was proposed by Boudt et. al. [1] to locate the robust measure of location and scatter for high dimensional data. The computational depth procedure for RRMD is as follows.

Let M_{MRCD}, and S_{MRCD} be the location and scatter matrix using the MRCD estimator, D_{MRCD} diagonal matrix, which consists of the diagonal elements of S_{MRCD} matrix. The Robust Regularized Mahalanobis Depth obtained from the regularized squared mahalanobis distance, $D(Y, M_{MRCD}, S_{MRCD}) = (Y - M_{MRCD})' D_{MRCD}{}^{-1}(Y - M_{MRCD})$,, and is given by

$$MD(Y, M_{MRCD}, S_{MRCD}) = [1 + D(Y, M_{MRCD}, S_{MRCD})]^{-1}$$

Let $Y = (Y_2, \dots, Y_d)$ be a d dimensional multivariate data set and y be a numerical vector whose depth is to be calculated. The following steps are carried out to find the proposed method.

i. By using the dataset calculate robust MRCD location (M_{MRCD}) and scatter estimators (S_{MRCD}). Obtain D_{MRCD} diagonal matrix, which consists of the diagonal elements of S_{MRCD} matrix.

ii. The Regularized Squared Mahalanobis distance can be calculated from (i)

ie., $D(Y, M_{MRCD}, S_{MRCD}) = (Y - \mu_{MRCD})' D_{MRCD}{}^{-1}(Y - \mu_{MRCD})$

iii. S_D be the sorted distance given in (ii)

iv. M_{S_D} be the median from the distance from (iii),

$M_{S_D} = Median(S_D)$

v. D_y be the difference between Regularized Squared Mahalanobis distance value from (ii) and median from (iv), ie., $D_y = D(Y, \mu_{MRCD}, \Sigma_{MRCD}) - M_{S_D}$

vi. $Abs(D_y)$ be the absolute value of the difference in (v)

vii. Now, the proposed depth procedure, Robust Regularized Mahalanobis Depth can be calculated by $RRMD_y = [1 + Abs(D_y)]^{-1}$

III. Experimental Results

The performance of the proposed RRMD procedure over the classical MD procedure, the experiments were conducted under actual and simulation environments by computing location measure corresponding to the deepest point in high-dimensional data and thus the obtained results are demonstrated in this sections.

I. Real data study

The proposed depth function can be used to find the location parameter in high-dimensional datasets. Two real data set is used here to evaluate the performance of the suggested algorithm compared with the existing method. The First one is the brain data which is also from the "rda" package of R software. The brain data contains two objects, namely the microarray expression data for 42 brain cancer samples, and the class labels for these samples. An expression data matrix (42x5597) and a class label vector for 42 samples. The second one is the NCI60 data which is obtained from the "ISLR" package of R software. The data contains expression levels on 6830 genes from 64 cancer cell lines. Due to the enormous dimensions of these datasets, this study only used the first p ($p > 3 * n$) variables for convenience. The results obtained from the real data study are summarized in the form of Table 1.

Tab le 1: *Deepest point and observation number of brain data and NCI60 data*

Methods		MD	RRMD
Brain data	Highest Depth Value	0.148 (0.103)	0.148 (0.103)
	Deepest Point	42 (16)	4 (4)
NCI60 data	Highest Depth Value	0.097 (0.086)	0.031 (0.029)
	Deepest Point	23 (5)	49 (49)

(.) – Without outlier

The suggested algorithm, Robust Regularized Mahalanobis Depth (RRMD) is obtained based on MRCD estimators and then the location parameter is calculated using the depth values for the high dimensional data set. Table 1 shows that the proposed method gives a same location under with/without outlier conditions, but the conventional method differs from it.

II. Simulation Study

Simulation study were carried out under two different dimensions along with various amounts of contamination are employed to compare the performance of the proposed methodology to the current approach. The experiments were carried out by computing the maximum depth values that correspond to location measure.

First generated data with dimension, 100×300, with mean vector $\mu = (0,0,\dots,0)_{1\times 300}$and covariance matrix$\Sigma = I_{300}$. Here n=100, and p=300. Further same experiments were performed under various levels of contaminations, such as ε= 10%, 15%, 20%, 30% (For Location, $\mu = (1.5, 1.5, \dots, 1.5)_{1\times 300}$ and $\Sigma = I_{300}$, Scale, $\mu = (0,0,\dots,0)_{1\times 300}$, and $\Sigma = 1.5I_{300}$, Location and Scale, $\mu = (1.5, 1.5, \dots, 1.5)_{1\times 300}\ and\ \Sigma = 1.5I_{300}$) are taken into account. Also, the same experiment is repeated for 200×400 dimensional data. The results obtained from the simulation study is given in table 2 and 3 respectively.

From the simulation study it is concluded that the suggested depth procedure, RRMD can tolerate and gives the same deepest point up to a certain level of contamination. The MD method fails to provide identical location measurements even if the data contamination is very low.

Table 2: *Deepest point and observation number* under various contamination models

Simulation Study 1 Dimension:100 × 300; n=100, p=300						
Highest Depth Values and the corresponding observation						
ε	**Location Contamination**		**Scale Contamination**		**Location-Scale Contamination**	
	MD	***RRMD***	***MD***	***RRMD***	***MD***	***RRMD***
0.10	0.099 (96)	0.019 (33)	0.870 (41)	0.018 (33)	0.044 (93)	0.019 (33)
0.15	0.034 (81)	0.019 (33)	0.056 (92)	0.018 (33)	0.086 (79)	0.019 (33)
0.20	0.126 (10)	0.019 (33)	0.108 (77)	0.018 (33)	0.032 (88)	0.019 (33)
0.30	0.014 (38)	0.019 (28)	0.082 (30)	0.018 (22)	0.046 (90)	0.019 (22)
(.) – Observation number						

Table 3: *Deepest point and observation number* under various contamination models

Simulation Study 2 Dimension:200 × 400; n=200, p=400						
Highest Depth Values and the corresponding observation						
ε	**Location Contamination**		**Scale Contamination**		**Location-Scale Contamination**	
	MD	RRMD	MD	RRMD	MD	RRMD
0.10	0.013 (5)	0.009 (17)	0.100 (20)	0.009 (17)	0.124 (86)	0.009 (17)
0.15	0.151(61)	0.009 (17)	0.162 (18)	0.009 (17)	0.073 (130)	0.010 (17)
0.20	0.356 (33)	0.010 (17)	0.406 (155)	0.009 (17)	0.042 (128)	0.010 (124)
0.30	0.126 (21)	0.010 (17)	0.035 (80)	0.009 (131)	0.112 (80)	0.010 (128)
(.) – Observation number						

IV. Conclusion

Conventional methods should work reasonably well if certain assumptions are true, however, they may not be trustworthy if one or more of these assumptions are erroneous. Both sample mean vector and covariance matrix are extremely susceptible to anomalies. As a result, when the data contains anomalies, the traditional Mahalanobis depth fails to generate reliable results. For non-normal conditions, a robust alternative is required to improve accuracy even when the data somewhat depart from the model assumptions. When robust estimators such as MCD and MRCD are used the analysis performs well compared to traditional methods. To find the location measure in high dimensional data, this paper proposed a new robust depth procedure namely RRMD. The proposed method is compared with the existing procedure and gives reliable results up to certain levels of contamination. Robust methods perform well even when the model assumptions are not met. The study came to the conclusion that for robust and affine equivariant location, the proposed depth procedure gives better results followed by the existing method. By finding the deepest point in a dataset instead of relying on a more conventional method of determining location, the research groups can find the best location with greater precision when using these methods.

References

[1] Boudt K., Rousseeuw P.J. and Vanduffel S. (2020). The minimum regularized covariance determinant estimator. *Statistics and Computing*, 30:113–128.

[2] Choral C.Y. and Rousseeuw P.J. (1992) Integrating a high-breakdown option into discriminant analysis in exploration geochemistry. Journal of Geochemical Exploration, 43:191-203.

[3] Donoho D.L. and Gasko M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness, *Annals of Statistics*, 2:1803–1827.

[4] Hubert M. and Debruyne M. (2009). Minimum covariance determinant. *WIREs Computational Statistics*, 2:36-43.

[5] Hubert M. Debruyne M. and Rousseeuw P.J. (2017). Minimum covariance determinant and extensions. *WIREs Computational Statistics*, 10: e-1421.

[6] Liu R.Y. (1992). Data depth and multivariate rank tests. *In: Dodge, Y. (ed.), L1-Statistics and Related Methods*, North-Holland (Amsterdam). 279–294.

[7] Liu R.Y., Parelius J.M. and Singh K. (1993.) Multivariate analysis by data depth: Descriptive Statistics, Graphics and Inference. *The Annals of Statistics*, 27:783-858.

[8] Liu R. Y. and Singh K. A. (2012). Quality Index Based on Data Depth and Multivariate Rank Tests. *Journal of the American Statistical Association,* 88:252-260.

[9] Mahalanobis P. (1936). On the generalized distance in statistics. *Proceedings of the National Academy India*, 12:49–55.

[10] Modarres R. (2014). Data Depth. *In: Lovric, M. (eds) International Encyclopedia of Statistical Science*. *Springer, Berlin, Heidelberg*, 334-336.

[11] Rousseeuw P. J. and Croux C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88:1273–1283.

[12] Stefan V.A. and Rousseeuw P. J. (2009). Minimum volume ellipsoid. *WIREs Computational Statistics*, 1:71-82.

[13] Tukey J. W. (1975). Mathematics and the picturing of data. *In: Proceeding of the International Congress of Mathematicians*, Vancouver, 523–531.

RT&A, No 4 (76)
Volume 18, December 2023

DISCRETE INVERSE GAMMA DISTRIBUTION BASED LOAD-SHARE MODEL WITH APPLICATION

Rachna Srivastava, Pramendra Singh Pundir
•
Department of Statistics, University of Allahabad, Prayagraj, 211002,
U.P., India
srivastava.rachna0607@gmail.com, pspundir@allduniv.ac.in

Abstract

In reliability engineering, the multi-component load-sharing models are being used to amplify system's reliability. This study consists of the k-component load-sharing parallel system model considering each component's failure time distribution as discrete inverse gamma. The classical and Bayesian analysis for this model is performed. The maximum likelihood estimates along with their standard errors for the parameters, system's reliability function, hazard rate function and reversed second rate of failure function are obtained. The asymptotic confidence intervals as well as two bootstrap intervals like bootstrap-p and bootstrap-t confidence intervals are constructed. Further, Bayes estimates along with their posterior standard errors and highest posterior density credible intervals for the parameters and system's reliability characteristics are obtained by using Markov Chain Monte Carlo techniques. A detailed simulation table is formed to demonstrate the effectiveness of the theory developed. Finally, a real data set is used to illustrate the applicability of the model.

Keywords: Load-share system model, Discrete inverse gamma distribution, Bootstrapping, Bayesian estimation, MCMC technique

1. Introduction

In today's engineering world, the products with high reliability are in demand. It can be accomplished with planned maintenance, improving components' reliabilities, reassignment of components etc. In this regard, manufacturers usually use redundancy techniques. While using redundancy techniques, it is assumed that the components within the system are performing independently. However, in practice, many multi-component systems work as load-sharing model which leads the independence assumption to be not valid any more. In load-sharing model (or dynamic system model), on a component's failure, the workload will be redistributed to the other working components, which imparts increased load on them. Originally, Daniels [1] developed the first load-sharing design to study textile strength.

In last two decades, various authors have discussed load-share modeling and estimated its parameters in different contextual aspects. Kim and Kvam [3] considered equal load-share rule for estimating the load-share parameters of multi-component system in parametric setup with failure time distribution as exponential. Singh et al. [12]

performed the classical and Bayesian inference for multi-component load-share system by assuming that the components have a combination of constant and linear failure rates. Park [5, 6] performed the classical and Bayesian inference for such models by assuming the underlying lifetime distribution as exponential or Weibull.

There are various real examples where load-share modeling can be used such as in mechanical and civil engineering for welded joints (enhances the stress on other joints), in textile engineering and materials science for crack growth induced by fatigue or material degradation (bigger cracks will grow faster than smaller cracks), in a power plant for electric generators (electrical load is shared), in a human body for two kidneys (simultaneously work together) and so forth.

There are some basic continuous lifetime distributions such as exponential, Weibull, Lindley and log-normal, which have been explored for analyzing load-share models. Some distributions are recently been used for this purpose like Chen, modified Weibull, and exponentiated Pareto distributions which are discussed by Pundir and Gupta [7], Singh and Goel [11] and Zhang et al. [16], respectively. However, there are various situations in which discrete distributions can perform well like number of shocks, number of patients in a ward, etc. Singh et al. [13] have dealt with discrete load-share modeling situation using geometric distribution.

Considering this, a load-share system model is developed and analyzed with discrete inverse gamma distribution (DIGD) in the present work. Hussain and Ahmad [2] obtained DIGD by discretizing the inverse gamma distribution (IGD). Its cumulative distribution function (CDF), reliability function, probability mass function (PMF), hazard rate function (HRF) and reversed second rate of failure (RSRF) are, respectively,

$$F(x) = \int_{1/x}^{\infty} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{\alpha-1} e^{-\frac{y}{\beta}} dy = \frac{\Gamma(\alpha, \frac{1}{\beta x})}{\Gamma(\alpha)} \quad ; x = 0, 1, \ldots ; \alpha > 0, \beta > 0$$

where $\Gamma(\alpha, \frac{1}{\beta x}) = \int_{1/x}^{\infty} \frac{1}{\beta^{\alpha}} y^{\alpha-1} e^{-\frac{y}{\beta}} dy$ and Γ represents the upper incomplete gamma function,

$$R(x) = 1 - F(x) = \int_{0}^{1/x} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{\alpha-1} e^{-\frac{y}{\beta}} dy = \frac{\gamma(\alpha, \frac{1}{\beta x})}{\Gamma(\alpha)} \quad ; x = 0, 1, \ldots ; \alpha > 0, \beta > 0$$

where $\gamma(\alpha, \frac{1}{\beta x}) = \int_{0}^{1/x} \frac{1}{\beta^{\alpha}} y^{\alpha-1} e^{-\frac{y}{\beta}} dy$ and γ represents the lower incomplete gamma function,

$$P[X = x] = p(x) = R(x) - R(x+1) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_{1/(x+1)}^{1/x} y^{\alpha-1} e^{-y/\beta} dy$$

$$= \frac{\gamma(\alpha, \frac{1}{\beta x}, \frac{1}{\beta(x+1)})}{\Gamma(\alpha)} \quad ; x = 0, 1, \ldots ; \alpha > 0, \beta > 0 \tag{1}$$

where $$\gamma\Big(\alpha, \frac{1}{\beta x}, \frac{1}{\beta(x+1)}\Big) = \gamma\Big(\alpha, \frac{1}{\beta x}\Big) - \gamma\Big(\alpha, \frac{1}{\beta(x+1)}\Big),$$

$$h(x) = \frac{p(x)}{R(x)} = \frac{\gamma(\alpha, \frac{1}{\beta x}, \frac{1}{\beta(x+1)})}{\gamma(\alpha, \frac{1}{\beta x})} \tag{2}$$

and

$$h_2(x) = log\left[\frac{F(x)}{F(x-1)}\right] = log\left[\frac{\Gamma(\alpha, \frac{1}{\beta x})}{\Gamma(\alpha, \frac{1}{\beta(x-1)})}\right]$$

where, α and β are shape and scale parameters.

Pundir et al. [8] obtained its reversed second rate of failure (RSRF) function and discussed its usefulness. They also discussed the classical and Bayesian inference methods for DIGD. DIGD can be applied to various applications where IGD is being used like in radar detection by Shang and Song [10], Stinco et al. [14] and in fading modeling by Yoo et al. [15], Ramirez-Epi'nosa and Lopez-Martinez [9], etc.

The current study deals with constructing a load-share parallel system model where the lifetime distribution of each component is considered as DIGD. In section 2, model description is presented. The reliability characteristics of the system are derived under section 3. In section 4, maximum likelihood (ML) estimation, bootstrapping and Bayesian estimation techniques are applied to obtain the estimates along with standard errors (SEs)/ posterior standard errors (PSEs) and confidence intervals (CIs)/ highest posterior density (HPD) credible intervals of the parameters and reliability characteristics of the load-share system model. Bayesian estimation is applied under informative as well as non-informative priors by using Markov Chain Monte Carlo (MCMC) techniques such as Gibbs sampler and Metropolis-Hastings (MH) algorithm. A simulation study is performed to compare the discussed estimation techniques and the results are given under section 5. Section 6 demonstrates the applicability of the proposed model to real data set. Finally, the article is concluded exhibiting some concluding remarks in section 7.

2. Model Description

A load-sharing system is modeled and analyzed under the following assumptions:

- A load-sharing system containing k-components is bearing a constant load which is equally shared by all its components.
- A test is prepared to record the failure times of n such i.i.d. parallel systems.
- Let T_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, k)$ be the failure time spacing between $(j-1)^{th}$ and j^{th} components in the i^{th} parallel system.
- On the successive failures of the components within the system, the load imposed on other remaining surviving components increases.
- The hazard rate of a remaining surviving component varies when the sharing load changes.
- The failure time distribution of each component in the system is independent.
- Initially, the hazard rate of each of the k-components is denoted by $h(t; \alpha, \beta_1)$. When the first component fails, the hazard rate of the remaining $(k-1)$ components changes to $h(t; \alpha, \beta_2)$ and so on. After the $(k-1)^{th}$ component failure, the hazard rate of the last surviving component changes to $h(t; \alpha, \beta_k)$.

- Each component is pertaining a failure time PMF and HRF given in equations (1) and (2), respectively.

Taking these assumptions into consideration, the hazard rate of the j^{th} component when $(j-1)$ components have already failed is

$$h(t_{ij}) = (k-j+1)h(t_{ij};\alpha,\beta_j) \;\; ;i=1,2,\ldots,n;j=1,2,\ldots,k$$

$$= (k-j+1)\frac{\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)}\right)}{\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)}$$

and the conditional failure time PMF for the j^{th} component in the i^{th} system is given by

$$p(t_{ij}) = \frac{(k-j+1)}{\Gamma(\alpha)^{(k-j+1)}}\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)}\right)\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)\right)^{k-j} \tag{3}$$

Thus, the likelihood function for the i^{th} system is

$$L_i(t_{i1},\ldots,t_{ik}|\alpha,\Lambda) = k!\prod_{j=1}^{k}\left[\frac{1}{\Gamma(\alpha)^{(k-j+1)}}\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)}\right)\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)\right)^{k-j}\right] \tag{4}$$

where, $\Lambda = (\beta_1,\beta_2,\ldots,\beta_k)$.

3. System Reliability Characteristics Measures

The reliability function of the 1-out-of-k load-share system is obtained by considering its mechanism i.e., the system works till its last component is functioning. Thus, the reliability of the system is given as

$R_s(t)$=P[at least one of k components is in operation]

=P[all k components are in operable state]
+P[any one component fails and remaining $(k-1)$ components survive]+...
+P[any $(k-1)$ components fail and last component is in operation]

$$= \left[\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_1 t}\right)\right]^k + k\left[1-\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_1 t}\right)\right]\left[\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_2 t}\right)\right]^{k-1} + \ldots$$
$$+k\left[1-\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_1 t}\right)\right]\left[1-\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_2 t}\right)\right]\ldots\left[1-\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_{k-1} t}\right)\right]\left[\frac{1}{\Gamma(\alpha)}\gamma\left(\alpha,\frac{1}{\beta_k t}\right)\right]$$

The CDF, PMF, HRF and RSRF function of the system's failure time T are respectively, given by

$$F_s(t) = 1 - R_s(t)$$

$$p_s(t) = R_s(t) - R_s(t-1)$$

$$h_s(t) = \frac{p_s(t)}{R_s(t)}$$

and

$$h_{2_s}(t) = log\frac{F_s(t)}{F_s(t-1)}.$$

4. Classical And Bayesian Inference

The likelihood function for n i.i.d. systems is obtained by using equation (4) as

$$L(T|\alpha,\Lambda) = (k!)^n \prod_{i=1}^{n}\prod_{j=1}^{k}\left[\frac{1}{\Gamma(\alpha)^{(k-j+1)}}\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)}\right)\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)\right)^{k-j}\right] \quad (5)$$

and the corresponding log-likelihood function is

$$logL = nlog(k!) + \sum_{i=1}^{n}\sum_{j=1}^{k}\Bigg[-(k-j+1)\log(\Gamma(\alpha)) + \log\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)}\right)\right) + (k-j)\log\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)\right)\Bigg] \quad (6)$$

4.1. Maximum Likelihood Estimation

The ML estimates $(\hat{\alpha},\hat{\Lambda})$ of (α,Λ) can be obtained on solving the following (k+1) normal equations:

$$\frac{\partial logL}{\partial\alpha} = -\frac{nk(k+1)}{2}\psi(\alpha) + \sum_{i=1}^{n}\sum_{j=1}^{k}\frac{1}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}\left[\frac{\partial}{\partial\alpha}\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)\right) - \frac{\partial}{\partial\alpha}\left(\gamma\left(\alpha,\frac{1}{\beta_j(t_{ij}+1)}\right)\right)\right] + \sum_{i=1}^{n}\sum_{j=1}^{k}(k-j)\frac{1}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}\cdot\frac{\partial}{\partial\alpha}\left(\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right)\right) = 0 \quad (7)$$

$$\frac{\partial logL}{\partial\beta_j} = \frac{-n\alpha(k-j+1)}{\beta_j} + \frac{1}{\beta_j}\sum_{i=1}^{n}\frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})} + \frac{(k-j)}{\beta_j}\sum_{i=1}^{n}\frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})} = 0 \quad (8)$$

where $\psi(\alpha) = \frac{\partial}{\partial\alpha}log(\Gamma(\alpha))$ is the digamma function and $\frac{\partial}{\partial\alpha}\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right) = \int_0^{\frac{1}{\beta_j t_{ij}}} v^{(\alpha-1)}e^{-v}log(v)dv$.

Since closed form solutions cannot be obtained from the equations (7) and (8). Therefore, any numerical iterative procedure such as Newton-Raphson method can be used here. By considering the invariance property of ML estimation, one can obtain the ML estimates of the reliability characteristics $R_s(t)$, $h_s(t)$ and $h_{2s}(t)$. The asymptotic sampling distribution of $(\hat{\alpha}-\alpha,\hat{\Lambda}-\Lambda)'$ is $N_{k+1}(0,\xi^{-1})$ with ξ representing the Fisher's information matrix whose elements are given as

$$\xi_{11} = -\left.\frac{\partial^2 logL}{\partial\alpha^2}\right|_{\alpha=\hat{\alpha}}$$

$$\xi_{ij} = 0; i \neq j$$

$$\xi_{1j} = -\left.\frac{\partial^2 logL}{\partial\alpha\partial\beta_j}\right|_{\alpha=\hat{\alpha},\beta_j=\hat{\beta}_j}$$

$$\xi_{jj} = -\left.\frac{\partial^2 logL}{\partial\beta_j^2}\right|_{\beta_j=\hat{\beta}_j}$$

Using equation (6), one can obtain the second-order derivatives of the log-likelihood function as

$$\frac{\partial^2 logL}{\partial\alpha^2} = \sum_{i=1}^{n}\sum_{j=1}^{k}\left[\frac{\frac{\partial^2}{\partial\alpha^2}\gamma(\alpha,\frac{1}{\beta_j t_{ij}}) - \frac{\partial^2}{\partial\alpha^2}\gamma(\alpha,\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})} - \left\{\frac{\frac{\partial}{\partial\alpha}\gamma(\alpha,\frac{1}{\beta_j t_{ij}}) - \frac{\partial}{\partial\alpha}\gamma(\alpha,\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}\right\}^2\right]$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{k}(k-j)\left[\frac{\frac{\partial^2}{\partial\alpha^2}\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})} - \left\{\frac{\frac{\partial}{\partial\alpha}\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}\right\}^2\right] - \frac{nk(k+1)}{2}\psi'(\alpha)$$

$$\frac{\partial^2 logL}{\partial\alpha\partial\beta_j} = \sum_{i=1}^{n}\frac{1}{\beta_j}\left[\frac{\frac{\partial}{\partial\alpha}\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})} - \frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})\frac{\partial}{\partial\alpha}\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})^2}\right]$$

$$+\sum_{i=1}^{n}\frac{(k-j)}{\beta_j}\left[\frac{\frac{\partial}{\partial\alpha}\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})} - \frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}})\frac{\partial}{\partial\alpha}\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})^2}\right] - \frac{n(k-j+1)}{\beta_j}$$

$$\frac{\partial^2 logL}{\partial\beta_j^2} = \frac{n(k-j+1)\alpha}{\beta_j^2} - \sum_{i=1}^{n}\frac{1}{{\beta_j}^2}\frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})} - \sum_{i=1}^{n}\frac{(k-j)}{\beta_j^2}\frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}$$

$$+\sum_{i=1}^{n}\frac{1}{\beta_j^2}\left[\frac{\gamma(\alpha+2,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)}) - \gamma(\alpha+1,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})} - \left\{\frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}},\frac{1}{\beta_j(t_{ij}+1)})}\right\}^2\right]$$

$$+\sum_{i=1}^{n}\frac{(k-j)}{\beta_j^2}\left[\frac{\gamma(\alpha+2,\frac{1}{\beta_j t_{ij}}) - \gamma(\alpha+1,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})} - \left\{\frac{\gamma(\alpha+1,\frac{1}{\beta_j t_{ij}})}{\gamma(\alpha,\frac{1}{\beta_j t_{ij}})}\right\}^2\right]$$

where, $\frac{\partial^2}{\partial\alpha^2}\gamma\left(\alpha,\frac{1}{\beta_j t_{ij}}\right) = \int_0^{\frac{1}{\beta_j t_{ij}}} v^{(\alpha-1)}e^{-v}(log(v))^2 dv.$

The asymptotic $100\times(1-\gamma')\%$ joint confidence ellipsoid for (α,Λ) is $(\hat{\alpha}-\alpha,\hat{\Lambda}-\Lambda)\xi^{-1}(\hat{\alpha}-\alpha,\hat{\Lambda}-\Lambda)' \leq \chi^2_{(k+1),\gamma'}$, where $\chi^2_{(k+1),\gamma'}$ is the $100\times\gamma'^{th}$ percentile of χ^2-distribution with (k+1) degrees of freedom. Moreover, the asymptotic distributions of the reliability characteristics $R_s(t)$, $h_s(t)$ and $h_{2s}(t)$ are $N(0,R'\xi^{-1}R)$, $N(0,h'\xi^{-1}h)$ and $N(0,h_2'\xi^{-1}h_2)$, respectively, where,

$$R' = \left(\frac{\partial R_s(t)}{\partial\alpha},\frac{\partial R_s(t)}{\partial\Lambda}\right),\ h' = \left(\frac{\partial h_s(t)}{\partial\alpha};\frac{\partial h_s(t)}{\partial\Lambda}\right)\ and\ h_2' = \left(\frac{\partial h_{2_s}(t)}{\partial\alpha};\frac{\partial h_{2_s}(t)}{\partial\Lambda}\right).$$

4.2. Bootstrap Method

The bootstrap method is a general resampling procedure for obtaining bootstrap CIs which are an alternative to the asymptotic CIs. Two types of CIs are being used here, i.e., percentile bootstrap (boot-p) and bootstrap-t (boot-t). The procedure given in the two algorithms for boot-p and boot-t methods will be employed for obtaining the bootstrap estimates and confidence intervals for the parameters, reliability, hazard rate and RSRF functions. The algorithms for both the methods are provided in appendix A.

4.3. Bayesian Estimation Using MCMC Approach

Bayesian estimation setup usually involves generating samples from the posterior distribution and using them to summarize the knowledge about the parameters. This makes the use of prior knowledge and the available data. When the prior knowledge is not available, one can make use of the non-informative prior. Here, both the informative and non-informative priors are considered. The prior distributions regarding the load-share parameters α and β_j are taken as gamma priors as

$$g(\alpha) \propto \alpha^{a-1} e^{-\alpha/b} \;\; ; \alpha, a, b > 0 \tag{9}$$

and

$$h(\beta_j) \propto \beta_j^{c_j-1} e^{-\beta_j/d_j} \;\; ; \beta_j, c_j, d_j > 0; j = 1, 2, \ldots, k \tag{10}$$

Using the priors given in equations (9) and (10) and the likelihood function given in equation (5), the joint distribution of the parameters and the dataset is

$$K(T, \alpha, \Lambda) = L(T|\alpha, \Lambda).g(\alpha).h(\Lambda)$$

$$= (k!)^n \prod_{i=1}^{n} \prod_{j=1}^{k} \left[\frac{1}{\Gamma(\alpha)^{(k-j+1)}} \gamma\Big(\alpha, \frac{1}{\beta_j t_{ij}}, \frac{1}{\beta_j (t_{ij}+1)}\Big) \Big(\gamma\Big(\alpha, \frac{1}{\beta_j t_{ij}}\Big)\Big)^{(k-j)} \right]$$

$$\times \alpha^{a-1} e^{-\alpha/b} \times \prod_{j=1}^{k} \beta_j^{c_j-1} e^{-\beta_j/d_j}$$

Here, to obtain Bayes estimates of the parameters, the marginal posterior densities of the parameters are required which are difficult to obtain. Therefore, the use of MCMC techniques like MH algorithm and Gibbs sampler (given in appendix B) will be followed. For that, the full conditional densities of the parameters are obtained as:

$$\pi_1(\alpha|T, \Lambda) \propto \alpha^{a-1} e^{-\alpha/b} \prod_{i=1}^{n} \prod_{j=1}^{k} \frac{1}{\Gamma(\alpha)^{(k-j+1)}} \gamma\Big(\alpha, \frac{1}{\beta_j t_{ij}}, \frac{1}{\beta_j (t_{ij}+1)}\Big) \left[\gamma\left(\alpha, \frac{1}{\beta_j t_{ij}}\right)\right]^{(k-j)} \tag{11}$$

$$\pi_2(\beta_j|T, \alpha) \propto \beta_j^{c_j-1} e^{-\beta_j/d_j} \prod_{i=1}^{n} \gamma\Big(\alpha, \frac{1}{\beta_j t_{ij}}, \frac{1}{\beta_j (t_{ij}+1)}\Big) \left[\gamma\left(\alpha, \frac{1}{\beta_j t_{ij}}\right)\right]^{(k-j)} ; j = 1, 2, \ldots, k \tag{12}$$

Note that, sampling from equations (11) and (12) is not easy to be done directly because of its complexity. Therefore, samples are generated by using the MH algorithm.

5. Simulation Study

In this section, a simulation study is conducted for analyzing the estimates of the parameters of the proposed model by using classical as well as Bayesian approach. Sample data of size n = 30, 50, 100 and 200 are generated for $k = 3, \alpha = 10$, $\beta_1 = 0.001$, $\beta_2 = 0.005$ and $\beta_3 = 0.01$. In classical approach, the ML estimates of the parameters along with their SEs and asymptotic CIs are obtained. Two bootstrapping techniques are applied and bootstrap confidence intervals based on B = 2000 bootstrap replications are also computed. Using Bayesian approach, Bayes estimates of the parameters and reliability characteristics along with their PSEs and HPD intervals are obtained with informative as

well as non-informative priors. For this, MCMC technique is employed using MH and Gibbs algorithms and $10,000$ realizations are generated from each posterior density given in (11) and (12). The burn-in period of 500 realizations are removed. From the generated chains, every 5^{th} value is taken for removing the autocorrelation among the values. The trace plots for the parameters are plotted in Figs. (1) - (4) to ensure the fine mixing of the chains. For informative prior, gamma prior is considered for all the parameters α, β_1, β_2 and β_3 setting the hyperparameters as $\alpha = a \times b, \beta_1 = c_1 \times d_1, \beta_2 = c_2 \times d_2$ and $\beta_3 = c_3 \times d_3$. The values of all the hyperparameters are taken as approximately 0 under non-informative or Jeffrey's prior. All the discussed estimates along with their SEs/PSEs and CIs/HPD intervals are summarized in Table (1).

After performing the simulation study, the following results are observed:

- All the obtained estimates of the parameters and reliability characteristics become more precise (closer to true values) with an increase in the sample size.
- The SEs/PSEs magnitude of the estimates and the widths of all the intervals decreases on increasing the sample size.
- Bayes estimation with gamma priors is more precise in terms of true values and SEs than the Bayes estimation with Jeffrey's prior as well as ML estimation and bootstrapping methods for different sample sizes.
- Boot-p and boot-t confidence intervals are more precise than the asymptotic CIs as they contain the parameters in smaller width for all the sample sizes.
- Boot-p CIs are providing slightly shorter widths than boot-t CIs for all the samples sizes.

6. Real Data Application

In this section, the applicability of the model is discussed through a real dataset of plasma display devices (PDPs) which is supposed to be a load-sharing model by Kvam and Pena [4]. For PDPs, the degradation is measured in luminosity and a PDP is considered as failed when the luminosity goes below a threshold. A test is conducted with 20 items and 3 sensors spaced evenly across the test device. The dataset contains the failure times for 3 sensors on each of 20 test items.

For fitting a discrete distribution , the integer parts of the data values are taken into consideration. Now, to check whether these failure times can be modeled using load-sharing models, we setup the following hypothesis:
H_0: Load-sharing behavior exists in the dataset i.e., $\beta_1 = \beta_2 = \beta_3$
H_1: Load sharing behavior does not exist in the dataset i.e., $\beta_1 \neq \beta_2 \neq \beta_3$.
The hypothesis can be tested using the following criteria:

- Akaike information criterion (AIC): $-2logL + 2p$
- Bayesian information criterion (BIC): $-2logL + plogn$
- Deviance test statistic: $d_n = -2[logL_{H_0}(T|\hat{\alpha}, \hat{\beta}) - logL_{H_1}(T|\hat{\alpha}, \hat{\beta_1}, \hat{\beta_2}, \hat{\beta_3})]$

Table 1: *Estimates of parameters with their SEs/PSEs and CI/HPD width for values of* $\alpha = 10$, $\beta_1 = 0.001$, $\beta_2 = 0.005$, $\beta_3 = 0.01$, $R(50) = 0.9851$ *and* $h(50) = 0.0038$ *and* $RSRF(50) = 0.2381$ *for varying sample size.*

sample size	para-meter	ML	SE	CI width	boot-p	SE	width-p	width-t	Bayes (J)	PSE	HPD width	Bayes (I)	PSE	HPD width
30	α	8.0143	0.9961	3.9047	8.0238	0.9923	3.5534	3.4182	9.9002	0.0354	0.1291	9.9031	0.0353	0.1121
	β_1	0.0012	1.5×10^{-4}	5.8×10^{-4}	0.0012	1.3×10^{-4}	5.1×10^{-4}	5.7×10^{-4}	0.0009	6.2×10^{-5}	1.7×10^{-4}	0.0009	5.8×10^{-5}	1.6×10^{-4}
	β_2	0.0059	9.8×10^{-4}	0.0038	0.0057	8.6×10^{-4}	0.0031	0.0034	0.0049	3.7×10^{-4}	0.0012	0.0049	4.9×10^{-4}	0.0016
	β_3	0.0158	0.0021	0.0082	0.0154	0.0018	0.0066	0.0069	0.0091	0.0003	0.0009	0.0091	0.0003	0.0009
	$R(50)$	0.9992	0.0293	0.1148	0.9991	0.0111	0.0531	0.0542	0.9876	0.0107	0.0273	0.9875	0.0104	0.0305
	$h(50)$	0.0027	0.0005	0.0019	0.0030	0.0006	0.0022	0.0021	0.0032	0.0003	0.0009	0.0032	0.0003	0.0008
	$RSRF(50)$	0.2950	0.0911	0.3571	0.3093	0.0902	0.3260	0.3293	0.2437	0.0254	0.0961	0.2434	0.0242	0.0734
50	α	8.1604	0.7694	3.0161	8.1501	0.8469	2.8107	2.9038	9.9003	0.0261	0.0441	9.9128	0.0288	0.0586
	β_1	0.0012	1.5×10^{-4}	5.8×10^{-4}	0.0012	1.4×10^{-4}	5.6×10^{-4}	5.8×10^{-4}	0.0009	4.4×10^{-5}	1.3×10^{-4}	0.0009	4.7×10^{-5}	1.2×10^{-4}
	β_2	0.0063	8.0×10^{-4}	0.0032	0.0062	8.1×10^{-4}	0.0032	0.0031	0.0046	3.1×10^{-4}	9.1×10^{-4}	0.0049	2.5×10^{-4}	7.4×10^{-4}
	β_3	0.0145	0.0022	0.0086	0.0143	0.0022	0.0084	0.0082	0.0091	0.0004	0.0013	0.0093	0.0004	0.0014
	$R(50)$	0.9956	0.0179	0.0702	0.9951	0.0229	0.0941	0.0952	0.9871	0.0085	0.0237	0.9867	0.0075	0.0234
	$h(50)$	0.0029	0.0004	0.0015	0.0031	0.0005	0.0018	0.0016	0.0032	0.0002	0.0005	0.0033	0.0002	0.0005
	$RSRF(50)$	0.2784	0.0467	0.1831	0.2777	0.0458	0.1577	0.1558	0.2433	0.0182	0.0526	0.2429	0.0195	0.0563
100	α	8.2093	0.6708	2.6295	8.2876	0.6128	2.5204	2.6778	9.9011	0.0246	0.0527	9.9207	0.0212	0.0357
	β_1	0.0011	9.1×10^{-5}	3.5×10^{-4}	0.0011	8.1×10^{-5}	3.1×10^{-4}	3.4×10^{-4}	0.0009	1.5×10^{-5}	5.7×10^{-5}	0.0009	2.2×10^{-5}	6.4×10^{-5}
	β_2	0.0057	5.2×10^{-4}	0.0021	0.0057	5.0×10^{-4}	0.0019	0.0021	0.0048	1.6×10^{-4}	4.9×10^{-4}	0.0049	2.0×10^{-4}	5.9×10^{-4}
	β_3	0.0132	0.0013	0.0051	0.0132	0.0012	0.0043	0.0049	0.0092	0.0003	0.0012	0.0096	0.0003	0.0011
	$R(50)$	0.9985	0.0162	0.0635	0.9983	0.0081	0.0341	0.0374	0.9875	0.0035	0.0105	0.9873	0.0023	0.0086
	$h(50)$	0.0031	0.0003	0.0011	0.0032	0.0003	0.0012	0.0010	0.0034	5.3×10^{-5}	0.0002	0.0035	7.1×10^{-5}	0.0002
	$RSRF(50)$	0.2446	0.0391	0.1532	0.2481	0.0353	0.1261	0.1284	0.2422	0.0091	0.0268	0.2414	0.0064	0.0233
200	α	8.3819	0.3894	1.5264	8.6394	0.3531	1.3197	1.1781	9.9014	0.0137	0.0327	9.9374	0.0141	0.0131
	β_1	0.0011	7.2×10^{-5}	2.8×10^{-4}	0.0011	7.3×10^{-5}	2.6×10^{-4}	2.5×10^{-4}	0.0009	9.3×10^{-6}	2.5×10^{-5}	0.0009	5.6×10^{-6}	2.1×10^{-5}
	β_2	0.0055	3.9×10^{-4}	0.0015	0.0054	4.3×10^{-4}	0.0018	0.0016	0.0049	1.1×10^{-4}	2.6×10^{-4}	0.0049	1.0×10^{-4}	2.5×10^{-4}
	β_3	0.0118	0.0010	0.0039	0.0114	0.0010	0.0037	0.0037	0.0102	0.0002	0.0007	0.0101	0.0001	0.0004
	$R(50)$	0.9969	0.0092	0.0361	0.9969	0.0091	0.0347	0.0362	0.9868	0.0008	0.0029	0.9865	0.0015	0.0039
	$h(50)$	0.0035	0.0002	0.0008	0.0035	0.0002	0.0008	0.0007	0.0035	1.8×10^{-5}	0.0001	0.0036	2.3×10^{-5}	0.0001
	$RSRF(50)$	0.2426	0.0126	0.0494	0.2452	0.0135	0.0498	0.0468	0.2415	0.0039	0.0106	0.2411	0.0023	0.0082

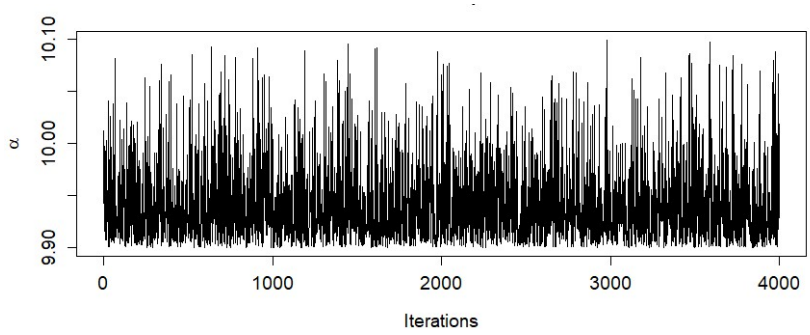


Figure 1: *Trace plot of* α.

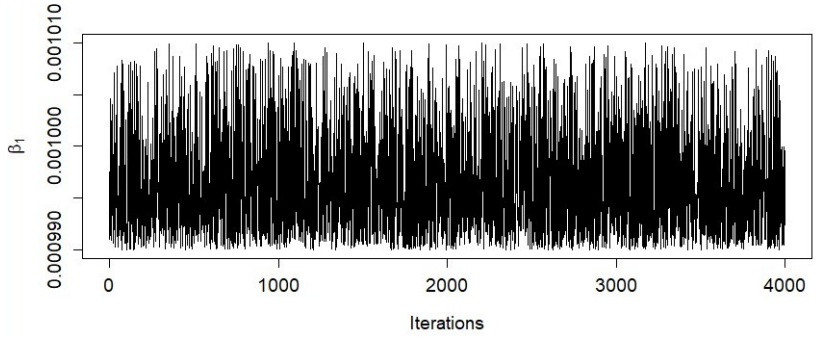


Figure 2: *Trace plot of* β_1.

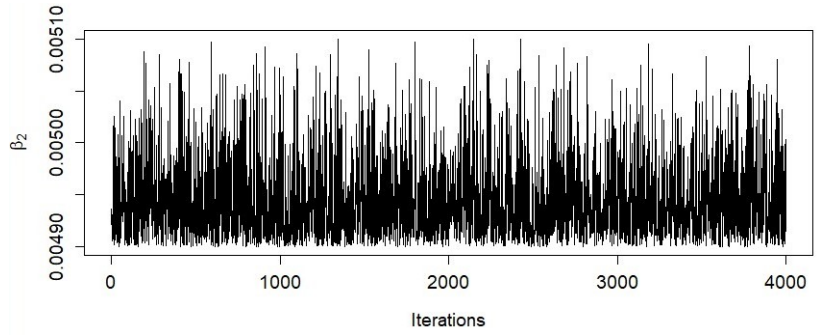


Figure 3: *Trace plot of* β_2.

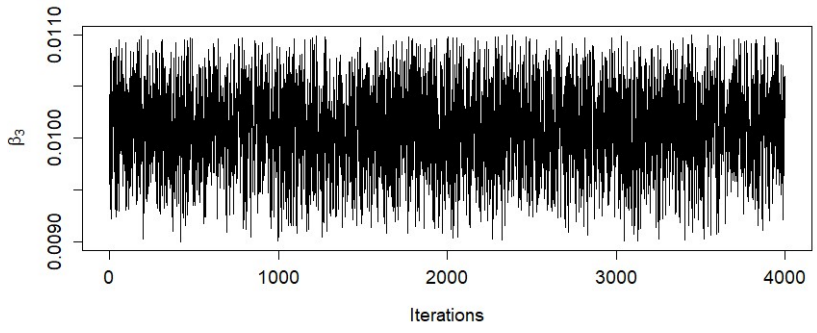


Figure 4: *Trace plot of* β_3.

The fitting summary of the dataset under both the models is provided in Table 2. The observed value of deviance test statistic d_n is 9.0385 with corresponding p-value 0.0108 (< 0.05) which suggests that H_0 cannot be accepted at 5% level of significance. The same is suggested by comparing AIC and BIC under H_0 and H_1. Hence, it is concluded that load-sharing behavior exists in the considered dataset.

Table 2: *Fitting summary of PDP dataset*

Model	-Log L	AIC	BIC	d_n
H_0	472.8236	949.6473	951.6388	9.0385
H_1	468.3044	944.6087	948.5917	

7. Concluding Remarks

In this research, a multi-component load-share parallel system is analyzed by assuming that the underlying failure time distribution of each component is DIGD. The classical as well as Bayesian estimation techniques are applied for estimating the parameters of the system. It is assumed that on a component's failure, the total workload imposed on the system will be redistributed to the other working components and this will affect their performance. Such systems exist in many engineering applications like fiber composites, power plants, manufacturing and many more. However, the study can be extended by considering non-identical components where each component is having different loads. Also, generalized IGD can be adopted instead of DIGD in future researches.

Conflict of interest: The authors declare that there is no conflict of interest.

References

[1] Daniels, H. E. (1945). The statistical theory of the strength of bundles of threads. I. *Proceedings Of The Royal Society Of London. Series A. Mathematical And Physical Sciences*, 183(995):405-435.

[2] Hussain, T. and Ahmad, M. (2012). Discrete inverse Gamma distribution. *Official Statistics And Its Impact On Good Governance And Economy Of Pakistan', 9th International Conference On Statistical Sciences, Islamic Countries Society Of Statistical Sciences, Lahore, Pakistan*, 381-394.

[3] Kim, H. and Kvam, P. H. (2004). Reliability estimation based on system data with an unknown load share rule. *Lifetime Data Analysis*, 10(1):83-94.

[4] Kvam, P. H. and Pena, E. A. (2005). Estimating load-sharing properties in a dynamic reliability system. *Journal Of The American Statistical Association*, 100(469):262-272.

[5] Park, C. (2010). Parameter estimation for the reliability of load-sharing systems. *IIE Transactions*, 42(10):753-765.

[6] Park, C. (2013). Parameter estimation from load-sharing system data using the expectation–maximization algorithm. *IIE Transactions*, 45(2):147-163.

[7] Pundir, P. S. and Gupta, P. K. (2018). Reliability estimation in load-sharing system model with application to real data. *Annals Of Data Science*, 5(1):69-91.

[8] Pundir, P. S., Srivastava, R., Agarwal, S. and Shrivastava, V. (2022). On the estimation problems for discrete inverse gamma lifetime model with real data applications. *Life Cycle Reliability And Safety Engineering*, 11(4):313-321.

[9] Ramírez-Espinosa, P. and Lopez-Martinez, F. J. (2021). Composite fading models based on inverse gamma shadowing: Theory and validation. *IEEE Transactions on Wireless Communications*, 20(8):5034-5045.

[10] Shang, X. and Song, H. (2011). Radar detection based on compound-Gaussian model with inverse gamma texture. *IET Radar, Sonar & Navigation*, 5(3):315-321.

[11] Singh, B. and Goel, R. (2019). MCMC estimation of multi-component load-sharing system model and its application. *Iranian Journal Of Science And Technology, Transactions A: Science*, 43(2):567-577.

[12] Singh, B., Sharma, K. and Kumar, A. (2008). A classical and Bayesian estimation of a k-components load-sharing parallel system. *Computational Statistics & Data Analysis*, 52(12):5175-5185.

[13] Singh, B., Sharma, K. and Kumar, A. (2009). Analyzing the dynamic system model with discrete failure time distribution. *Statistical Methods And Applications*, 18(4):521-542.

[14] Stinco, P., Greco, M. and Gini, F. (2011). Adaptive detection in compound-Gaussian clutter with inverse-gamma texture. *Proceedings Of 2011 IEEE CIE International Conference On Radar*, 1:434-437.

[15] Yoo, S. K., Bhargav, N., Cotton, S. L., Sofotasios, P. C., Matthaiou, M., Valkama, M. and Karagiannidis, G. K. (2017). The $\kappa - \mu$/inverse gamma and $\eta - \mu$/inverse gamma composite fading models: Fundamental statistics and empirical validation. *IEEE Transactions On Communications*, 69(8):5514-5530.

[16] Zhang, Z., Yang, Y. and Li, D. (2022). Estimation of parameters for load-sharing parallel systems under exponential Pareto distribution. *Proceedings Of The Institution Of Mechanical Engineers, Part O: Journal Of Risk And Reliability*, 236(2):248-255.

Appendix A

Boot-p Method

1. Generate a sample $T_{ij}(i = 1, 2, \ldots, n; j = 1, 2, \ldots, k)$ of size n by using equation (3).
2. Now, regenerate B bootstrap samples $T^*_{ij}(i = 1, 2, \ldots, n; j = 1, 2, \ldots, k)$ of size n from the original sample T_{ij} to compute B bootstrap estimates $(\hat{\alpha}^*, \hat{\Lambda}^*) \equiv \hat{\Theta}^*$ of $(\alpha, \Lambda) \equiv \Theta$.
3. Let $\hat{\Theta}^*_{(1)}, \ldots, \hat{\Theta}^*_{(B)}$ be the ordered statistics of the estimates $\hat{\Theta}^*_1, \ldots, \hat{\Theta}^*_B$. Then, $100 \times (1 - \gamma)\%$ boot-p CI is: $(\hat{\Theta}^*[\gamma B/2], \hat{\Theta}^*[1 - \gamma B/2])$
4. Finally, the bootstrap estimates and their corresponding variances are obtained.

Boot-t Method

1. On the basis of the genereted sample T_{ij}^*, compute the following pivots as:

$$\kappa_1^* = \frac{\hat{\alpha}^* - \hat{\alpha}}{\sqrt{\hat{V}(\hat{\alpha}^*)}} \qquad and \qquad \kappa_2^* = \frac{\hat{\Lambda}^* - \hat{\Lambda}}{\sqrt{\hat{V}(\hat{\Lambda}^*)}}$$

2. Now, repeat the step 1, B-times.

3. Consider $S_1(x) = P(\kappa_1^* \leq x)$ and $S_2(x) = P(\kappa_2^* \leq x)$ as the CDFs of κ_1^* and κ_2^*, respectively. Let for a given value x,
$\hat{\alpha}_{boot-t}(x) = \hat{\alpha}^* + \sqrt{\hat{V}(\hat{\alpha}^*)} S_1^{-1}(x)$ and $\hat{\Lambda}_{boot-t}(x) = \hat{\Lambda} + \sqrt{\hat{V}(\hat{\Lambda})} S_2^{-1}(x)$.

4. The $100 \times (1-\gamma)\%$ boot-t CIs for α and Λ are
$\hat{\alpha}_{boot-t}(\gamma/2), \hat{\alpha}_{boot-t}(1-\gamma/2)$ and $\hat{\Lambda}_{boot-t}(\gamma/2), \hat{\Lambda}_{boot-t}(1-\gamma/2)$.

Appendix B

Gibbs Algorithm

1. Generate α and β_j $(j = 1, 2, \ldots, k)$ from $\pi_1(\alpha|T, \Lambda)$ and $\pi_2(\beta_j|T, \alpha)$ as given in equations (11) and (12), respectively.

2. Repeat the above step, M times. To remove the effect of starting values, record the generated sequence of the parameters after some $N(< M)$ burn-in draws.

3. Bayes estimates and their corresponding posterior variances of the parameters as well as system reliability function, hazard rate function and RSRF function are computed by considering the means and variances of the generated values of the parameters and these three reliability characteristics.

4. Now, considering the ordered sequence of the parameters and reliability characteristics, the $100 \times (1-\gamma)\%$ HPD intervals are constructed.

Metropolis-Hastings Algorithm

1. Start with an initial value x_0 from the support of the prior distribution and consider $i = 1$.

2. Now, generate a proposal x_{prop} by using the proposal density q$(x_i|x_{i-1})$.

3. Calculate the acceptance probability as

$$P_\alpha(x_{prop}|x_{i-1}) = min\left[1, \frac{q(x_{i-1}|x_{prop}) f(x_{prop})}{q(x_{prop}|x_{i-1}) f(x_{i-1})}\right]$$

4. Generate a random variable U from uniform distribution on $(0, 1)$.

5. The proposal point will be accepted if $u < P_\alpha$ by considering $x_i = x_{prop}$, otherwise, reject it and set $x_i = x_{i-1}$.

6. Set $i = i + 1$ and return to step 2.

RT&A, No 4 (76)
Volume 18, December 2023

SURVIVAL ANALYSIS OF A MULTI-STATE SEMI-MARKOV MODEL ON INFECTIOUS DISEASE CONSIDERING VARIOUS LEVELS OF SEVERITY

Sujata Sukhija*

•

Research Scholar, Department of Mathematics, Maharshi Dayanand University, Rohtak
sukhijasujata860@gmail.com

Rajeev Kumar

•

Professor, Department of Mathematics, Maharshi Dayanand University, Rohtak
profrajeevmdu@gmail.com

*Corresponding Author

Abstract

The aim of the paper is to carry out survival analysis of a novel multi-state model on infectious disease considering various levels of severity using semi-Markov processes. Various levels of severity of the disease over time and transitions between these severity levels have been considered. Transition probabilities and expected waiting times are derived. Expressions for mean survival time, expected total time in home isolation, and expected total time in hospital are obtained. The analysis of the proposed model is carried out through numerical computation and plotting several graphs. Important conclusions are drawn. The modelling framework proposed here can be used to model any infectious disease irrespective of disease states. The study will be helpful in designing effective measures to control the infectious disease and selecting the appropriate intervention policies.

Keywords: Multi-state model, Markov property, Semi-Markov process, Transition probabilities

1. Introduction

Modelling infectious diseases has always been an area of interest for researchers in various fields for the sake of prevention and control of these diseases. According to World Health Organization 2019 report [17], infectious diseases are still in the top ten leading causes of death. In low income countries, six of the top ten causes of death are infectious diseases including neonatal conditions, lower respiratory infections, diarrhoeal diseases, malaria, tuberculosis, HIV/AIDS [17]. As per the above report, there was 50% drop in Disability Adjusted Life Years (DALYs) since 2000 due to infectious, maternal, perinatal and nutritional conditions.

Multi-state Markov models have been frequently used to study the progression of diseases such as cancer [7, 11, 15], HIV infection [4, 20], renal disease [9] and many more. For a Markov model, it is assumed that the holding time in a state is exponentially distributed. For many real-world situations, such as time to failure and time to discover a fault, the exponential distribution may be acceptable. In fact, the exponential distributions have memoryless property. In some cases,

the memoryless property could be seen as a problematic assumption. For example, patients who respond well to a treatment are likely to respond well to the treatment in the future, violating the Markov property [5]. To overcome this limitation, semi-Markov processes came into existence.

Semi-Markov processes are very important generalizations of Markov processes. While Markov processes assume that holding time in a state is exponentially distributed, semi-Markov processes relax the assumption allowing any arbitrary distributions for holding time in a state. Semi-Markov processes were defined by Levy [12] and Smith [14]. Since then, semi-Markov process concepts have been applied to solve various problems like electronics and missile related problems, to improve reliability of various systems, for cost-benefit analysis of a system, for economic decision making problem and so on. The field of biomedical science is not an exception to this, for example, see [1, 2, 6, 8, 10, 13, 16]. Weiss and Zelen [16] applied the theory of semi-Markov processes to the construction of a stochastic model for interpreting data obtained from clinical trials on patients with acute leukemia. Kao [10] derived results for computing the mean and variance of times in transient states and times to absorption in a transient semi-Markov process. Davidov [6] developed expressions for the steady-state probabilities for regenerative semi-Markov processes. Castelli et al. [2] performed cost-effective analysis to compare the follow-up strategies in colorectal cancer study. Goshu and Dessie [8] analysed hospital data obtained from a cohort of AIDS patients who have been under antiretroviral therapy follow-up and estimated the conditional probability of transitions between two states for a finite time period. Cao et al. [1] developed a semi-Markov model to analyse the long-term cost effectiveness of heart failure management programmes. Ramezankhani et al. [13] applied a multi-state semi-Markov model to estimate the number of years of life lost due to diabetes with and without cardiovascular disease.

However, the majority of the literature studies were focused on the disease's progressive stages and omitted the transitions back to normal state. We tried to bridge this gap through our article. There were some studies which included the transitions back to normal state however their main focus was to understand the threshold dynamics of the disease, for example, see [18], [19]. Further, since infectious diseases typically necessitate isolation, such as measles, cholera, diphtheria, infectious tuberculosis, plague, smallpox, yellow fever, and viral hemorrhagic fevers [3], home isolation is one of the possible states of our model. Moreover, through our model, we estimated the expected length of stay in home isolated state which has not been reported in previous studies. Besides, it is a well known fact that elderly patients, pregnant women and patients with co-morbidities are at risk of developing severe and critical illness and transition rates would be different in each category of severity illness. Thus, it becomes necessary to consider separate states for each category of severity. Keeping this in mind, we have considered the four states as mild disease state, moderate disease state, severe disease state and critical disease state. This brings another novelty to this model.

Keeping these in mind, a novel multi-state model for infectious disease based on the theory of semi-Markov processes is proposed. Various levels of severity of the disease over time have been considered. Thus, our model included every transition that a patient who is infected might experience. As in the model, the general scenario for an infectious disease have been considered, the model can be used to study and gain insights about any infectious disease. The paper is organized as follows. The newly developed multi-state semi-Markov model is described in Section 2. Transition probabilities and expected waiting times are derived and theoretical expressions regarding mean survival time, expected total time in home isolation and expected total time in hospital are obtained in Section 3. Numerical computations are performed in Section 4. Finally, conclusions are presented in Section 5.

2. Model Formulation

A semi-Markov model is proposed considering a person having an infectious disease showing the transition between various states. There are nine states in the model in which a healthy individual has the possibility to transit (see Figure 1). Infected persons can experience a range of clinical

manifestations, from no symptoms to critical illness. Infected persons can generally be divided into categories based on the severity of their illnesses: mild illness, moderate illness, severe illness, and critical illness. In light of this, we have taken into account the corresponding four states in increasing order of illness severity. Transitions from mild illness to moderate illness, moderate illness to severe illness, and so forth are permitted since a patient is at risk of developing severe and critical illness. Keeping this in mind, we have taken into consideration the deteriorating rates from mild illness to moderate illness, moderate illness to severe illness and so on. More details are described in the notations below.

Other assumptions made in the model are as under:
(i) All normal persons are exposed to the disease.
(ii) Testing of all infected persons is done.
(iii) Clinical testing is not perfect, i.e. there may be an error in testing.
(iv) Patient is home isolated if test results are false negative while patient is hospitalised if the test results are positive.
(v) All random variables are independent of each other.

The following states are considered in the model:

S_0	Normal state
S_1	Asymptomatic state
S_2	Symptomatic state
S_3	Home isolated state
S_4	Mild disease state
S_5	Moderate disease state
S_6	Severe disease state
S_7	Critical disease state
S_8	Death state

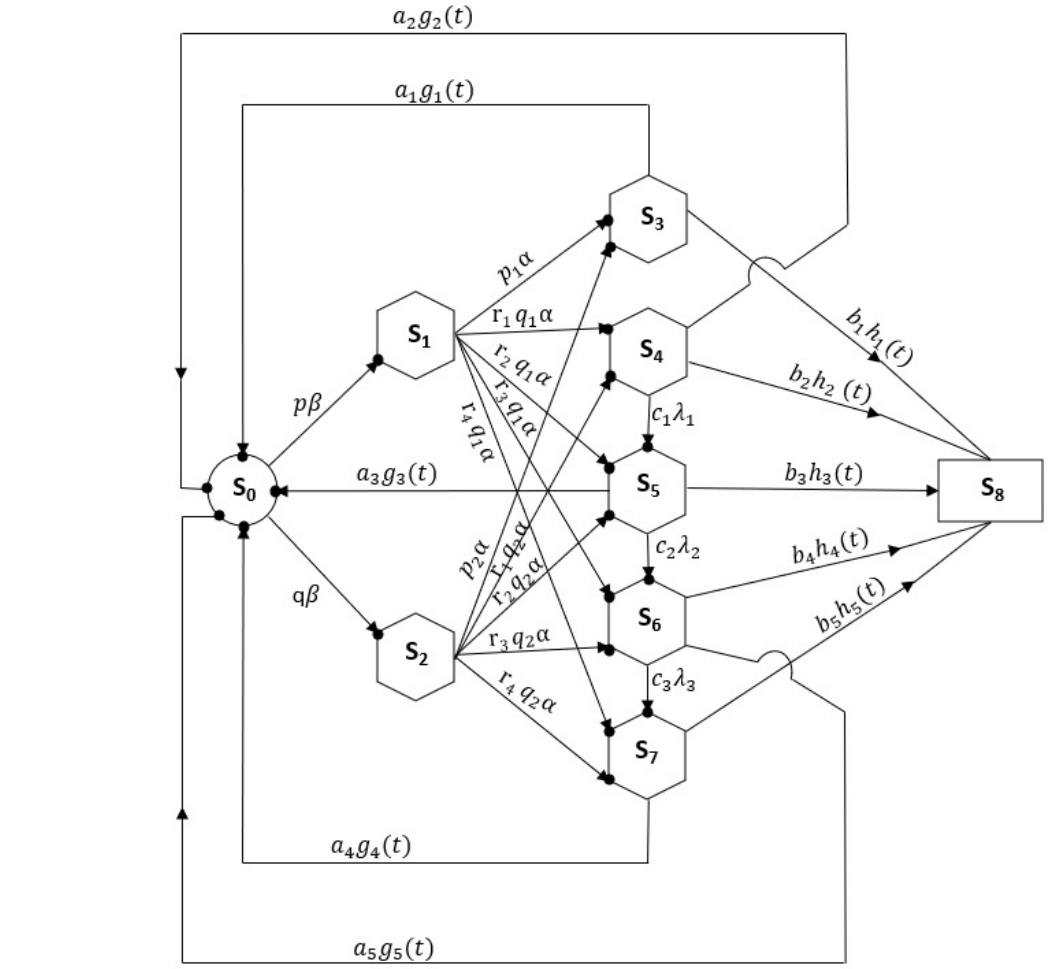
The following notation is used:

β	incidence rate
α	testing rate
$\lambda_1/\lambda_2/\lambda_3$	deteriorating rate from mild illness to moderate illness/moderate illness to severe illness/severe illness to critical illness
$g_1(t)/g_2(t)/g_3(t)/g_4(t)/g_5(t)$	probability density function of recovery time in home isolated state/mild disease state/moderate disease state/severe disease state/critical disease state
$G_1(t)/G_2(t)/G_3(t)/G_4(t)/G_5(t)$	cumulative distribution function of recovery time in home isolation/mild disease state/moderate disease state/severe disease state/critical disease state
$h_1(t)/h_2(t)/h_3(t)/h_4(t)/h_5(t)$	probability density function of time to death in home isolation/mild disease state/moderate disease state/severe disease state/critical disease state
$H_1(t)/H_2(t)/H_3(t)/H_4(t)/H_5(t)$	cumulative distribution function of time to death in home isolation/mild disease state/moderate disease state/severe disease state/critical disease state
$W_i(t)$	probability that the patient is in home isolation at instant t without passing through any other state

p/q	probability of an infected person to be asymptomatic/ symptomatic ($p+q=1$)
p_1/q_1	probability of an asymptomatically infected person to be tested false negative/positive ($p_1+q_1=1$)
p_2/q_2	probability of an symptomatically infected person to be tested false negative/positive ($p_2+q_2=1$)
$r_1/r_2/r_3/r_4$	probability that a patient is diagnosed with mild illness/ moderate illness/severe illness/critical illness ($r_1+r_2+r_3+r_4=1$)
a_2/b_2	probability that an home isolated person will recover/ move to death state ($a_1+b_1=1$)
$a_2/b_2/c_1$	probability that a person with mild illness will recover to normal state/move to death/deteriorate to moderate illness ($a_2+b_2+c_1=1$)
$a_3/b_3/c_2$	probability that a person with moderate illness will recover to normal state/move to death/deteriorate to severe illness($a_3+b_3+c_2=1$)
$a_4/b_4/c_3$	probability that a person with severe illness will recover to normal state/move to death/deteriorate to critical illness ($a_4+b_4+c_3=1$)
a_5/b_5	probability that a person with critical illness will recover to normal state/move to death state ($a_5+b_5=1$)

The following symbols/abbreviations are used:

SpO_2	Pulse oximetry
R.R.	Respiratory Rate
*	Laplace Transform symbol
**	Laplace Transform symbol
©	Laplace Convolution symbol
Ⓢ	Laplace-Stieltjes Convolution symbol



Normal State

Infected State

Death State

Figure 1: *State Transition Diagram. Possible states which an individual may occupy are depicted in the figure.*

3. Analysis

Let $q_{ij}(t)/Q_{ij}(t)$ represents probability density function/cumulative distribution function of first passage time from state S_i to state S_j without visiting any other state in $(0,t]$. Thus, the time dependent transition probabilities are given by

$$
\begin{array}{lll}
q_{01}(t) = p\beta e^{-\beta t} & q_{02}(t) = q\beta e^{-\beta t} & q_{13}(t) = p_1\alpha e^{-\alpha t} \\
q_{14}(t) = r_1q_1\alpha e^{-\alpha t} & q_{15}(t) = r_2q_1\alpha e^{-\alpha t} & q_{16}(t) = r_3q_1\alpha e^{-\alpha t} \\
q_{17}(t) = r_4q_1\alpha e^{-\alpha t} & q_{23}(t) = p_2\alpha e^{-\alpha t} & q_{24}(t) = r_1q_2\alpha e^{-\alpha t} \\
q_{25}(t) = r_2q_2\alpha e^{-\alpha t} & q_{26}(t) = r_3q_2\alpha e^{-\alpha t} & q_{27}(t) = r_4q_2\alpha e^{-\alpha t} \\
q_{30}(t) = a_1g_1(t) & q_{38}(t) = b_1h_1(t) & q_{40}(t) = a_2g_2(t) \\
q_{45}(t) = c_1\lambda_1 e^{-\lambda_1 t} & q_{48}(t) = b_2h_2(t) & q_{50}(t) = a_3g_3(t) \\
q_{56}(t) = c_2\lambda_2 e^{-\lambda_2 t} & q_{58}(t) = b_3h_3(t) & q_{60}(t) = a_4g_4(t) \\
q_{67}(t) = c_3\lambda_3 e^{-\lambda_3 t} & q_{68}(t) = b_4h_4(t) & q_{70}(t) = a_5g_5(t) \\
q_{78}(t) = b_5h_5(t) & &
\end{array}
$$

The steady state transition probabilities, $p_{ij} = \lim_{t\to\infty} \int_0^t q_{ij}(t)\,dt$ are obtained as

$$
\begin{array}{lllll}
p_{01} = p & p_{02} = q & p_{13} = p_1 & p_{14} = r_1q_1 & p_{15} = r_2q_1 \\
p_{16} = r_3q_1 & p_{17} = r_4q_1 & p_{23} = p_2 & p_{24} = r_1q_2 & p_{25} = r_2q_2 \\
p_{26} = r_3q_2 & p_{27} = r_4q_2 & p_{30} = a_1 & p_{38} = b_1 & p_{40} = a_2 \\
p_{45} = c_1 & p_{48} = b_2 & p_{50} = a_3 & p_{56} = c_2 & p_{58} = b_3 \\
p_{60} = a_4 & p_{67} = c_3 & p_{68} = b_4 & p_{70} = a_5 & p_{78} = b_5
\end{array}
$$

Let T_i denote the waiting time in state S_i then the expected waiting time in state S_i is given by $\mu_i = \int_0^\infty P(T_i > t)\,dt$. Thus, the expected waiting times are obtained as

$$
\begin{array}{ll}
\mu_0 = \dfrac{1}{\beta} & \mu_1 = \dfrac{1}{\alpha} \\
\mu_2 = \dfrac{1}{\alpha} & \mu_3 = -a_1g_1^{*'}(0) - b_1h_1^{*'}(0) \\
\mu_4 = \dfrac{c_1}{\lambda_1} - a_2g_2^{*'}(0) - b_2h_2^{*'}(0) & \mu_5 = \dfrac{c_2}{\lambda_2} - a_3g_3^{*'}(0) - b_3h_3^{*'}(0) \\
\mu_6 = \dfrac{c_3}{\lambda_3} - a_4g_4^{*'}(0) - b_4h_4^{*'}(0) & \mu_7 = -a_5g_5^{*'}(0) - b_5h_5^{*'}(0)
\end{array}
$$

The expected waiting time in state S_i given that the next state visited is S_j, is defined as $m_{ij} = \int_0^\infty tq_{ij}(t)\,dt = -q_{ij}^{*'}(0)$. Thus, the following relations are satisfied:

$$
\begin{aligned}
&m_{01} + m_{02} = \mu_0 \\
&m_{13} + m_{14} + m_{15} + m_{16} + m_{17} = \mu_1 \\
&m_{23} + m_{24} + m_{25} + m_{26} + m_{27} = \mu_2 \\
&m_{30} + m_{38} = \mu_3 \\
&m_{40} + m_{45} + m_{48} = \mu_4 \\
&m_{50} + m_{56} + m_{58} = \mu_5 \\
&m_{60} + m_{67} + m_{68} = \mu_6 \\
&m_{70} + m_{78} = \mu_7
\end{aligned}
$$

Theorem 1. If T_0 is the mean survival time for the patient starting in state S_0 then

$$T_0 = \frac{N}{D},$$

where

$$N = \mu_0 + p_{01}\mu_1 + p_{02}\mu_2 + p_{01}p_{13}\mu_3 + p_{01}p_{14}\mu_4 + p_{01}p_{15}\mu_5 + p_{01}p_{16}\mu_6 + p_{01}p_{17}\mu_7 + p_{02}p_{23}\mu_3 + p_{02}p_{24}\mu_4 + p_{02}p_{25}\mu_5 + p_{02}p_{26}\mu_6 + p_{02}p_{27}\mu_7 + p_{01}p_{14}p_{45}\mu_5 + p_{01}p_{15}p_{56}\mu_6 + p_{01}p_{16}p_{67}\mu_7 + p_{02}p_{24}p_{45}\mu_5 + p_{02}p_{25}p_{56}\mu_6 + p_{02}p_{26}p_{67}\mu_7 + p_{01}p_{14}p_{45}p_{56}\mu_6 + p_{02}p_{24}p_{45}p_{56}\mu_6 + p_{01}p_{15}p_{56}p_{67}\mu_7 + p_{02}p_{25}p_{56}p_{67}\mu_7 + p_{01}p_{14}p_{45}p_{56}p_{67}\mu_7 + p_{02}p_{24}p_{45}p_{56}p_{67}\mu_7$$

and

$$D = 1 - p_{01}p_{13}p_{30} - p_{01}p_{14}p_{40} - p_{01}p_{15}p_{50} - p_{01}p_{16}p_{60} - p_{01}p_{17}p_{70} - p_{02}p_{23}p_{30} - p_{02}p_{24}p_{40} - p_{02}p_{25}p_{50} - p_{02}p_{26}p_{60} - p_{02}p_{27}p_{70} - p_{01}p_{14}p_{45}p_{50} - p_{01}p_{15}p_{56}p_{60} - p_{01}p_{16}p_{67}p_{70} - p_{02}p_{24}p_{45}p_{50} - p_{02}p_{25}p_{56}p_{60} - p_{02}p_{26}p_{67}p_{70} - p_{01}p_{14}p_{45}p_{56}p_{60} - p_{01}p_{15}p_{56}p_{67}p_{70} - p_{02}p_{24}p_{45}p_{56}p_{60} - p_{02}p_{25}p_{56}p_{67}p_{70} - p_{01}p_{14}p_{45}p_{56}p_{67}p_{70} - p_{02}p_{24}p_{45}p_{56}p_{67}p_{70}.$$

Proof. Let $\phi_i(t)$ denote the cumulative distribution function of passage time from state S_i to the absorbing state.

The individual in state S_0 at $t = 0$ can reach the absorbing state at time t in two possible ways:
(i) The individual transited from state S_0 to state S_1 in time τ ($\tau < t$) and reached the absorbing state in $t - \tau$ time.
(ii) The individual transited from state S_0 to state S_2 in time τ ($\tau < t$) and reached the absorbing state in $t - \tau$ time.

Thus, we obtain
$\phi_0(t) = Q_{01}(t)\textcircled{S}\phi_1(t) + Q_{02}(t)\textcircled{S}\phi_2(t)$

Similarly, the following equations are obtained:
$\phi_1(t) = Q_{13}(t)\textcircled{S}\phi_3(t) + Q_{14}(t)\textcircled{S}\phi_4(t) + Q_{15}(t)\textcircled{S}\phi_5(t) + Q_{16}(t)\textcircled{S}\phi_6(t) + Q_{17}(t)\textcircled{S}\phi_7(t)$
$\phi_2(t) = Q_{23}(t)\textcircled{S}\phi_3(t) + Q_{24}(t)\textcircled{S}\phi_4(t) + Q_{25}(t)\textcircled{S}\phi_5(t) + Q_{26}(t)\textcircled{S}\phi_6(t) + Q_{27}(t)\textcircled{S}\phi_7(t)$
$\phi_3(t) = Q_{30}(t)\textcircled{S}\phi_0(t) + Q_{38}(t)$
$\phi_4(t) = Q_{40}(t)\textcircled{S}\phi_0(t) + Q_{45}(t)\textcircled{S}\phi_5(t) + Q_{48}(t)$
$\phi_5(t) = Q_{50}(t)\textcircled{S}\phi_0(t) + Q_{56}(t)\textcircled{S}\phi_6(t) + Q_{58}(t)$
$\phi_6(t) = Q_{60}(t)\textcircled{S}\phi_0(t) + Q_{67}(t)\textcircled{S}\phi_7(t) + Q_{68}(t)$
$\phi_7(t) = Q_{70}(t)\textcircled{S}\phi_0(t) + Q_{78}(t)$

Taking Laplace-Stieltjes transform of the above system of equations, rearranging the terms and solving the above system of equations for $\phi_0^{**}(s)$, we obtain

$$\phi_0^{**}(s) = \frac{N_1(s)}{D_1(s)}$$

where

$$N_1(s) = \begin{vmatrix} 0 & -Q_{01}^{**}(s) & -Q_{02}^{**}(s) & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -Q_{13}^{**}(s) & -Q_{14}^{**}(s) & -Q_{15}^{**}(s) & -Q_{16}^{**}(s) & -Q_{17}^{**}(s) \\ 0 & 0 & 1 & -Q_{23}^{**}(s) & -Q_{24}^{**}(s) & -Q_{25}^{**}(s) & -Q_{26}^{**}(s) & -Q_{27}^{**}(s) \\ Q_{38}^{**}(s) & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ Q_{48}^{**}(s) & 0 & 0 & 0 & 1 & -Q_{45}^{**}(s) & 0 & 0 \\ Q_{58}^{**}(s) & 0 & 0 & 0 & 0 & 1 & -Q_{56}^{**}(s) & 0 \\ Q_{68}^{**}(s) & 0 & 0 & 0 & 0 & 0 & 1 & -Q_{67}^{**}(s) \\ Q_{78}^{**}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

and

$$D_1(s) = \begin{vmatrix} 1 & -Q_{01}^{**}(s) & -Q_{02}^{**}(s) & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -Q_{13}^{**}(s) & -Q_{14}^{**}(s) & -Q_{15}^{**}(s) & -Q_{16}^{**}(s) & -Q_{17}^{**}(s) \\ 0 & 0 & 1 & -Q_{23}^{**}(s) & -Q_{24}^{**}(s) & -Q_{25}^{**}(s) & -Q_{26}^{**}(s) & -Q_{27}^{**}(s) \\ -Q_{30}^{**}(s) & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -Q_{40}^{**}(s) & 0 & 0 & 0 & 1 & -Q_{45}^{**}(s) & 0 & 0 \\ -Q_{50}^{**}(s) & 0 & 0 & 0 & 0 & 1 & -Q_{56}^{**}(s) & 0 \\ -Q_{60}^{**}(s) & 0 & 0 & 0 & 0 & 0 & 1 & -Q_{67}^{**}(s) \\ -Q_{70}^{**}(s) & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

Solving the above determinants, we get

$$\begin{aligned} N_1(s) = {} & Q_{01}^{**}(s)Q_{13}^{**}(s)Q_{38}^{**}(s) + Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{48}^{**}(s) + Q_{02}^{**}(s)Q_{23}^{**}(s)Q_{38}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{48}^{**}(s) + Q_{01}^{**}(s)Q_{15}^{**}(s)Q_{58}^{**}(s) + Q_{01}^{**}(s)Q_{16}^{**}(s)Q_{68}^{**}(s) \\ & + Q_{01}^{**}(s)Q_{17}^{**}(s)Q_{78}^{**}(s) + Q_{02}^{**}(s)Q_{25}^{**}(s)Q_{58}^{**}(s) + Q_{02}^{**}(s)Q_{26}^{**}(s)Q_{68}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{27}^{**}(s)Q_{78}^{**}(s) + Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{45}^{**}(s)Q_{58}^{**}(s) \\ & + Q_{01}^{**}(s)Q_{15}^{**}(s)Q_{56}^{**}(s)Q_{68}^{**}(s) + Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{45}^{**}(s)Q_{58}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{25}^{**}(s)Q_{56}^{**}(s)Q_{68}^{**}(s) + Q_{01}^{**}(s)Q_{16}^{**}(s)Q_{67}^{**}(s)Q_{78}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{26}^{**}(s)Q_{67}^{**}(s)Q_{78}^{**}(s) + Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{68}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{68}^{**}(s) + Q_{01}^{**}(s)Q_{15}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{78}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{25}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{78}^{**}(s) \\ & + Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{78}^{**}(s) \\ & + Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{78}^{**}(s) \end{aligned}$$

$$\begin{aligned} D_1(s) = {} & 1 - Q_{01}^{**}(s)Q_{13}^{**}(s)Q_{30}^{**}(s) - Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{40}^{**}(s) - Q_{02}^{**}(s)Q_{23}^{**}(s)Q_{30}^{**}(s) \\ & - Q_{01}^{**}(s)Q_{15}^{**}(s)Q_{50}^{**}(s) - Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{40}^{**}(s) - Q_{01}^{**}(s)Q_{16}^{**}(s)Q_{60}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{25}^{**}(s)Q_{50}^{**}(s) - Q_{01}^{**}(s)Q_{17}^{**}(s)Q_{70}^{**}(s) - Q_{02}^{**}(s)Q_{26}^{**}(s)Q_{60}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{27}^{**}(s)Q_{70}^{**}(s) - Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{45}^{**}(s)Q_{50}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{45}^{**}(s)Q_{50}^{**}(s) - Q_{01}^{**}(s)Q_{15}^{**}(s)Q_{56}^{**}(s)Q_{60}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{25}^{**}(s)Q_{56}^{**}(s)Q_{60}^{**}(s) - Q_{01}^{**}(s)Q_{16}^{**}(s)Q_{67}^{**}(s)Q_{70}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{26}^{**}(s)Q_{67}^{**}(s)Q_{70}^{**}(s) - Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{60}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{60}^{**}(s) - Q_{01}^{**}(s)Q_{15}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{70}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{25}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{70}^{**}(s) \\ & - Q_{01}^{**}(s)Q_{14}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{70}^{**}(s) \\ & - Q_{02}^{**}(s)Q_{24}^{**}(s)Q_{45}^{**}(s)Q_{56}^{**}(s)Q_{67}^{**}(s)Q_{70}^{**}(s) \end{aligned}$$

Mean survival time for the patient starting in state S_0 is given by

$$T_0 = lim_{s\to 0}\frac{1-\phi_0^{**}(s)}{s}$$

Using the above value of $\phi_0^{**}(s)$, we obtain

$$T_0 = \frac{N}{D},$$

where

$$N = \mu_0 + p_{01}\mu_1 + p_{02}\mu_2 + p_{01}p_{13}\mu_3 + p_{01}p_{14}\mu_4 + p_{01}p_{15}\mu_5 + p_{01}p_{16}\mu_6 + p_{01}p_{17}\mu_7 + p_{02}p_{23}\mu_3 + p_{02}p_{24}\mu_4 + p_{02}p_{25}\mu_5 + p_{02}p_{26}\mu_6 + p_{02}p_{27}\mu_7 + p_{01}p_{14}p_{45}\mu_5 + p_{01}p_{15}p_{56}\mu_6 + p_{01}p_{16}p_{67}\mu_7 + p_{02}p_{24}p_{45}\mu_5 + p_{02}p_{25}p_{56}\mu_6 + p_{02}p_{26}p_{67}\mu_7 + p_{01}p_{14}p_{45}p_{56}\mu_6 + p_{02}p_{24}p_{45}p_{56}\mu_6 + p_{01}p_{15}p_{56}p_{67}\mu_7 + p_{02}p_{25}p_{56}p_{67}\mu_7 + p_{01}p_{14}p_{45}p_{56}p_{67}\mu_7 + p_{02}p_{24}p_{45}p_{56}p_{67}\mu_7$$

and

$$D = 1 - p_{01}p_{13}p_{30} - p_{01}p_{14}p_{40} - p_{01}p_{15}p_{50} - p_{01}p_{16}p_{60} - p_{01}p_{17}p_{70} - p_{02}p_{23}p_{30} - p_{02}p_{24}p_{40} - p_{02}p_{25}p_{50} - p_{02}p_{26}p_{60} - p_{02}p_{27}p_{70} - p_{01}p_{14}p_{45}p_{50} - p_{01}p_{15}p_{56}p_{60} - p_{01}p_{16}p_{67}p_{70} - p_{02}p_{24}p_{45}p_{50} - p_{02}p_{25}p_{56}p_{60} - p_{02}p_{26}p_{67}p_{70} - p_{01}p_{14}p_{45}p_{56}p_{60} - p_{01}p_{15}p_{56}p_{67}p_{70} - p_{02}p_{24}p_{45}p_{56}p_{60} - p_{02}p_{25}p_{56}p_{67}p_{70} - p_{01}p_{14}p_{45}p_{56}p_{67}p_{70} - p_{02}p_{24}p_{45}p_{56}p_{67}p_{70}.$$

■

Theorem 2. Expected total time in home isolation for the patient starting in state S_0 is given by

$$\frac{\mu_3(p_{01}p_{13} + p_{02}p_{23})}{D},$$

where D has been already specified in Theorem 1.

Proof. Let $\psi_i(t)$ denote the probability that the patient is in home isolation at instant t, given that the patient entered state S_i at $t = 0$. Proceeding on similar lines shown in Theorem 1, we obtained the following recursive relations:

$\psi_0(t) = q_{01}(t)©\psi_1(t) + q_{02}(t)©\psi_2(t)$
$\psi_1(t) = q_{13}(t)©\psi_3(t) + q_{14}(t)©\psi_4(t) + q_{15}(t)©\psi_5(t) + q_{16}(t)©\psi_6(t) + q_{17}(t)©\psi_7(t)$
$\psi_2(t) = q_{23}(t)©\psi_3(t) + q_{24}(t)©\psi_4(t) + q_{25}(t)©\psi_5(t) + q_{26}(t)©\psi_6(t) + q_{27}(t)©\psi_7(t)$
$\psi_3(t) = W_3(t) + q_{30}(t)©\psi_0(t)$
$\psi_4(t) = q_{40}(t)©\psi_0(t) + q_{45}(t)©\psi_5(t)$
$\psi_5(t) = q_{50}(t)©\psi_0(t) + q_{56}(t)©\psi_6(t)$
$\psi_6(t) = q_{60}(t)©\psi_0(t) + q_{67}(t)©\psi_7(t)$
$\psi_7(t) = q_{70}(t)©\psi_0(t)$
where
$W_3(t) = 1 - a_1G_1(t) - b_1H_1(t)$

Taking Laplace transform of the above system of equations and solving for $\psi_0^*(s)$, we obtain

$$\psi_0^*(s) = \frac{N_2(s)}{D_1(s)},$$

where
$N_2(s) = W_3^*(s)(q_{01}^*(s)q_{13}^*(s) + q_{02}^*(s)q_{23}^*(s))$
and $D_1(s)$ has been already specified in Theorem 1.

Expected total time in home isolation for the patient starting in state S_0 is given by

$$\int_0^\infty \psi_0(t)\,dt = lim_{s\to 0}\ \psi_0^*(s) = \frac{\mu_3(p_{01}p_{13} + p_{02}p_{23})}{D},$$

where D has been already specified in Theorem 1. ■

Theorem 3. Expected total time in hospital for the patient starting in state S_0 is given by

$$\frac{1}{D}(p_{01}p_{14}\mu_4 + p_{01}p_{15}\mu_5 + p_{01}p_{16}\mu_6 + p_{01}p_{17}\mu_7 + p_{02}p_{24}\mu_4 + p_{02}p_{25}\mu_5 + p_{02}p_{26}\mu_6 + p_{02}p_{27}\mu_7 + p_{01}p_{14}p_{45}\mu_5 + p_{01}p_{15}p_{56}\mu_6 + p_{01}p_{16}p_{67}\mu_7 + p_{02}p_{24}p_{45}\mu_5 + p_{02}p_{25}p_{56}\mu_6 + p_{02}p_{26}p_{67}\mu_7 + p_{01}p_{14}p_{45}p_{56}\mu_6 + p_{02}p_{24}p_{45}p_{56}\mu_6 + p_{01}p_{15}p_{56}p_{67}\mu_7 + p_{02}p_{25}p_{56}p_{67}\mu_7 + p_{01}p_{14}p_{45}p_{56}p_{67}\mu_7 + p_{02}p_{24}p_{45}p_{56}p_{67}\mu_7),$$

where D has been already specified in Theorem 1.

Proof. Let $\chi_i(t)$ denote the probability that the patient is in hospital at instant t, given that the patient entered state S_i at $t = 0$. Proceeding on similar lines shown in Theorem 1, we obtained the following recursive relations.

$\chi_0(t) = q_{01}(t)\copyright\chi_1(t) + q_{02}(t)\copyright\chi_2(t)$
$\chi_1(t) = q_{13}(t)\copyright\chi_3(t) + q_{14}(t)\copyright\chi_4(t) + q_{15}(t)\copyright\chi_5(t) + q_{16}(t)\copyright\chi_6(t) + q_{17}(t)\copyright\chi_7(t)$
$\chi_2(t) = q_{23}(t)\copyright\chi_3(t) + q_{24}(t)\copyright\chi_4(t) + q_{25}(t)\copyright\chi_5(t) + q_{26}(t)\copyright\chi_6(t) + q_{27}(t)\copyright\chi_7(t)$
$\chi_3(t) = q_{30}(t)\copyright\chi_0(t)$
$\chi_4(t) = W_4(t) + q_{40}(t)\copyright\chi_0(t) + q_{45}(t)\copyright\chi_5(t)$
$\chi_5(t) = W_5(t) + q_{50}(t)\copyright\chi_0(t) + q_{56}(t)\copyright\chi_6(t)$
$\chi_6(t) = W_6(t) + q_{60}(t)\copyright\chi_0(t) + q_{67}(t)\copyright\chi_7(t)$
$\chi_7(t) = W_7(t) + q_{70}(t)\copyright\chi_0(t)$
where
$W_4(t) = 1 - a_2G_2(t) - c_1(1 - e^{-\lambda_1 t}) - b_2H_2(t)$
$W_5(t) = 1 - a_3G_3(t) - c_2(1 - e^{-\lambda_2 t}) - b_3H_3(t)$
$W_6(t) = 1 - a_4G_4(t) - c_3(1 - e^{-\lambda_3 t}) - b_4H_4(t)$
$W_7(t) = 1 - a_5G_5(t) - b_5H_5(t)$

Taking Laplace transform of the above system of equations and solving for $\chi_0^*(s)$, we obtain

$$\chi_0^*(s) = \frac{N_3(s)}{D_1(s)},$$

where

$$\begin{aligned} N_3(s) = {} & W_4^*(s)(q_{01}^*(s)q_{14}^*(s) + q_{02}^*(s)q_{24}^*(s)) + W_5^*(s)(q_{01}^*(s)q_{15}^*(s) + q_{02}^*(s)q_{25}^*(s) \\ & + q_{01}^*(s)q_{14}^*(s)q_{45}^*(s) + q_{02}^*(s)q_{24}^*(s)q_{45}^*(s)) + W_6^*(s)(q_{01}^*(s)q_{16}^*(s) + q_{02}^*(s)q_{26}^*(s) \\ & + q_{01}^*(s)q_{15}^*(s)q_{56}^*(s) + q_{02}^*(s)q_{25}^*(s)q_{56}^*(s) + q_{01}^*(s)q_{14}^*(s)q_{45}^*(s)q_{56}^*(s) \\ & + q_{02}^*(s)q_{24}^*(s)q_{45}^*(s)q_{56}^*(s)) + W_7^*(s)(q_{01}^*(s)q_{17}^*(s) + q_{02}^*(s)q_{27}^*(s) \\ & + q_{01}^*(s)q_{16}^*(s)q_{67}^*(s) + q_{02}^*(s)q_{26}^*(s)q_{67}^*(s) + q_{01}^*(s)q_{15}^*(s)q_{56}^*(s)q_{67}^*(s) \\ & + q_{02}^*(s)q_{25}^*(s)q_{56}^*(s)q_{67}^*(s) + q_{01}^*(s)q_{14}^*(s)q_{45}^*(s)q_{56}^*(s)q_{67}^*(s) \\ & + q_{02}^*(s)q_{24}^*(s)q_{45}^*(s)q_{56}^*(s)q_{67}^*(s)) \end{aligned}$$

and $D_1(s)$ has been already specified in Theorem 1.

Expected total time in hospital for the patient starting in state S_0 is given by

$$\begin{aligned} \int_0^\infty \chi_0(t)\,dt &= lim_{s\to 0}\chi_0^*(s) \\ &= lim_{s\to 0}\frac{N_3(s)}{D_1(s)} \end{aligned}$$

Using the above value of $N_3(s)$ and simplifying we get the required result. ■

4. Numerical Computations

Numerical computations for the mean survival time, expected total time in home isolation, and expected total time in hospital have been performed. For illustrating our model results, the waiting time distributions are assumed as exponentials, as follows: $g_i(t) \sim exp(\gamma_i)$ and $h_i(t) \sim exp(\delta_i)$ where i=1,2,...,5.

The severity levels of the disease are defined as under.

mild illness	$SpO_2 \geq 94\%$ on room air and no shortness of breath
moderate illness	$90\% \leq SpO_2 < 94\%$ on room air or 24 <R.R.$\leq$ 30 breaths per minute
severe illness	$SpO_2 < 90\%$ on room air or R.R.>30 breaths per minute
critical illness	Respiratory failure or septic shock or multiple organ dysfunction or requires life sustaining treatment

In addition, the following values for parameters are assumed:
β=0.031/day, α=0.5/day, γ_1=0.074/day, γ_2=0.071/day, γ_3=0.055/day, γ_4=0.034/day, γ_5=0.023/day, λ_1=0.12/day, λ_2=0.15/day, λ_3=0.20/day, δ_1=0.0011/day, δ_2=0.0012/day, δ_3=0.0015/day, δ_4=0.0018/day, δ_5=0.0020/day, p=0.7, q=0.3, p_1=0.74, q_1=0.26, p_2=0.18, q_2=0.82, r_1=0.83, r_2=0.07, r_3=0.06, r_4=0.04, a_1=0.98, b_1=0.02, a_2=0.85, b_2=0.05, c_1=0.10, a_3=0.75, b_3=0.08, c_2=0.17, a_4=0.65, b_4=0.15, c_3=0.20, a_5=0.1, b_5=0.9.

For the above values of parameters, we obtained mean survival time, expected total time in home isolation and expected total time in hospital and analysed how these parameters vary corresponding to perturbations in transmission rates, deteriorating rates, and death rates. The results obtained are depicted below. Figure 2-6 forecasts how variations in the transmission rate and death rates will affect the mean survival time. Figure 7 illustrates how the expected total time in home isolation changes as the recovery rate and death rate vary. Figure 8-11 predicts how perturbations in the recovery rates and deteriorating rates will affect the expected total time in hospital.

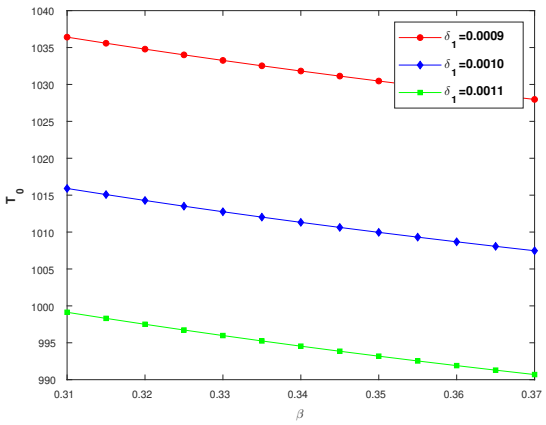


Figure 2: *A plot of T_0 with varied β and δ_1. The behaviour of the mean survival time (T_0) with respect to the transmission rate (β) for different values of death rate (δ_1) is demonstrated. From the figure, it can be seen that the mean survival time (T_0) decreases as the transmission rate (β) increases and gives lower values for higher values of death rate (δ_1).*

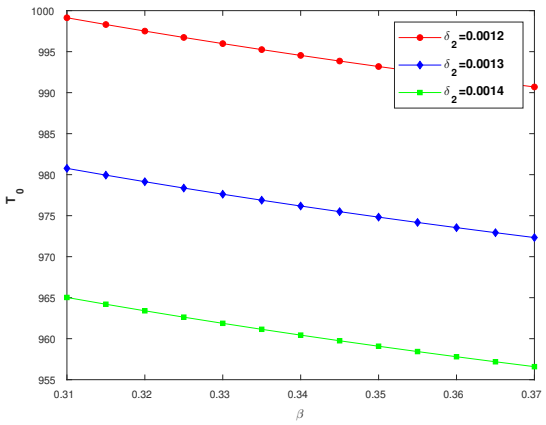


Figure 3: *A plot of T_0 with varied β and δ_2. The behaviour of the mean survival time (T_0) with respect to the transmission rate (β) for different values of death rate (δ_2) is demonstrated. From the figure, it can be seen that the mean survival time (T_0) decreases as the transmission rate (β) increases and gives lower values for higher values of death rate (δ_2).*

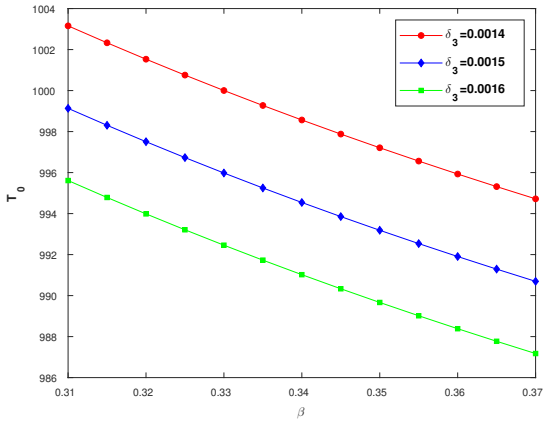


Figure 4: *A plot of T_0 with varied β and δ_3. The behaviour of the mean survival time (T_0) with respect to the transmission rate (β) for different values of death rate (δ_3) is demonstrated. From the figure, it can be seen that the mean survival time (T_0) decreases as the transmission rate (β) increases and gives lower values for higher values of death rate (δ_3).*

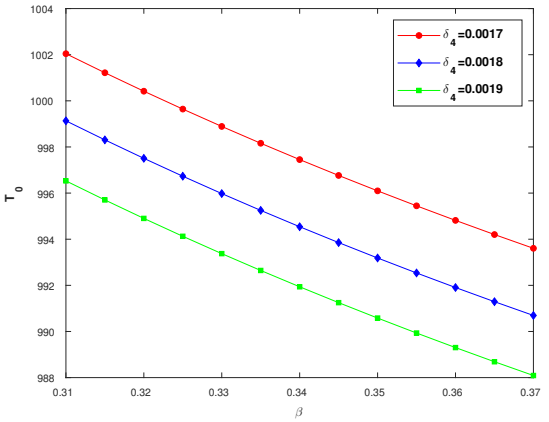


Figure 5: *A plot of T_0 with varied β and δ_4. The behaviour of the mean survival time (T_0) with respect to the transmission rate (β) for different values of death rate (δ_4) is demonstrated. From the figure, it can be seen that the mean survival time (T_0) decreases as the transmission rate (β) increases and gives lower values for higher values of death rate (δ_4).*

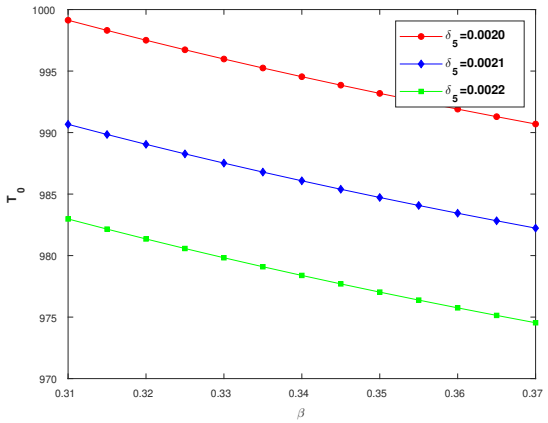


Figure 6: *A plot of T_0 with varied β and δ_5. The behaviour of the mean survival time (T_0) with respect to the transmission rate (β) for different values of death rate (δ_5) is demonstrated. From the figure, it can be seen that the mean survival time (T_0) decreases as the transmission rate (β) increases and gives lower values for higher values of death rate (δ_5).*

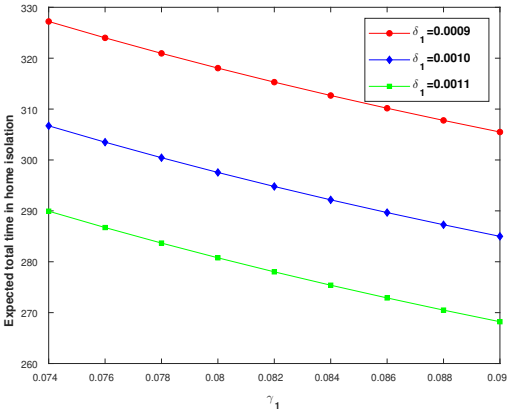


Figure 7: *A plot of expected total time in home isolation with varied γ_1 and δ_1. The behaviour of the expected total time in home isolation with respect to the recovery rate (γ_1) for different values of death rate (δ_1) is demonstrated. From the figure, it can be seen that the expected total time in home isolation decreases as the recovery rate (γ_1) increases and gives lower values for higher values of death rate (δ_1).*

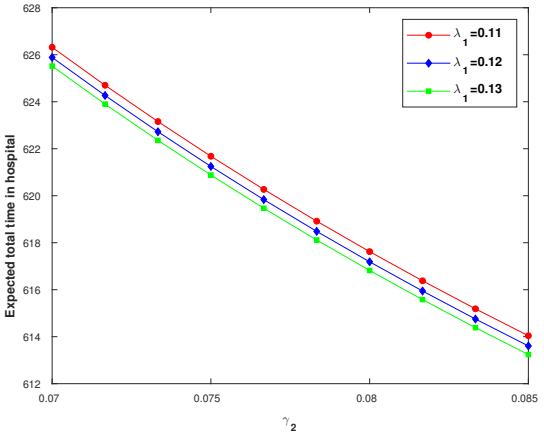


Figure 8: *A plot of expected total time in hospital with varied γ_2 and λ_1. The behaviour of the expected total time in hospital with respect to the recovery rate (γ_2) for different values of deteriorating rate (λ_1) is demonstrated. From the figure, it can be seen that the expected total time in hospital decreases as the recovery rate (γ_2) increases and gives lower values for higher values of deteriorating rate (λ_1).*

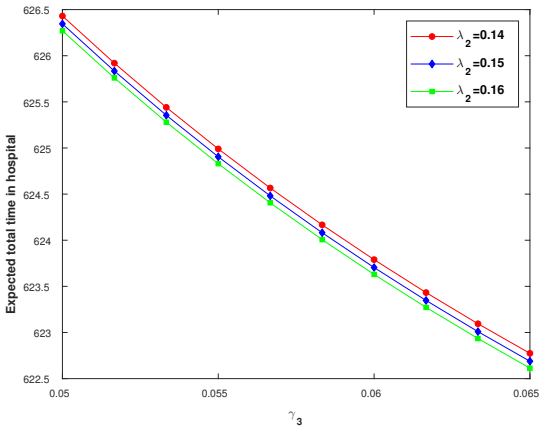


Figure 9: *A plot of expected total time in hospital with varied γ_3 and λ_2. The behaviour of the expected total time in hospital with respect to the recovery rate (γ_3) for different values of deteriorating rate (λ_2) is demonstrated. From the figure, it can be seen that the expected total time in hospital decreases as the recovery rate (γ_3) increases and gives lower values for higher values of deteriorating rate (λ_2).*

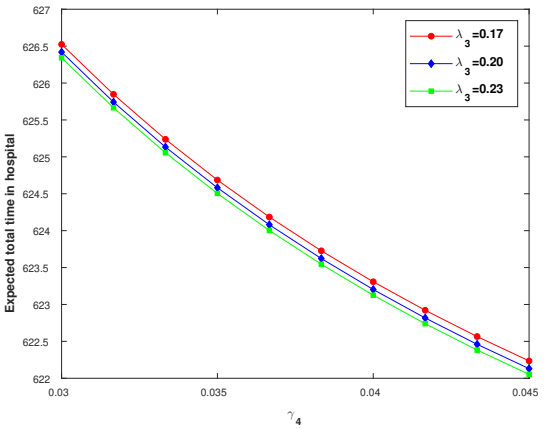


Figure 10: *A plot of expected total time in hospital with varied γ_4 and λ_3. The behaviour of the expected total time in hospital with respect to the recovery rate (γ_4) for different values of deteriorating rate (λ_3) is demonstrated. From the figure, it can be seen that the expected total time in hospital decreases as the recovery rate (γ_4) increases and gives lower values for higher values of deteriorating rate (λ_3).*

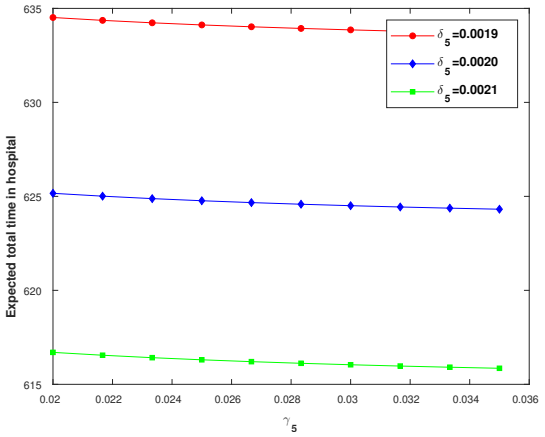


Figure 11: *A plot of expected total time in hospital with varied γ_5 and δ_5. The behaviour of the expected total time in hospital with respect to the recovery rate (γ_5) for different values of death rate (δ_5) is demonstrated. From the figure, it can be seen that the expected total time in hospital decreases as the recovery rate (γ_5) increases and gives lower values for higher values of death rate (δ_5).*

5. Conclusion

Designing prevention strategies and infection control policies can be benefitted using mathematical models of infectious diseases. On the basis of the idea of semi-Markov process, a new framework for modelling infectious diseases have been presented. The analysis of the model aids in examining the effects of various parameters on various system measures. According to the analysis presented,

it is concluded that the mean survival time declines as the disease's transmission rate rises and has lower values for greater values of death rate. The expected total time in home isolation reduces with rising recovery rates and has lower values for higher death rates. The expected total time in hospital decreases as the recovery rate increases and gives lower values for higher values of deteriorating rate. Through this article, the use and significance of semi-Markov models in understanding infectious diseases trends is demonstrated. This study may be helpful in selecting the optimal intervention tactics and creating effective infection control measures.

Declaration of Competing Interest

None.

Acknowledgement

The author Sujata Sukhija delightedly acknowledges Human Resource Development Group of Council of Scientific & Industrial Research (CSIR), India, for providing fellowship through file number 09/382(0258)/2020-EMR-I.

References

[1] Cao, Q., Buskens, E., Feenstra, T., Jaarsma, T., Hillege, H., & Postmus, D. (2016). Continuous time semi-markov models in health economic decision making: An illustrative example in heart failure disease management. *Medical Decision Making*, *36*(1), 59–71. https://doi.org/10.1177/0272989x15593080

[2] Castelli, C., Combescure, C., Foucher, Y., & Daures, J.-P. (2007). Cost-effectiveness analysis in colorectal cancer using a semi-markov model. *Statistics in Medicine*, *26*(30), 5557–5571. https://doi.org/10.1002/sim.3112

[3] Centers for Disease Control and Prevention. (n.d.). Quarantine and isolation. Retrieved December 2, 2022, from https://www.cdc.gov

[4] Chakraborty, H., Hossain, A., & Latif, M. A. (2019). A three-state continuous time Markov chain model for HIV disease burden. *Journal of Applied Statistics*, *46*(9), 1671–1688. https://doi.org/10.1080/02664763.2018.1555573

[5] Claris, S., & Delson, C. (2018). Time-homogeneous markov process for hiv/aids progression under a combination treatment therapy: Cohort study, south africa. *Theoretical Biology and Medical Modelling*, *15*(1), 1–14. https://doi.org/10.1186/s12976-017-0075-4

[6] Davidov, O. (1999). The steady-state probabilities for regenerative semi-markov processes with application to prevention and screening. *Applied Stochastic Models and Data Analysis*, *15*(1), 55–63. https://doi.org/https://doi.org/10.1002/(SICI)1099-0747(199903)15:1<55::AID-ASM358>3.0.CO;2-4

[7] Farahani, M. V., et al. (2020). Application of multi-state model in analyzing of breast cancer data. *Journal of research in health sciences*, *19*(4), 1–5. https://pubmed.ncbi.nlm.nih.gov/32291364

[8] Goshu, A. T., & Dessie, Z. G. (2013). Modelling progression of HIV/AIDS disease stages using semi-markov processes. *Journal of Data Science*, *11*(2), 269–280. https://doi.org/10.6339/JDS.2013.11(2).1136

[9] Grover, G., Sabharwal, A., Kumar, S., & Thakur, A. K. (2019). A multi-state markov model for the progression of chronic kidney disease. *Turkiye Klinikleri Journal of Biostatistics*, *11*(1). https://doi.org/10.5336/biostatic.2018-62156

[10] Kao, E. P. C. (1974). A note on the first two moments of times in transient states in a semi-markov process. *Journal of Applied Probability*, *11*(1), 193–198. https://doi.org/10.2307/3212598

[11] Kay, R. (1986). A markov model for analysing cancer markers and disease states in survival studies. *Biometrics*, *42*(4), 855–865. https://doi.org/10.2307/2530699

[12] Levy, P. (1954). Processus semi-markoviens. *Proc. Int. Congress. Math. (Amsterdam)*, *3*, 416–426.

[13] Ramezankhani, A., Azizi, F., Hadaegh, F., & Momenan, A. A. (2018). Diabetes and number of years of life lost with and without cardiovascular disease: A multi-state homogeneous semi-markov model. *Acta Diabetologica*, *55*(3), 253–262. https://doi.org/10.1007/s00592-017-1083-x

[14] Smith, W. L. (1955). Regenerative stochastic processes. *Proc. Roy. Soc. Ser. A*, *232*, 6–31.

[15] Uhry, Z., Hédelin, G., Colonna, M., Asselain, B., Arveux, P., Rogel, A., Exbrayat, C., Guldenfels, C., Courtial, I., Soler-Michel, P., Molinié, F., Eilstein, D., & Duffy, S. (2010). Multi-state markov models in cancer screening evaluation: A brief review and case study. *Statistical Methods in Medical Research*, *19*(5), 463–486. https://doi.org/10.1177/0962280209359848

[16] Weiss, G. H., & Zelen, M. (1965). A semi-markov model for clinical trials. *Journal of Applied Probability*, *2*(2), 269–285. https://doi.org/10.2307/3212194

[17] World Health Organization. (n.d.). The top 10 causes of death. Retrieved May 4, 2022, from https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[18] Wu, P., Zhang, R., & Din, A. (2023). Mathematical analysis of an age-since infection and diffusion HIV/AIDS model with treatment adherence and dirichlet boundary condition. *Mathematics and Computers in Simulation*, *214*, 1–27. https://doi.org/https://doi.org/10.1016/j.matcom.2023.06.018

[19] Yang, J., Chen, Z., Tan, Y., Liu, Z., & Cheke, R. A. (2023). Threshold dynamics of an age-structured infectious disease model with limited medical resources. *Mathematics and Computers in Simulation*, *214*, 114–132. https://doi.org/https://doi.org/10.1016/j.matcom.2023.07.003

[20] Zvifadzo, M. Z., F, C. T., Jim, T., & Eustasius, M. (2019). HIV disease progression among antiretroviral therapy patients in zimbabwe: A multistate markov model. *Frontiers in public health*, *7*, 326. https://doi.org/10.3389/fpubh.2019.00326

ON AN IMPATIENT CONSUMER QUEUE WITH SECONDARY SERVICE, MULTIPLE VACATIONS AND SERVER BREAKDOWNS

K. Jyothsna[1,*], P. Vijaya Kumar[2], P. Vijaya Laxmi[3],

•

[1,*]Department of Mathematics, GSS, GITAM (Deemed to be University),
Visakhapatnam - 530045, Andhra Pradesh, India.
mail2jyothsnak@yahoo.co.in, drjyothsnak1984@gmail.com
[2]Department of Mathematics, GSS, GITAM (Deemed to be University),
Visakhapatnam - 530045, Andhra Pradesh, India.
vprathi@gitam.edu
[3]Department of Applied Mathematics, Andhra University,
Visakhapatnam- 530003, Andhra Pradesh, India.
vijayalaxmiau@gmail.com

Abstract

This study presents a limited buffer secondary service queue with multiple vacations and server breakdowns. The model under consideration includes two types of impatient policies: balking and reneging. After the completion of the essential primary service, only few consumers choose to proceed with secondary service with a certain probability. During the active period of the server, it is subject to breakdown and the broken down server is immediately sent for repair. Further, the server will go on vacation as soon as there are no waiting consumers in the queue. On returning from a vacation, if the system is still empty the server leaves for another vacation and continues to do so until atleast one consumer is found at a vacation termination epoch. The model is analyzed under steady-state conditions and the explicit expressions of various performance indices are evaluated. A few numerical results illustrate how the model parameters have an effect on the performance metrics.

Keywords: Balking, reneging, secondary service, multiple vacations, breakdown, repair.

1. Introduction

Vacation models address a very significant category of the real-world congestion scenarios that are seen in both day-to-day living and in industrial settings. In the context of queueing, the period of time during which the server is not available is referred to as a vacation. During the active period, the server operates at maximum capacity, but while it is on vacation, it does not carry out any tasks. The numerous adaptable implementations encourage us to investigate queuing systems with server vacations that can be exploited in some beneficial method when coping with congestion problems in various frameworks. These systems may be used in a variety of contexts. Over the course of more than two decades, number of scholars and practitioners have investigated vacation models of many sorts. Their goals have been either to find solutions to specific queueing issues at hand or to acquire a knowledge of the stochastic processes that evolve as a result of these models. Excellent studies on these vacation models have been done by Doshi [8, 9], Takagi [21], Tian and Zhang [22], Jau-Chaun Ke et al. [15], Panta et al. [18], etc.

Consumers typically have less patience when waiting for service because they value their time. In the research that has been done on queueing, impatience of consumers has been examined mostly in the context of consumers abandoning the queue because of either a prolonged

wait that they had previously experienced or a lengthy wait that they predicted would occur upon arrival. In many service or operational contexts, such impatience is frequently noticed through the acts of consumers “balking" or “reneging" from waiting in a queue. Altman and Yechiali [1] undertook an analysis of consumer dissatisfaction in server vacation queues. A study on priority queues with impatient consumers has been presented by Foad and Baris [10]. Yue et al. [27] conducted research on the effects of synchronised vacations and impatient consumers in a multi-server queue. Jamol Pender [14] computed a novel approximation for single-server queues with abandonment based on the truncated normal distribution. Ammar [3] came up with the time dependent results of an $M/M/1$ vacation queue that included an anticipating server and impatient consumers. Sampath and Liu [19] conducted the research to determine how impatience of customers affected the performance of an $M/M/1$ queueing system subject to waiting server and differential vacations. A Bernoulli feedback queueing system with K-variant vacations, waiting server and impatient consumers has been considered by Amina and Guendouzi [2]. Mathematical evaluation of the $M/M/C$ vacation queuing model with a waiting server and dissatisfied consumers has been carried out by Ganesh and Ghimire [11].

Numerous malfunctions in the service providing facility are a major source of service interruptions in various manufacturing processes. In such circumstances, service will not be provided to the waiting consumers until the service providing facility is repaired. Such breakdowns in service are typical in commercial settings like factories and phone booths, as well as in the use of mechanical technologies like electronic computers. William et al. [26] examined a queueing model with vacations where the service station may experience a breakdown while it is in operation. Queuing system with fixed capacity and vacancies as well as server breakdowns has been dealt by Ghimire and Ritu [12]. An $M^X/G/1$ queue with server breakdowns and repairs has been analyzed by Djamila et al. [7]. Hanumantha Rao et al. [13] has studied an impatient consumer two-phase queuing system with server breakdowns and delayed repair. Under the T-policy, the investigation of two phase queue with breakdowns and vacations has been carried out by Khalid and Lotfi [17]. A bulk service queue with server breakdowns and repairs had been investigated by Bharathidass et al. [4]. Srinivas et al. [20] researched a server breakdown queueing system with repairs and vacations.

In the modelling of a wide variety of congestion issues that arise during real-life activities, queueing systems that include the provision of a secondary service (SS) play an essential role. In queueing models that include SS, the server offers the primary service (PS) to all of the arriving consumers. However, after the PS has been completed, only few consumers choose to receive SS, according to a predetermined probability. Take the banking sector as an illustration; among the primary duties imposed upon any bank are the receipt and dispensing funds in the form of deposits, withdrawals, loans, advances, etc. Printing of passbooks, issue of checks, lockers, etc., fall under the category of secondary services that banks conduct in addition to their primary duties. Kalyanaraman and Pazhani [16] analyzed a single server queue with optional service and server vacations. Uma and Punniyamoorthy [23] have investigated a single-server bulk queue with vacations, balking and secondary service. A bulk arrival queue with SS and server breakdowns has been researched by Charan and Sandeep [5]. Charan et al. [6] have studied the effects of secondary services and service disruptions on bulk queues. The time dependent behaviour of a bulk service queueing system with optional service and impatient consumers has been studied by Vijaya Laxmi and Andwilile [24]. A Markovian secondary service queue operating under triadic policy has been considered by Vijaya Laxmi et al. [25].

The current article deals with a finite capacity Markovian queue with secondary service, multiple vacations, breakdowns and impatient consumers. The paradigm under consideration has many real-world uses in places as diverse as communications networks, manufacturing systems, cloud computing, customer service centres, etc. Consider the following example of a call centre that provides services to customers. The call centre employs customer support agents to handle inbound calls received at the call centre from customers seeking assistance. Customers after receiving the required assistance and are prepared to pay an extra price can take advantage of the premium service option that is available through the contact centre. During times of

low call volume or off-peak hours, it is possible that the server can schedule for a vacation during which they will be offline in order to conserve money and resources. Occasionally, there may be brief disruptions in service as a result of server experiencing technical problems or being required to undergo maintenance. It is always possible that customers contacting the call centre may be impatient and reluctant to wait in the queue for lengthy periods of time. In this practical application, the call centre aims to manage server breakdowns effectively, offers an optional service to customers who value quicker support and allows server vacations during low demand to optimize resource usage. By doing so, the call centre enhances the overall customer experience and reduces the likelihood of impatient customers abandoning the queue.

Owing to the practical application as one mentioned above, we study an $M/M/1/N$ multiple vacation queue with secondary service, breakdowns and impatient consumers. The vacation durations, secondary service durations and breakdown times are assumed to be follow exponential distribution. Balking and reneging are the two forms of consumer impatience which have been included in the current article. Both the forms of consumer impatience are considered to be state dependent. Using iterative approach, the model's steady-state results are achieved. The expected system size, expected balking rate, expected reneging rate and other performance parameters are reported. Through a limited number of numerical experiments, the parameter influence on the performance indices is demonstrated.

The remaining sections of the paper are structured as follows. A detailed explanation of the model has been provided in Section 2. In Section 3, we reported the results of the steady-state model under discussion. Section 4 provides different metrics by which the model's effectiveness may be evaluated. In Section 5, some numerical findings illustrating the impact of the model parameters on the performance metrics are shown and in Section 6, conclusions are drawn.

2. Model Overview

Consider an impatient consume $M/M/1/N$ queue with multiple vacations, secondary service, and server breakdowns.

- Consumers arrive one at a time according to Poisson process with rate λ. Arriving consumers make a decision whether to be a part of the queue or not depending on the queue size. Let b_n be the probability of joining the queue and $1 - b_n$ be the probability of not joining the queue, where n denotes the number of consumers in the system. We also, assume that $b_0 = 1, b_{n+1} \leq b_n$ and $b_N = 0$.

- Consumers who join the queue wait for certain period of time, T, which follows exponential distribution with parameter α. If the service does not start before this time, he may leave due to impatience. The average reneging rate of a consumer is taken as $(n-1)\alpha$ $(n \geq 0)$, where n represents the number of consumers in the system.

- Consumers who enter the system are served according to FCFS discipline by a single server. All the consumers receive a primary service (PS) and exit from the system with probability ω while only few consumers may opt for secondary service (SS) with probability $\bar{\omega} = 1 - \omega$. The service durations during PS and SS follow exponential distribution with parameters μ_1 and μ_2, respectively.

- Under the multiple vacation policy, the server will take a series of breaks in the form of vacations in between two consecutive busy times, and it will continue to do so until it locates a waiting consumer in the system. The vacation durations are also assumed to follow exponential distribution with parameter σ.

- The server is subject to breakdown both during PS and SS with rate β. The broken down server is immediately sent for repair. The repair times are exponentially distributed with rate δ.

3. Steady-state analysis

At steady-state, let

- $\pi_{n,0}$ - Probability of n consumers in the system and the server in vacation,
- $\pi_{n,1W}$ - Probability of n consumers in the system and server in working state during PS,
- $\pi_{n,1B}$ - Probability of n consumers in the system and server in breakdown state during PS,
- $\pi_{n,2W}$ - Probability of n consumers in the system and server in working state during SS,
- $\pi_{n,2B}$ - Probability of n consumers in the system and server in breakdown state during SS.

Using the Markov theory, the set of steady-state equations may be obtained as

$$\begin{aligned}
\lambda\pi_{0,0} &= \omega\mu_1\pi_{1,1W} + \mu_2\pi_{1,2W}, && (1)\\
(\lambda b_n + \sigma + (n-1)\alpha)\,\pi_{n,0} &= \lambda b_{n-1}\pi_{n-1,0} + n\alpha\pi_{n+1,0}, 1 \le n \le N-1, && (2)\\
(\sigma + (N-1)\alpha)\,\pi_{N,0} &= \lambda b_{N-1}\pi_{N-1,0}, && (3)\\
(\lambda b_1 + \beta + \mu_1)\,\pi_{1,1W} &= (\omega\mu_1 + \alpha)\,\pi_{2,1W} + \sigma\pi_{1,0} + \mu_2\pi_{2,2W} + \delta\pi_{1,1B}, && (4)\\
(\lambda b_n + \beta + \mu_1 + (n-1)\alpha)\,\pi_{n,1W} &= (\omega\mu_1 + n\alpha)\,\pi_{n+1,1W} + \mu_2\pi_{n+1,2W} + \sigma\pi_{n,0} + \delta\pi_{n,1B}, \\
&\quad + \lambda b_{n-1}\pi_{n-1,1W}, 2 \le n \le N-1, && (5)\\
(\beta + \mu_1 + (N-1)\alpha)\,\pi_{N,1W} &= \sigma\pi_{N,0} + \delta\pi_{N,1B} + \lambda b_{N-1}\pi_{N-1,1W}, && (6)\\
(\lambda b_1 + \beta + \mu_2)\,\pi_{1,2W} &= \alpha\pi_{2,2W} + \bar{\omega}\mu_1\pi_{1,1W} + \delta\pi_{1,2B}, && (7)\\
(\lambda b_n + \beta + \mu_2 + (n-1)\alpha)\,\pi_{n,2W} &= n\alpha\pi_{n+1,2W} + \bar{\omega}\mu_1\pi_{n,1W} + \delta\pi_{n,2B} + \lambda b_{n-1}\pi_{n-1,2W}, \\
&\quad 2 \le n \le N-1, && (8)\\
(\beta + \mu_2 + (N-1)\alpha)\,\pi_{N,2W} &= \bar{\omega}\mu_1\pi_{N,1W} + \delta\pi_{N,2B} + \lambda b_{N-1}\pi_{N-1,2W}, && (9)\\
(\lambda b_1 + \delta)\,\pi_{1,1B} &= \alpha\pi_{2,1B} + \beta\pi_{1,1W}, && (10)\\
(\lambda b_n + \delta + (n-1)\alpha)\,\pi_{n,1B} &= n\alpha\pi_{n+1,1B} + \beta\pi_{n,1W} + \lambda b_{n-1}\pi_{n-1,1B}, 2 \le n \le N-1, && (11)\\
(\delta + (N-1)\alpha)\,\pi_{N,1B} &= \beta\pi_{N,1W} + \lambda b_{N-1}\pi_{N-1,1B}, && (12)\\
(\lambda b_1 + \delta)\,\pi_{1,2B} &= \alpha\pi_{2,2B} + \beta\pi_{1,2W}, && (13)\\
(\lambda b_n + \delta + (n-1)\alpha)\,\pi_{n,2B} &= n\alpha\pi_{n+1,2B} + \beta\pi_{n,2W} + \lambda b_{n-1}\pi_{n-1,2B}, 2 \le n \le N-1, && (14)\\
(\delta + (N-1)\alpha)\,\pi_{N,2B} &= \beta\pi_{N,2W} + \lambda b_{N-1}\pi_{N-1,2B}. && (15)
\end{aligned}$$

The steady-state probabilities are obtained by solving the above system of equations recursively as shown below.

$$\begin{aligned}
\pi_{n,0} &= r_n\pi_{N,0}, 1 \le n \le N,\\
\pi_{n,2B} &= (d_n + s_n k_{13} + t_n k_{14} + z_n k_{15} + \gamma_n k_{16})\,\pi_{N,0}, 1 \le n \le N,\\
\pi_{n,2W} &= (l_n + y_n k_{13} + w_n k_{14} + x_n k_{15} + m_n k_{16})\,\pi_{N,0}, 1 \le n \le N,\\
\pi_{n,1W} &= (g_n + p_n k_{13} + o_n k_{14} + f_n k_{15} + h_n k_{16})\,\pi_{N,0}, 1 \le n \le N,\\
\pi_{n,1B} &= (q_n + \chi_n k_{13} + c_n k_{14} + v_n k_{15} + u_n k_{16})\,\pi_{N,0}, 1 \le n \le N,
\end{aligned}$$

where

$$\begin{aligned}
r_N &= s_N = w_N = f_N = u_N = 1,\\
t_N &= z_N = d_N = \gamma_N = x_N = y_N = l_N = m_N = g_N = h_N = o_N = p_N = v_N = q_N = c_N = 0,\\
z_{N-1} &= d_{N-1} = \gamma_{N-1} = m_{N-1} = l_{N-1} = o_{N-1} = p_{N=1} = q_{N-1} = c_{N-1} = \chi_N = \chi_{N-1} = 0\\
d_{N-2} &= \gamma_{N-2} = p_{N-2} = c_{N-2} = \chi_{N-2} = \chi_{N-3} = 0,\\
r_{N-1} &= \frac{\sigma + (N-1)\alpha}{\lambda b_{N-1}},\ s_{N-1} = \frac{\delta + (N-1)\alpha}{\lambda b_{N-1}},\ t_{N-1} = -\frac{\beta}{\lambda b_{N-1}}, w_{N-1} = \frac{\beta + \mu_2 + (N-1)\alpha}{\lambda b_{N-1}},
\end{aligned}$$

$$
\begin{aligned}
x_{N-1} &= -\frac{\bar{\omega}\mu_1}{\lambda b_{N-1}}, u_{N-1} = \frac{\delta+(N-1)\alpha}{\lambda b_{N-1}}, v_{N-1} = -\frac{\beta}{\lambda b_{N-1}},\ f_{N-1} = \frac{\beta+\mu_1+(N-1)\alpha}{\lambda b_{N-1}},\\
g_{N-1} &= -\frac{\sigma}{\lambda b_{N-1}},\ h_{N-1} = -\frac{\delta}{\lambda b_{N-1}}/\\
r_n &= \left(\frac{\lambda b_{n+1}+\sigma+n\alpha}{\lambda b_n}\right) r_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) r_{n+2}, n = N-2, N-3, \ldots, 1,\\
s_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right) s_{n+1} - \left(\frac{\beta}{\lambda b_n}\right) y_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) s_{n+2}, n = N-2, N-3, \ldots, 1,\\
t_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right) t_{n+1} - \left(\frac{\beta}{\lambda b_n}\right) w_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) t_{n+2}, n = N-2, N-3, \ldots, 1,\\
z_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right) z_{n+1} - \left(\frac{\beta}{\lambda b_n}\right) x_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) z_{n+2}, n = N-2, N-3, \ldots, 1,\\
d_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right) d_{n+1} - \left(\frac{\beta}{\lambda b_n}\right) l_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) d_{n+2}, n = N-2, N-3, \ldots, 1,\\
\gamma_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right) \gamma_{n+1} - \left(\frac{\beta}{\lambda b_n}\right) m_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) \gamma_{n+2}, n = N-2, N-3, \ldots, 1,\\
w_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_2+n\alpha}{\lambda b_n}\right) w_{n+1} - \left(\frac{\delta}{\lambda b_n}\right) t_{n+1} - \left(\frac{\bar{\omega}\mu_1}{\lambda b_n}\right) o_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) w_{n+2},\\
&\quad n = N-2, N-3, \ldots, 1,\\
x_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_2+n\alpha}{\lambda b_n}\right) x_{n+1} - \left(\frac{\delta}{\lambda b_n}\right) z_{n+1} - \left(\frac{\bar{\omega}\mu_1}{\lambda b_n}\right) f_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) x_{n+2},\\
&\quad n = N-2, N-3, \ldots, 1,\\
y_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_2+n\alpha}{\lambda b_n}\right) y_{n+1} - \left(\frac{\delta}{\lambda b_n}\right) s_{n+1} - \left(\frac{\bar{\omega}\mu_1}{\lambda b_n}\right) p_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) y_{n+2},\\
&\quad n = N-2, N-3, \ldots, 1,\\
l_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_2+n\alpha}{\lambda b_n}\right) l_{n+1} - \left(\frac{\delta}{\lambda b_n}\right) d_{n+1} - \left(\frac{\bar{\omega}\mu_1}{\lambda b_n}\right) g_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) l_{n+2},\\
&\quad n = N-2, N-3, \ldots, 1,\\
m_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_2+n\alpha}{\lambda b_n}\right) m_{n+1} - \left(\frac{\delta}{\lambda b_n}\right) \gamma_{n+1} - \left(\frac{\bar{\omega}\mu_1}{\lambda b_n}\right) h_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) m_{n+2},\\
&\quad n = N-2, N-3, \ldots, 1,\\
f_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_1+n\alpha}{\lambda b_n}\right) f_{n+1} - \left(\frac{\omega\mu_1+(n+1)\alpha}{\lambda b_n}\right) f_{n+2} - \left(\frac{\mu_2}{\lambda b_n}\right) x_{n+2} - \left(\frac{\delta}{\lambda b_n}\right) v_{n+1},\\
&\quad n = N-2, N-3, \ldots, 1,\\
g_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_1+n\alpha}{\lambda b_n}\right) g_{n+1} - \left(\frac{\omega\mu_1+(n+1)\alpha}{\lambda b_n}\right) g_{n+2} - \left(\frac{\mu_2}{\lambda b_n}\right) l_{n+2} - \left(\frac{\delta}{\lambda b_n}\right) q_{n+1},\\
&\quad - \left(\frac{\sigma}{\lambda b_n}\right) r_{n+1}, n = N-2, N-3, \ldots, 1,\\
h_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_1+n\alpha}{\lambda b_n}\right) h_{n+1} - \left(\frac{\omega\mu_1+(n+1)\alpha}{\lambda b_n}\right) h_{n+2} - \left(\frac{\mu_2}{\lambda b_n}\right) m_{n+2} - \left(\frac{\delta}{\lambda b_n}\right) u_{n+1},\\
&\quad n = N-2, N-3, \ldots, 1,\\
o_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_1+n\alpha}{\lambda b_n}\right) o_{n+1} - \left(\frac{\omega\mu_1+(n+1)\alpha}{\lambda b_n}\right) o_{n+2} - \left(\frac{\mu_2}{\lambda b_n}\right) w_{n+2} - \left(\frac{\delta}{\lambda b_n}\right) c_{n+1},\\
&\quad n = N-2, N-3, \ldots, 1,\\
p_n &= \left(\frac{\lambda b_{n+1}+\beta+\mu_1+n\alpha}{\lambda b_n}\right) p_{n+1} - \left(\frac{\omega\mu_1+(n+1)\alpha}{\lambda b_n}\right) p_{n+2} - \left(\frac{\mu_2}{\lambda b_n}\right) y_{n+2} - \left(\frac{\delta}{\lambda b_n}\right) \chi_{n+1},\\
&\quad n = N-2, N-3, \ldots, 1,\\
u_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right) u_{n+1} - \left(\frac{(n+1)\alpha}{\lambda b_n}\right) u_{n+2} - \left(\frac{\beta}{\lambda b_n}\right) h_{n+1}, n = N-2, N-3, \ldots, 1,
\end{aligned}
$$

$$
\begin{aligned}
v_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right)v_{n+1}-\left(\frac{(n+1)\alpha}{\lambda b_n}\right)v_{n+2}-\left(\frac{\beta}{\lambda b_n}\right)f_{n+1}, n=N-2,N-3,\ldots,1,\\
q_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right)q_{n+1}-\left(\frac{(n+1)\alpha}{\lambda b_n}\right)q_{n+2}-\left(\frac{\beta}{\lambda b_n}\right)g_{n+1}, n=N-2,N-3,\ldots,1,\\
c_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right)c_{n+1}-\left(\frac{(n+1)\alpha}{\lambda b_n}\right)c_{n+2}-\left(\frac{\beta}{\lambda b_n}\right)o_{n+1}, n=N-2,N-3,\ldots,1,\\
\chi_n &= \left(\frac{\lambda b_{n+1}+\delta+n\alpha}{\lambda b_n}\right)\chi_{n+1}-\left(\frac{(n+1)\alpha}{\lambda b_n}\right)\chi_{n+2}-\left(\frac{\beta}{\lambda b_n}\right)p_{n+1}, n=N-2,N-3,\ldots,1,\\
k_1 &= \frac{\alpha z_2+\beta x_1-(\lambda b_1+\delta)z_1}{(\lambda b_1+\delta)\gamma_1-\alpha\gamma_2-\beta m_1},\ k_2=\frac{\alpha d_2+\beta l_1-(\lambda b_1+\delta)d_1}{(\lambda b_1+\delta)\gamma_1-\alpha\gamma_2-\beta m_1},\\
k_3 &= \frac{\alpha t_2+\beta w_1-(\lambda b_1+\delta)t_1}{(\lambda b_1+\delta)\gamma_1-\alpha\gamma_2-\beta m_1},\ k_4=\frac{\alpha s_2+\beta y_1-(\lambda b_1+\delta)s_1}{(\lambda b_1+\delta)\gamma_1-\alpha\gamma_2-\beta m_1},\\
k_5 &= \alpha u_2+\beta h_1-(\lambda b_1+\delta)u_1,\ k_6=\frac{k_2k_5+\alpha q_2+\beta g_1-(\lambda b_1+\delta)q_1}{(\lambda b_1+\delta)v_1-\alpha v_2-\beta f_1-k_1k_5},\\
k_7 &= \frac{k_3k_5+\alpha c_2+\beta o_1-(\lambda b_1+\delta)c_1}{(\lambda b_1+\delta)v_1-\alpha v_2-\beta f_1-k_1k_5},\ k_8=\frac{k_4k_5+\alpha\chi_2+\beta p_1-(\lambda b_1+\delta)\chi_1}{(\lambda b_1+\delta)v_1-\alpha v_2-\beta f_1-k_1k_5},\\
k_9 &= \alpha x_2+\bar{\omega}\mu_1 f_1+\delta z_1-(\lambda b_1+\mu_2+\beta)x_1,\ k_{10}=\alpha m_2+\bar{\omega}\mu_1 h_1+\delta\gamma_1-(\lambda b_1+\mu_2+\beta)m_1,\\
k_{11} &= \frac{k_9k_6+\alpha l_2+\bar{\omega}\mu_1 g_1+\delta d_1+k_2k_{10}+k_1k_6k_{10}-(\lambda b_1+\mu_2+\beta)l_1}{(\lambda b_1+\mu_2+\beta)w_1-\alpha w_2-\bar{\omega}\mu_1 o_1-\delta t_1-k_9k_7-k_3k_{10}-k_1k_7k_{10}},\\
k_{12} &= \frac{k_9k_8+\alpha y_2+\bar{\omega}\mu_1 p_1+\delta s_1+k_4k_{10}+k_1k_8k_{10}-(\lambda b_1+\mu_2+\beta)y_1}{(\lambda b_1+\mu_2+\beta)w_1-\alpha w_2-\bar{\omega}\mu_1 o_1-\delta t_1-k_9k_7-k_3k_{10}-k_1k_7k_{10}},\\
num &= (k_6+k_7k_{11})\left((\omega\mu_1+\alpha)f_2+\mu_2x_2+\delta v_1-(\lambda b_1+\beta+\mu_1)f_1\right)\\
&\quad+(k_2+k_3k_{11}+k_1k_6+k_1k_7k_{11})\left((\omega\mu_1+\alpha)h_2+\mu_2m_2+\delta u_1-(\lambda b_1+\beta+\mu_1)h_1\right)\\
&\quad+(\omega\mu_1+\alpha)\left(g_2+o_2k_{11}\right)+\mu_2\left(l_2+w_2k_{11}\right)+\delta\left(q_1+c_1k_{11}\right)\\
&\quad+\sigma r_1-(\lambda b_1+\beta+\mu_1)\left(g_1+o_1k_{11}\right),\\
den &= (k_8+k_7k_{12})\left((\lambda b_1+\beta+\mu_1)f_1-(\omega\mu_1+\alpha)f_2-\mu_2x_2-\delta v_1\right)+(k_4+k_3k_{12}\\
&\quad+k_8k_1+k_1k_7k_{12})\left((\lambda b_1+\beta+\mu_1)h_1-(\omega\mu_1+\alpha)h_2-\mu_2m_2-\delta u_1\right)+(\lambda b_1+\beta+\mu_1)\\
&\quad\times(p_1+o_1k_{12})-(\omega\mu_1+\alpha)\left(p_2+o_2k_{12}\right)-\mu_2\left(y_2+w_2k_{12}\right)-\delta\left(\chi_1+c_1k_{12}\right)\\
k_{13} &= \frac{num}{den},\\
k_{14} &= k_{11}+k_{12}k_{13},\ k_{15}=k_6+k_7k_{14}+k_8k_{13},\ k_{16}=k_2+k_1k_{15}+k_3k_{14}+k_4k_{13}
\end{aligned}
$$

Finally, $\pi_{0,N}$ is computed from $\sum_{n=0}^{N}\pi_{n,0}+\sum_{n=1}^{N}\left(\pi_{n,1W}+\pi_{n,1B}+\pi_{n,2W}+\pi_{n,2B}\right)=1$ as

$$
\begin{aligned}
\pi_{0,N} &= 1/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n+w_n+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)+\\
&\quad k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n]).
\end{aligned}
$$

4. Performance Indices

Under this section various performance indices are presented. The expected system length ($E[L]$) is given by

$$
\begin{aligned}
E[L] &= \sum_{n=1}^{N} n\left(\pi_{n,0}+\pi_{n,1W}+\pi_{n,1B}+\pi_{n,2W}+\pi_{n,2B}\right)\\
&= \sum_{n=1}^{N} n\big((r_n+d_n+l_n+g_n+q_n)+k_{13}(s_n+y_n+p_n+\chi_n)+k_{14}(t_n+w_n+o_n+c_n)\\
&\quad+k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)\big)/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)
\end{aligned}
$$

$$+k_{14}(t_n+w_n+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n+k_{16}(\gamma_n+m_n+h_n+u_n)$$
$$+d_n+l_n+g_n+q_n+r_n)$$

The probabilities that the server is in vacation (P_v), busy with PS (P_{1w}), in breakdown state during PS (P_{1b}), busy with SS (P_{2w}) and the probability that the server is in breakdown state during SS (P_{2b}) are, respectively, given by

$$\begin{aligned}
P_v &= \sum_{n=0}^{N}\pi_{n,0},\\
&= \sum_{n=0}^{N} r_n/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n+w_n+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)\\
&\quad +k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n),\\
P_{1b} &= \sum_{n=1}^{N}\pi_{n,1b}\\
&= \sum_{n=1}^{N}(q_n+\chi_n k_{13}+c_n k_{14}+v_n k_{15}+u_n k_{16})/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n+w_n+\\
&\quad o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n),\\
P_{1w} &= \sum_{n=1}^{N}\pi_{n,1w}\\
&= \sum_{n=1}^{N}(g_n+p_n k_{13}+o_n k_{14}+f_n k_{15}+h_n k_{16})/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n+w_n\\
&\quad +o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n),\\
P_{2w} &= \sum_{n=1}^{N}\pi_{n,2w}\\
&= \sum_{n=1}^{N}(l_n+y_n k_{13}+w_n k_{14}+x_n k_{15}+m_n k_{16})/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n+w_n+\\
&\quad o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n),\\
P_{2b} &= \sum_{n=1}^{N}\pi_{n,2b}\\
&= \sum_{n=1}^{N}(d_n+s_n k_{13}+t_n k_{14}+z_n k_{15}+\gamma_n k_{16}))/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n+w_n\\
&\quad +o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n).
\end{aligned}$$

The expected balking rate (B_r), expected reneging rate (R_r) and the expected rate of losing a consumer (L_r) are given by

$$\begin{aligned}
B_r &= \sum_{n=1}^{N}\lambda(1-b_n)(\pi_{n,0}+\pi_{n,1W}+\pi_{n,1B}+\pi_{n,2W}+\pi_{n,2B})\\
&= \sum_{n=1}^{N}\lambda(1-b_n)\big((d_n+l_n+g_n+q_n)+k_{13}(s_n+y_n+p_n+\chi_n)+k_{14}(t_n+w_n+o_n+c_n)\\
&\quad +k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)\big)/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n\\
&\quad +w_n+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n),
\end{aligned}$$

$$
\begin{aligned}
R_r &= \sum_{n=1}^{N}(n-1)\alpha\left(\pi_{n,0}+\pi_{n,1W}+\pi_{n,1B}+\pi_{n,2W}+\pi_{n,2B}\right)\\
&= \sum_{n=1}^{N}(n-1)\alpha\big((d_n+l_n+g_n+q_n)+k_{13}(s_n+y_n+p_n+\chi_n)+k_{14}(t_n+w_n+o_n+c_n)\\
&\quad +k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)\big)/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)+k_{14}(t_n\\
&\quad +w_n+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n+q_n+r_n),\\
L_r &= B_r+R_r\\
&= \sum_{n=1}^{N}(\lambda(1-b_n)+(n-1)\alpha)\left((d_n+l_n+g_n+q_n)+k_{13}(s_n+y_n+p_n+\chi_n)+k_{14}(t_n+w_n\right.\\
&\quad \left.+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n)+k_{16}(\gamma_n+m_n+h_n+u_n)\right)/(r_0+\sum_{n=1}^{N}[k_{13}(s_n+y_n+\chi_n)\\
&\quad +k_{14}(t_n+w_n+o_n+c_n)+k_{15}(z_n+x_n+f_n+v_n+k_{16}(\gamma_n+m_n+h_n+u_n)+d_n+l_n+g_n\\
&\quad +q_n+r_n).
\end{aligned}
$$

5. Numerical Results

The impact of the various model parameters on the performance indices is presented in this section. The arbitrary choice of the model parameters for the purpose of numerical results is $N=10, \lambda=1.9, \omega=0.3, \alpha=0.5, \mu_1=2.9, \mu_2=2.5, \sigma=2.0, \beta=1.5, \delta=1.2$. Table 1 displays the steady-state probabilities for the above chosen set of parameters. The table also presents the corresponding performance measures $E[L], P_v, P_{1w}, P_{1b}, P_{2w}, P_{2b}, B_r, R_r$ and L_r.

Table 1: *Steady-state probability distributions*

n	$\pi_{n,0}$	$\pi_{n,1B}$	$\pi_{n,1W}$	$\pi_{n,2B}$	$\pi_{n,2W}$
0	0.043328	–	–	–	–
1	0.022608	0.027649	0.038451	0.015641	0.019548
2	0.010840	0.055023	0.052865	0.037740	0.036276
3	0.004760	0.064685	0.050222	0.050595	0.041409
4	0.001911	0.055825	0.037802	0.048113	0.035437
5	0.000702	0.038317	0.023580	0.035631	0.024374
6	0.000236	0.021832	0.012510	0.021605	0.013991
7	0.000073	0.010615	0.005749	0.011075	0.006871
8	0.000021	0.004492	0.002317	0.004912	0.002940
9	0.000000	0.001682	0.000819	0.001923	0.001112
10	0.000000	0.000572	0.000236	0.00068	0.000379
$E[L]=3.145050,\ P_v=0.084489,\ P_{1W}=0.224555,$ $P_{1b}=0.280694,\ P_{2w}=0.182339,\ P_{2b}=0.227923,$ $B_r=0.062196,\ R_r=1.095870,\ L_r=1.158070$					

The impact of arrival rate (λ) on $P_v, P_{1w}, P_{1b}, P_{2w}$ and P_{2b} is depicted in Figure 1. From the figure, it is evident that except P_v, all other values P_{1w}, P_{1b}, P_{2w} and P_{2b} increase with the increase of λ. The reason behind this nature is that as λ increases the number of consumers in the system increase due to which the server cannot leave for a vacation.

The changes in $P_v, P_{1w}, P_{1b}, P_{2w}$ and P_{2b} with β is shown in Figure 2. The probability of the server being in breakdown state both during PS and SS, P_{1b} and P_{2b}, respectively, increase with the increase of β which is obvious. Due to this the remaining probabilities P_v, P_{1w} and P_{2w} decrease with the growth of β as evident from the graph.

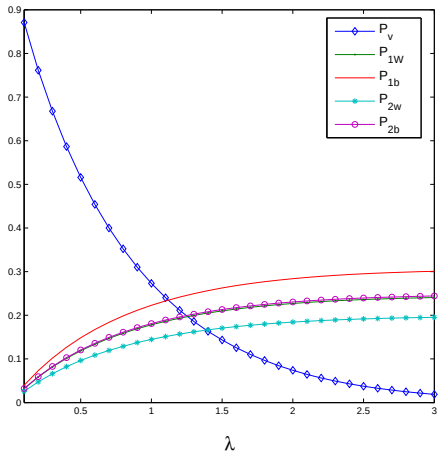


Figure 1: *Impact of* λ *on* $P_v, P_{1w}, P_{1b}, P_{2w}, P_{2b}$

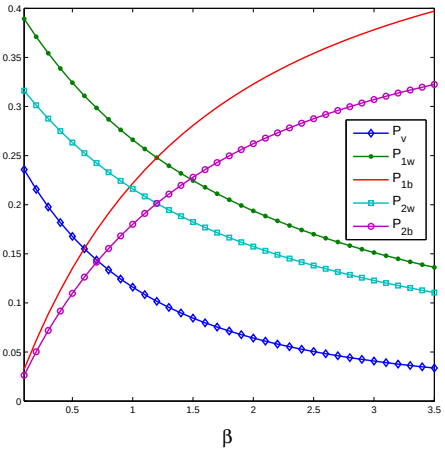


Figure 2: *Changes in* $P_v, P_{1w}, P_{1b}, P_{2w}, P_{2b}$ *with* β

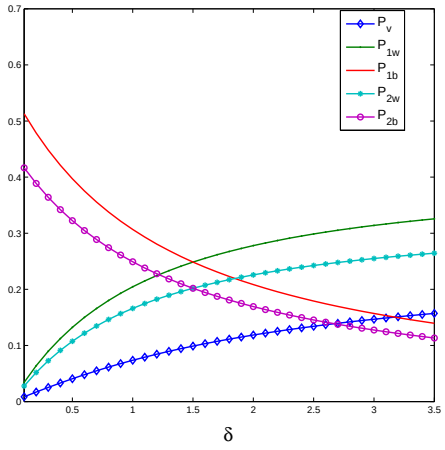


Figure 3: *Influence of* δ *on* $P_v, P_{1w}, P_{1b}, P_{2w}, P_{2b}$

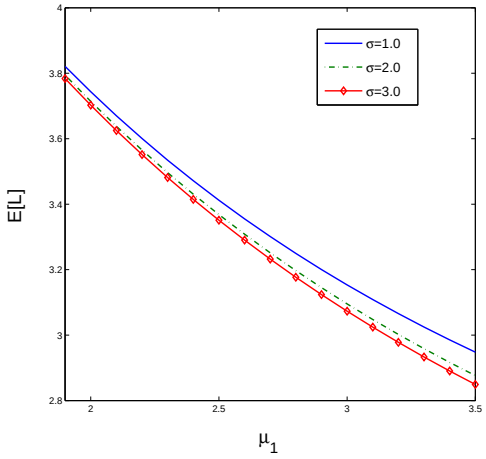


Figure 4: *Impact of μ_1 on $E[L]$ for different σ*

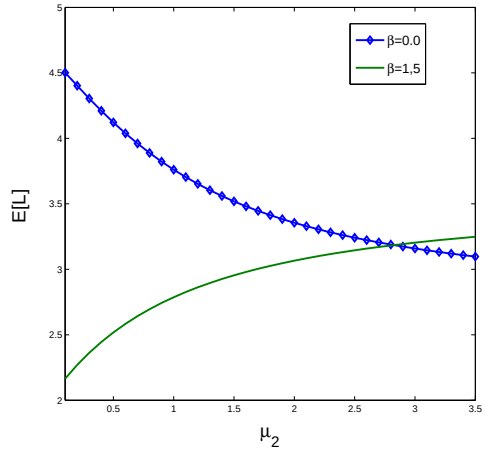


Figure 5: *Effect of μ_2 on $E[L]$ with and without server breakdowns*

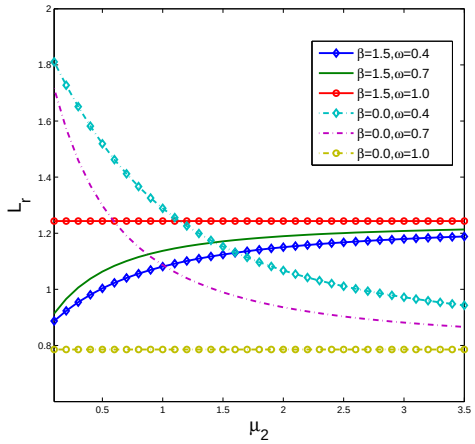


Figure 6: *Effect of μ_2 on L_r for different ω with and without server breakdowns*

The influence of δ on $P_v, P_{1w}, P_{1b}, P_{2w}$ and P_{2b} is displayed in Figure 3. In contrary to Figure 2, P_{1b} and P_{2b} decrease with the increase of δ while P_v, P_{1w} and P_{2w} increase with δ.

The effect of the service rate during PS (μ_1) on the expected system length ($E[L]$)is depicted in Figure 4 for different vacation rates σ. As evident from the figure, the expected system length decline with the increment of μ_1 for any choice of vacation rate (σ). Further, for a constant μ_1 value, $E[L]$ diminish with the growth in σ.

Figure 5 exhibits the effect of μ_2 on the expected system length ($E[L]$) in models with ($\beta = 1.5$) and without ($\beta = 0.0$) server breakdowns. In models without server breakdown, $E[L]$ diminishes with μ_2 as intuitively expected. However, this trend is reversed in models with server breakdown because there will no service during breakdown period leading to the increase in system length.

The impact of μ_2 on the expected rate of losing a consumer (L_r) in models with ($\beta = 1.5$) and without ($\beta = 0.0$) server breakdowns for different ω is revealed in Figure 6. It may be perceived that, for any ω ($\neq 1$), L_r drops with the growth of μ_2 in models without breakdown ($\beta = 0.0$) while L_r grows with the growth of μ_2 in models with breakdown ($\beta = 1.5$). Further, $\omega = 1.0$ implies that no client is opting for SS and hence, μ_2 has no impact on L_r as a result L_r remains the same for any μ_2 in both the models. Furthermore, with the increase of ω, the number of consumers leaving the system increase resulting in the decrease of L_r in models without breakdown while this trend gets reversed in models with breakdown.

6. Conclusions

An impatient consumer queue with secondary service, vacations and server breakdowns has been examined in this research. A wide range of real-time systems, including production and manufacturing systems, computer and communication networks, inventory and distribution systems, and others, may employ the described approach. The state-dependent nature of consumers' impatience in a secondary service queue with server breakdowns and vacations are the key contributions of this article. We used a recursive approach to obtain the steady-state probabilities. To illustrate the effect of the system parameters, numerical data is presented as table and graphs. This study might be expanded to include a renewal input impatient consumer queue with SS and working vacations and is a topic for further research.

References

[1] Altman, E. and Yechiali, U. (2006). Analysis of customers' impatience in queues with server vacations. *Queueing Systems*, 52:261-279.

[2] Amina Angelika Bouchentouf and Guendouzi, A. (2021). Single server batch arrival Bernoulli feedback queueing system with waiting server, K-variant vacations and impatient customers. Operations Research Forum, 2(1):Article 14.

[3] Ammar, S. (2017). Transient solution of an $M/M/1$ vacation queue with a waiting server and impatient customers. Journal of the Egyptian Mathematical Society, 25(3): 337-342.

[4] Bharathidass, S., Arivukkarasu, V. and Ganesan, V. (2018). Bulk service queue with server breakdown and repairs. *International Journal of Statistics and Applied Mathematics*, 3(1):136-142.

[5] Charan Jeet Singh and Sandeep Kaur (2018). $M^X/G/1$ queue with optional service and server breakdowns. *Performance Prediction and Analytics of Fuzzy, Reliability and Queuing Models*, Chapter:177-189.

[6] Charan Jeet Singh, Sandeep Kaur and Madhu Jain (2019). Analysis of bulk queue with additional optional service, vacation and unreliable server. *International Journal of Mathematics in Operational Research*, 14(4):517-540.

[7] Djamila Zirem, Mohamed Boualem and Djamil Aissani (2016). $M^x/G/1$ queueing system with breakdowns and repairs. Proceedings of the *International Conference on Advances in Information Processing and Communication Technology*, Chapter:62-65.

[8] Doshi, B. T. (1986). Queueing systems with vacations—a survey. *Queueing Systems*, 1(1): 29-66.

[9] Doshi, B. T. (1990). Single server queues with vacations. *Stochastic Analysis of Computer and Communication Systems*, H. Takag, Ed:217-265.

[10] Foad Iravani and Baris Balcioglu (2008). On priority queues with impatient customers. *Queueing Systems*, 58:239-260.

[11] Ganesh Sapkota and Ghimire, R. P. (2022). Mathematical analysis of $M/M/C$ vacation queueing model with a waiting server and impatient customers. *Journal of Mathematics and Statistics*, 18:36-48.

[12] Ghimire, R.P. and Ritu Basnet (2009). Finite capacity queueing system with vacations and server breakdowns. *IJE Transactions A: Basics*, 24(4):387-394.

[13] Hanumantha Rao, S. and Vasanta Kumar, V., Srinivasa Kumar, B. and Srinivasa Rao, T. (2017). Analysis of two-phase queuing system with impatient customers, server breakdowns and delayed repair. *International Journal of Pure and Applied Mathematics*, 115(4):651-663.

[14] Jamol Pender (2015). The truncated normal distribution: Applications to queues with impatient customers. *Operations Research Letters*, 43(1):40-45.

[15] Jau-Chuan Ke, Chia-Huang Wu and Zhe George Z. (2010). Recent developments in vacation queueing models: A short survey. *International Journal of Operations Research*, 7(4):3-8.

[16] Kalyanaraman, R. and Pazhani Bala Murugan, S. (2008). A single server queue with additional optional service in batches and server vacation. *Applied Mathematical Sciences*, 2(56):2765-2776.

[17] Khalid A., Lotfi T., (2018). Analysis of two phases queue with vacations and breakdowns under T-policy, *Encyclopedia of Information Science and Technology*, Fourth Edition:1570-1583.

[18] Panta, A. P. and Ghimire, R. P. Panthi, D. and Pant, S. R. (2021). A review of vacation queueing models in different framework. *Annals of Pure and Applied Mathematics*, 24(2):99-121.

[19] Sampath, M. I. G. S. and Liu, J. (2020). Impact of customers' impatience on an $M/M/1$ queueing system subject to differentiated vacations with a waiting server. *Quality Technology and Quantitative Management*, 17(2):125-148.

[20] Srinivas R. C., Shruti and Rakhee K. (2020). A queueing model with server breakdowns, repairs, vacations, and backup server. Operations Research Perspectives, 7:Article:100131.

[21] Takagi, H. (1991). Queueing Analysis: A Foundation of Performance Analysis, vol. 1 of Vacation and Priority Systems, part 1, Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

[22] Tian, N. and Zhang, Z. G. (2006). Vacation Queueing Models: Theory and Applications. Springer, New York, USA.

[23] Uma, S. and Punniyamoorthy, K. (2016). Single server bulk queue with second optional service, balking and compulsory vacation. *International Journal of Engineering Research and Application*, 6(10):15-20.

[24] Vijaya Laxmi, P. and Andwilile A. G. (2020). Transient analysis of batch service queue with second optional service and reneging. *International Research Journal on Advanced Science Hub*, 2(10):29-38.

[25] Vijaya Laxmi, P., Girija Bhavani, E. and Jyothsna, K. (2023). Analysis of Markovian queueing system with second optional service operating under the triadic policy. *OPSEARCH*, 60:256-275.

[26] William J. Gray, Pu Patrick Wang and Mec Kinley Scott (2000). A vacation queueing model with service breakdowns. *Applied Mathematical Modelling*, 24(5-6):391-400.

[27] Yue, D., Yue, W. and Zhao, G. Analysis of an $M/M/c$ queueing system with impatient customers and synchronous vacations. *Journal of Applied Mathematics*, Volume 2014:Article ID 893094.

MODELING OF RELIABILITY AND SURVIVAL DATA WITH EXPONENTIATED GENERALIZED INVERSE LOMAX DISTRIBUTION

Sule Omeiza Bashiru[1] and Ibrahim Ismaila Itopa[2]

•

[1,2]Department of Mathematical Sciences, Prince Abubakar Audu University, Anyigba, Kogi State, Nigeria

Email: [1]bash0140@gmail.com; [2]tripplei1975@yahoo.com

Abstract

In this paper, a new four-parameter distribution is developed and studied by combining the properties of the exponentiated generalized-G family of distributions and the features of the Inverse Lomax distribution. The newly developed distribution is called the exponentiated generalized inverse Lomax distribution that extends the classical inverse Lomax distribution. The shape of the hazard rate function is very flexible because it possesses increasing, decreasing, and inverted (upside-down) bathtub shapes. Some important characteristics of the exponentiated generalized inverse Lomax distribution are derived, including moments, moment generating function, survival function, hazard function and order statistics. The method of maximum likelihood estimation is used to obtain estimates of the unknown parameters of the new model. The application of the new model is based on two real-life data sets used to show the modeling potential of the proposed distribution. The exponentiated generalized inverse Lomax distribution turns out to be the best by capturing important details in the structure of the data sets considered.

Keywords: MLE, carbon fibers, non-linear, inverse Lomax, cancer

I. Introduction

Different distributions have been put out to extend existing distributions and act as models for numerous applications on data from various real-life scenarios. There are many ways to accomplish this, but the most popular one is to give the baseline distributions more flexibility by employing families of distributions.

The Lomax (Lx) distribution, also known as the Pareto type II distribution, is a special case of the generalized beta distribution of the second kind [1]. The distribution can be observed in many application areas, including actuarial science, economics, biological sciences, engineering, lifetime and reliability modeling, and so on [2]. According to [3], this distribution can be used as a substitute for survival issues and life-testing in engineering and survival analysis.

A significant lifetime distribution that can be used as a good substitute for well-known distributions like gamma, inverse Weibull, Weibull, and Lomax distributions is the Inverse Lomax (ILx) distribution. Because its hazard rate can be decreasing and upside-down bathtub shaped, it has a variety of uses in modeling diverse sorts of data, including economics and actuarial sciences data. The Inverse Lomax (ILx) distribution's unique characteristics and applications make it a valuable tool for statisticians, data analysts, and researchers dealing with a wide range of data sets.

A random variable X is said to have an ILx distribution if its cdf and pdf are given as

$$G(x)=\left[1+\frac{c}{x}\right]^{-d} \tag{1}$$

$$g(x)=cdx^{-2}\left[1+\frac{c}{x}\right]^{-d-1} \tag{2}$$

It was demonstrated in [4] that the ILx distribution is a member of the inverted family of distributions. They also discovered that the ILx distribution is a very flexible model in analyzing situations with a realized non-monotonic failure rate. For more details on the extensions of ILx distribution, readers are referred to [5], [6], [7], [8], [9], [10], [11], [12], [13], among others.

A new class of exponentiated generalized distribution that extends the exponentiated-G class was proposed by [14]. Given a continuous cdf $G(x)$, the cdf and pdf of exponentiated generalized class of distributions are given as

$$F(x)=\left[1-\left[1-G(x)\right]^{a}\right]^{b} \tag{3}$$

$$f(x)=abg(x)\left[1-G(x)\right]^{a-1}\left[1-\left[1-G(x)\right]^{a}\right]^{b-1} \tag{4}$$

In real-world applications, it is often impractical to generate data with a high failure rate, which can limit the applicability of the inverse Lomax distribution. Consequently, it becomes necessary to introduce additional flexibility into the inverse Lomax distribution to accommodate various hazard function shapes. The primary objective of this extension is to leverage the advantages offered by the exponentiated generalized-G class of distributions, as introduced by [14]. This extension is designed to enhance the modeling capabilities of the baseline distribution, ultimately leading to improved goodness-of-fit. When compared to the baseline model, the newly extended model demonstrates a wider range of hazard function shapes and provides superior fits for diverse types of datasets.

II. Methods

2.1 Exponentiated Generalized Inverse Lomax (EGILx) Distribution

This section introduces a novel continuous probability distribution function known as the Exponentiated Generalized Inverse Lomax (EGILx) distribution. It includes plots of its probability density function (pdf) and cumulative distribution function (cdf) to evaluate the characteristics of this new distribution. By substituting (1) into (3) and (2) into (4), we obtain the cdf and pdf of the EGILx distribution, respectively, as follows:

$$F(x)=\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b} \tag{5}$$

$$f(x)=abcdx^{-2}\left[1+\frac{c}{x}\right]^{-d-1}\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a-1}\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b-1} \tag{6}$$

Where $x\geq 0$, $c>0$ is the scale parameter and $a,b,d>0$ are the shape parameters.

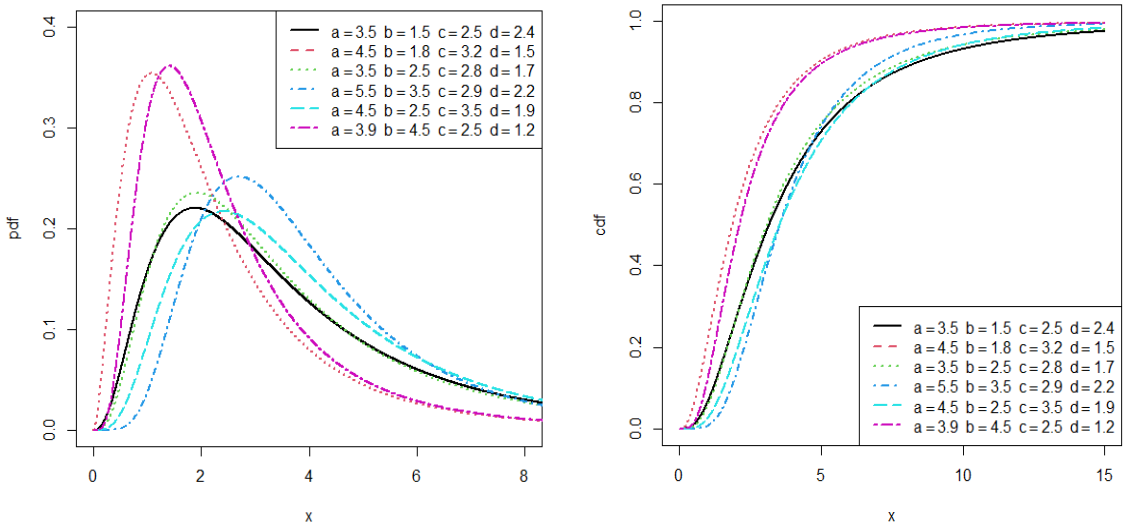


Figure 1: *Plots of pdf and cdf of EGILx distribution*

2.2 Expansion of Density

Using the generalized bionomial expansion

$$(1-x)^{p-1}=\sum_{i=0}^{\infty}\frac{(-1)^i\Gamma(p)}{i!\Gamma(p-i)}x^i \tag{7}$$

The pdf of EGILx distribution in (6) can be expressed in mixture form in terms of Lx densities as

$$f(x)=abcdx^{-2}\left[1+\frac{c}{x}\right]^{-d-1}\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a-1}\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b-1}$$

$$\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b-1}=\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{ai}$$

$$\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a(i+1)-1}=\sum_{j=0}^{\infty}(-1)^j\binom{a(i+1)-1}{j}\left[1+\frac{c}{x}\right]^{-dj}$$

$$f(x)=abcd\sum_{i,j=0}^{\infty}(-1)^{i+j}\binom{b-1}{i}\binom{a(i+1)-1}{j}x^{-2}\left[1+\frac{c}{x}\right]^{-d(j+1)-1} \tag{8}$$

where

$$\int_0^{\infty}x^{r-2}\left[1+\frac{c}{x}\right]^{-d(j+1)-1}dx=c^rB\left[(1-r),d(j+1)+r\right]$$

On expanding the cdf, we have

$$\left[F(x)\right]^h=\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{bh}$$

$$\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{bh}=\sum_{k=0}^{h}(-1)^{k}\binom{bh}{k}\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{ak}$$

$$\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{ak}=\sum_{w=0}^{\infty}(-1)^{w}\binom{ak}{w}\left[1+\frac{c}{x}\right]^{-dw}$$

$$\left[F(x)\right]^{h}=\sum_{k=0}^{h}\sum_{w=1}^{\infty}(-1)^{k+w}\binom{ak}{w}\binom{bh}{k}\left[1+\frac{c}{x}\right]^{-dw} \tag{9}$$

2.3 Properties of EGILx Distribution

This section derives some statistical properties of the EGILx distribution including moments, moment generating function, survival function, hazard function, quantile functions, and order statistics.

2.3.1 Moments

Using moments, some of the most important features and characteristics of a distribution such as tendency, dispersion, skewness and kurtosis can be studied. The rth ordinary moment of the EGILx distribution can be written from (8) as

$$E\left(X^{r}\right)=\int_{0}^{\infty}x^{r}f\left(x\right)dx \tag{10}$$

$$=abcd\sum_{i,j=0}^{\infty}(-1)^{i+j}\binom{b-1}{i}\binom{a(i+1)-1}{j}\int_{0}^{\infty}x^{r-2}\left[1+\frac{c}{x}\right]^{-d(j+1)-1}dx$$

$$\int_{0}^{\infty}x^{r-2}\left[1+\frac{c}{x}\right]^{-d(j+1)-1}dx=c^{r}B\left[(1-r),d(j+1)+r\right]$$

$$E\left(X^{r}\right)=abc^{r}d\sum_{i,j=0}^{\infty}(-1)^{i+j}\binom{b-1}{i}\binom{a(i+1)-1}{j}B\left[(1-r),d(j+1)+r\right] \tag{11}$$

the mean of the EGILx distribution can be obtained by setting r = 1 in (11)

2.3.2 Moment generating function (mgf)

Following the process of moments, the mgf is obtained as

$$M_{x}(t)=abd\sum_{m=0}^{\infty}\sum_{i,j=0}^{\infty}(-1)^{i+j}\binom{b-1}{i}\binom{a(i+1)-1}{j}\frac{(tc)^{m}}{m!}B\left[(1-m),d(j+1)+m\right] \tag{12}$$

2.3.3 Reliability function (rf)

$$\mathrm{R}(x)=1-\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b} \tag{13}$$

2.3.4 Hazard Function (hf)

$$H(x)=\frac{abcdx^{-2}\left[1+\frac{c}{x}\right]^{-d-1}\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a-1}\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b-1}}{1-\left[1-\left[1-\left[1+\frac{c}{x}\right]^{-d}\right]^{a}\right]^{b}} \tag{14}$$

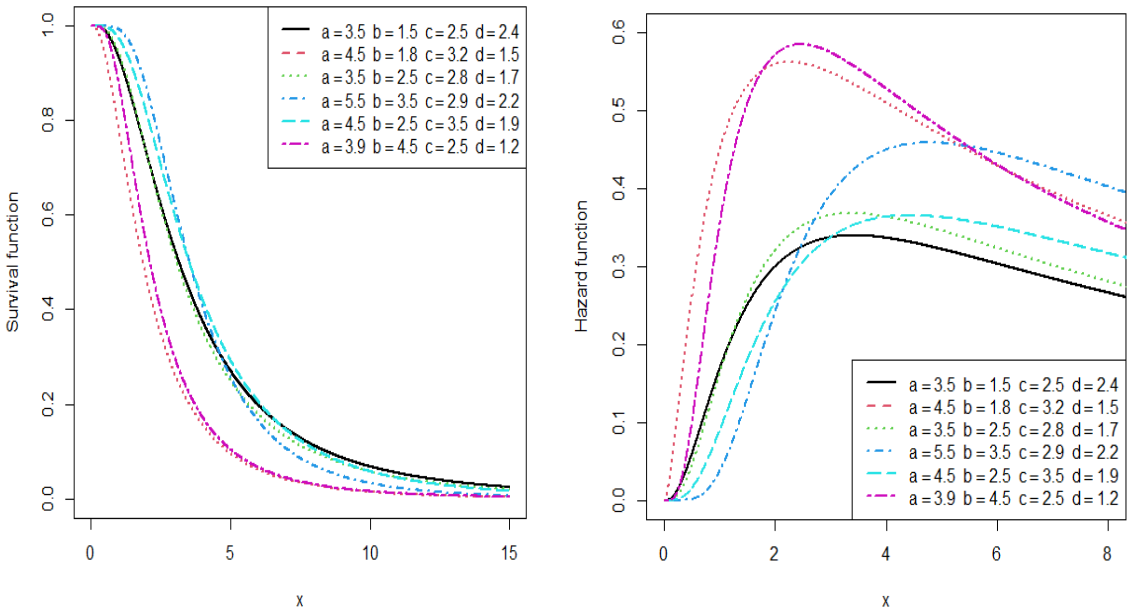


Figure 2: *Plots of survival function and hazard function of EGILx distribution*

2.3.5 Quantile Function (qf)

The quantile function (qf) of a probability distribution is the real solution of the equation $F(x)=u$, for $0\le u\le 1$, and $F(x)$ is the cumulative distribution function (CDF) of the random variable x
For EGILx distribution the quantiles are given by

$$x=Q(u)=\frac{c}{1-\left[1-\left[1-u^{\frac{1}{b}}\right]^{\frac{1}{a}}\right]^{\frac{-1}{d}}} \tag{15}$$

On setting $u=0.5$ in (15), the median of the EGILx distribution can be obtained.

2.3.6 Order Statistics

Let $X_1, X_2, ...X_n$ be n independent random variable from the EGILx distribution and let $X_{(1)}, X_{(2)}, ...X_{(n)}$ be their corresponding order statistic. Let $F_{r:n}(x)$ and $f_{r:n}(x)$, $r = 1,2,3,...,n$ denote the cdf and pdf of the r^{th} order statistics $X_{r:n}$ respectively. The pdf of the r^{th} order statistics of $X_{r:n}$ is given as

$$f_{r:n}(x) = \frac{f(x)}{B(r, n-r+1)} \sum_{v=0}^{n-r} (-1)^v \binom{n-r}{v} F(x)^{v+r-1} \tag{16}$$

The pdf of rth order statistic for distribution is obtained by replacing h with v+r-1 in cdf expansion. We have

$$f_{r:n}(x) = abcdx^{-2} \frac{1}{B(r, n-r+1)} \sum_{v=0}^{n-r} \sum_{i,j,w=0}^{\infty} \sum_{k=0}^{v+r-1} (-1)^{i+j+w+k+v} \binom{n-r}{v} \binom{b-1}{i} \binom{a(i+1)-1}{j} \binom{ak}{w} \binom{b(v+r-1)}{k} \left[1+\frac{c}{x}\right]^{-d(j+1+w)-1} \tag{17}$$

The pdf of the minimum order statistic of the distribution is obtained by setting r=1 in (17) as

$$f_{1:n}(x) = abcdx^{-2} n \sum_{v=0}^{n-1} \sum_{i,j,w=0}^{\infty} \sum_{k=0}^{v} (-1)^{i+j+w+k+v} \binom{n-1}{v} \binom{b-1}{i} \binom{a(i+1)-1}{j} \binom{ak}{w} \binom{bv}{k} \left[1+\frac{c}{x}\right]^{-d(j+1+w)-1} \tag{18}$$

Also, the pdf of the maximum order statistic of the distribution is obtained by setting r = n in (17) as

$$f_{n:n}(x) = abcdx^{-2} n \sum_{k=0}^{v+n-1} \sum_{i,j,w=0}^{\infty} (-1)^{i+j+w+k+v} \binom{b-1}{i} \binom{a(i+1)-1}{j} \binom{ak}{w} \binom{b(v+n-1)}{k} \left[1+\frac{c}{x}\right]^{-d(j+1+w)-1} \tag{19}$$

2.4 Maximum Likelihood Estimation (MLE)

In this section, the estimation of the unknown parameters for the EGILx distribtuion by MLE method. Let $X_1, X_2, ...X_n$ be random variables of the EGILx distribution of size n. Then the sample log-likelihood function of the EGILx distribution is obtained as

$$\begin{aligned} \log(L) = {} & n\log(a) + n\log(b) + n\log(c) + n\log(d) - 2\sum_{i=1}^{n} \log(x_i) - (d+1)\sum_{i=1}^{n} \log\left[1+\frac{c}{x_i}\right] \\ & + (a-1)\sum_{i=1}^{n} \log\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right] + (b-1)\sum_{i=1}^{n} \log\left[1-\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a}\right] \end{aligned} \tag{20}$$

Differentiating (20) with respect to each parameter and equationg them zero yields the MLE as

$$\frac{\partial L}{\partial a}=\frac{n}{a}+\sum_{i=1}^{n}\log\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]-(b-1)\sum_{i=1}^{n}\frac{\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a}\log\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]}{1-\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a}}=0 \tag{21}$$

$$\frac{\partial L}{\partial b}=\frac{n}{b}+\sum_{i=1}^{n}\log\left[1-\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a}\right]=0 \tag{22}$$

$$\frac{\partial L}{\partial c}=\frac{n}{c}-(d+1)\sum_{i=1}^{n}\frac{1}{\left[1+\frac{c}{x_i}\right]x_i}+d(a-1)\sum_{i=1}^{n}\frac{\left[1+\frac{c}{x_i}\right]^{-d-1}}{\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]x_i}+ad(b-1)\sum_{i=1}^{n}\frac{\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a-1}\left[1+\frac{c}{x_i}\right]^{-d-1}}{\left[1-\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a}\right]x_i}=0 \tag{23}$$

$$\frac{\partial L}{\partial d}=\frac{n}{d}-\sum_{i=1}^{n}\log\left[1+\frac{c}{x_i}\right]+(a-1)\sum_{i=1}^{n}\frac{\left[1+\frac{c}{x_i}\right]^{-d}\log\left[1+\frac{c}{x_i}\right]}{1-\left[1+\frac{c}{x_i}\right]^{-d}}+a(b-1)\sum_{i=1}^{n}\frac{\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a-1}\left[1+\frac{c}{x_i}\right]^{-d}\log\left[1+\frac{c}{x_i}\right]}{1-\left[1-\left[1+\frac{c}{x_i}\right]^{-d}\right]^{a}}=0 \tag{24}$$

Equations (21) to (24) lack a straightforward analytical representation, rendering them intractable. Consequently, we must employ non-linear parameter estimation techniques through iterative methods.

III. Results

3.1 Applications

In this section, we provide two applications to real-life data sets to illustrate the importance and flexibility of the EGILx distribution compared to some classical distributions such as inverse Weibull (IW), inverse Lomax (ILx) and Lomax (Lx) distributions.

The first data set represents the breaking stress of carbon fibers of 50 mm length (GPa) was reported by [15]. This data was used by [16]. The data are: 0.39, 0.85, 1.08, 1.25, 1.47, 1.57, 1.61, 1.61, 1.69, 1.80, 1.84, 1.87, 1.89, 2.03, 2.03, 2.05, 2.12, 2.35, 2.41, 2.43, 2.48, 2.50, 2.53, 2.55, 2.55, 2.56, 2.59, 2.67, 2.73, 2.74, 2.79, 2.81, 2.82, 2.85, 2.87, 2.88, 2.93, 2.95, 2.96, 2.97, 3.09, 3.11, 3.11, 3.15, 3.15, 3.19, 3.22, 3.22, 3.27, 3.28, 3.31, 3.31, 3.33, 3.39, 3.39, 3.56, 3.60, 3.65, 3.68, 3.70, 3.75, 4.20, 4.38, 4.42, 4.70, 4.90.

The second data set was given by [17] and it represents the remission times (in months) of a random sample of one hundred and twenty-eight (128) bladder cancer patients. The data set is as follows: 0.08, 2.09, 3.48, 4.87, 6.94 , 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46 , 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

Table 1: *The models' MLEs and performance requirements based on data set 1*

Models	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$
EGILx	68.2326	1.2151	4.6461	4.2988
IW	-	-	2.0356	1.6479
Lx	-	-	0.2884	1.5594
ILx	-	-	0.0139	164.8019

Table 2: *The models' MLEs and performance requirements based on data set 2*

Models	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$
EGILx	2.2598	63.3711	21.6622	0.0938
IW	-	-	3.2589	0.7522
Lx	-	-	0.0083	13.9032
ILx	-	-	2.0033	2.4610

Table 3: *The Performance requirements based on data set 1*

Models	ll	AIC	AICc	BIC	HQIC
EGILx	-89.2197	186.4394	187.951	195.1900	189.9003
IW	-121.1949	246.3898	246.5803	250.7697	248.1203
Lx	-149.8545	303.7089	303.8994	308.0882	305.43.94
ILx	-136.1274	276.2548	276.4453	280.6342	277.9853

Table 4: *The Performance requirements based on data set 2*

Models	ll	AIC	AICc	BIC	HQIC
EGILx	-414.2379	836.4759	836.8017	847.8840	841.1111
IW	-444.0008	892.0015	892.0975	897.7056	894.3191
Lx	-416.8329	837.6658	837.7618	848.0126	842.0654
ILx	-424.6757	853.3514	853.4474	859.0555	855.6690

IV. Discussion

In this paper, we introduce a novel four-parameter distribution called the Exponentiated Generalized Inverse Lomax distribution and conduct a comprehensive study of its properties. We delve into various mathematical and statistical aspects of this newly developed model, including moments, moment generating functions, reliability functions, quantile functions, and order statistics. Furthermore, we investigate and present the probability density functions of both the maximum and minimum order statistics. To estimate the unknown parameters of this distribution, we employ the maximum likelihood estimation method, which allows us to determine their values effectively. To demonstrate the practical utility of our proposed model, we provide insights into its performance by applying it to two real-life datasets. Our analysis reveals that the Exponentiated Generalized Inverse Lomax distribution outperforms other lifetime models considered in this study when applied to the provided datasets, underscoring its potential for improved modeling and prediction in real-world applications.

References

[1] Kleiber, C. and Kotz, S. (2003). Statistical size distributions in economics and actuarial sciences. *John Wiley, Sons Inc*, Hoboken, New Jersey.

[2] Al-Zaharani, B. and Al-Sobhi, M. (2013). On parameters estimation of Lomax distribution under general progressive censoring. *Journal of Quality and Survival Engineering* 2013: 1-7.

[3] Hassan, A.S. and Al-Ghamdi, A.S. (2009). Optimum step stress accelerated life testing for Lomax distribution. *Journal of Applied Sciences Research*, 5: 2153-2164.

[4] Singh, S.K., Singh, U. and Kumar, D. (2012). Bayes estimators of the survival function and parameters of inverted exponential distribution using informative and non-informative priors. *Journal of Statistical computation and simulation*, 83(12): 2258-2269.

[5] Ogunde, A.A., Chukwu A.U. and Oseghale I.O. (2023). The Kumaraswamy Generalized Inverse Lomax distribution and applications to reliability and survival data. *Scientific African*, 19.

[6] Maxwell, O., Kayode, A.A., Onyedikachi, I.P., Obi-Okpala, I. and Victor, E.U. (2019). Useful generalization of the inverse Lomax distribution: Statistical Properties and Application to Lifetime Data. *America Journal of Biomedical Science & Research,* 6(3): 258-265.

[7] Hassan, A.S., Al-Omar, A.I., Ismail, D.M. and Al-Anzi, A. (2021). A new generalization of the inverse Lomax distribution with statistical properties and applications. *International Journal of Advanced and Applied Sciences,* 8(4): 89-97.

[8] Hassan, A. S. and Abd-Allah, M. (2019). On the inverse power Lomax distribution. *Annals of Data Science*, *6(2)*, 259-278.

[9] Hassan, A. S., and Mohamed, R. E. (2019). Weibull inverse lomax distribution. *Pakistan Journal of Statistics and Operation Research*, *15(3)*, 587-603.

[10] Maxwell, O., Chukwu, A. U., Oyamakin, O. S. and Khaleel, M. (2019). The Marshall–Olkin inverse lomax distribution (MO-ILD) with application on cancer stem cell. *Journal of Advances in Mathematics and Computer Science*, 33(4), 1-12.

[11] ZeinEldin, R.A., Haq, M. A., Hashmi, S. and Elsehety, M. (2020). Alpha power trans- formed inverse lomax distribution with different methods of estimation and applications. *Complexity*, *2020(1)*, 1-15.

[12] Almarashi, A.M. (2021). A New Modified Inverse Lomax Distribution: Properties, Estimation and Applications to Engineering and Medical Data. *Computer Modeling in Engineering & Sciences,* 127(2):621-643.

[13] Kumar, D., Yadav, A.S., Kumar, P., Kumar, P., Singh, S.K. and Singh, U. (2021). Transmuted Inverse Lomax Distribution and its Properties. *International Journal of Agricultural and Statistical Sciences,* 17(1):1-8.

[14] Cordeiro, G.M., Ortega, E.M., and da Cunha, D. C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, 11(1), 1-27.

[15] Nicholas, M.D. and Padgett W.J. (2006). A bootstrap control chart for weibull percentiles. *Quality and Reliability Engineering International*, 22:141-151.

[16] Yousof, H.M., Alizadeh, M., Jahanshahi, S.M.A., Ramires, T.G., Ghosh, I. and Hamedani G.G. (2017). The transmuted Topp-leone g family of distributions: Theory, characterizations and applications. *Journal of Data Science*, 15:723-740.

[17] Lee, E.T. and Wang, J. W. (2003). Statistical methods for survival data analysis (3rd Edition), *John Wiley and Sons,* New York, USA, 535 Pages, ISBN 0-471-36997-7.

Reliability Analysis of the Shaft Subjected to Twisting Moment and Bending Moment for Normally Distributed Strength and Stress

Md. Yakoob Pasha[1], M. Tirumala Devi[2], T. Sumathi Uma Maheswari[3].

•

[1, 2, 3]Department of Mathematics
Kakatiya University, Warangal-506009, Telangana, India.
[1]mypasha139@gmail.com, [2]oramdevi@yahoo.com, [3]tsumathiuma@gmail.com

Abstract

Shaft is the rotating component that transmits power from one place to another. The shafts are commonly subject to torsional and bending moments and combinations of these moments. In general, shafts are subjected to a combination of torsional and bending stresses. The design of a shaft is essential, subject to its strength and stress. This paper presents the ***reliability analysis of the shaft*** *subjected to (a) twisting moment, (b) bending moment and (c) combined twisting and bending moment for which* ***stress and strength follow the normal distribution****.*

Keywords: Reliability, normal distribution, twisting moment, bending moment, round solid shaft, maximum shear stress theory, maximum normal stress theory.

1. Introduction

Shaft is the most vital component used in almost every mechanical system and machine. Out of all power transmission components, the shaft is the main component that must be designed carefully for the efficient working of the machine. The shafts are designed based on their strength and rigidity considerations. Based strength includes shafts subjected to bending moments, twisting moments, combined bending and twisting moments, and fluctuating loads. While considering rigidity, torsional and lateral rigidity are considered for design.

Adekunle A. A. et al. [1] studied the development of CAD software for shafts under various loading conditions. Dr. Edward E. Osakue et al. [2] studied fatigue shaft design verification for bending and torsion. Dr. Edward E. Osakue et al. [3] studied the probabilistic fatigue design of shafts for bending and torsion. Frydrysek K. et al. [4] studied performance-based design applied to a shaft subjected to combined stress. Gowtham et al. [5] studied drive shaft design and analysis. Kamboh, M. S., et al. [7] discussed the design and analysis of the drive shaft with a critical review of advanced composite materials and the root causes of shaft failure. Kececioglu D. et al. [8] studied the reliability analysis of mechanical components and systems. Misra, A., et al. [9] studied the reliability analysis of drilled shaft behaviour using the finite difference method and Monte Carlo simulation. Munoz-Abella, B., et al. [10] discussed the determination of the critical speed of a cracked shaft from experimental data. Nayek et al. [11]

studied the reliability approximation for a solid shaft under a gamma setup. Patel, B., et al. [12] studied a critical review of the design of a shaft with multiple discontinuities and combined loadings. Villa-Covarrubias, B., et al. [14] discussed the probabilistic method to determine the shaft diameter and design reliability.

2. Statistical model

The normal distribution takes the well-known bell shape. This distribution is symmetrical about its mean value. **The probability density function** for a normally distributed stress χ and normally distributed strength ξ is given by [4]

$$f_{\chi}(\chi) = \frac{1}{\sigma_{\chi}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\chi-\mu_{\chi}}{\sigma_{\chi}}\right)^2\right] \text{ for } -\infty < \chi < \infty$$

$$f_{\xi}(\xi) = \frac{1}{\sigma_{\xi}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi-\mu_{\xi}}{\sigma_{\xi}}\right)^2\right] \text{ for } -\infty < \xi < \infty$$

where μ_{χ} = mean value of the stress

σ_{χ} = standard deviation of the stress

μ_{ξ} = mean value for the strength

σ_{ξ} = standard deviation of the strength

Let us define $y = \xi - \chi$. It is well known that the random variable y is normally distributed with mean of $\mu_y = \mu_{\xi} - \mu_{\chi}$ and standard deviation of $\sigma_y = \sqrt{\sigma_{\xi}^2 + \sigma_{\chi}^2}$.

The reliability R can be expressed in terms of y as

$$R = P(y > 0) = \frac{1}{\sigma_y\sqrt{2\pi}} \int_0^{\infty} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right] dy$$

If we let $z = (y - \mu_y)/\sigma_y$, then $\sigma_y\, dz = dy$.

When $y = 0$, the lower limit of z is given by [4]

$$z = \frac{0-\mu_y}{\sigma_y} = -\frac{(\mu_{\xi}-\mu_{\chi})}{\sqrt{\sigma_{\xi}^2+\sigma_{\chi}^2}}$$

and $y \to +\infty$, the upper limit of $z \to +\infty$.

Then the reliability is given by [4]

$$R = \frac{1}{2\pi} \int_{-\frac{(\mu_{\xi}-\mu_{\chi})}{\sqrt{\sigma_{\xi}^2+\sigma_{\chi}^2}}}^{\infty} \exp\left(\frac{-z^2}{2}\right) dz \tag{1}$$

The random variable $z = \frac{(y-\mu_y)}{\sigma_y}$ is the standard normal variable.

a. Shaft subjected to twisting moment only

If the shaft is subjected to a twisting moment only, then torsion equation is given by [6]

$$\frac{T}{J} = \frac{\sigma_t}{r}$$

where T = Twisting moment acting upon the shaft

$J =$ Polar moment of inertia of the shaft about the axis of rotation

$\sigma_t =$ Torsional shear stress and

$r =$ Distance from neutral axis to the outer-most fibre

$= \frac{d}{2}$ where d is the diameter of the shaft

For round solid shaft, polar moment of inertia is given by [6]

$$J = \frac{\pi}{32} \times d^4$$

From the torsion equation twisting moment is

$$T = \frac{\pi}{16} \times \sigma_t \times d^3 \text{ and}$$

The shear stress due to twisting moment is the normally distributed stress

$$\mu_\chi = \sigma_t = \frac{16 \times T}{\pi \times d^3}$$

Then the reliability of the shaft subjected to twisting moment for normally distributed strength and stress is

$$R_t = \frac{1}{2\pi} \int_{-\frac{\left(\mu_\xi - \frac{16\times T}{\pi \times d^3}\right)}{\sqrt{\sigma_\xi^2 + \sigma_\chi^2}}}^{\infty} \exp\left(\frac{-z_t^2}{2}\right) dz \tag{2}$$

b. Shaft subjected to bending moment only

If the shaft is subjected to a bending moment only, then bending equation is given by [6]

$$\frac{M}{I} = \frac{\sigma_b}{y}$$

where $M =$ Bending moment

$I =$ Moment of inertia of cross-sectional area of the shaft about the axis of rotation

$\sigma_b =$ Bending stress and

$y =$ Distance from neutral axis to the outer-most fibre

For solid shaft, moment of inertia is given by [6]

$$I = \frac{\pi}{64} \times d^4 \text{ and } y = \frac{d}{2}$$

From the bending equation bending moment is given by [8]

$$M = \frac{\pi}{32} \times \sigma_b \times d^3$$

The shear stress due to bending moment is the normally distributed stress

$$\mu_\chi = \sigma_\mathrm{b} = \frac{32 \times M}{\pi \times d^3}$$

Then the reliability of the shaft subjected to bending moment for normally distributed strength and stress is

$$R_b = \frac{1}{2\pi} \int_{-\frac{\left(\mu_\xi - \frac{32\times M}{\pi \times d^3}\right)}{\sqrt{\sigma_\xi^2 + \sigma_\chi^2}}}^{\infty} \exp\left(\frac{-z_b^2}{2}\right) dz \tag{3}$$

c. Shaft subjected to combined twisting moment and bending moment

When the shaft is subjected to combined twisting moment and bending moment, then the shaft must be designed on the basis of the two moments simultaneously.

According to **maximum shear stress theory**, the maximum shear stress in the shaft is

$$\sigma_{t_{max}} = \frac{1}{2}\sqrt{(\sigma_b)^2 + 4(\sigma_t)^2}$$

where σ_t = shear stress induced due to twisting moment

σ_b = bending stress induced due to bending moment

Therefore

$$\sigma_{t_{max}} = \frac{16}{\pi \times d^3}\sqrt{M^2 + T^2}$$

The expression $\sqrt{M^2 + T^2}$ is known as equivalent twisting moment and is denoted by T_e.

The equivalent twisting moment may be defined as that twisting moment, which when acting alone, produces same shear stress (σ_t) as the actual twisting moment.

By limiting the maximum shear stress ($\sigma_{t_{max}}$) equal to the allowable shear stress (σ_{ts}) for the material is the normally distributed stress

$$\mu_{\chi s} = \sigma_{ts} = \frac{16 \times T_e}{\pi \times d^3}$$

Then the reliability of the shaft subjected to combined twisting moment and bending moment for the normally distributed strength and stress is

$$R_{ts} = \frac{1}{2\pi} \int_{-\frac{\left(\mu_\xi - \frac{16\times T_e}{\pi \times d^3}\right)}{\sqrt{\sigma_\xi^2 + \sigma_\chi^2}}}^{\infty} \exp\left(\frac{-z_{ts}^2}{2}\right) dz \tag{4}$$

According to **maximum normal stress theory**, the maximum normal stress in the shaft is

$$\sigma_{b_{max}} = \frac{1}{2}\sigma_b + \frac{1}{2}\sqrt{(\sigma_b)^2 + 4(\sigma_t)^2}$$

where σ_t = shear stress induced due to twisting moment

σ_b = bending stress induced due to bending moment

Therefore

$$\sigma_{b_{max}} = \frac{32}{\pi \times d^3}\left[\frac{1}{2}\left(M + \sqrt{M^2 + T^2}\right)\right]$$

The expression $\frac{1}{2}\left(M + \sqrt{M^2 + T^2}\right)$ is known as equivalent bending moment and is denoted by M_e.

The equivalent bending moment may be defined as that moment which when acting alone produces the same tensile or compressive stress (σ_b) as the actual bending moment.

By limiting the maximum normal stress ($\sigma_{b_{max}}$) equal to the allowable bending stress (σ_{bn}) for the material is the normally distributed stress

$$\mu_{\chi n} = \sigma_{bn} = \frac{32\times M_e}{\pi \times d^3}$$

Then the reliability of the shaft subjected to combined twisting moment and bending moment for normally distributed strength and stress is

$$R_{bn} = \frac{1}{2\pi} \int_{-\frac{\left(\mu_\xi - \frac{32\times M_e}{\pi \times d^3}\right)}{\sqrt{\sigma_\xi^2 + \sigma_\chi^2}}}^{\infty} \exp\left(\frac{-z_{bn}^2}{2}\right) dz \tag{5}$$

3. Numerical results and discussion

a. For twisting moment only

Table-1: $d = 110\ mm$, $\mu_\xi = 119.6584\ N/mm^2$.

T	z_t	R_t
100000	-1.993596824	0.976901934
2500000	-1.834327549	0.966697306
5000000	-1.659291647	0.951471481
7500000	-1.478628009	0.930380119
10000000	-1.296096181	0.902528825
13000000	-1.079424515	0.859800736
18000000	-0.736112792	0.769168971
20000000	-0.607809459	0.728343073
25000000	-0.313849897	0.623182477
30000000	-0.059258971	0.523627080

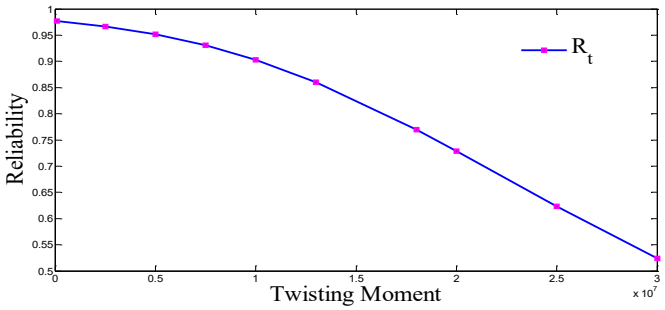


Figure 1: *Reliability and Twisting Moment*

Table-2: $T = 10^4\ N\text{-}mm$, $\mu_\xi = 119.6584\ N/mm^2$.

d	z_t	R_t
18	-0.437034097	0.668956690
20	-0.826725375	0.795803632
24	-1.323221625	0.907119157
30	-1.664313463	0.951975098
32	-1.725837528	0.957811677
40	-1.862933958	0.968764221
50	-1.930809365	0.973246684
58	-1.955923837	0.974762937
70	-1.975040282	0.975868212

82	-1.984508267	0.976400396

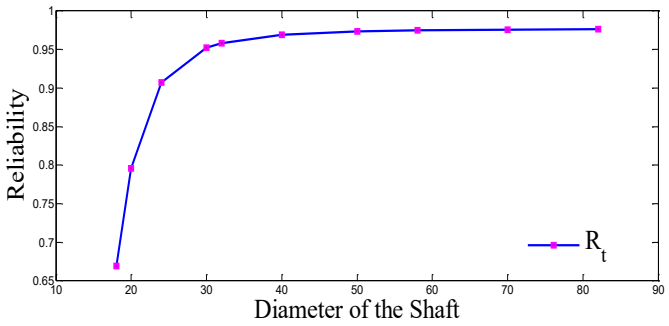


Figure 2: *Reliability and Diameter of the Shaft*

b. For bending moment only

Table-3: $d = 110\ mm$, $\mu_\xi = 119.6584\ N/mm^2$.

M	z_b	R_b
100000	-1.98717341	0.976548408
1200000	-1.841158255	0.967200815
2000000	-1.730194966	0.958202276
2400000	-1.673552215	0.952890682
3800000	-1.471338225	0.929400165
4800000	-1.325278842	0.907460658
6500000	-1.079424515	0.859800736
7800000	-0.897319098	0.815225666
8600000	-0.789017221	0.784949029
12000000	-0.369480494	0.644115195

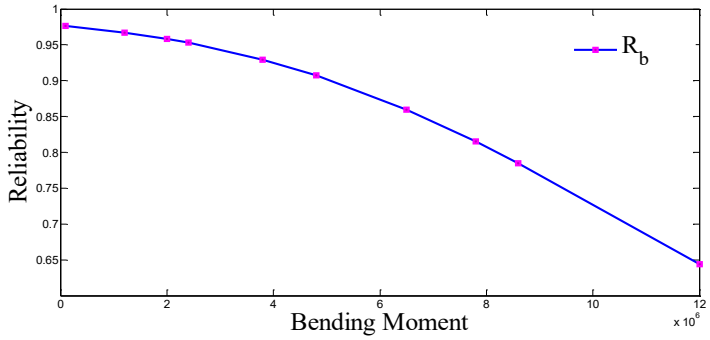


Figure 3: *Reliability and Bending Moment*

Table-4: $M = 10^4$ *N-mm*, $\mu_\xi = 119.6584 N/mm^2$.

d	z_b	R_b
22	-0.313849897	0.623182477
26	-0.928635716	0.823461047
32	-1.433116930	0.924087788
36	-1.608706553	0.946159739
40	-1.718965160	0.957189642
52	-1.875536297	0.969640510
62	-1.927365578	0.973032957
82	-1.968900245	0.975517726
100	-1.982910074	0.976311262
110	-1.987173410	0.976548408

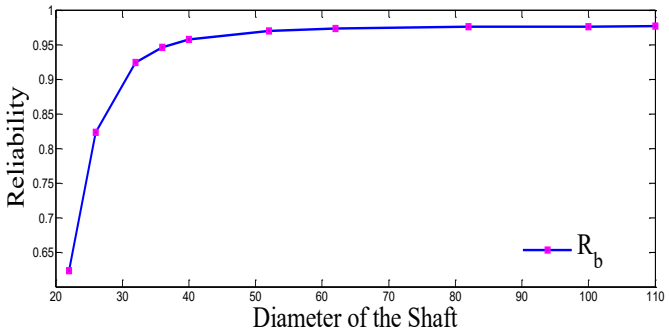


Figure 4: *Reliability and Diameter of the Shaft*

c. For combined twisting moment and bending moment

Table-5: $M = 10^4$ *N-mm*, $d = 110$ *mm*, $\mu_\xi = 119.6584$ N/mm^2.

T	z_{ts}	z_{bn}	R_{ts}	R_{bn}
100000	-1.990938603	-1.984506864	0.976756181	0.976400318
2500000	-1.834190835	-1.827344797	0.966687164	0.966176028
5000000	-1.659220259	-1.652076253	0.951464292	0.950740496
7500000	-1.478579423	-1.471289619	0.930373622	0.929393595
10000000	-1.296059727	-1.288770095	0.902522546	0.901260987
13000000	-1.079397107	-1.072274996	0.859794630	0.858201733
18000000	-0.736094575	-0.729543208	0.769163428	0.767165276
20000000	-0.607793789	-0.601533305	0.728337875	0.726257582
25000000	-0.31383893	-0.308363433	0.623178312	0.621097098
30000000	-0.059251123	-0.054549707	0.523623955	0.521751396

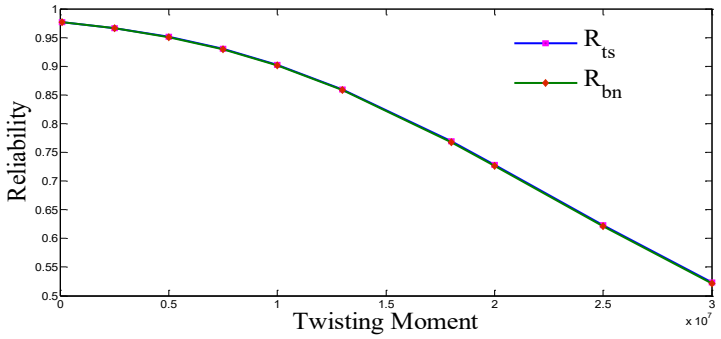


Figure 5: *Reliability and Twisting Moment*

Table-6: $T = 10^4$ *N-mm*, $d = 110$ *mm*, $\mu_\xi = 119.6584$ *N/mm²*.

M	z_{ts}	z_{bn}	R_{ts}	R_{bn}
100000	-1.990938603	-1.984506864	0.976756181	0.976400318
1200000	-1.921595252	-1.840874437	0.972671647	0.967180019
2000000	-1.868157213	-1.730019145	0.969129920	0.958186573
2400000	-1.841016165	-1.673403924	0.967190405	0.952876097
3800000	-1.744153719	-1.471242305	0.959433855	0.929387200
4800000	-1.673478046	-1.325202811	0.952883388	0.907448053
6500000	-1.551268837	-1.079369702	0.939581364	0.859788524
7800000	-1.456703987	-0.897275084	0.927400946	0.815213926
8600000	-1.398287099	-0.788978462	0.918986565	0.784937703
12000000	-1.151035866	-0.369456983	0.875141260	0.644106434

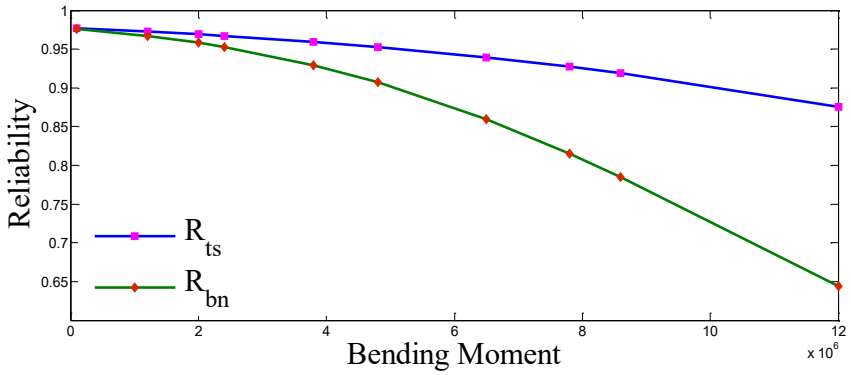


Figure 6: *Reliability and Bending Moment*

Table-7: $T = 10^4\ N-mm,\ M = 15 \times 10^4$ *N-mm,* $\mu_\xi = 119.6584$ *N/mm²*.

d	z_{ts}	z_{bn}	R_{ts}	R_{bn}
24	-0.778536422	-0.208090258	0.781873578	0.582420753
28	-1.228383905	-0.802787694	0.890348557	0.788951272
32	-1.491547172	-1.195487826	0.932091052	0.884051755
36	-1.649076771	-1.442464991	0.950434046	0.925414380
40	-1.747806723	-1.599826782	0.959751250	0.945181492
52	-1.888095775	-1.823839541	0.970493453	0.965911833
62	-1.934633552	-1.897558264	0.973482360	0.971122852
82	-1.971987598	-1.956288280	0.975694489	0.974784398
100	-1.984601679	-1.976010500	0.976405597	0.975923206
110	-1.988441857	-1.982002340	0.976618578	0.976260511

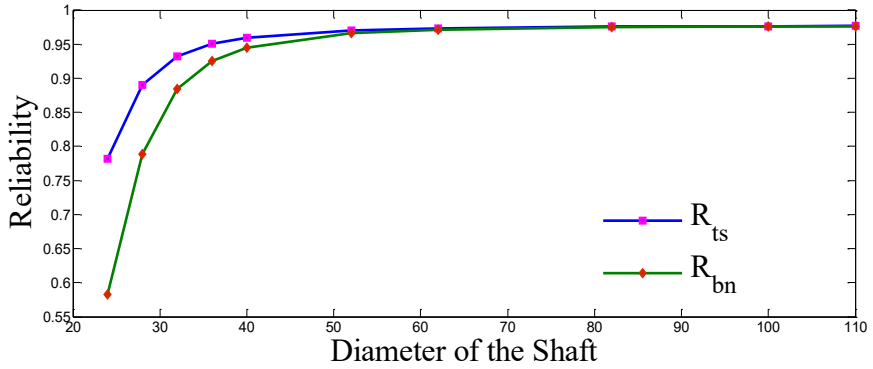


Figure 7: *Reliability and Diameter of the Shaft*

Table-8: $T = 4 \times 10^4\ \ N-mm, M = 2 \times 10^4$ *N-mm,* $d = 110$ *mm.*

μ_ξ	z_{ts}	z_{bn}	R_{ts}	R_{bn}
2.5	-0.521239161	-0.013920898	0.698899912	0.505553455
3.5	-0.918703097	-0.477954094	0.820874555	0.683658561
6.5	-1.425158981	-1.157220313	0.922944375	0.87640882
10.5	-1.652400548	-1.487687204	0.950773538	0.931583298
20.5	-1.826770123	-1.745803964	0.966132830	0.959577489
40.5	-1.913822863	-1.874254579	0.972178603	0.969552328
60.5	-1.942677032	-1.916561189	0.973972403	0.972353149
120.5	-1.971410740	-1.958499434	0.975661543	0.974914281
210.5	-1.983682381	-1.976343042	0.976354371	0.975942031
350.5	-1.990215750	-1.985824871	0.976716413	0.976473613

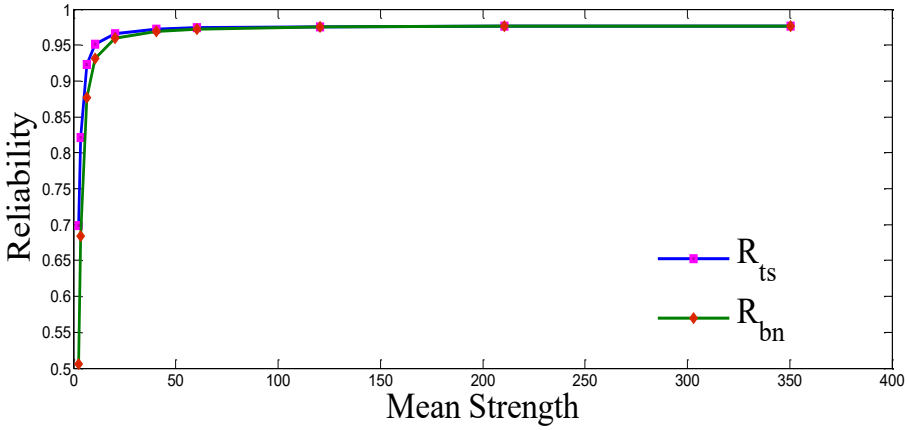


Figure 8: *Reliability and Diameter of the Shaft*

4. Conclusion

The reliability of the shaft is derived from the twisting and bending moments and the combined twisting and bending moments. According to the computations, the reliability of the shaft decreases when the twisting and bending moments increase. Reliability increases when the diameter and strength of the shaft increase.

References

[1] Adekunle, A. A., Adejuyigbe, S. B., & Arulogun, O. T. (2012). Development of CAD software for shaft under various loading conditions. Procedia Engineering, 38, 1962-1983.

[2] Dr. Edward E. Osakue et al., "Fatigue Shaft Design Verification for Bending and Torsion" International Journal of Engineering Innovation and Research, pp.197-206, 2015.

[3] Dr. Edward E. Osakue, "Probabilistic Fatigue Design of Shaft for Bending and Torsion", International Journal of Research in Engineering and Technology, Volume: 03, Issue: 09, pp. 370-386, 2014, DOI: 10.15623/ijret.2014.0309059.

[4] Frydrysek, K. Performance-Based design applied for a Shaft subjected to combined stress.

[5] Gowtham, S., Raaghul, V., Sreeharan, B. N., Kumar, R. P., & Kasim, S. M. (2021, April). Drive Shaft Design and Analysis for Mini Baja. In IOP Conference Series: Materials Science and Engineering (Vol. 1123, No. 1, p. 012015). IOP Publishing.

[6] K. C. Kapur and L. R. Lamberson, "Reliability in Engineering Design", Wiley India Private Limited, New Delhi, 1997.

[7] Kamboh, M. S., Machhi, M. A., & Kamboh, M. F. (2020). Design and Analysis of Drive Shaft with a Critical Review of Advance Composite Materials and the Root Causes of Shaft Failure.

[8] Kececioglu, D. (1972)., Reliability analysis of mechanical components and systems. Nuclear Engineering and Design, 19*(2), 259-290.*

[9] Misra, A., Roberts, L.A. & Levorson, S.M. Reliability analysis of drilled shaft behaviour using finite difference method and Monte Carlo simulation. Geotech Geol Eng 25, 65–77 (2007). https://doi.org/10.1007/s10706-006-0007-2.

[10] Munoz-Abella, B., Montero, L., Rubio, P., & Rubio, L. (2022). Determination of the Critical Speed of a Cracked Shaft from Experimental Data. Sensors, 22(24), 9777.

[11] Nayek, S., Seal, B., & Roy, D. (2014). Reliability approximation for solid shaft under Gamma setup. Journal of Reliability and Statistical Studies, 11-17.

[12] Patel, B., Prajapati, H. R., & Thakar, D. B. (2014). Critical Review on design of shaft with multiple discontinuities and combined loadings ICCIET–2014.

[13] R. S. Khurmi and J. K. Gupta "A Textbook of Machine Design", Eurasia Publishing House (PVT.), Limited, Ram Nagar, New Delhi, 2005.

[14] Villa-Covarrubias, B., Baro-Tijerina, M., & Pina-Monarrez, M. R. Probabilistic Methodology to Determine the Shaft's Diameter and Designed Reliability.

GROUP RUNS AND MODIFIED GROUP RUNS CONTROL CHARTS FOR MONITORING LINEAR REGRESSION PROFILES

Onkar Ghadge[1] and Vikas Ghute[2*]

•

[1,2]Department of Statistics, Punyashlok Ahilyadevi Holkar Solapur University,
Solapur(MS) - 413255, India
onkarghadge1710@gmail.com, vbghute_stats@rediffmail.com
* Corresponding Author

Abstract

Profile monitoring is a critical tool for manufacturing industries to evaluate and maintain quality performance, as well as detect faults. The process of profile monitoring involves observing how variables interact with one another throughout a given period. This enables the understanding of any changes in their functional relationship over time. Generally, control charts are used for monitoring profiles. This paper proposes two new methods to enhance the monitoring of simple and multiple linear regression profiles in Phase II. The proposed methods are based on group runs (GR) and modified group runs (MGR) control charting schemes. The procedure to obtain optimal design parameters for the proposed methods is discussed in detail. The effectiveness of the suggested techniques is assessed through the ARL standard. The study found that the proposed GR and MGR monitoring methods displayed superior performance compared to other available monitoring methods in the literature. A real-life example is illustrated using proposed GR and MGR charting schemes.

Keywords: Statistical process control**,** Control chart, Average run length, Linear regression models, profile monitoring.

1. Introduction

Statistical process control (SPC) plays a crucial role in enhancing the quality and productivity of manufacturing processes by creating effective methods. The control chart serves as a valuable tool in SPC for monitoring the quality of manufactured products. In particular, Shewhart-type, exponentially weighted moving average (EWMA) charts and cumulative sum (CUSUM) control charts are frequently utilized to monitor the quality of single-variable characteristics. In various instances, the manufacturing product quality depends on multiple factors; thus, a multivariate control chart is employed for such scenarios. Some of the commonly used multivariate control charts include Hotelling's T^2 chart, multivariate EWMA chart, and multivariate CUSUM chart. In many practical situations, there is a functional relationship between a variable being studied and other factors that can explain it. In these situations, the quality of the manufacturing product is effectively monitored by using a profile. Profile monitoring involves the surveillance of how the study variable relates to one or several explanatory variables. When the response variable has a linear relationship with only one explanatory variable, it is a simple linear regression profile. However, if the response variable has a linear relationship with more than one explanatory

variable, it is a multiple linear regression profile. The practice of profile monitoring utilizes control charts and is implemented to ascertain whether any alterations have occurred in the established functional relation, with respect to time. The use of profile monitoring in both Phase I and II have become increasingly important in various industries as it allows for the early detection of shifts in a process and contributes to the maintenance of quality control. During the initial stage, denoted as Phase I, there is a lack of knowledge regarding the parameters governing each profile. As such, these values are appraised through a reliable and consistent method. Moving on to Phase II, the profiles' parameters have been distinctly determined beforehand and serve as reference points for detecting any alterations in the process being analyzed.

There are various charting techniques in academic literature to track simple and multiple linear regression data for Phase I and II scenarios. Kang and Albin [1] used simple linear profiles to calibrate semiconductor manufacturing equipment. They monitored the relationship between gas flow and pressure in the chamber Kang and Albin [1] proposed common charting techniques to supervise linear profiles by employing the multivariate Hotelling's T^2 and EWMA/R methods. Kim et al. [2] devised a Phase II monitoring strategy, known as the EWMA-3 chart, which employs three individual univariate EWMA regression control charts to monitor the intercept, slope, and variance of errors in linear profiles. Their findings indicate that the EWMA-3 method is more effective than the methods of Kang and Albin [1] in Phase II analysis. Woodall et al. [3] addressed the key concerns regarding using control charts for monitoring process and product quality profiles. They have also given a review of the literature on profile monitoring in SPC. Gupta et al. [4] presented a Shewhart-based simple linear profile method called the Shewhart-3 chart. Saghaei et al. [5] suggested CUSUM-3 method to monitor simple linear profile parameters. Woodall [6] presented a review of linear profile monitoring methods. Riaz et al. [7] developed the Assorted-3 control charting strategy for monitoring parameters of simple linear profiles. To monitor multiple linear profiles, several methods have been developed by researchers, including Zou et al. [8], Zhang et al. [9], Zou et al. [10] and Amiri et al. [11]. Further, the T^2 control chart method by Kang and Albin [1] can also be used for Phase II monitoring of multiple linear regression profiles. Maleki et al. [12] provided an overview of profile monitoring papers published from 2008 to 2018.

Various methods have been suggested to improve conventional control charts in the literature of SPC. The synthetic control chart is one such approach that seeks to enhance the effectiveness of traditional control charts. Wu and Spedding [13] first developed synthetic chart technique which combines the features of both Shewhart chart and CRL chart. Bourke [14] developed the CRL control chart for monitoring nonconforming fractions. Wu and Spedding [13] demonstrated that the synthetic $\bar{X}$ chart outperforms the Shewhart $\bar{X}$ chart. Synthetic control charts have proven to be effective in detecting shifts in both univariate and multivariate processes, which has led many researchers to focus on designing them to enhance their shift detection capabilities. Numerous research papers on synthetic control charts exist in the literature. Some of these are Chen and Huang [15], Huang and Chen [16], Ghute and Shirke [14], Ghute and Shirke [18], Ghute and Shirke [19], Aparisi and de Luna [20], Rajmanya and Ghute [21]. Rakitzis et al. [22] reviewed over 100 scholarly articles based on synthetic-type control charts. In literature, the Group Runs (GR) method is recommended in literature as an upgraded version of the synthetic method for identifying changes in process parameters. Gadre and Rattihalli [23] devised GR chart to enhance shift detection ability of Shewhart $\bar{X}$ chart and synthetic $\bar{X}$ chart. The GR control chart scheme combines Shewhart $\bar{X}$ chart and an extended version of the CRL chart and it shows more acceptable performance than the Shewhart $\bar{X}$ chart and synthetic $\bar{X}$ chart. More related work on GR charts can be seen in Gadre et al. [24], Gadre and Kakade [25], Gadre and Kakade [26], Chong et al. [27], Khilare and Shirke [28]. Gadre and Rattihalli [29] developed Modified Group Runs (MGR) scheme to detect the shift in process mean of the normally distributed process. The results of their study suggest that the MGR chart is more efficient than Shewhart, synthetic, and GR charts. Gadre

and Kakade [26] introduced multivariate GR and MGR control charts for monitoring process mean vector of normally distributed process. It was shown that, the proposed multivariate versions of GR and MGR charts detect changes in the process mean vector faster than the Hotelling's T^2 chart and synthetic T^2 chart proposed by Ghute and Shirke [19].

By getting motivation of improved performance from GR and MGR charts, we implement these charting schemes to the T^2 method proposed by Kang and Albin [1] for monitoring simple and multiple linear regression profiles. The objective of this paper is to propose GR and MGR charting methods for improved monitoring of linear profiles and we expect that the proposed methods will provide a better option for the detection of a wide range of shifts in the model parameters of both simple and multiple linear regression models in Phase II monitoring. In this paper, we propose group runs based T^2 method (denoted as GR-T^2) and modified group runs based T^2 method (denoted as MGR-T^2) for monitoring simple and multiple linear profiles in Phase II. The performance of the proposed methods for monitoring simple linear profiles is compared with the T^2 method by Kang and Albin [1] and Shewhart-3 method of Gupta et al. [4]. There are other methods in the literature based on EWMA and CUSUM control charts for monitoring profiles. But for a fair comparison, we have selected only T^2 and Shewhart-3 methods because our proposed method GR-T^2 method is Shewhart based method and Shewhart-3 approach is Shewhart type also Hotelling's T^2 chart is multivariate extension of the univariate $\bar{X}$ chart. The performance of the proposed methods for monitoring multiple linear profiles is also compared with the T^2 method. The performance of the proposed GR-T^2 and MGR-T^2 methods is evaluated using Monte Carlo simulations in terms of average run length (ARL) criterion under sustained shifts of different magnitudes in the regression parameters and error standard deviation of simple and multiple linear profile monitoring.

The rest of this paper is organized as follows: The model for simple linear profiles under consideration is presented in Section 2. In Section 3, we discuss multiple linear regression model to be considered in this study for monitoring profiles in Phase II. Our proposed GR-T^2 and MGR-T^2 methods for monitoring simple and multiple linear profiles in Phase II are described in Sections 4 and 5 respectively. Section 6 of this study evaluates the effectiveness of our suggested approach in comparing with traditional methods for monitoring simple linear and multiple linear profiles, using the average run length standard. A demonstrative example is provided in Section 7. The conclusions are given in Section 8.

2. Simple linear regression profile model

This section will analyze a simple linear regression model that utilizes two parameters to demonstrate how the response variable Y is a function of an explanatory variable X. Let Y_{ij} is the i^{th} observation of the response variable in the j^{th} profile and X_i is the corresponding explanatory variable, which is assumed to be known constant in each profile ($i = 1,2, \dots, n; j = 1,2, \dots$). If the relationship between the response variable and explanatory variable can be accurately shown through a simple linear regression model, then the model should be created in the following way,

$$Y_{ij} = A_{0j} + A_{1j}X_i + \varepsilon_{ij} \quad (1)$$

where, $\beta_0, \beta_1 and\ \varepsilon_{ij}$ represents intercept, slope and error terms respectively. It is assumed that $\varepsilon_{ij}{}'$s are independent and identically distributed (i.i.d.) normal random variables with mean 0 and variance σ^2.The regression coefficients, $\beta_0\ and\ \beta_1$ are assumed to be known. Furthermore, it is presupposed that the X values remain constant and invariant in every sample. The ordinary least squares estimators of $\beta_0\ and\ \beta_1$ for the j^{th}sample are given by

$$\hat{\beta}_{0j} = \overline{Y}_j - \hat{\beta}_{1j}\overline{X} \text{ and } \hat{\beta}_{1j} = \frac{S_{xy(j)}}{S_{xx}} \quad (2)$$

where, $\overline{Y_j} = \frac{1}{n}\sum_{i=1}^{n} Y_{ij}, \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X, S_{xy(j)} = \sum_{i=1}^{n}(X_i - \bar{X})Y_{ij}$ and $S_{xx} = \sum_{i=1}^{n}(X_i - \bar{X})^2$. The least

square estimators $\hat{\beta}_{0j}$ and $\hat{\beta}_{1j}$ have a bivariate normal distribution with mean vector $\boldsymbol{U} = (\beta_0, \beta_1)'$ and variance-covariance matrix,

$$\Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \quad (3)$$

$\sigma_0^2 = \sigma^2(n^{-1} + \frac{\bar{X}^2}{S_{xx}})$ and $\sigma_1^2 = \sigma^2/S_{xx}$ are the variances of $\hat{\beta}_{0j}$ and $\hat{\beta}_{1j}$ respectively and $\sigma_{01} = -\sigma^2\bar{X}^2/S_{xx}$ is the covariance between $\hat{\beta}_{0j}$ and $\hat{\beta}_{1j}$.

Kang and Albin [1] proposed the Hotelling's T^2 chart for monitoring intercept and slope of simple linear regression model. The vector of sample estimators $\boldsymbol{Z_j} = (\hat{\beta}_{0j}, \hat{\beta}_{1j})'$ for sample j is computed and then T^2 statistic is computed as

$$T_j^2 = (\boldsymbol{Z_j} - \boldsymbol{U})'\boldsymbol{\Sigma^{-1}}(\boldsymbol{Z_j} - \boldsymbol{U}) \quad (4)$$

where, $\boldsymbol{Z_j} = (\hat{\beta}_{0j}, \hat{\beta}_{1j})'$ is vector of estimated values of intercept and slope for j^{th} sample, $\boldsymbol{U} = (\beta_0, \beta_1)'$ and Σ is defined by Eq. (3). When the process is under control, T_j^2 follows a chi-square distribution with 2 degrees of freedom. Therefore, the upper control limit for the chart is $UCL = \chi_{2,\alpha}^2$, where $\chi_{2,\alpha}^2$ is the upper 100α percentage point of chi-square distribution with 2 degrees of freedom. When the process is not stable, the Hotelling's T^2statistic follows a non-central chi-square distribution with non-centrality parameter $\tau = n(\lambda\sigma + \delta\sigma\bar{X})^2 + (\delta\sigma)^2 S_{xx}$, where λ and δ are the shifts in the intercept and slope of the model given in Eq. (1).

3. Multiple linear regression profile model

In this section, the profile is represented by a multiple linear regression model. We consider a multiple linear regression model in $(p+1)$ parameters to model the response variable Yas a function of p explanatory variables $X_1, X_2, \ldots, X_p$. Let us assume that for the j^{th} random sample collected over time, we have n observations given as $X_{i1}, X_{i2}, \ldots, X_{ip}$ and Y_{ij} $i = 1,2, \ldots, n$, where p is the number of explanatory variables. The relation between response variable and explanatory variables is characterized by a multiple linear regression model as

$$\boldsymbol{Y_j} = \boldsymbol{X\beta_j} + \boldsymbol{\varepsilon_j} \quad (5)$$

where $\boldsymbol{Y_j}$ is $n \times 1$ vector of response variables for the j^{th} sample, $\boldsymbol{X}$ **is** $n \times (p+1)$ matrix of explanatory variables, $\boldsymbol{\beta_j}$ is $(p+1) \times 1$ vector of regression parameters and $\boldsymbol{\varepsilon_j}$ is a $n \times 1$ vector of error terms which are assumed to be independently and identically distributed normal variables with mean zero and known variance σ^2. When $p = 1$, the model of multiple linear profiles reduces to a simple linear profile. In addition, the X values are assumed to be fixed and constants for each sample. The least squares estimator of $\boldsymbol{\beta_j}$ is given as

$$\widehat{\boldsymbol{\beta_j}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \quad (6)$$

and variance-covariance matrix of sample estimates of regression parameters is given as

$$\Sigma = cov(\widehat{\boldsymbol{\beta}}) = (\boldsymbol{X'X})^{-1}\sigma^2 \quad (7)$$

The Hotelling's T^2 statistic for j^{th} sample is computed by Eq. (4), where $\boldsymbol{Z_j} = (\hat{\beta}_{0j}, \hat{\beta}_{1j}, \ldots, \hat{\beta}_{pj})'$ is the vector of estimated regression parameters and $U = (\beta_0, \beta_1, \ldots, \beta_p)'$ is the vector of known regression parameters. The upper control limit of Hotelling's T^2 chart is given as $\chi_{p+1,\alpha}^2$. The theoretical ARL values of the T^2chart calculated as follows,

$$ARL = \frac{1}{P(T_j^2 > \chi_{p+1,\alpha}^2)} \quad (8)$$

4. The GR control chart for monitoring linear regression profiles

In this section, we present the design structure of the proposed GR-T^2 chart for monitoring simple and multiple linear regression profiles. Following the work of Gadre and Ratihalli [23], in order to increase the detection ability of the Hotelling's T^2 chart, GR-T^2 chart combines the T^2 chart with an

improved version of the CRL chart. The T^2 chart has only upper control limit (UL) and CRL chart has only the lower control limit L. Let Y_r be the r^{th} group based CRL and L be the lower limit of the CRL chart.

Operation of GR-T^2 Chart

Following Gadre and Rattihalli [23], the operation of the GR-T^2 chart is outlined by the following steps

1. Decide the upper control limit UL for the T^2 chart and lower control limit L of the CRL chart.
2. Take a group of n items (sample of size n) at each inspection point j and compute the chart statistic say T_j^2.
3. If $T_j^2 \leq UL$, then the group is considered as a conforming group and control flow goes back to step 2. Otherwise, the group is considered as nonconforming and control flow goes to the next step.
4. To determine the CRL (Y_r), count the number of T_j^2 samples between nonconforming groups.
5. If $Y_r > L$ the process is thought to be under control, and control flow goes back to step 2. If $Y_1 \leq L$ or two successive $Y_r \leq L$ and $Y_{r+1} \leq L$, for $r = 2,3, \ldots$ for the first time, the process is thought to be out-of-control, and control flow proceeds to the next step.
6. Indicate the out-of-control signal.
7. An assignable cause should be searched and take corrective action should be taken to remove it.

Let the expected number of groups (samples) required for a GR-T^2 chart to detect a shift of magnitude d in process mean vector be denoted by $ARL_{GR}(d)$. Following Gadre and Rattihalli [23], the ARL measure for the GR-T^2 chart is as follows:

$$ARL_{GR}(d) = \frac{1}{P(d)} \times \frac{1}{\{1-[1-P(d)]^L\}^2} \tag{9}$$

where, $P(d)$ is probability of detecting nonconforming group (sample) when shift of magnitude d is occurred.

$$P(d) = P[T^2 > UL/d \neq 0] = 1 - F_{(p+1,\lambda^2)}(UL) \tag{10}$$

where, $F_{(p+1,\lambda^2)}(.)$ is the cumulative distribution function of chi-square distribution with $p+1$ degrees of freedom and non-centrality parameter λ^2. The value of λ^2 is given as

$$\lambda^2 = (\mu_{1-}\mu)'\Sigma^{-1}(\mu_{1-}\mu) \tag{11}$$

where, $\boldsymbol{\mu_1}$ is vector of shifted regression parameters and $\boldsymbol{\mu}$ is vector of in-control regression parameters.

If $d = 0$, the in-control ARL of the GR-T^2 chart is given as

$$ARL_{GR}(0) = \frac{1}{P(0)} \times \frac{1}{\{1-[1-P(0)]^L\}^2} \tag{12}$$

where, $$P(0) = P[T^2 > UL/d = 0] = 1 - F_{(p+1)}(UL) \tag{13}$$

The optimal design of GR-T^2chart depends on the desired in-control ARL, $ARL_{GR}(0)$ and out-of-control ARL, $ARL_{GR}(d^*)$. Here, d^* is the magnitude of shift considered large enough to seriously impair the quality of the products; the corresponding $ARL_{GR}(d^*)$ should be as small as possible. The GR control chart is designed by solving an optimization problem. The objective function to be minimized is

$$ARL_{GR}(d^*) = \frac{1}{P(d^*)} \times \frac{1}{\{1-[1-P(d^*)]^L\}^2} \tag{14}$$

subject to the equality constraint in Eq. (12).

Optimal Design Procedure of GR-T^2 Chart

We present the optimal design to obtain the optimal parameters (L, UL) for the GR-T^2 chart that result in minimum $ARL_{GR}(d^*)$ value, subject to in-control ARL $ARL_{GR}(0)$ which is at least equal to 200.

The optimal design procedure for the GR-T^2 chart is described as follows:

1. Specify p, d^*, Σ and $ARL_{GR}(0)$.
2. Obtain $P(0)$ by solving Eq. (12) numerically. From this value of $P(0)$, obtain the value of UL

using Eq. (13).

3. From the current values of L and UL compute $P(d^*)$ using Eq. (10) and then compute $ARL_{GR}(d^*)$ using Eq. (14)
4. If $ARL_{GR}(d^*)$ has been reduced then increase L by 1 and go back to Step 3. Otherwise, go to the next step.
5. Take the current L and UL as the final values for which $ARL_{GR}(d^*)$ is minimum.

5. The MGR control chart for monitoring linear regression profiles

The MGR chart proposed by Gadre and Rattihalli [29] is an extension of the GR chart with the inclusion of warning limit in the extended CRL scheme. Following Gadre and Rattihalli [29], we develop the procedure related to our proposed MGR-T^2 chart. The MGR-T^2 chart is divided into two parts. The first part assesses group conformity using a T^2-based technique, while the second analyzes process status through a group runs approach. This component has two levels of inspection.

- T^2 based procedure: If the value of T^2 statistic for a group of n units falls outside the upper control limit UL, declare the group as nonconforming; otherwise it is treated as a conforming group. based procedure: If the value of T^2 statistic for a group of n units falls outside the upper control limit UL, declare the group as nonconforming; otherwise it is treated as a conforming group.
- Group runs based procedure: Let Y_r denote the r^{th} (r=1,2,…) group based CRL. In other words, it is the number of groups inspected between $(r-1)^{th}$ and r^{th} nonconforming group.
- The group runs based procedure declares the process out-of-control if $Y_1 \leq L_2$ or for some $r(>1)$, $Y_r \leq L_1$ and $Y_{r+1} \leq L_2$ for the first time. Here L_1 is a warning limit and L_2 is lower limit of the CRL sub-chart chart.

Following Gadre and Rattihalli [29], the ARL expression for MGR-T^2 chart to detect a shift of magnitude d in the process mean vector is given as

$$ARL_{MGR}(d) = \frac{1}{P(d)} \times \frac{[1+Q(d)^{L_2}-Q(d)^{L_1}]}{[1-Q(d)^{L_1}][1-Q(d)^{L_2}]} \quad (15)$$

where, $Q(d) = 1 - P(d)$ and $P(d)$ is given in Eq. (10).

If $d = 0$, the In-Control ARL of the MGR-T^2 chart is given as

$$ARL_{MGR}(0) = \frac{1}{P(0)} \times \frac{[1+Q(0)^{L_2}-Q(0)^{L_1}]}{[1-Q(0)^{L_1}][1-Q(0)^{L_2}]} \quad (16)$$

The MGR-T^2 chart is designed by solving an optimization problem. The objective function to be minimized is

$$ARL_{MGR}(d^*) = \frac{1}{P(d^*)} \times \frac{[1+Q(d^*)^{L_2}-Q(d^*)^{L_1}]}{[1-Q(d^*)^{L_1}][1-Q(d^*)^{L_2}]} \quad (17)$$

subject to the equality constraint in Eq. (16).

Optimal design procedure of MGR-T^2 Chart

We present the optimal design to obtain the optimal parameters $(L_1, L_2, UL$ for the MGR-T^2 chart that result in minimum $ARL_{MGR}(d)$ value, subject to In-Control $ARL_{MGR}(0)$ which is at least equal to 200.

The optimal design procedure for the MGR-T^2 chart is described as follows

1. Specify p, d^*, Σ and $ARL_{MGR}(0)$.
2. Initialize L_1 as 1.
3. Initialize L_2 as 1.
4. Obtain $P(0)$ by solving Eq. (16) numerically. From this value of $P(0$ obtain the value of UL using Eq. (13).
5. From the current values of L_1, L_2 and UL compute $P(d^*)$ using Eq. (10) and then compute $ARL_{MGR}(d^*)$ using Eq. (17).
6. If $ARL_{MGR}(d^*)$ has been reduced then increase L_2 by 1 and go back to Step 4. Otherwise, go to

the next step.

7. If L_1 or the value of $ARL_{MGR}(d^*)$ has been reduced then increase L_1 by 1 and go back to Step 3; else go to the next step.
8. Take the values of L_1, L_2 and UL as the final values for which $ARL_{MGR}(d^*)$ is minimum.

6. Performance comparison

In this section, the proposed GR-$T^2$and MGR-T^2control chart methods for Phase II monitoring of linear regression profile processes were evaluated through simulation results, which demonstrate their effectiveness in detecting out-of-control conditions in simple linear and multiple linear regression profile processes. ARL is a performance measure used in this study for the evaluation of the proposed GR T^2 and MGR T^2 methods.

6.1. Performance comparison of simple linear regression profiles

In sub-section 6.1 of study, we aim to evaluate and compare the average run length performance of two proposed control chart methods for monitoring simple linear regression profiles, namely the GR T^2 and MGR T^2, with two existing methods. These existing methods include the control chart methods proposed by Kang and Albin [1], as well as the Shewhart-3 method suggested by Gupta et al. in [4], which are used as benchmarks for comparison. This study aims to provide valuable insights into the effectiveness of these control chart methods for monitoring simple linear regression profiles, which will aid practitioners in choosing an appropriate method for their specific needs. To compare the performance of proposed methods with above mentioned methods, the GR-T^2 and MGR-T^2 control charts are designed such that in-control ARL is approximately 200. For our study purpose, we have used in-control simple linear profile model which is given by Kang and Albin [1] as,

$$Y = 3 + 2X + \varepsilon \tag{18}$$

with ε follows i.i.d. normal random variables with mean zero and variance one and values of explanatory variable are fixed as $X = 2,4,6\ and\ 8$, following Kang and Albin [1]. The optimal design parameters and control limits for the proposed GR T^2 and MGR T^2 methods under $p = 1, n = 4$ and $d^* = 1$are given in Table 1 in order to achieve an overall ARL of 200.

Table 1: *Optimal design parameters for GR-T^2 and MGR-T^2control chart*

Control chart	Optimal design parameters
GR T^2	$L = 16, UL = 6.9248,$
MGR T^2	$L_1 = 1, L_2 = 31, UL = 6.2459$

The shifts in the β_0, β_1 and σ considered in the study are presented in Table 2. These are same as the example discussed by Kang and Albin [1]. For performance evaluation of proposed GR-T^2 and MGR-T^2 methods to monitor simple regression profiles, we used 50000 simulation runs to obtain out-of-control ARL under different shifts in the β_0, β_1 and σ of the model given in Eq. (18). The ARL performance of above mentioned methods for monitoring simple linear regression profiles is given by Riaz et al. [7]. The out-of-control ARL values of proposed GR T^2 and MGR T^2 and other methods for detecting shifts in regression parameters of a simple linear regression model are presented in Table 3.

From Table 3, we observe that under the intercept shift from β_0 to $\beta_0 + \lambda\sigma$, the proposed GR-T^2 and MGR-T^2 methods produces smaller out-of-control ARL than the $T^2$and Shewhart-3 methods for entire range of shifts in the β_0 parameter. Similarly, under the slope shift from β_1 to $\beta_1 + \delta\sigma$, the proposed GR-T^2 and MGR-T^2 methods consistently produces smaller out-of-control ARL than the T^2method and Shewhart-3 method for entire range of shifts in

the β_1 parameter. Finally, for monitoring the error standard deviation shifts from $\sigma\ to\ \gamma\sigma$, the proposed GR-T^2 and MGR-T^2 methods consistently produces shorter out-of-control average run lengths than the T^2 and Shewhart-3 methods for entire range of shifts in the slope parameter. Therefore, we conclude that both the GR-T^2 and MGR-T^2 methods are suitable for monitoring different shifts in the β_0, β_1 and σ of the simple linear regression model.

Table 2: *Shifts considered for various methods*

Type of shift	Notation	Values of the shift
Simple linear profiles		
Shift in β_0	β_0 to $\beta_0 + \lambda\sigma$	$\lambda = 0.2,0.4,0.6, \ldots ,2.0$
Shift in β_1	β_1 to $\beta_1 + \delta\sigma$	$\delta = 0.025,0.05,0.075, \ldots ,0.25$
Shift in σ	$\sigma\ to\ \gamma\sigma$	$\gamma = 1.2,1.4,1.6, \ldots ,3.0$
Multiple linear profiles		
Shift in β_0	β_0 to $\beta_0 + \lambda_0\sigma$	$\lambda_0 = 0.2,0.4,0.6, \ldots ,2.0$
Shift in β_1	β_1 to $\beta_1 + \lambda_1\sigma$	$\lambda_1 = 0.02,0.04,0.06, \ldots ,0.20$
Shift in σ	$\sigma\ to\ \gamma\sigma$	$\gamma = 1.1,1.2,1.3, \ldots ,2.0$

Table 3: *Performance comparison under the shifts in intercept, slope and error variance*

Method	Shift in $\beta_0(\lambda)$									
	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
T^2	137.7	63.5	28.0	13.2	6.9	4.0	2.6	1.8	1.5	1.2
Shewhart-3	151.4	78.0	33.3	15.5	7.7	4.3	2.7	1.9	1.5	1.2
GR-T^2	106.8	30.4	10.0	4.8	2.9	2.0	1.6	1.3	1.1	1.1
MGR-T^2	89.7	17.9	6.5	3.8	2.5	1.8	1.5	1.2	1.1	1.1
Method	Shift in $\beta_1(\delta)$									
	0.025	0.05	0.075	0.1	0.125	0.15	0.175	0.2	0.225	0.25
T^2	166.0	105.6	60.7	34.5	20.1	12.2	7.8	5.2	3.7	2.7
Shewhart-3	178.3	125.0	79.2	46.7	27.9	17.1	10.9	7.1	5.0	3.6
GR-T^2	146.2	68.5	28.4	12.9	7.0	4.5	3.2	2.4	1.9	1.6
MGR-T^2	133.7	50.0	16.5	7.8	5.1	3.6	2.7	2.2	1.8	1.5
Method	Shift in $\sigma(\gamma)$									
	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
T^2	39.6	14.9	7.9	5.1	3.8	3.0	2.5	2.2	2.0	1.8
Shewhart-3	40.1	13.5	6.5	4.0	2.8	2.2	1.8	1.6	1.5	1.4
GR-T^2	18.2	6.5	3.9	2.9	2.4	2	1.8	1.7	1.6	1.5
MGR-T^2	10.8	4.9	3.4	2.6	2.2	1.9	1.7	1.6	1.5	1.4

6.2. Performance comparison of multiple linear regression profiles

In this section, we extend the T^2 method of Kang and Albin [1] to monitor multiple linear regression profiles. Further to improve performance of T^2 method, we apply the concept of GR and MGR charting schemes to the T^2 method. We compare ARL performance of the proposed GR-T^2, MGR-T^2 and T^2 to monitor multiple linear profiles in Phase II.

To compare the performance of proposed GR-T^2 and MGR-T^2 methods with T^2 method all the methods are designed such that in-control ARL is approximately 200. For our study purpose, we have used in-control multiple linear profile model used by Amiri et al. [11] as

$$Y = 3 + 2X_{1+}X_2 + X_3 + \varepsilon \tag{19}$$

where, $\beta_0 = 3, \beta_1 = 2, \beta_2 = 1$ and $\beta_3 = 1$. The error terms are independent normal random

variables with mean 0 and known variance of $\sigma^2 = 1$. Following Amiri et al. [11], the values of explanatory variables X_1, X_2 and X_3 are given in the following transpose matrix.

$$X' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 4 & 6 & 8 & 2 & 4 & 6 & 8 \\ 1 & 4 & 3 & 2 & 1 & 4 & 3 & 2 \\ 1 & 3 & 2 & 4 & 4 & 3 & 2 & 4 \end{bmatrix}$$

The optimal design parameters of the proposed GR-T^2 and MGR-T^2 methods under $p = 3, n = 8$ and $d^* = 1$ are provided in Table 4 in order to achieve an overall ARL of 200.

Table 4: *Optimal design parameters for GR-T^2 and MGR-T^2 control chart*

Control chart	Optimal design parameters
GR T^2	$L = 20, UL = 10.2978$
MGR T^2	$L_1 = 1, L_2 = 31, UL = 6.2459,$

The shifts in regression parameters β_0, β_1 and σ considered in the study for multiple linear regression profiles are also presented in Table 2. These are same as the example discussed by Amiri et al. [11].For performance evaluation of proposed GR-T^2 and MGR-T^2 methods to monitor multiple regression profiles, we used 50000 simulation runs to obtain out-of-control ARL under different shifts in the regression parameters and error standard deviation of the model given in Eq. (19). The out-of-control ARL values of T^2, GR-T^2 and MGR-T^2 methods for detecting shifts in regression parameters and error standard deviation of a multiple linear regression model are presented in Table 5.

Table 5: The simulated out-of-control ARL values under the shifts in regression parameters of multiple linear regression model when $p = 3, n = 8$

Method	λ_0 (shift in β_0)										
	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
T^2	200.0	126.0	48.0	17.5	7.2	3.5	2.1	1.5	1.2	1.1	1.0
GR-T^2	200.0	89.2	19.8	6.4	3.2	2.0	1.4	1.2	1.1	1.0	1.0
MGR-T^2	200.0	64.8	11.5	5.2	2.8	1.8	1.3	1.1	1.0	1.0	1.0
Method	λ_1 (shift in β_1)										
	0.0	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
T^2	200.0	172.0	116.3	69.2	39.4	22.6	13.3	8.2	5.3	3.7	2.7
GR-T^2	200.0	153.2	77.8	33.6	15.3	8.1	5.1	3.5	2.6	2.0	1.6
MGR-T^2	200.0	138.4	53.5	18.6	9.5	6.2	4.3	3.1	2.4	1.9	1.5
Method	γ (shift in σ)										
	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
T^2	200.0	64.6	27.1	14.2	8.4	5.3	3.6	2.6	2.0	1.6	1.3
GR-T^2	200.0	32.5	11.4	6.3	4.4	3.3	2.7	2.3	2.0	1.8	1.7
MGR-T^2	200.0	18.3	7.9	5.2	3.8	3.0	2.5	2.1	1.9	1.7	1.6

From Table 5, we observe that the proposed GR-T^2 and MGR-T^2 methods are superior methods in detecting shifts in the parameters of multiple linear regression model. Under the error variance shifts, proposed methods are better than T^2 method except in very large shifts in which performance of all methods is approximately same.

7. An example

In this section, an example is illustrated by using proposed GR-T^2 and MGR-T^2 methods. Dataset is taken from Gupta et al. [4]. The dataset contains measurements of line widths on ten photo mask reference standards. These measurements are used to keep track of the linear calibration profiles of optical imaging systems. Simple linear regression profile for monitoring data given in Table 6 is as follows: $Y_{ij} = 0.2817 + 0.9767X_i$. Residual standard deviation is 0.06826 micrometers. The estimates of regression coefficients β_{0j} and β_{1j} are calculated using ordinal least square method.

Table 6: *Line-Width Measurements and* T^2 *values*

Day	Position	X	Y	$\hat{\beta}_{0j}$	$\hat{\beta}_{1j}$	T^2
1	L	0.76	1.12			
1	M	3.29	3.49	0.3194	0.9862	4.73
1	U	8.89	9.11			
2	L	0.76	0.99			
2	M	3.29	3.53	0.2891	0.9693	0.80
2	U	8.89	8.89			
3	L	0.76	1.05			
3	M	3.29	3.46	0.2726	0.9824	0.40
3	U	8.89	9.02			
4	L	0.76	0.76			
4	M	3.29	3.75	0.1149	1.0406	38.45
4	U	8.89	9.3			
5	L	0.76	0.96			
5	M	3.29	3.53	0.2279	0.9935	2.36
5	U	8.89	9.05			
6	L	0.76	1.03			
6	M	3.29	3.52	0.2847	0.9827	0.81
6	U	8.89	9.02			

GR-T^2 and MGR-T^2 control charts for data given in Table 6 are plotted in Figure 1. From Figure 1, it is noted that on the fourth day, the value of T^2 statistic is 38.45 which is greater than UL of both the GR-T^2 and MGR-T^2control charts. First conforming run length (Y_1) does not satisfy the criteria of in-control for both the GR-T^2 and MGR-T^2control charts. Therefore, charts give an out-of-control signal on fourth day.

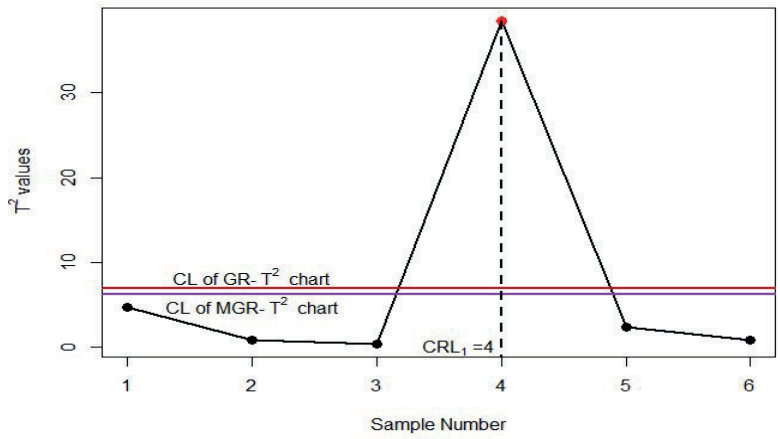


Figure 1: *GR-*T^2 *and MGR-*T^2 *control chart for Line-Width Measurements*

8. Conclusions

In this paper, we proposed two methods namely, GR-T^2 and MGR-T^2 to monitor simple and multiple linear regression profiles in Phase II. The performance of the proposed methods under simple linear profile monitoring is compared with the existing Shewhart-type methods namely T^2 method by Kang and Albin [1] and Shewhart-3 method by Gupta et al. [4] in terms of average run length criterion. From the numerical results, it is shown that the GR-T^2 and MGR-T^2 methods are very effective for detecting shifts in intercept, slope and error standard deviation. In addition, the performance of proposed methods in detecting shifts in the regression parameters and error standard deviation of multiple linear regression profiles is better than the T^2 method except very large shifts in which performance of all the methods is approximately same. Furthermore, the MGR-T^2 method has better performance than the GR-T^2 method for monitoring simple and multiple linear regression profiles. Hence, due to the effectiveness of the MGR-T^2 method, it can be more suitable for monitoring simple and multiple linear regression profiles.

References

[1] Kang, L. and Albin S. L. (2000). On-line monitoring when the process yields a linear profile. *Journal of Quality Technology*, 32(4): 418-426.

[2] Kim, K., Mahmoud, M. A. and Woodall, W. (2003). On the monitoring of linear profiles. *Journal of Quality Technology*, 35(3): 317-328.

[3] Woodall, W. H., Spitzner, D. J., Montgomery, D. C. and Gupta, S. (2004). Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, 36(3): 309-320.

[4] Gupta, S., Montgomery, D. and Woodall, W. (2006). Performance evaluation of two methods for online monitoring of linear calibration profiles, *International Journal of Production Research*, 44(10): 1927-1942.

[5] Saghaei, A., Mehrjoo, M. and Amiri, A. (2009). A CUSUM-based method for monitoring simple linear profiles. *The International Journal of Advanced Manufacturing Technology*, 45(11): 1252-1260.

[6] Woodall, W. H. (2007). Current research on profile monitoring, Production, 17(3): 420-425.

[7] Riaz, M., Saeed, U., Mahmood, T., Abbas, N. and Abbasi, S. A. (2020). An improved control chart for monitoring linear profiles and its application in thermal conductivity. *IEEE Access*, 8: 120679-120693.

[8] Zou, C., Tsung, F. and Wang, Z. (2007). Monitoring general linear profiles using multivariate exponentially weighted moving average schemes. *Technometrics*, 49(4): 395-408.

[9] Zhang, J., Li, Z. and Wang, Z. (2009). Control chart based on likelihood ratio for monitoring linear profiles. *Computational statistics and data analysis*, 53(4): 1440-1448.

[10] Zou, C., Jiang, W. and Tsung, F. (2011). A LASSO-based diagnostic framework for multivariate statistical process control. *Technometrics*, 53(3): 297-309.

[11] Amiri, A., Eyvazian, M., Zou, C., and Noorossana, R. (2012). A parameters reduction method for monitoring multiple linear regression profiles. *The International Journal of Advanced Manufacturing Technology*, 58(5): 621-629.

[12] Maleki, M. R., Amiri, A. and Castagliola, P. (2018). An overview on recent profile monitoring papers (2008–2018) based on conceptual classification scheme. *Computers and Industrial Engineering*, 126: 705-728.

[13] Wu, Z. and Spedding, T. A. (2000). A synthetic control chart for detecting small shifts in the process mean. *Journal of Quality Technology*, 32(1): 32-38.

[14] Bourke, P.D. (1991). Detecting a shift in fraction non-conforming using run-length control

charts with 100% inspection. *Journal of Quality Technology*, 23(3): 225-238.

[15] Chen, F.L. and Huang, H.J. (2005). A synthetic control chart for monitoring process dispersion with sample range. *The International Journal of Advanced Manufacturing Technology*,26(7-8): 842-851.

[16] Huang, H. J. and Chen, F. L. (2005). A synthetic control chart for monitoring process dispersion with sample standard deviation. *Computers and Industrial Engineering,* 49(2): 221-240.

[17] Ghute, V. B. and Shirke, D. T. (2007). Joint monitoring of multivariate process using synthetic control charts. *International Journal of Statistics and Management System,* 2(1-2): 129-141.

[18] Ghute, V. B. and Shirke, D. T. (2008a). A multivariate synthetic control chart for monitoring process mean vector. *Communications in Statistics-Theory and Methods,* 37(13): 2136-2148.

[19] Ghute, V. B. and Shirke, D. T. (2008b). A multivariate synthetic control chart for process dispersion. *Quality Technology and Quantitative Management,* 5(3): 271-288.

[20] Aparisi, F. and de Luna, M. (2009). The design and performance of the multivariate synthetic-T^2 control chart. *Communications in Statistics-Theory and Methods*, 38(2): 173-192.

[21] Rajmanya, S. V. and Ghute, V. B. (2014). A synthetic control chart for monitoring process variability. *Quality and Reliability Engineering International,* 30(8): 1301-1309.

[22] Rakitzis, A. C., Chakraborti, S., Shongwe, S. C., Graham, M. A. and Khoo, M. B. C. (2019). An overview of synthetic-type control charts, Techniques and methodology. *Quality and Reliability Engineering International,* 35(7): 2081-2096.

[23] Gadre, M. P. and Rattihalli, R. N. (2004). A group runs control chart for detecting shifts in the process mean. *Economic Quality Control*, 19(1): 29-43.

[24] Gadre, M. P., Rattihalli, R. N. and Shewade, A. V. (2005). Robustness of group runs control chart to non-normality. *Economic Quality Control*, 20(1): 81-95.

[25] Gadre, M. P. and Kakade, V. C. (2014). A nonparametric group runs control chart to detect shift in the process median. *IAPQR Transactions*, 39(1): 29-53.

[26] Gadre, M. P. and Kakade, V. C. (2016). Some group runs based multivariate control charts for monitoring the process mean vector. *Open Journal of Statistics*, 6(6): 1098.

[27] Chong, Z. L., Khoo, M. B., Lee, M. H. and Chen, C. H. (2017). Group runs revised m-of-k runs rule control chart. *Communications in Statistics-Theory and Methods*, 46(14): 6916-6935.

[28] Khilare, S. K. and Shirke, D. T. (2023). A Nonparametric Group Runs Control Chart for Location Using Sign Statistic. *Thailand Statistician*, 21(1): 165-179.

[29] Gadre, M. P. and Kakade, V. C. (2016). Some group runs based multivariate control charts for monitoring the process mean vector. *Open Journal of Statistics*, 6(6): 1098.

RT&A, No 4 (76)
Volume 18, December 2023

ON A CLASS OF LORENTZIAN PARA-KENMOTSU MANIFOLDS ADMITTING QUARTER-SYMMETRIC METRIC CONNECTION

S. Sunitha Devi[1], K. L. Sai Prasad[2]

•

Department of Mathematics
[1] Vardhaman College of Engineering, Hyderabad, Telangana, 501218, INDIA
[2] Gayatri Vidya Parishad College of Engineering for Women, Madhurawada, Visakhapatnam, 530 048, INDIA
sunithamallakula@yahoo.com[1,*] klsprasad@yahoo.com[2]

Abstract

In this present paper, a class of Lorentzian almost paracontact metric manifolds known as the LP-Kenmotsu (Lorentzian para-Kenmotsu) is considered that accepts a connection of quarter-symmetric. In this work, it was found that an LP-Kenmotsu manifold is either ϕ-symmetric or concircular ϕ-symmetric with respect to quarter-symmetric metric connection if and only if it is symmetric with respect to the Riemannian connection, provided the scalar curvature of Riemannian connection is constant.

Keywords: Lorentzian para-Kenmotsu manifold, quarter- symmetric metric connection, concircular curvature tensor.
2010 Mathematics Subject Classification: 53C05, 53C20, 53C50

I. Introduction

In 1989, Matsumoto [4] introduced the notion of Lorentzian paracontact and in particular, Lorentzian para-Sasakian (*LP*-Sasakian) manifolds. Later, these manifolds have been widely studied by many geometers such as Matsumoto and Mihai [5], Mihai and Rosca [6], Mihai, Shaikh and De [7], Venkatesha and C. S. Baggewadi [13], Venkatesha, Pradeep Kumar and Bagewadi [14, 15] and obtained several results on these manifolds.

In 1995, Sinha and Sai Prasad [11] defined a class of almost paracontact metric manifolds namely para-Kenmotsu (briefly *P*-Kenmotsu) and special para-Kenmotsu (briefly *SP*- Kenmotsu) manifolds in similar to *P*-Sasakian and *SP*- Sasakian manifolds. In 2018, Abdul Haseeb and Rajendra Prasad [1] defined a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly *LP*- Kenmotsu) manifolds. As an extension, many geometers have studied these Lorentzian para-Kenmotsu manifolds [8, 10, 12]. Sai Prasad, Sunitha Devi and Deekshitulu have considered *LP*-Kenmotsu manifolds admitting the Weyl-projective curvature tensor W_2 and shown that these manifolds admitting a Weyl-flat projective curvature tensor, an irrotational Weyl-projective curvature tensor and a conservative Weyl-projective curvature tensor are an Einstein manifolds of constant scalar curvature [9].

A linear connection $\widetilde{\nabla}$ in an n-dimensional differentiable manifold is said to be a quarter-symmetric connection [3] if its torsion tensor T is of the form

$$\begin{aligned} T(X,Y) &= \widetilde{\nabla}_X Y - \widetilde{\nabla}_Y X - [X,Y] \\ &= \eta(Y)\phi X - \eta(X)\phi Y, \end{aligned} \tag{1}$$

where η is a 1-form and ϕ is a tensor field of type (1,1). In particular, if we replace ϕX by X and ϕY by Y, then the quarter-symmetric connection reduces to the semi-symmetric connection [2]. Thus, the notion of quarter-symmetric connection generalizes the idea of semi-symmetric connection, and if quarter-symmetric linear connection $\widetilde{\nabla}$ satisfies the condition $\left(\widetilde{\nabla}_X g\right)(Y,Z)=0$ for all $X,Y,Z\in\chi(M_n)$, where $\chi(M_n)$ is the Lie algebra of vector fields on the manifold M_n, then $\widetilde{\nabla}$ is said to be a quarter-symmetric metric connection.

Motivated by these studies, in the present paper, we study the geometry of Lorentzian para-Kenmotsu (*LP*-Kenmotsu) manifolds with respect to quarter-symmetric metric connection. The present paper is organized as follows. Section 2 is equipped with some prerequisites about Lorentzian para-Kenmotsu manifolds.

Further on, in relation to the quarter-symmetric metric connection in an Lorentzian para-Kenmotsu manifold, we derive the relations for the Ricci tensor $\tilde{S}$ and the Riemannian curvature tensor $\tilde{R}$ in section 3.

Further in sections 4 and 5, we study the ϕ-symmetry and concircular ϕ-symmetry of Lorentzian para-Kenmotsu manifolds with respect to quarter-symmetric metric connection respectively.

II. Preliminaries

An n-dimensional differentiable manifold M_n admitting a (1,1) tensor field ϕ, contravariant vector field ξ, a 1-form η and the Lorentzian metric $g(X,Y)$ satisfying

$$\eta(\xi)=-1, \tag{2}$$
$$\phi^2X=X+\eta(X)\xi, \tag{3}$$
$$g(\phi X,\phi Y)=g(X,Y)+\eta(X)\eta(Y), \tag{4}$$
$$g(X,\xi)=\eta(X), \tag{5}$$
$$\phi\xi=0,\quad \eta(\phi X)=0,\quad rank\phi=n-1. \tag{6}$$

is called Lorentzian almost paracontact manifold [4].

In a Lorentzian almost paracontact manifold, we have

$$\Phi(X,Y)=\Phi(Y,X)\quad where\quad \Phi(X,Y)=g(\phi X,Y). \tag{7}$$

A Lorentzian almost paracontact manifold M_n is called Lorentzian para-Kenmotsu manifold if [1]

$$(\nabla_X\phi)Y=-g(\phi X,\ Y)\xi-\eta(Y)\phi X, \tag{8}$$

for any vector fields X and Y on M_n, and ∇ is the operator of covariant differentiation with respect to the Lorentzian metric g.

It can be easily seen that in a *LP*-Kenmotsu manifold M_n, the following relations hold [1]:

$$\nabla_X\xi=-\phi^2X=\ -X-\eta(X)\xi, \tag{9}$$
$$(\nabla_X\eta)Y=\ -g(X,\ Y)\xi-\ \eta(X)\ \eta(Y), \tag{10}$$

for any vector fields X and Y on M_n.

Also, in an *LP*-Kenmotsu manifold, the following relations hold [1]:

$$g(R(X,Y)Z,\xi)=\eta(R(X,Y)Z)=g(Y,Z)\eta(X)-g(X,Z)\eta(Y) \tag{11}$$

$$R(\xi, X)Y = g(X, Y)\xi - \eta(Y)X, \tag{12}$$

$$R(X, Y)\xi = \eta(Y)X - \eta(X)Y, \tag{13}$$

$$S(X, \xi) = (n-1)\eta(X), \tag{14}$$

$$S(\phi X, \phi Y) = S(X, Y) + (n-1)\eta(X)\eta(Y) \tag{15}$$

for any vector fields X, Y and Z, where R is the Riemannian curvature tensor and S is the Ricci tensor of M_n.

Definition 1. An LP- Kenmotsu manifold M_n is said to be symmetric if

$$(\nabla_W R)(X,\ Y)Z = 0, \tag{16}$$

for all vector fields X, Y, Z and W.

Definition 2. An LP-Kenmotsu manifold M_n is said to be ϕ-symmetric if

$$\phi^2 (\nabla_W R)(X,\ Y)Z = 0, \tag{17}$$

for all vector fields X, Y, Z and W.

Definition 3. An LP-Kenmotsu manifold M_n is said to be concircular symmetric if

$$\left(\nabla_W \tilde{C}\right)(X,\ Y)Z = 0, \tag{18}$$

for all vector fields X, Y and Z. Here $\tilde{C}$ is the concircular curvature tensor and is given by [16]

$$\tilde{C}(X,\ Y)Z = R(X, Y)Z - \frac{r}{n(n-1)}\left[g(Y, Z)X - g(X, Z)Y\right], \tag{19}$$

for all vector fields X, Y and Z, where R and r are the Riemannian curvature tensor and scalar curvature respectively.

Definition 4. An LP-Kenmotsu manifold M_n is said to be concircular ϕ-symmetric if

$$\phi^2 \left(\nabla_W \tilde{C}\right)(X, Y)Z = 0, \tag{20}$$

for all vector fields X, Y, Z and W.

III. Expression of $\tilde{R}(X, Y)Z$ in terms of $R(X, Y)Z$

In this section we express $\tilde{R}(X, Y)Z$, the curvature tensor with respect to quarter-symmetric metric connection, in terms of $R(X, Y)Z$ which is the curvature tensor with respect to Riemannian connection.

Let $\widetilde{\nabla}$ be a linear connection and ∇ be a Riemannian connection of an almost contact metric manifold M_n such that

$$\widetilde{\nabla}_X Y = \nabla_X Y + U(X, Y), \tag{21}$$

where U is a tensor of type (1, 1). For $\widetilde{\nabla}$ to be a quarter-symmetric metric connection in M_n, we have

$$U(X, Y) = \frac{1}{2}\left[T(X, Y) + T'(X, Y) + T'(Y, X)\right] \tag{22}$$

and

$$g(T'(X, Y), Z) = g(T(Z, X), Y). \tag{23}$$

From (1) and (23), we get

$$T'(X,Y) = \eta(Y)\phi X - g(\phi X, Y)\xi. \tag{24}$$

Using (1) and (24) in (22), we obtain

$$U(X,Y) = \eta(Y)\phi X - g(\phi X, Y)\xi.$$

Thus the quarter-symmetric metric connection $\widetilde{\nabla}$ in an *LP*-Kenmotsu manifold is given by

$$\widetilde{\nabla}_X Y = \nabla_X Y + \eta(Y)\phi X - g(\phi X, Y)\xi, \tag{25}$$

which is the relation between Riemannian connection and the quarter-symmetric metric connection on Lorentzian para-Kenmotsu manifolds.

Similarly, on simplication, we obtain the relation between the curvature tensor $\tilde{R}(X,Y)Z$ of M_n with respect to the quarter-symmetric metric connection $\widetilde{\nabla}$ and the curvature tensor $R(X,Y)Z$ of Riemannian connection ∇ as follows:

$$\begin{aligned}\tilde{R}(X,Y)Z = R(X,Y)Z &+ [g(\phi Y, Z) + g(Y,Z)\xi]\phi X] \\ &- [g(\phi X, Z) + g(X,Z)\xi]\phi Y \\ &+ Xg(\phi Y, Z) - Yg(\phi X, Z).\end{aligned} \tag{26}$$

Then from (26), it follows that

$$\tilde{S}(Y,Z) = S(Y,Z), \tag{27}$$

where $\tilde{S}$ and S are the Ricci tensors of the connections $\widetilde{\nabla}$ and ∇ respectively.

Further contracting (27), we get

$$\tilde{r} = r, \tag{28}$$

where $\tilde{r}$ and r are the scalar curvatures of the connections $\widetilde{\nabla}$ and ∇ respectively.

IV. Symmetry of *LP*-Kenmotsu manifold with respect to quarter-symmetric metric connection

By the definition of symmetric *LP*-Kenmotsu manifold with respect to Riemannian connection, we define a symmetric *LP*-Kenmotsu manifold with respect to quarter-symmetric metric connection by

$$\left(\widetilde{\nabla}_W \tilde{R}\right)(X,\ Y)Z = 0, \tag{29}$$

where

$$\begin{aligned}\left(\widetilde{\nabla}_W \tilde{R}\right)(X,\ Y)Z = \widetilde{\nabla}_W\left(\tilde{R}(X,\ Y)Z\right) &- \tilde{R}\left(\left(\widetilde{\nabla}_W X, Y\right)Z\right) \\ &- \tilde{R}\left(\left(X,\ \widetilde{\nabla}_W Y\right)Z\right) - \tilde{R}\left((X,Y)\widetilde{\nabla}_W Z\right),\end{aligned} \tag{30}$$

for all vector fields X, Y, Z and W.

$$\begin{aligned}\widetilde{\nabla}_W\left(\tilde{R}(X,\ Y)Z\right) = \nabla_W\left(\tilde{R}(X,\ Y)Z\right) &+ \eta\left(\tilde{R}(X,\ Y)Z\right)\phi W \\ &- g\left(\phi W,\ \left(\tilde{R}(X,\ Y)Z\right)\right)\xi,\end{aligned} \tag{31}$$

$$\begin{aligned}\tilde{R}\left(\left(\widetilde{\nabla}_W X, Y\right)Z\right) = \tilde{R}\left(\nabla_W X, Y\right)Z + \eta(X)\tilde{R}(\phi W, Y)Z \\ - g(\phi W,\ X)\tilde{R}(\xi,\ Y)Z,\end{aligned} \tag{32}$$

$$\begin{aligned}\tilde{R}\left(\left(X,\ \widetilde{\nabla}_W Y\right)Z\right) = \tilde{R}(X,\ \nabla_W Y)Z + \eta(Y)\tilde{R}(X,\ \phi W)Z \\ - g(\phi W,\ Y)\tilde{R}(X,\xi)Z,\end{aligned} \tag{33}$$

$$\begin{aligned}\tilde{R}\left((X,Y)\,\widetilde{\nabla}_W Z\right) =& \tilde{R}\,(X,Y)\,\nabla_W Z + \eta\,(Z)\,\tilde{R}\,(X,Y)\,\phi W \\ & g\,(\phi W,\ Z)\,\tilde{R}\,(X,\ Y)\,\xi,\end{aligned} \tag{34}$$

$$\tilde{R}(\xi,Y)Z = g(Y,Z)\xi - \eta(Z)Y - \eta(Z)\phi Y\xi + g(\phi Y,Z)\xi, \tag{35}$$

$$\tilde{R}(X,\xi)Z = \eta(Z)X - g(X,Z)\xi - \eta(Z)\phi X\xi + g(\phi X,Z)\xi, \tag{36}$$

$$\tilde{R}(X,Y)\xi = \eta(Y)X - \eta(X)Y + \eta(Y)\phi X\xi - \eta(X)\phi Y\xi. \tag{37}$$

By using (25), (31) to (37) in (30), we get

$$\begin{aligned}\left(\widetilde{\nabla}_W\tilde{R}\right)(X,\ Y)\,Z = & \left(\nabla_W\tilde{R}\right)(X,\ Y)\,Z + \eta\left(\tilde{R}\,(X,Y)\,Z\right)\phi W \\ & - g\left(\phi W,\ \tilde{R}\,(X,Y)\,Z\right)\xi - \eta\,(X)\,\tilde{R}\,(\phi W,Y)\,Z \\ & - \eta\,(Y)\,\tilde{R}\,(X,\phi W)\,Z + g\,(W,\ \phi X)\,\tilde{R}\,(\xi,Y)\,Z \\ & + g\,(W,\ \phi Y)\,\tilde{R}\,(X,\xi)\,Z + g\,(W,\ \phi Z)\,\tilde{R}\,(X,Y)\,\xi \\ & - \eta\,(Z)\,\tilde{R}\,(X,Y)\,\phi W.\end{aligned} \tag{38}$$

Then by differentiating (26) with respect to W and on using (6), (7) and (10), we get

$$\begin{aligned}(\nabla_W\tilde{R})(X,Y)Z = & (\nabla_W R)(X,Y)Z - [g(\phi W,Y)\eta(Z) + g(\phi W,Z)\eta(Y) \\ & + Wg(Y,Z) + \eta(W)g(Y,Z)\xi]\phi X \\ & + [g(\phi W,X)\eta(Z) + g(\phi W,Z)\eta(X) \\ & + Wg(X,Z) + \eta(W)g(X,Z)\xi]\phi Y \\ & + g(\phi X,Z)[g(\phi W,Y)\xi + \eta(Y)\phi W] \\ & - g(\phi Y,Z)[g(\phi W,X)\xi + \eta(X)\phi W] \\ & + [\eta(Y)g(X,Z)\xi - \eta(X)g(Y,Z)\xi]\phi W \\ & + [Yg(\phi W,X) - Xg(\phi W,Y)]\eta(Z) \\ & + g(\phi W,Z)[\eta(X)Y - \eta(Y)X].\end{aligned} \tag{39}$$

Therefore, by using (2), (8) and (39) in (38), we obtain

$$\left(\widetilde{\nabla}_W\tilde{R}\right)(X,\ Y)\,Z = (\nabla_W R)\,(X,\ Y)\,Z. \tag{40}$$

Thus we can state the following:

Theorem 1. An *LP*-Kenmotsu manifold is symmetric with quarter-symmetric metric connection $\widetilde{\nabla}$ if and only if it is symmetric with respect to Riemannian connection ∇.

Corollary 1. An *LP*-Kenmotsu manifold is ϕ-symmetric with respect to quarter-symmetric metric connection $\widetilde{\nabla}$ if and only if it is symmetric with respect to Riemannian connection ∇.

V. Concircular symmetry of *LP*-Kenmotsu manifold with respect to quarter-symmetric metric connection

An *LP*-Kenmotsu manifold M_n is said to be a concircular symmetric with respect to quarter-symmetric metric connection if

$$(\widetilde{\nabla}_W\widetilde{\widetilde{C}})(X,Y)Z = 0, \tag{41}$$

for all vector fields X, Y, Z and W. Here the concircular curvature tensor $\widetilde{\widetilde{C}}$ with respect to quarter-symmetric metric connection is given by

$$\widetilde{\widetilde{C}}(X,Y)Z = \tilde{R}(X,Y)Z - \frac{\widetilde{r}}{n(n-1)}[g(Y,Z)X - g(X,Z)Y], \tag{42}$$

where $\widetilde{R}$ is the Riemannian curvature tensor and $\widetilde{r}$ is the scalar curvature of the quarter-symmetric metric connection $\widetilde{\nabla}$.

Using (25), we can write

$$\begin{aligned}(\widetilde{\nabla}_W\widetilde{\widetilde{C}})(X,Y)Z &= (\nabla_W\widetilde{\widetilde{C}})(X,Y)Z+\eta(\widetilde{\widetilde{C}}(X,Y)Z)\phi W\\ &\quad -g(\phi(\widetilde{\widetilde{C}}(X,Y)Z,W))\xi-\eta(X)\widetilde{\widetilde{C}}(\phi W,Y)Z\\ &\quad -\eta(Y)\widetilde{\widetilde{C}}(X,\phi W)Z-\eta(Z)\widetilde{\widetilde{C}}(X,Y)\phi W\\ &\quad +g(\phi W,X)\widetilde{\widetilde{C}}(\xi,Y)Z+g(\phi W,Y)\widetilde{\widetilde{C}}(X,\xi)Z\\ &\quad +g(\phi W,Z)\widetilde{\widetilde{C}}(X,Y)\xi.\end{aligned} \tag{43}$$

Now on differentiating (42) with respect to W, we obtain

$$(\nabla_W\widetilde{\widetilde{C}})(X,Y)Z=(\nabla_W\widetilde{R})(X,Y)Z-\frac{\nabla_W\widetilde{r}}{n(n-1)}[g(Y,Z)X-g(X,Z)Y]. \tag{44}$$

Therefore, by using of (28) and (39), we get from (44) that

$$\begin{aligned}(\nabla_W\widetilde{\widetilde{C}})(X,Y)Z &= (\nabla_WR)(X,Y)Z-[g(\phi W,Y)\eta(Z)+g(\phi W,Z)\eta(Y)\\ &\quad +Wg(Y,Z)+\eta(W)g(Y,Z)\xi]\phi X+[g(\phi W,X)\eta(Z)\\ &\quad +g(\phi W,Z)\eta(X)+Wg(X,Z)+\eta(W)g(X,Z)\xi]\phi Y\\ &\quad +g(\phi X,Z)[g(\phi W,Y)\xi+\eta(Y)\phi W]\\ &\quad -g(\phi Y,Z)[g(\phi W,X)\xi+\eta(X)\phi W]\\ &\quad +[\eta(Y)g(X,Z)\xi-\eta(X)g(Y,Z)\xi]\phi W\\ &\quad +[Yg(\phi W,X)-Xg(\phi W,Y)]\eta(Z)\\ &\quad +g(\phi W,Z)[\eta(X)Y-\eta(Y)X]\\ &\quad -\frac{\nabla_Wr}{n(n-1)}[g(Y,Z)X-g(X,Z)Y].\end{aligned} \tag{45}$$

Then by making use of (19), we rewrite the above equation (45) as

$$\begin{aligned}(\nabla_W\widetilde{\widetilde{C}})(X,Y)Z &= (\nabla_W\widetilde{C})(X,Y)Z-[g(\phi W,Y)\eta(Z)+g(\phi W,Z)\eta(Y)\\ &\quad +Wg(Y,Z)+\eta(W)g(Y,Z)\xi]\phi X+[g(\phi W,X)\eta(Z)\\ &\quad +g(\phi W,Z)\eta(X)+Wg(X,Z)+\eta(W)g(X,Z)\xi]\phi Y\\ &\quad +g(\phi X,Z)[g(\phi W,Y)\xi+\eta(Y)\phi W]\\ &\quad -g(\phi Y,Z)[g(\phi W,X)\xi+\eta(X)\phi W]\\ &\quad +[\eta(Y)g(X,Z)\xi-\eta(X)g(Y,Z)\xi]\phi W\\ &\quad +[Yg(\phi W,X)-Xg(\phi W,Y)]\eta(Z)\\ &\quad +g(\phi W,Z)[\eta(X)Y-\eta(Y)X].\end{aligned} \tag{46}$$

Using (2), (6) and (46) in (43), we get

$$(\widetilde{\nabla}_W\widetilde{\widetilde{C}})(X,Y)Z=(\nabla_W\widetilde{C})(X,Y)Z. \tag{47}$$

Hence we can state the following:

Theorem 2. An *LP*-Kenmotsu manifold is concircular symmetric with respect to $\widetilde{\nabla}$ if and only if it is so with respect to Riemannian connection ∇.

Corollary 2. An *LP*-Kenmotsu manifold is concircular ϕ-symmetric with respect to $\widetilde{\nabla}$ if and only if it is so with respect to Riemannian connection ∇.

Now taking (2), (6) and (45) in (43), we get

$$(\widetilde{\nabla}_W\widetilde{\widetilde{C}})(X,Y)Z = (\nabla_W R)(X,Y)Z - \frac{\nabla_W r}{n(n-1)}[g(Y,Z)X - g(X,Z)Y]. \tag{48}$$

If scalar curvature r is constant, the above equation (48) reduces to

$$(\widetilde{\nabla}_W\widetilde{\widetilde{C}})(X,Y)Z = (\nabla_W R)(X,Y)Z. \tag{49}$$

Thus we have the following assertion.

Theorem 3. An *LP*-Kenmotsu manifold is concircular symmetric with respect to quarter-symmetric metric connection $\widetilde{\nabla}$ if and only if it is symmetric with respect to Riemannian connection ∇, provided the scalar curvature r is constant.

Corollary 3. An *LP*-Kenmotsu manifold is concircular ϕ-symmetric with respect to quarter-symmetric metric connection $\widetilde{\nabla}$ if and only if it is symmetric with respect to Riemannian connection ∇, provided the scalar curvature r is constant.

VI. Conclusion

We explore a class of Lorentzian almost paracontact metric manifolds known as the Lorentzian para-Kenmotsu that accepts a quarter-symmetric connection. In relation to the quarter-symmetric metric connection, the relations for the Ricci tensor and the Riemannian curvature tensor in a Lorentzian para-Kenmotsu manifold were derived. Further, it was found that an *LP*-Kenmotsu manifold is either ϕ-symmetric or concircular ϕ-symmetric with respect to quarter-symmetric metric connection if and only if it is symmetric with respect to the Riemannian connection, provided the scalar curvature of Riemannian connection is constant. The paper ends with a handful of bibliography.

Acknowledgements: The authors acknowledge Dr. A. Kameswara Rao, Assistant Professor of G. V. P. College of Engineering for Women for his valuable suggestions in preparation of the manuscript.

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Abdul Haseeb and Rajendra Prasad. Certain results on Lorentzian para-Kenmotsu manifolds, Bulletin of Parana's Mathematical Society. (2018) doi.10.5269/bspm.40607.

[2] Friedmann, A. and Schouten, J. A. Uber die Geometric der halbsymmetrischen Ubertragung, Math. Zeitschr. 21(1924), pp.211-223.

[3] Golab, S. On semi-symmetric and Quarter-symmetric linear connections, Tensor, N. S. 29(1975), pp. 249-254.

[4] Matsumoto, K. On Lorentzian paracontact manifold, Bulletin of the Yamagata University Natural Science. 12(2)(1989), pp. 151-156.

[5] Matsumoto, K. and Mihai, I. On a certain transformation in a Lorentzian para-Sasakian manifold, Tensor, N.S. 47(1988), pp.189-197.

[6] Mihai, I. and Rosca, R. On Lorentzian p-Sasakian manifolds, Classical Analysis, World Scientific Publ. Singapore. (1992) pp.155-169.

[7] Mihai, I. Shaikh, A. A. and De, U. C. On Lorentzian para-Sasakian manifolds, Rendiconti del Seminario Matematico di Messina. Serie II. (1999).

[8] Rajendra Prasad, Abdul Haseeb and Umesh Kumar Gautam, On ϕ-semisymmetric *LP*-Kenmotsu manifolds with a QSNM-connection admitting Ricci solitons, Kragujevac Journal of Mathematics. 45 (5)(2021), pp. 815-827.

[9] Sai Prasad, K. L. Sunitha Devi, S. and Deekshitulu, G. V. S. R. On a class of Lorentzian para-Kenmotsu manifolds admitting the Weyl-projective curvature tensor of type (1,3), Italian Journal of Pure and Applied Mathematics. 45(2021), pp. 990-1001.

[10] Sai Prasad, K. L. Sunitha Devi, S. and Deekshitulu, G. V. S. R. On a class of Lorentzian paracontact metric manifolds, Italian Journal of Pure and Applied Mathematics. 49(1)(2023), pp. 514-527.

[11] Sinha, B. B. and Sai Prasad, K. L. A class of almost paracontact metric manifold, Bulletin of the Calcutta Mathematical Society. 87(1995), pp.307-312.

[12] Sunitha Devi, S. Sai Prasad, K. L. and Satyanarayana, T. Certain curvature conditions on Lorentzian para-Kenmotsu manifolds, Reliability- Theory & Applications. Vol. 17, 2(68)(2022), pp. 413-421.

[13] Venkatesha and Bagewadi, C. S. On concircular ϕ-recurrent LP-Sasakian manifolds, Differ. Geom. Dyn. Syst. 10(2008), pp. 312-319.

[14] Venkatesha, Pradeep Kumar and Bagewadi, C. S. On Quarter-symmetric metric connection in a Lorentzian para-Sasakian manifold, Ajerbaijan Journal of Mathematics. I.(2015), pp. 3-12.

[15] Venkatesha, Pradeep Kumar and Bagewadi, C. S. On Lorentzian para-Sasakian manifolds satisfying W_2 curvature tensor, IOSR J. of Mathematics. 9(6)(2014), pp. 124-127.

[16] Yano, K. Concircular geometry I Concircular transformations, proc. Imp. Acad, Tokyo. 16(1970), pp. 195-200.

EVALUATION OF SAMPLE SIZE AND EFFICIENT FIELD SAMPLING PLAN IN HDP APPLE ORCHARDS

Tabasum Mushtaq[1], Mushtaq A. Lone[2], S. A. Mir[3], Sonali Kedar Powar[4], Aafaq A. Rather[5,*], Adil H. Khan[6], Faizan Danish[7]

•

[1,2,3]FOH, SKUAST –K, Jammu and Kashmir, India
[4]Department of Computer Science, Faculty of Science & Technology, Vishwakarma University, Pune
[5, *]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[6]Department of Statistics, University of Kashmir, Srinagar-190006, J&K, India
[7]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Amravati, Andhra Pradesh-522237, India
[1]tabumushtaq2010@gmail.com, [2]mushtaqstat11@gmail.com, [3]mir_98@msn.com,
[4]sonali.powar@vupune.ac.in, [5,*]aafaq7741@gmail.com, [6]khanadil_192@yahoo.com,
[7]danishstat@gmail.com

Abstract

An essential stage in research is choosing an adequate sample size and sampling strategy. In order to obtain the most accurate estimates possible when surveying high density apple orchards, this paper provides the proper procedure for selecting the sample and an effective sampling strategy. For this study, primary information gathered during a two-year period from the SKUAST-Kashmir exotic apple block Plate I was employed. This investigation was conducted using the TCSA of exotic apple trees of the Gala and Fuji types. The sample was obtained using a variety of sampling techniques in order to find the parameters of population. Findings revealed that using proportional allocation of a stratified sample technique, in both the varieties, produces the most efficient population parameter estimates.

Keywords: Sample size, Stratification, Proportional Allocation, Gain in Efficiency.

1. Introduction

After data have been gathered, the most crucial goal of data analysis is to make generalizations about the population based on sample data. One of the most commonly requested questions by the investigators is "How big a sample is necessary." Inferences from the research cannot accurately reveal the reality of entire population when the sample size isn't calculated properly. The accuracy and caliber of research are impacted by inappropriate, insufficient, or excessive sample numbers. Choosing a sample size and addressing non-response bias are crucial in a quantitative survey design. The ability of quantitative methods to draw conclusions about vast populations that would be too expensive to research using smaller populations is one of their significant advantages [1]. Sample size is one of the four associated characteristics of a survey design that can impact the recognition of important disparity, correlations, or relations [2]. Precision, risk, and variability in the variables assessed are typically three criteria that need to be specified besides the research's objective and size of population to allot the right size of sample.

The "degree of precision" is used to find out the size of sample. The permitted discrepancy between the estimated value and the population value is known as the "degree of precision." Particularly, it is a measurement of how closely an estimate resembles the population's real distribution of a property [3]. You may refer to the degree of precision as sampling error. Cochran asserts that the level of mistakes that one is ready to accept in the sample estimates can be used to achieve the required precision. The SE is the discrepancy between the statistic & the relevant parameter of population [4]. It depends on how much hazard a researcher is eager to take when the data is used to inform their choices. Frequently it is stated as a percentage. If the degree of uncertainty for sampling is ±5 (%), and 70 (%) of the sample members ascribe a certain criterion, it can be inferred that 65% to 75% of members of the population also attribute that criterion. Higher expense and greater sample sizes are necessary for high levels of precision [5].

For the level of confidence or risk level the Central Limit Theorem's provides the basis [6]. Approximately 95% of the sample values in a normal distribution are within two standard deviations of the mean, or the real value of the population.

The degree of variability in the attributes being measured is referred to the distribution of the traits in the population. The sample size increases with population heterogeneity in order to achieve a particular degree of precision. Smaller sample sizes are ideal for populations that are less variable (more homogeneous). Remember that a proportion of 50% denotes a higher degree of variability than one of 20% or 80%. In order to calculate more cautious sample size, a proportion of 0.5 is frequently employed because it represents the population's maximum variability.

The choice of an appropriate survey plan, is a key component of the collection of information in any decision that is supported by science. A solid sampling strategy is essential to assure; the data are adequate to support the necessary inferences. It's critical to have reliable information while making science-based decisions. Creating a sample plan that appropriately reflects the issue under investigation is essential for obtaining valid and reliable data. Sampling is crucial in survey designs necessitating the human population and is receiving more and more heed from sociologists, pharmacologists, designers, accountants, physiologists, and medical professionals involved in business dealings, as well as from those in the fields of education, public administration, biostatistics, and even sociology, accounting, economics, anthropology, and political science [8]. Sampling plans are provided in order to end in fundamental investigation challenges, particularly in the social science disciplines and their uses [9]. In order to bring out interpretations about the population value from the sample value, the sampling plan also incorporates the assignments of estimate & selection [10]. With error-free measures, there are issues when extrapolating survey results from one group to another and often a larger population. The standard errors would differ depending on the sample designs. The main goal of sampling design is to choose the least error-prone option. An efficient sampling approach within a population display an adequate elicitation of relevant data, giving a significant understanding of the key elements of the population [7]. A sampling method that is cost-effective, effortless to use, bring in fair estimates, and lessens the consequences of sample-related volatility is therefore required. Multiple demographic parameters are frequently estimated in sample surveys; these parameters may be mutually exclusive. For the purpose of ensuring that samples represent all relevant points of view, stratified sampling has been established. The whole diversified population of interest is divided into homogenous classes, sometimes called as strata, each of which is similar within itself, using the stratification method of sample design. Depending on how important each stratum is to the population as a whole, the sample size changes for each stratum. Then, individual stratum sampling will be carried out [11]. There are several uses for stratified sample designs. These include obtaining more accurate estimates for fascinating domains. They also include increasing the true value of estimates for the entire population being gathered in the investigation [12]. The purpose of the experimentation is to establish sample size and choose the best sampling plan for surveys of apple orchards with a high population

density. For this investigation, the high-density apple block (HDP, plate-1) of SKUAST-primary Kashmir's data on a variety of tree/fruit properties were employed. Gala and Fuji were two types of high-density apple trees selected in the study based on trunk cross-sectional area.

2. Methodology

According to Cochran (1977), the sample size n can be estimated by mentioning the margins of error for the survey items that are thought to be most important. For each of these significant factors, an estimation of the required sample size is first made. The data can be used to determine the sampling strategy. The largest can be used as the sample size if the values for the variables of interest are close together, and one can be sure that the sample size will produce the desired findings. Keeping in mind that a sample size that is too small will result in more estimation errors for demographic parameters and a sample size that is too large will result in higher survey costs to calculate the necessary sample size, a compromise between estimation accuracy and cost has to be made. Cochran's formula for sample size is:

$$n_0 = \frac{z^2\sigma^2}{e^2} \tag{1}$$

Where n_o is the sample size (without fpc factor), z = value for selected alpha level (1.96), σ = standard deviation = 0.17 and e= margin of error = 0.05.

Hence, for the population of 270, the calculated sample size is 47.704. But the sample size surpasses 5% of the population (270*0.05=13.5), Cochran's [4]. Correction formula must be incorporated to evaluate the optimum sample size. Put in the fpc factor out-turn in optimum sample size n, enumerated as in formula:

$$n = \frac{n_0 N}{n_0 + (N-1)} \tag{2}$$

$$n = \frac{47.704 * 270}{47.704 + (270-1)} = 41 \tag{3}$$

Where population size =270, n_o = required return sample size = 47.704, n_o = optimum return sample size.

Therefore, by using the Cochran's 1977 formula, the optimum sample size selected for study is 41 with the margin of error set up to 5% (0.05) and an alpha level of 1.96. In our study, sample selection methods included both stratified random sampling and plain random sampling. The TCA of each type of exotic apple plant was divided into three strata (TCA). The following guiding principles were followed in the stratification process: I The strata, or range of TCAs, do not overlap and make up the entire population as a whole. (ii) Within each stratum, or range of TCA, there is homogeneity [13].

The proportional allocation method was used to allocate the samples throughout the several TCA ranges. The sample fraction, n/N, in this approach is constant throughout all strata. With the help of this allocation, a sample that can estimate sample size more quickly and precisely was created. The allocation of a given sample of size n to different stratum was done in proportion to their sizes i.e. in the i^{th} stratum, $n_i = n\frac{N_i}{N}$ where n represents sample size, N_i refers to population size of the i^{th} strata and N is the size of population. In this investigation, population size (N) = 270; sample size (n) = 41

The variance of estimate of population mean in StRS (proportional allocation) and SRS (*wor*) is given by

$$Var(\overline{y_{st}})_{Prop} = \left(\frac{1}{n} - \frac{1}{N}\right)\sum_{i=1}^{k} W_i S_i^2 \tag{4}$$

$$Var(\overline{y}_n)_{Rand} = \left(\frac{1}{n} - \frac{1}{N}\right) S^2 \tag{5}$$

$$where \qquad S^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Y_{ij} - \overline{Y_N}\right)^2 \tag{6}$$

$$S^2 = \sum_{i=1}^{k}(N_i - 1)S_i^{\,2} + \sum_{i=1}^{k} N_i\left(\overline{Y}_{Ni} - \overline{Y}_N\right)^2 \tag{7}$$

The sample size in every stratum proportionate allocation differs relying on the total quantity of plants in each classified TCA, for duo exotic varieties. The quantitative values are SHOWN in Tables 1 and 2.

Table 1: *Grouping of TCA for Gala and Fuji: an empirical study*

Variety	Strata	N_i	n_i	$\overline{Y}_{Ni}$	$\overline{Y}_{Ni}{}^2$	$N_i\overline{Y}_{Ni}$	$N_i\overline{Y}_{Ni}{}^2$	S_i	$S_i{}^2$	N_iS_i	$N_iS_i{}^2$	$(N_i-1)S_i{}^2$
Gala	1	77	1	6.03	36.36	464.31	2799.79	0.86	0.72	66.22	56.95	56.21
	2	150	3	8.88	78.85	1332	11828.16	1.32	1.724	198.00	261.36	259.62
	3	43	07	12.78	163.33	549.54	7023.12	1.48	2.19	63.64	94.19	92.00
	Total	270	41	27.69	278.54	2345.85	21651.07	3.66	4.67	327.86	412.50	407.82
Fuji	1	48	07	3.16	9.99	151.68	479.31	0.98	0.96	47.04	46.10	72.99
	2	153	23	6.45	41.60	986.85	6365.18	1.81	3.28	276.93	501.24	488.14
	3	69	11	10.69	114.28	737.61	7885.05	2.27	5.15	156.63	355.55	216.42
	Total	270.00	41	20.30	165.86	1876.14	14729.54	5.06	9.39	480.60	902.89	777.55

Table 2: *Calculation of variances*

Sampling Strategy	HDP Variety	
	Gala	Fuji s
$Var(\overline{y}_n)_{Rand}$	0.128	0.189
$Var(\overline{y_{st}})_{Prop}$	0.032	0.069

Table 2 lists variations found in Gala and Fuji as a result of SRS &. StRS. Gala and Fuji hold variances of 0.118 and 0.189 under SRS, respectively, and .032 and .069 under StRS with proportionate allocation.

It is essential to examine the gain in efficiency (GE) brought on by separate kinds of allocations in order to determine how they affect sample size [16]. The efficiency improvement attributed to proportional allocation over SRS (wor) is given in equation below

$$(E-1) = \frac{Var(\overline{y}_n)_{Rand} - Var(\overline{y_{st}})_{Prop}}{Var(\overline{y_{st}})_{Prop}} \tag{8}$$

Table 3: *Gain in Efficiency and Percentage gain in Efficiency Due to Stratification*

Variety	Gain in Efficiency	% Gain in Efficiency
Gala	3.05	305
Fuji	1.34	134

According to Table 3, the efficiency gain (GE) for Gala and Fuji is 3.05 (305%) and 1.34 (134%) correspondingly.

3. Conclusion

The purpose of sample size formulas "is not merely to give a samples number," rather than to determine whether tens, hundreds, or thousands of plants are needed, but to scrutinize the study design, including a review of the validity and reliability of data collecting. Finally, it is concluded that proportional allocation in stratified sampling provides the best accurate estimates over simple random sampling for population parameters estimation in both of the HDP exotic apple types under investigation. This has been determined after using different sampling plans to obtain the sample statistic for population parameters.

References

[1] Singh, H., Caste and Premarket, (2022). Discrimination: Access to Civic Amenities and Healthcare Facilities in Rural Punjab, *Contemporary Voice of Dalit*, p. 2455328X2211069, doi: 10.1177/2455328X221106908.

[2] Gupta, S., Haq, A., and Varshney, R., (2022). Problem of Compromise Allocation in Multivariate Stratified Sampling Using Intuitionistic Fuzzy Programming, *Annals of Data Science 2022*, pp. 1–20, doi: 10.1007/S40745-022-00410-Y.

[3] Li, C., Wang, J., Wang, L. Hu, L. and Gong, P., (2014). Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery, *Remote Sensing 2014, Vol. 6, Pages 964-983*, vol. 6, no. 2, pp. 964–983, doi: 10.3390/RS6020964.

[4] Varshney, R., Khan, M. G. M., Fatima, U., and Ahsan, M. J., (2015). Integer compromise allocation in multivariate stratified surveys, *Ann Oper Res*, vol. 226, no. 1, pp. 659–668, doi: 10.1007/S10479-014-1734-Z/TABLES/2.

[5] Miranda, C., Santesteban, L. G., Urrestarazu, J., Loidi, M. and Royo, J. B., (2018). Sampling Stratification Using Aerial Imagery to Estimate Fruit Load in Peach Tree Orchards, *Agriculture 2018, Vol. 8, Page 78*, vol. 8, no. 6, p. 78, doi: 10.3390/AGRICULTURE8060078.

[6] McCravy, K. W., (2018). A Review of Sampling and Monitoring Methods for Beneficial Arthropods in Agroecosystems, *Insects 2018, Vol. 9, Page 170*, vol. 9, no. 4, p. 170, doi: 10.3390/INSECTS9040170.

[7] Varshney, R., Najmussehar, Ahsan, M. J., Varshney, R. and Ahsan, M. J., (2011). An optimum multivariate stratified double sampling design in presence of non-response, *Optimization Letters 2011 6:5*, vol. 6, no. 5, pp. 993–1008, doi: 10.1007/S11590-011-0329-8.

[8] Peers, I., (2006). Statistical Analysis for Education and Psychology Researchers: Tools for researchers in education and psychology, *Statistical Analysis for Education and Psychology Researchers*, doi: 10.4324/9780203985984.

[9] Ranjan, R., Sinha, R., Khot, L. R., Hoheisel, G. A., Grieshop, M., and Ledebuhr, M., (2021). Spatial Distribution of Spray from a Solid Set Canopy Delivery System in a High-Density Apple Orchard Retrofitted with Modified Emitters, *Applied Sciences 2021, Vol. 11, Page 709*, vol. 11, no. 2, p. 709, doi: 10.3390/APP11020709.

[10] Mukherjee, S., Baral, M. M., Pal, S. K., Chittipaka, V., Roy, R., and Alam, K., (2022). Humanoid robot in healthcare: A Systematic Review and Future Research Directions," *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, pp. 822–826, doi: 10.1109/COM-IT-CON54601.2022.9850577.

[11] Roy, R., Baral, M. M., Pal, S. K., Kumar, S., Mukherjee S., and Jana, B., (2022). Discussing the present, past, and future of Machine learning techniques in livestock farming: A systematic literature review, *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, pp. 179–183, doi: 10.1109/COM-IT-CON54601.2022.9850749.

[12] Vishwakarma, D. K., Kumar, R., Kumar, A., Kushwaha, N. L., Kushwaha, K. S., and Elbeltagi, A., (2022). Evaluation and development of empirical models for wetted soil fronts under drip irrigation in high-density apple crop from a point source," *Irrig Sci*, vol. 1, pp. 1–24, doi: 10.1007/S00271-022-00826-7/FIGURES/17.

[13] Ranjan, R., Sinha, R., Khot, L. R., Hoheisel, G. A., Grieshop, M. J. and Ledebuhr, M., (2021). Effect of Emitter Modifications on Spray Performance of a Solid Set Canopy Delivery System in a High-Density Apple Orchard, *Sustainability 2021, Vol. 13, Page 13248*, vol. 13, no. 23, p. 13248, doi: 10.3390/SU132313248.

RT&A, No 4 (76)
Volume 18, December 2023

Estimation of Stress-Strength Reliability Based on KME Model

Kavya P. and Manoharan M.

•

University of Calicut
kavyapnair90@gmail.com, manumavila@gmail.com

Abstract

In reliability theory the estimation of stress-strength reliability is an important problem. It has many applications in engineering and physics areas. In many practical situations, the assumption of identical strength distributions may not be quite realistic because components of a system are of different structure. Here we establish the estimation of stress-strength reliability of the KM-Exponential (KME) distribution. In this article, we consider the case that the stress-strength variables are independent. KME distribution is parsimonious in parameter and has decreasing failure rate. The stress-strength reliability based on KME model is estabished and using maximum likelihood estimation method, the estimation of the stress-strength reliability is derived and also the asymptotic distribution. Simulation method is used to show the performance of the parameters and the 95% confidence interval is also calculated. With the help of simulated data, we depicts the application of the stress-strength reliability of KME distribution.

Keywords: KM-Exponential Stress-Strength reliability Estimation Simulation study.

1. Introduction

The problem of estimation of stress-strength reliability has great attention in reliability theory. The term stress is defined as a failure inducing variable. That means the stress (load) which tends to produce a failure of a component or of a device of a material. For example, environment, pressure, load, velocity, resistance, temperature, humidity, vibrations, and voltage etc. The term strength is defined as it is failure resisting variable. The ability of component, device or a material to accomplish its required function (mission) satisfactorily without failure when subjected to the external loading and environment.

The stress-strength reliability model depicts the life of a component or item with a random strength X and is subjected to a random stress Y. If the stress on the component surpasses the strength, it fails instantaneously. Whenever $Y < X$ the item functions satisfactorily. The component reliability is defined as

$$R = P(Y < X) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f(x,\ y)\ dy\ dx,$$

where $f(x,\ y)$ is the joint pdf of X and Y. Suppose the random variable X and Y are independent, then R can be written as

$$R = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f(x)\ g(y)\ dy\ dx,$$

where $f(x)$ and $g(y)$ are the marginal pdfs of X and Y. This is also can be written as

$$R = \int_{-\infty}^{\infty} f(x)\ G_y(x)\ dx.$$

where $G_y(x)$ is the cdf of $g(y)$.

The germ of this idea was proposed by Birnbaum [1] and was developed by Birnbaum and McCarty [2]. The formal term stress-strength firstly appears in the title of Church and Harris [3]. Based on certain parametric assumptions regarding X and Y, the first attempt to study R was undertaken by Owen et al. [4]. They also calculated the confidence interval for R when X and Y are independent or dependent normally distributed random variables. The estimation of R for major distributions like normal (Church and Harris [3], Downton [5];, Woodward and Kelley [6]), exponential (Kelly et al [7], Tong [8]), Pareto (Beg and Singh [9]), and exponential families (Tong [10]) was derived by the end of seventies. Enis and Geisser [11] contribute the Bayes estimation of R for exponentially or normally distributed X and Y. The other major works of the seventies include the introduction of a non-parametric empirical Bayes estimation of R by Ferguson [12] and Hollander and Korwar [13], and the study of system reliability (Bhattacharya and Johnson [14]).

Both stress and strength depend on some known covariates, Guttman et al. [15] and Weerahandi and Johnson [16] discussed the estimation and associated confidence interval of R. Using Bayesian approach Sun et al. [17] estimated the stress-strength reliability. Raqab and Kundu [18] carried out the estimation of stress-strength reliability, when Y and X two independent scaled Burr type X distribution. A comprehensive treatment of the different stress-strength models till 2001 can be found in the excellent monograph by Kotz et al. [19]. Some of the work on the estimation of stress-strength reliability can be obtained in Kundu and Gupta ([20], [21]), Kundu and Raqab [22], Krishnamoorthy et al. [23], Raqab et al. [24], Rezaei et al. [25], and Baklizi [26]. Baklizi and Eidous [27] introduced an estimator of stress-strength reliability based on kernel estimators. Estimation of stress-strength reliability using empirical likelihood method was studied by Jing et al. [28].

Basirat et al. [29] studied the estimation of stress-strength parameter using record values from proportional hazard model. Estimation of stress-strength reliability based on the generalized exponential distribution was developed by Asgharzadeh et al.[30]. Bai et al. [31] considered reliability inference of stress-strength model under progressively Type-II censored samples when stress and strength have truncated proportional hazard rate distributions. Bi and Gui [32] derived Bayesian estimation of R using inverse Weibull distribution. Ghitany et al. [33] discussed inference on stress-strength reliability based on power Lindley distribution. Sharma [34] proposed an upside-down bathtub shape distribution and estimate of stress-strength reliability of inverse Lindley distribution.

This paper is organized as follows. Preliminaries of the KME model are given in Section 2. In Section 3 the stress-strength reliability for the KME model is derived. The estimation of stress-strength reliability R is explained in Section 4. In Section 5 the asymptotic distribution and confidence interval are given. Simulation study and applications are discussed in Section 6 and Section 7 respectively. Finally we concluded the present work in Section 8.

2. Preliminaries of KME model

We obtain the KME model using the cumulative distribution function (cdf) of exponential distribution in KM transformation given in Kavya and Manoharan [36]. The probability distribution function (pdf) and cdf of the KME distribution are

$$f(x) = \frac{\lambda e^{-\lambda x} e^{e^{-\lambda x}}}{e-1}, \quad x > 0, \ \ \lambda > 0,$$
$$F(x) = \frac{e}{e-1}[1 - e^{-(1-e^{-\lambda x})}], \quad x > 0, \ \ \lambda > 0,$$

3. Stress-strength reliability based on the model

The stress-strength reliability model depicts the life of a component or item with a random strength X and is subjected to a random stress Y. If the stress on the component surpasses the strength, it fails instantaneously. Whenever $Y < X$ the item functions satisfactorily. The component reliability is defined as $R = P(Y < X)$. It has applications in engineering fields such as failure of aircraft structures, deterioration of rocket motors, and the aging of concrete pressure vessels.

Suppose X and Y are two independent random variables. If $X \sim \text{KME}(\lambda_1)$ and $Y \sim \text{KME}(\lambda_2)$, then the stress-strength reliability is obtained as

$$\begin{aligned} R = P(Y < X) = &\int_0^\infty \frac{\lambda_1 e^{-\lambda_1 x} e^{e^{-\lambda_1 x}}}{e-1}\left[\frac{e}{e-1}\left(1 - e^{-(1-e^{-\lambda_2 x})}\right)\right]dx \\ = &\frac{\lambda_1 e}{(e-1)^2}\sum_{m=0}^{\infty}\frac{1}{m!}\Big[I_1 + I_2\Big] \end{aligned} \tag{1}$$

where $I_1 = \int_0^\infty e^{-\lambda_1 x(m+1)}dx$ and $I_2 = \int_0^\infty e^{-\lambda_1 x(m+1)}e^{-(1-e^{-\lambda_2 x})}dx$. After integration, we get the values of $I_1 = \frac{1}{\lambda_1(m+1)}$ and $I_2 = \sum_{n=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^{n+i}}{n!}\binom{n}{i}\frac{1}{\lambda_1(m+1)+\lambda_2 i}$. Substituting these values in (1), the stress-strength reliability based on the KME model is obtained as

$$\begin{aligned} R &= P(Y < X) \\ &= \frac{\lambda_1 e}{(e-1)^2}\sum_{m=0}^{\infty}\frac{1}{m!}\left[\frac{1}{\lambda_1(m+1)} - \sum_{n=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^{n+i}}{n!}\binom{n}{i}\frac{1}{\lambda_1(m+1)+\lambda_2 i}\right]. \end{aligned} \tag{2}$$

4. Estimation of R

Suppose we drawn a random sample $x_1, x_2, ..., x_n$ of size p from KME(λ_1) and $y_1, y_2, ..., y_n$ of size q from KME(λ_2). The likelihood function is obtained as

$$L = \left(\frac{1}{e-1}\right)^p \lambda_1^p e^{-\lambda_1\sum_{i=1}^p x_i} e^{\sum_{i=1}^p e^{-\lambda_1 x_i}} \left(\frac{1}{e-1}\right)^q \lambda_2^q e^{-\lambda_2\sum_{j=1}^q y_j} e^{\sum_{j=1}^q e^{-\lambda_2 y_j}} \tag{3}$$

The log likelihood function is

$$\begin{aligned} \log\ L = &p\log(\frac{1}{e-1}) + p\log(\lambda_1) - \lambda_1\sum_{i=1}^p x_i + \sum_{i=1}^p e^{-\lambda_1 x_i} \\ &q\log(\frac{1}{e-1}) + q\log(\lambda_2) - \lambda_2\sum_{j=1}^q y_j + \sum_{j=1}^q e^{-\lambda_2 y_j} \end{aligned} \tag{4}$$

The partial derivatives of the log likelihood function with respect to λ_1 and λ_2 are

$$\frac{\partial \log\ L}{\partial \lambda_1} = \frac{p}{\lambda_1} - \sum_{i=1}^p x_i\left(1 + e^{-\lambda_1 x_i}\right)$$

and

$$\frac{\partial \log\ L}{\partial \lambda_2} = \frac{q}{\lambda_2} - \sum_{j=1}^q y_j\left(1 + e^{-\lambda_2 y_j}\right)$$

The maximum likelihood estimates of the parameters are obtained as the solution of the above non-linear equations.

The second partial derivatives of the log likelihood function with respect to λ_1 and λ_2 are

$$\frac{\partial^2 \log\ L}{\partial \lambda_1^2} = \frac{-p}{\lambda_1^2} + \sum_{i=1}^{p} x_i^2 e^{-\lambda_1 x_i}$$

and

$$\frac{\partial^2 \log\ L}{\partial \lambda_2^2} = \frac{-q}{\lambda_2^2} + \sum_{j=1}^{q} y_j^2 e^{-\lambda_2 y_j}$$

The maximum likelihood estimate of Stress-Strength reliability R is

$$\widehat{R}_{ML} = \frac{\widehat{\lambda_1} e}{(e-1)^2} \sum_{m=0}^{\infty} \frac{1}{m!} \left[\frac{1}{\widehat{\lambda_1}(m+1)} - \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{n+i}}{n!} \binom{n}{i} \frac{1}{\widehat{\lambda_1}(m+1) + \widehat{\lambda_2} i} \right] \tag{5}$$

We obtain the expression of the maximum likelihood estimate of Stress-Strength reliability R by substituting the estimated parameters in the Equation (2).

5. Asymptotic distribution and confidence interval

In this section we focused on the asymptotic distribution and confidence interval of the maximum likelihood estimate of R. To obtain the asymptotic variance of the maximum likelihood estimate of R, we consider the Fisher information matrix of λ and is denoted as I.

$$I = - \begin{pmatrix} E(\frac{\partial^2 \log\ L}{\partial \lambda_1^2}) & E(\frac{\partial^2 \log\ L}{\partial \lambda_1 \partial \lambda_2}) \\ E(\frac{\partial^2 \log\ L}{\partial \lambda_2 \partial \lambda_1}) & E(\frac{\partial^2 \log\ L}{\partial \lambda_2^2}) \end{pmatrix}$$

Using the standard method of asymptotic properties of maximum likelihood estimate, we derive the asymptotic normality of R as

$$d(\lambda) = \left(\frac{\partial R}{\partial \lambda_1}, \frac{\partial R}{\partial \lambda_2} \right)' = (d_1, d_2)'$$

Here

$$\frac{\partial R}{\partial \lambda_1} = -\frac{e}{e-1} \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{n+i}}{n!} \binom{n}{i} \frac{\widehat{\lambda_2} i}{(\widehat{\lambda_1}(m+1) + \widehat{\lambda_2} i)^2}$$

and

$$\frac{\partial R}{\partial \lambda_2} = -\frac{\widehat{\lambda_1} e}{(e-1)^2} \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{n+i}}{n!} \binom{n}{i} \frac{i}{(\widehat{\lambda_1}(m+1) + \widehat{\lambda_2} i)^2}$$

Now we obtain the asymptotic distribution of $\widehat{R}_{ML}$ as

$$\sqrt{p+q}(\widehat{R}_{ML} - R) \to^d N(0, d'(\lambda) I^{-1} d(\lambda)).$$

The asymptotic variance of the $\widehat{R}_{ML}$ is

$$\begin{aligned} AV(\widehat{R}_{ML}) =& \frac{1}{p+q} 0, d'(\lambda) I^{-1} d(\lambda) \\ =& V(\widehat{\lambda_1}) d_1^2 + V(\widehat{\lambda_2}) d_2^2 + 2 d_1 d_2 Cov(\widehat{\lambda_1}, \widehat{\lambda_2}). \end{aligned}$$

Hence an asymptotic $100(1-\xi)\%$ confidence interval for R can be obtained as

$$\widehat{R}_{ML} \pm Z_{\frac{\xi}{2}} \sqrt{AV(\widehat{R}_{ML})},$$

where $Z_{\frac{\xi}{2}}$ is the upper $\frac{\xi}{2}$ quantile function of the standard normal distribution.

6. Simulation study

In this section we check the performance of estimators in R using simulation technique. For this purpose we generate 1000 pseudo random samples using Newton-Raphson method. The random samples are generated for different population parameters of (λ_1, λ_2) as (0.5,1), (0.9,0.5), and (1.0,0.9) and sample sizes (p, q) as (10,10),(15,25),(20,20),(30,30),(40,40), and (50,50). The maximum likelihood estimates, their mean square error (MSE) and 95% confidence interval (CI) are calculated and the results are given in the following tables.

Table 1: *The ML estimates, MSEs and confidence interval of different estimators of R when* $\lambda_1 = 0.5$ *and* $\lambda_2 = 1.0$

(p, q)	ML estimates	MSEs	CI
(10,10)	λ_1 =0.57077	0.24634	(0.08795, 1.05359)
	λ_2 =1.14347	0.06295	(0.80201, 1.26686)
(15,25)	λ_1 =0.53939	0.20729	(0.13312, 0.94566)
	λ_2 =1.03657	0.12229	(0.80124, 1.06082)
(20,20)	λ_1 =0.53241	0.32406	(-0.10276, 1.16757)
	λ_2 =1.05715	0.04907	(0.96097, 1.15333)
(30,30)	λ_1 =0.52109	0.13712	(0.25233, 0.78985)
	λ_2 =1.04110	0.01237	(0.80140, 1.28079)
(40,40)	λ_1 =0.51019	0.22229	(0.07450, 0.94588)
	λ_2 =1.02810	0.02570	(0.97773, 1.07847)
(50,50)	λ_1 =0.51125	0.19753	(0.12409, 0.89841)
	λ_2 =1.02751	0.02640	(0.97577, 1.07925)

Table 2: *The ML estimates, MSEs and confidence interval of different estimators of R when* $\lambda_1 = 0.9$ *and* $\lambda_2 = 0.5$

(p, q)	ML estimates	MSEs	CI
(10,10)	λ_1 =1.00947	0.29601	(0.42928, 1.58966)
	λ_2 =0.55836	0.10123	(0.35994, 0.75677)
(15,25)	λ_1 =0.96646	0.12012	(0.73102, 1.2019)
	λ_2 =0.52620	0.01846	(0.35258, 0.52656)
(20,20)	λ_1 =0.95902	0.08818	(0.78620, 1.13185)
	λ_2 =0.53293	0.01209	(0.43045, 0.53541)
(30,30)	λ_1 =0.94372	0.06563	(0.81508, 1.07237)
	λ_2 =0.51937	0.00126	(0.49567, 0.54307)
(40,40)	λ_1 =0.92773	0.00185	(0.82411, 0.93135)
	λ_2 =0.51179	0.00033	(0.41115, 0.51243)
(50,50)	λ_1 =0.91653	0.00167	(0.89133, 0.91980)
	λ_2 =0.51249	0.00017	(0.45121, 0.51283)

Results from the simulation study reveals that sample sizes p and q increase, the estimated parameter values tends to population parameter values. Also the MSEs are decreasing with increase in sample sizes (p, q).

7. Application

In this section we have generated two data sets ($p = q = 20$) using KME model with parameter values $\lambda_1 = 1$ and $\lambda_2 = 0.5$. Therefore the value of R is obtained as 0.17633. The data points are adjusted in two decimal points and the data sets are presented in the following tables.

Table 3: *The ML estimates, MSEs and confidence interval of different estimators of R when* $\lambda_1 = 1.5$ *and* $\lambda_2 = 0.9$

(p, q)	ML estimates	MSEs	CI
(10,10)	λ_1 =1.70398	0.11763	(1.06944, 1.73853)
	λ_2 =1.01417	0.04031	(0.13567, 1.01478)
(15,25)	λ_1 =1.62178	0.09751	(1.10266, 1.64089)
	λ_2 =0.93769	0.03302	(0.27297, 1.00240)
(20,20)	λ_1 =1.59399	0.08241	(1.23654, 1.65144)
	λ_2 =0.95547	0.00594	(0.43819, 0.96712)
(30,30)	λ_1 =1.56481	0.08071	(1.39415, 1.54773)
	λ_2 =0.93981	0.00556	(0.63479, 0.98262)
(40,40)	λ_1 =1.55355	0.05150	(1.39382, 1.71329)
	λ_2 =0.92423	0.00625	(0.71198, 0.93647)
(50,50)	λ_1 =1.53238	0.00228	(1.32790, 1.53686)
	λ_2 =0.92832	0.00279	(0.81279, 0.94385)

Table 4: *Data set I*

4.02	0.44	1.43	0.09	0.49	0.27	0.54	0.02	0.48	1.77
3.04	4.30	0.94	3.08	1.42	0.09	3.05	2.17	0.21	0.64

Table 5: *Data set II*

0.09	0.26	0.20	0.11	2.08	1.48	0.85	2.57	1.04	0.26
0.01	0.40	1.37	0.71	0.29	1.10	0.81	0.13	1.73	2.25

In this case the maximum likelihood estimates of λ_1 and λ_2 are obtained respectively as 0.854 and 0.542. Here the estimated value of R, $\widehat{R}_{ML}$ is obtained as 0.21336. The corresponding 95% confidence interval based on asymptotic distribution is (0.19153, 0.23519).

8. Conclusion

In this paper we consider the estimation of the stress-strength reliability for the KME model for independent stress and strength random variables when the parameters are unknown. The maximum likelihood estimators of the unknown parameters are calculated. Then provide the asymptotic distributions of the maximum likelihood estimators, which have been used to construct the asymptotic confidence intervals. Simulation study is carried out to examine the performance of the estimators. The study reveals that MSEs are decreasing with increase in sample sizes. Using a simulated data set, we find the estimates of the parameters, $\widehat{R}_{ML}$ value and 95% confidence interval.

References

[1] Birnbaum, Z.W. (1956). On a use of the Mann-Whitney statistic. Contributions to the theory of statistics. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California, Vol. 1, pp. 13-17.

[2] Birnbaum, Z. W., McCarty, B.C. (1958). A distribution-free upper confidence bounds for $Pr(Y < X)$ based on independent samples of X and Y. *The Annals of Mathematical Statistics*, Vol. 29(2), pp. 558-562.

[3] Church, J. D., Harris, B. (1970). The estimation of reliability from stress-strength relationships. *Technometrics*, vol. 12 (1), pp. 49-54.
[4] Owen, D. B., Craswell, K. J., Hanson, D. L. (1964). Nonparametric upper confidence bounds for $PrY < X$ and confidence limits for $PrY < X$ when X and Y are normal. *Journal of the American Statistical Association*, Vol. 59(307), pp. 906-924. https://doi.org/10.1080/01621459.1964.10480739.
[5] Downton, F. (1973). The estimation of $P(X > Y)$ in the normal case. *Technometrics*, Vol. 15, pp. 551-558.
[6] Woodward, W. A., Kelley, G. D (1977). Minimum variance unbiased estimation of $P[Y < X]$ in the normal case. *Technometrics*, Vol. 19(1), pp. 95-98. DOI: 10.1080/00401706.1977.10489505.
[7] Kelley, G. D, Kelley, J. A, Schucany, W. R. (1976). Efficient estimation of $P(Y < X)$ in the exponential case *Technimetrics*, pp. 359-360
[8] Tong, H. (1974). A note on the estimation of $PrY < X$ in the exponential case. *Technimetrics*, Vol. 16(4).
[9] Beg, M. A., Singh, N. (1979). Estimation of $Pr(Y < X)$ for the pareto distribution. *IEEE Transaction on Reliability*, Vol. 28(5), pp. 411-414. DOI: 10.1109/TR.1979.5220665
[10] Tong, H. (1977). On the estimation of $P(Y < X)$ exponential families. *IEEE Transaction on Reliability*, Vol. 26(1), pp. 54-56.
[11] Enis, P., Geisser, s. (1971). Estimation of the probability that $Y < X$. *Journal of the American Statistical Association*, Vol. 66, pp. 162-168. https://doi.org/10.1080/01621459.1971.10482238.
[12] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Annals of Statistics*, Vol. 1(2), pp. 209-230. https://doi.org/10.1214/aos/1176342360.
[13] Hollander, M., Korwar, R. M. (1976). Nonparametric empirical bayes estimation of the probability that $x \geq y$. *Communications in Statistics-Theory and Methods*, Vol. 5(14), pp. 1369-1383. https://doi.org/10.1080/03610927608827448.
[14] Bhattacharya, G. K., Johnson, R. A. (1974). Estimation of reliability in a multicomponent stress-strength model. *Journal of the American Statistical Association*, Vol. 69, pp. 966-970.
[15] Guttman, I., Johnson, R. A., Bhattacharyya, G. K., Reiser, B. (1988). Confidence limits for stress-strength models with explanatory variables. *Technometrics*, Vol. 30(2), pp. 161-168.
[16] Weerahandi, S., Johnson, R. A. (1992). Testing reliability in a stress-strength model when X and Y are normally distributed. *Technometrics*, Vol. 34(1), pp. 83-91.
[17] Sun, D., Ghosh, M., Basu, A. P. (1998). Bayesian analysis for a stress-strength system under non-informative priors. *The Canadian Journal of Statistics*, Vol. 26(2), pp. 323-332.
[18] Raqab, M.Z., Kundu, D. (2005). Comparison of different estimators of $P[Y < X]$ for a scaled Burr type X distribution. *Communications in Statistics-Simulation and Computation*, Vol. 34, pp. 465-483.
[19] Kotz, S., Lumelskii, Y., Pensky, M. (2003). The Stress-Strength Model and its Generalizations: Theory and Applications. Singapore: World Scientific Press.
[20] Kundu, D., Gupta, R.D. (2005). Estimation of $P[Y < X]$ for generalized exponential distribution. *Metrika*, Vol. 61, pp. 291-308.
[21] Kundu, D., Gupta, R.D. (2006). Estimation of $R = P[Y < X]$ for Weibull distributions. *IEEE Transactions on Reliability*, Vol. 55, pp. 270-280.
[22] Kundu, D., Raqab, M.Z. (2009). Estimation of $R = P(Y < X)$ for three-parameter Weibull distribution. *Statistics and Probability Letters*, Vol. 79, pp. 1839-1846.
[23] Krishnamoorthy, K., Mukherjee, S., Guo, H. (2007). Inference on reliability in two-parameter exponential stress-strength model. *Metrika*, Vol. 65, pp. 261-273.
[24] Raqab, M. Z, Madi, T., Kundu, D. (2008). Estimation of $P(Y < X)$ for the three-parameter generalized exponential distribution. *Communications in Statistics- Theory and Methods*, Vol. 37, pp. 2854-2865.
[25] Rezaei, S., Tahmasbi, R., Mahmoodi, M. (2010). Estimation of $P(Y < X)$ for generalized Pareto distribution. *Journal of Statistical Planning and Inference*, Vol. 140(2), pp. 480-494.
[26] Baklizi, A. (2012). Inference on $P(X < Y)$ in the two-parameter Weibull model based on records. *ISRN Probability and Statistics*, pp. 1-11.

[27] Baklizi, A., Eidous, O. (2006). Nonparametric estimation of $P(X < Y)$ using kernel methods. *Metron-International Journal of Statistics*, Vol. 64(1), pp. 47-60.

[28] Jing, B.Y., Yuan, J., Zhou, W. (2009). Jackknife empirical likelihood. *Journal of American Statistical Association*, Vol. 104, pp. 1224-1232.

[29] Basirat, M., Baratpour, S., Ahmadi, J. (2016). On estimation of stress- strength parameter using record values from proportional hazard rate models. *Communications in Statistics-Theory and Methods*, Vol. 45(19), pp. 5787-5801.

[30] Asgharzadeh, A., Valiollahi, R., Raqab, M. Z. (2017). Estimation of $P(Y < X)$ for the two-parameter of generalized exponential records. *Communications in Statistics- Simulation and Computation*, Vol. 46(1), pp. 379-394.

[31] Bai, X., Shi, Y., Liu, Y., Liu, B. (2018). Reliability inference of stress- strength model for the truncated proportional hazard rate distribution under progressively Type-II censored samples. *Applied Mathematical Modelling*, Vol.65, pp. 377-389.

[32] Bi, Q., Gui, W. (2017). Bayesian and classical estimation of stress- strength reliability for Inverse Weibull lifetime models. *Algorithms*, Vol. 10(2), pp. 1-16.

[33] Ghitany, M. E., Al-Mutairi, D. K., Aboukhamseen, S. M. (2014). Estima- tion of the reliability of a stress-strength system from power Lindley dis- tributions. *Communications in Statistics-Theory and Methods*, Vol. 44(1), pp. 118-136.

[34] Sharma, V. K. (2014). Bayesian analysis of head and neck cancer data using generalized inverse Lindley stress-strength reliability model. *Communications in Statistics- Theory and Methods*, Vol. 47(5), pp. 1155-1180.

[35] R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

[36] Kavya, P., Manoharan, M. (2021). Some parsimonious models for lifetimes and applications. *Journal of Statistical Computation and Simulation*, vol. 91, no. 18, 3693-3708.

Redundancy Optimization for a System Comprising One Operative Unit and N Hot Standby Units

Parveen

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
parveenkumar5893@gmail.com

Dalip Singh

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
dsmdur@gmail.com

Anil Kumar Taneja*

•

Department of Mathematics, Sh. L. N. Hindu College, Rohtak, Haryana, India
dranilkrtaneja@gmail.com

*Corresponding Author

Abstract

In many industries and applications, downtime or failure can have serious consequences, such as financial losses, safety hazards, or reputational damage. A hot standby unit can help minimize the impact of such events by providing a backup that can quickly and seamlessly take over in the event of a failure. Further, the question of as to how many hot standby units should be used also needs to be addressed. So, an N+1-Unit-system is investigated wherein N units are on hot standby, whereas one unit is operational and the system is such that the hot standby units can take over seamlessly if the single operative unit fails. The system breaks down completely when all the units fail. It is assumed that the failure rates of all the operational units and the redundant units will vary exponentially. To get different performability measurements, the regenerative point technique has been applied to optimize the value of N.

Keywords: Redundancy, Optimization, N Hot Standby Units, One Operative Unit, Profit Analysis, Regenerative Point Technique

1. Introduction

A crucial and difficult issue facing the manufacturing industry is reliability. Complexity makes it more challenging to successfully manage and run a system. Various system analyzers have had various issues as a result of expensive and unreliable components. Therefore, creating efficient modeling, monitoring, and control techniques is crucial for increasing the reliability of systems. To increase the dependability of a system, redundancy is necessary. In redundant systems, one or more than one unit runs while backups step in to take over as needed. Active redundancy and passive redundancy are the two types of redundancy. In the past, active redundancy has generally got greater attention. However, in actuality, a specific system design may include both active and cold-standby redundancies. Therefore, the challenge is to determine for each subsystem the appropriate redundancy approach, component, and level to maximize system

dependability within system-level restrictions. Other considerations in redundancy optimization may include selecting the most appropriate type of standby units, such as identical or functionally equivalent units, and designing a failover mechanism that can quickly and seamlessly switch from one mode to the other. Reliability Models for different systems' dependability that take the idea into consideration of standby by taking some operative units have been developed by various researchers. For a 2-unit cold standby system, Ritu and Malhotra [13] discussed a stochastic analysis wherein both units might become operational depending on the demand. Wang et al. [16] considered the optimization of cold-standby systems that are subject to periodic inspection and maintenance, as well as the evaluation of system reliability and expected cost using an approximation method. Behboudi et al. [21] suggested a periodic switching approach for analyzing and evaluating the overall reliability of two-unit cold standby systems utilizing the idea of the virtual age. Shenyang et al. [23] presented an optimization model that employs a novel methodology using an imprecise and nonperiodic different types of switching strategy to improve the performance and efficiency of warm standby systems and various reliability functions are obtained using a recursive method. Zhang et al. [5] analyzed the performance of warm standby systems in which k units out of M+N must be active to provide complete system functioning with r repair facilities, and system availability and reliability are derived by considering a specific example. M. N. Gopalan [1] is focused on the availability and reliability analysis of a system with n units, single repair facility, and (n-1) warm standbys and discusses the cases n=2 and 3. Papageorgiou and Kokolakis [6], [10] studied n units parallel system in which 2 units are operative simultaneously and n-2 are in standby mode. When either operative unit fails, the other (n-2) warm/cold standbys immediately take over. By using recursive relations, the system's dependability is evaluated, unlike the majority of preceding findings, which offer boundaries for joint pdfs with limited information.

The reliability models considering hot standby systems have also been developed by various researchers. Goel et al. [1] analyzed three different sorts of failure mechanisms in a two-unit hot standby system with a single repair facility. A number of dependability characteristics that are important to system designers and operations managers have been assessed. Fujii and Sandoh [4] focused on the evaluation of the reliability of a 2-unit hot standby redundancy system using a Bayesian approach by taking both units identically. Rizwan, et al. [11] presented a theoretical analysis or mathematical model for the reliability enhancement of the hot standby industrial system and discussed various factors such as failure rates, repair times, availability, and downtime in the context of the system's reliability. Additionally, they provide simulation outcomes or case studies to demonstrate the effectiveness of the suggested reliability analysis method. Ebrahimipour et al. [9] addressed Optimizing Redundancy in Multi-State Series-Parallel k-out-of-n Systems with Hot Standby and discussed mathematical models, optimization algorithms, or simulation results to show that the suggested strategy is successful at maximizing redundancy and improving system reliability. Manocha et al. [14] focused on the profit analysis of a 2-unit hot standby database system to compare the costs associated with implementing and maintaining the redundant system with the potential benefits in terms of improved reliability, reduced downtime, and enhanced system availability. Venkatachalam and Parvatham [2] focused on the stochastic behavior analysis of a 1-server 2-unit hot standby system that deals with the probabilistic aspects of the system's performance and reliability. Shuhang and Zhang [8] developed a Markov model to analyze and evaluate the reliability of the hot standby repairable supply system and provide simulation results or case studies to validate the Markov model's effectiveness in assessing reliability.

Cold or warm standby redundancy can still be suitable for certain applications where downtime and response time are less critical or cost considerations are paramount, hot standby redundancy offers the highest level of reliability, availability, and quick recovery. It ensures continuous operation, minimizes disruptions, and provides a seamless user experience, making it essential for systems that demand high performance, real-time response, and minimal downtime.

Batra and Taneja [17] developed reliability models and optimized the number of hot standby units in a system having one or two operational units to ensure the desired level of system reliability

while considering factors such as cost, availability, and other constraints by employing approaches for regenerative point techniques and the Markov process. However, more than two hot standby units may also be taken into consideration in a system after carrying out an optimum analysis as to how many hot standby units should be used in order to have a more reliable/available and economically viable system. So, this paper refers to a system design where there are N+1 units, with one unit being actively used (the single operative unit) and the remaining N units serving as hot standby backups. In an N+1 system, the extra unit provides an additional level of redundancy, which can increase system availability and reduce the likelihood of downtime. To get different performability measurements, the regenerating point technique has been applied to optimize the value of N which is laid out as follows. The system's assumptions and characterizations are covered in Section 2. The nomenclature used in this analysis is described in Section 3. The system's description is covered in Section 4. The model of the system, transition densities, and mean sojourn time are covered in Section 5. To increase the number of standby units that will be deployed, the Generalised Results for the different Measures of System Effectiveness and Profit Equation have been derived in Sections 6 and 7. For some specific circumstances, Section 8 gives graphical interpretation and numerical results. The study's conclusions are presented in Section 9.

2. Assumptions and System Characterizations

1. The system is designed so that only one operational unit is required at any given moment, therefore N additional units are kept on hot standby.
2. Both the active and standby unit's failure rates are modeled by an exponential distribution.
3. The item is as good as new after being repaired.
4. With the system, there is just one repairman.
5. Even with only one active unit, the system continues to function.
6. First come first serve service is followed.

3. Nomenclature

The terminology for various transition densities and probabilities is as follows:

λ_o/α_1	Rate of failure/ repair of operative unit
λ_1/α_2	Rate of failure/ repair of standby unit
Hs	Hot standby unit
Op	Operative unit.
NF_r	N Operative units are failed .
NF_{rh}	N standby units are failed .
HS_N	N units are on hot standby .
F_r	Operative unit under repair
Fr_h	Standby unit under repair
F_{wr}	A failed operational unit is awaiting repair.
F_{wr_h}	Failed hot standby unit waiting for the repair.

For the notation $\phi_i(t)$, $q_{ij}(t), Q_{ij}(t)$, $A_i(t)$, $Bi(t)$, $V_i(t)$, $M_i(t)$ one may refer [17].

4. System Description

Here we develop a reliability model consisting of one primary operating unit responsible for handling the system's operations. Hot standby units remain ready so that one of them takes over in case the primary unit fails or becomes unavailable. The hot standby units are fully functional and synchronized with the primary unit, ensuring a seamless transition in the event of a failure. The model examined here may be used in a variety of actual scenarios, including power

distribution systems, network router systems, emergency power supply systems, and navigator components.

5. Transition Densities and Mean Sojourn Times

In this model, there are a total N^2+3N+1 Number of states in the model out of which N^2+N+1 states are operative states and 2N number of states are failed states. Possible states of the system along with transitions are represented in Figure 1 and Figure 2

Repersentatinos of the States of the System;

$U_1(s,k)$: Operative state in which 1 represents the Main operative unit failed before any hot standby unit and then going for under repair and (s-2) the number of failed hot standby unit waiting for repair and k denotes the number of failed operative unit waiting for the repair and remaining [(N+1)-(s+k)] unit are on standby mode.

$U_2(s,k)$: Operative state in which 2 represent hot standby unit failed before Main operative unit and then going for under repair and (s-2) a number of failed hot standby unit waiting for repair and k denote the number of failed operative unit waiting for the repair and remaining [(N+1)-(s+k)] unit is on standby mode.

$F_1(s,k)$: Failed state in which one failed operating unit is under repair and (s-2) number of failed hot standby units awaiting repair and k+1 denotes the number of failed operative unit waiting for the repair and remaining [N-(s+k)] unit are on standby mode.

$F_2(s,k)$: Failed state in which one failed hot standby unit is under repair and (s-2) a number of failed hot standby units awaiting repair and k+1 denotes the number of failed operative unit waiting for the repair and remaining [N-(s+k)] unit are on standby mode.

Here, it has been assumed that if the value [(N+1)-(s+k)] and [N-(s+k)] is negative it means there is no hot standby unit available in the system.

The states where the main operative unit fails before any standby unit

State $U_1(2,0)$: (Op, F_r, HS_{N-1}); State $U_1(2,1)$: $(Op, F_r, F_{wr}, HS_{N-2})$;
State $U_1(2,2)$: $(Op, F_r, 2F_{wr}, HS_{N-3})$; . . . State $U_1(2,N-3)$: $(Op, F_r, (N-3)F_{wr}, HS_2)$;
State $U_1(2,N-2)$: $(Op, F_r, (N-2)F_{wr}, WS_1)$; State $U_1(2,N-1)$: $(Op, F_r, (N-1)F_{wr})$;
State $U_1(3,0)$: $(Op, F_r, F_{wr_h}, HS_{N-2})$; State $U_1(3,1)$: $(Op, F_r, F_{wr}, F_{wr_h}, HS_{N-3})$;
State $U_1(3,2)$: $(Op, F_r, 2F_{wr}, F_{wr_h}, HS_{N-4})$; . . . State $U_1(3,N-3)$: $(Op, F_r, (N-3)F_{wr}, F_{wr_h}, HS_1)$;
State $U_1(3,N-2)$: $(Op, F_r, (N-2)F_{wr}, F_{wr_h})$;
State $U_1(4,0)$: $(Op, F_r, 2F_{wr_h}, HS_{N-3})$; State $U_1(4,1)$: $(Op, F_{wr}, F_r, 2F_{wr_h}, HS_{N-4})$;
State $U_1(4,2)$: $(Op, 2F_{wr}, F_r, 2F_{wr_h}, HS_{N-5})$;. . . State $U_1(4,N-3)$: $(Op, F_r, (N-3)F_{wr}, 2F_{wr_h})$;
.
.
.
State $U_2(N-1,0)$: $(Op, F_r, (N-3)F_{wr_h}, HS_2)$; State $U_2(N-1,1)$: $(Op, F_r, F_{wr}, (N-3)F_{wr_h}, HS)$;
State $U_2(N-1,2)$: $(Op, F_r, 2F_{wr}, (N-3)F_{wr_h})$;
State $U_1(N,0)$: $(Op, F_r, (N-2)F_{wr_h}, HS)$; State $U_1(N,1)$: $(Op, F_r, F_{wr}, (N-2)F_{wr_h})$;
State $U_1(N+1,0)$: $(Op, F_r, (N-1)F_{wr_h})$;
State $F_1(N+1,0)$: $(F_r, (N-1)F_{wr_h})$; State $F_1(N,1)$: $(F_r, 2F_{wr}, (N-2)F_{wr_h})$;
State $F_1(N-1,2)$: $(F_r, 3F_{wr}, (N-3)F_{wr_h})$;. . . State $F_1(4,N-3)$: $(F_r, (N-2)F_{wr}, 2F_{wr_h})$;
State $F_1(3,N-2)$: $(F_r, (N-1)F_{wr}, F_{wr_h})$; State $F_1(2,N-1)$: (F_r, NF_{wr});

The states where the first hot standby unit fails before the main operative unit

State $U_2(2,0)$: (Op, F_{r_h}, HS_{N-1}); State $U_2(2,1)$: $(Op, F_{r_h}, F_{wr}, HS_{N-2})$;
State $U_2(2,2)$: $(Op, F_{r_h}, 2F_{wr}, HS_{N-3})$; . . . State $U_2(2,N-3)$: $(Op, F_{r_h}, (N-3)F_{wr}, HS_2)$;
State $U_2(2,N-2)$: $(Op, F_{r_h}, (N-2)F_{wr}, HS_1)$; State $U_2(2,N-1)$: $(Op, F_{r_h}, (N-1)F_{wr})$;
State $U_2(3,0)$: $(Op, F_{r_h}, F_{wr_h}, HS_{N-2})$; State $U_2(3,1)$: $(Op, F_{r_h}, F_{wr}, F_{wr_h}, HS_{N-3})$;
State $U_2(3,2)$: $(Op, F_{r_h}, 2F_{wr}, F_{wr_h}, HS_{N-4})$; . . . State $U_2(3,N-3)$: $(Op, F_{r_h}, (N-3)F_{wr}, F_{wr_h}, HS_1)$;
State $U_2(3,N-2)$: $(Op, F_{r_h}, (N-2)F_{wr}, F_{wr_h})$;

State $U_2(4,0)$: $(Op, F_{r_h}, 2F_{wr_h}, HS_{N-3})$;
State $U_2(4,1)$: $(Op, F_{wr}, F_{r_h}, 2F_{wr_h}, HS_{N-4})$;
State $U_2(4,2)$: $(Op, 2F_{wr}, F_{r_h}, 2F_{wr_h}, HS_{N-5})$;...
State $U_2(4,N-3)$: $(Op, F_{r_h}, (N-3)F_{wr}, 2F_{wr_h})$;

.
.
.

State $U_2(N-1,0)$: $(Op, F_{r_h}, (N-3)F_{wr_h}, HS_2)$;
State $U_2(N-1,1)$: $(Op, F_{r_h}, F_{wr}, (N-3)F_{wr_h}, HS)$;
State $U_2(N-1,2)$: $(Op, F_{r_h}, 2F_{wr}, (N-3)F_{wr_h})$;
State $U_2(N,0)$: $(Op, F_{r_h}, (N-2)F_{wr_h}, HS)$;
State $U_2(N,1)$: $(Op, F_{r_h}, F_{wr}, (N-2)F_{wr_h})$;
State $U_1(N+1,0)$: $(Op, F_{r_h}, (N-1)F_{wr_h})$;
State $F_2(N+1,0)$: $(F_{r_h}, (N-1)F_{wr_h})$;
State $F_2(N,1)$: $(F_{r_h}, 2F_{wr}, (N-2)F_{wr_h})$;
State $F_2(N-1,2)$: $(F_{r_h}, 3F_{wr}, (N-3)F_{wr_h})$;. . .
State $F_2(4,N-3)$: $(F_{r_h}, (N-2)F_{wr}, 2F_{wr_h})$;
State $F_2(3,N-2)$: $(F_{r_h}, (N-1)F_{wr}, F_{wr_h})$;
State $F_2(2,N-1)$: (F_{r_h}, NF_{wr});

Transition Diagram of the Model

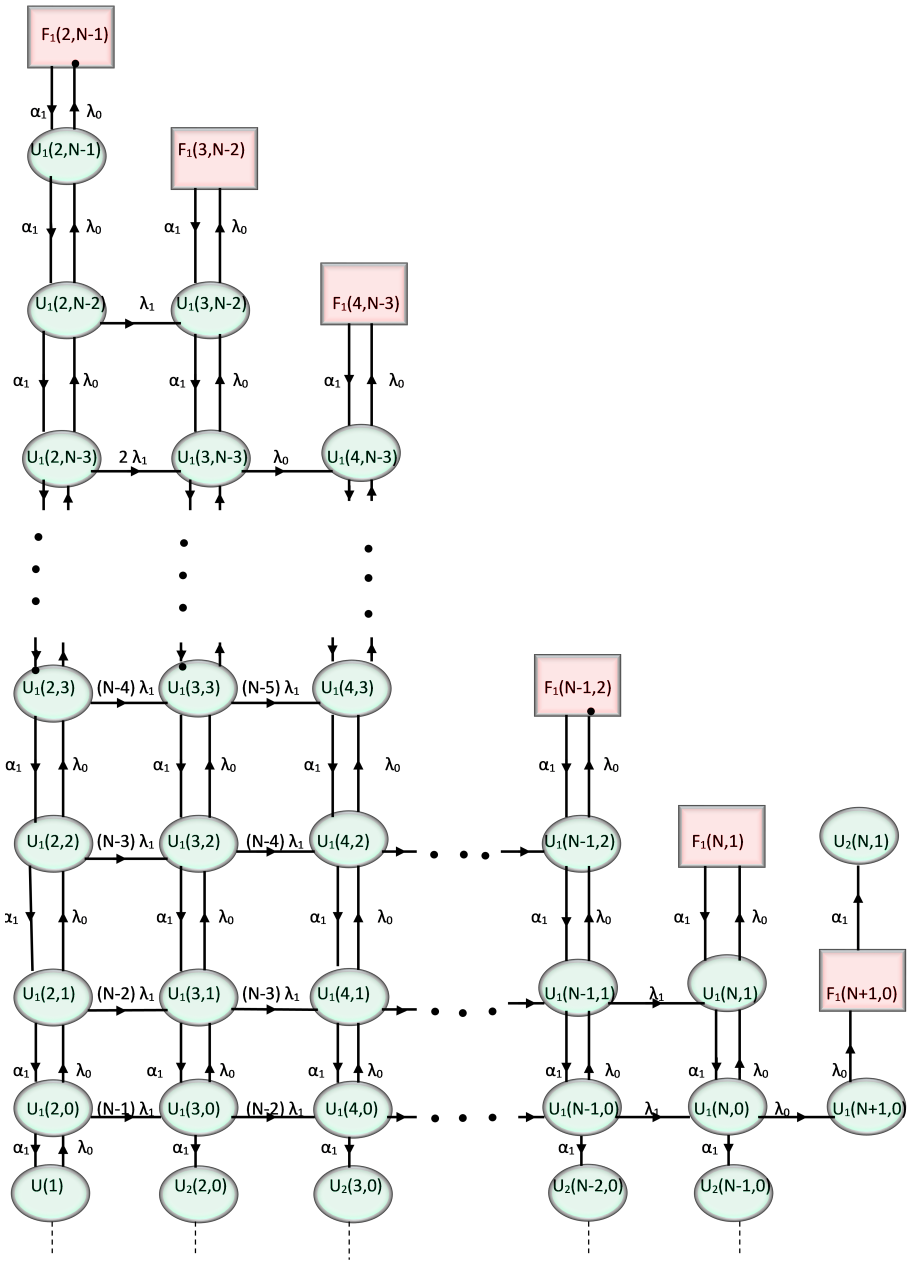


Figure 1: *State Transition Diagram (When main operative unit failed before any standby unit)*

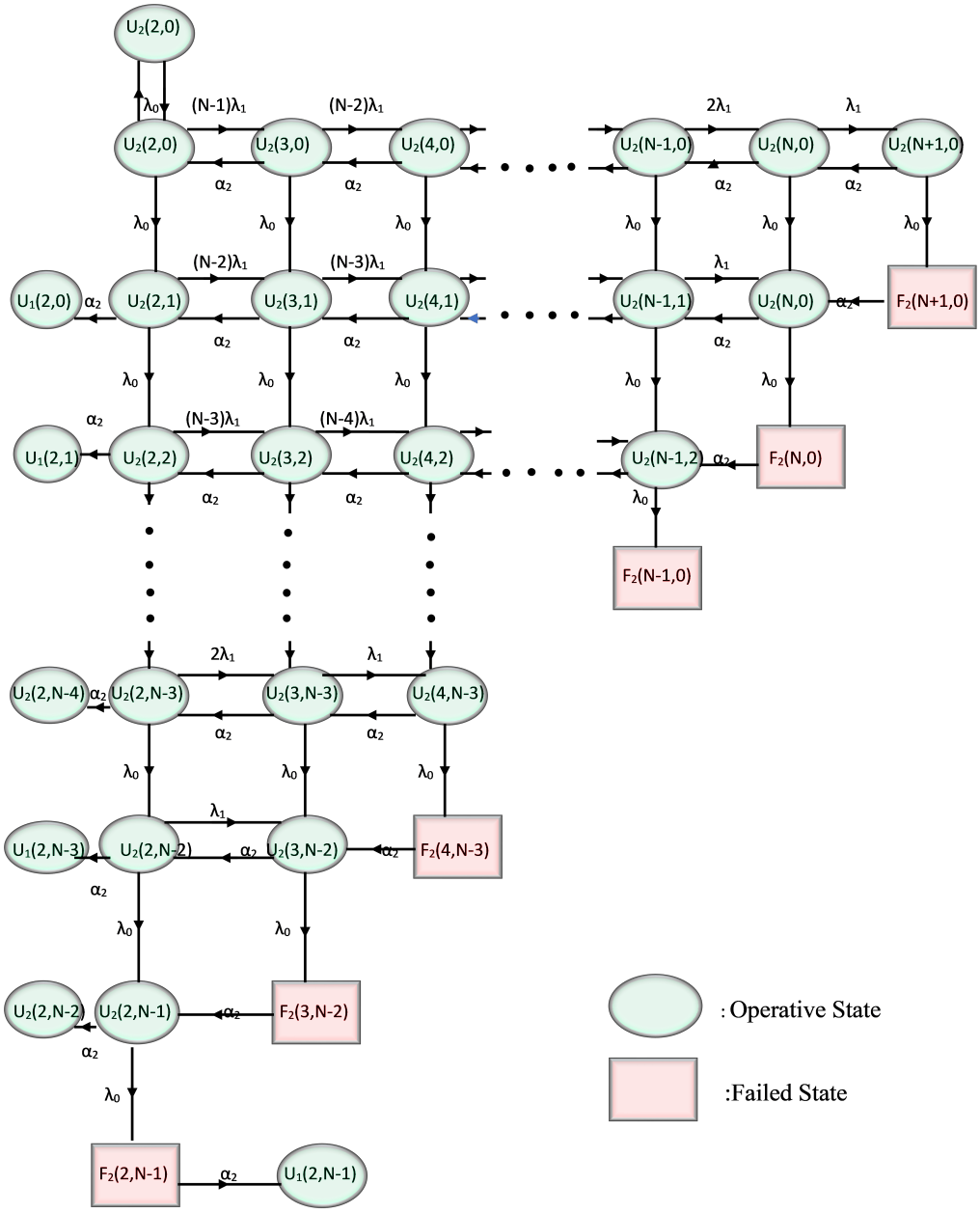


Figure 2: *State Transition Diagram (When first Standby unit failed before operational unit)*

The densities $q_{ij}(t)$ for transiting from state i to j are given by

$$q_{U(1),U_1(2,0)}(t) = \lambda_0 e^{-(\lambda_0+N\lambda_1)t}, \quad q_{U(1),U_2(2,0)}(t) = N\lambda_1 e^{-(\lambda_0+N\lambda_1)t}, \tag{1}$$

$$\begin{cases} q_{U_i(s,0),U_i(s,1)}(t) = \lambda_0 e^{-(\lambda_0+\alpha_i+[(M+1)-s]\lambda_1)t} \\ q_{U_i(s,0),U_i(s+1,0)}(t) = (M+1-s)\lambda_1 e^{-(\lambda+\alpha_i+[(M+1)-s]\lambda_1)t} \\ q_{U_i(s,0),U_2(s-1,0)}(t) = \alpha_i e^{-(\lambda+\alpha_i+[(M+1)-s]\lambda_1)t} \\ where\ i = 1\ or\ 2,\ 3 \le s \le N, 2 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{2}$$

$$\begin{cases} q_{U_i(2,s),U_i(2,s+1)}(t) = \lambda_0 e^{-(\lambda_0+\alpha_i+[M-1-s]\lambda_1)t} \\ q_{U_i(2,s),U_i(3,s+1)}(t) = [M-1-s)]\lambda_1 e^{-(\lambda_0+\alpha_i+[M-1-s]\lambda_1)t} \\ q_{U_i(2,s),U_1(2,s-1)}(t) = \alpha_i e^{-(\lambda_0+\alpha_i+[M-1-s]\lambda_1)t} \\ where\ i = 1\ or\ 2,\ 0 \le s \le N-2, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{3}$$

$$\begin{cases} q_{U_i(s,k),U_i(s,k+1)}(t) = \lambda_0 e^{-(\lambda_0+\alpha_i+[(M+1)-(s+k)]\lambda_1)t} \\ q_{U_i(s,k),U_i(s+1,k)}(t) = [(M+1)-(s+k)]\lambda_1 e^{-(\lambda_0+\alpha_i+[(M+1)-(s+k)]\lambda_1)t} \\ q_{U_1(s,k),U_1(s,k-1)}(t) = \alpha_1 e^{-(\lambda_0+\alpha_1+[(M+1)-(s+k)]\lambda_1)t} \\ q_{U_2(s,k),U_2(s-1,k)}(t) = \alpha_2 e^{-(\lambda_0+\alpha_2+[(M+1)-(s+k)]\lambda_1)t} \\ \text{where } i = 1 \text{ or } 2,\ 3 \le s \le N-1,\ 1 \le k \le N-s, 3 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{4}$$

From operative states to failed states;

$$\begin{cases} q_{U_i(s,N+1-s),F_i(s,N+1-s)}(t) = \lambda_0 e^{-(\lambda_0+\alpha_i)t}, \quad \text{where } i = 1 \text{ or } 2,\ 2 \le s \le N+1 \\ q_{U_1(s,M+1-s),U_1(s,M-s)}(t) = \alpha_1 e^{-(\lambda_0+\alpha_1)t} \quad ; 2 \le s \le N, 1 < M \le N \text{ and } M \text{ is an integer} \\ q_{U_1(N+1,0),U_2(N,0)}(t) = \alpha_1 e^{-(\lambda_0+\alpha_1)t} \end{cases} \tag{5}$$

$$\begin{cases} q_{U_2(s,M+1-s),U_2(s-1,M+1-s)}(t) = \alpha_i e^{-(\lambda_0+\alpha_i)t} \\ q_{U_2(2,N-1),U_1(2,N-2)}(t) = \alpha_2 e^{-(\lambda_0+\alpha_2)t} \\ \text{where } 3 \le s \le N+1, 1 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{6}$$

From failed states to operative states;

$$\begin{cases} q_{F_1(s,M+1-s),U_1(s,M+1-s)}(t) = \alpha_1 e^{-\alpha_1 t}, \quad q_{F_1(N+1,0),U_2(N,1)}(t) = \alpha_1 e^{-\alpha_1 t} \\ \text{where } 2 \le s \le N, 1 < M \le N \text{ and } M \text{ is an integer} \\ q_{F_2(s,M+1-s),U_2(s-1,M+2-s)}(t) = \alpha_2 e^{-\alpha_2 t}, \quad q_{F_2(2,N-1),U_1(2,N-2)}(t) = \alpha_2 e^{-\alpha_2 t} \\ \text{where } 3 \le s \le N+1, 1 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{7}$$

Thus, $p_{ij} = \lim_{s\to o} q^*_{ij}(s)$ ase given as:

$$p_{U(1),U_1(2,0)} = \frac{\lambda_0}{(\lambda_0+N\lambda_1)}, \quad p_{U(1),U_2(2,0)} = \frac{N\lambda_1}{(\lambda_0+N\lambda_1)}, \tag{8}$$

$$\begin{cases} p_{U_i(s,0),U_i(s,1)} = \frac{\lambda_0}{(\lambda_0+\alpha_i+[(M+1)-s]\lambda_1)} M, \quad p_{U_i(s,0),U_i(s+1,0)} = \frac{(M+1-s)\lambda_1}{(\lambda+\alpha_i+[(M+1)-s]\lambda_1)} \\ p_{U_i(s,0),U_2(s-1,0)} = \frac{\alpha_i}{(\lambda+\alpha_i+[(M+1)-s]\lambda_1)} \\ \text{where } i = 1 \text{ or } 2,\ 3 \le s \le N, 2 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{9}$$

$$\begin{cases} p_{U_i(s,k),U_i(s,k+1)} = \frac{\lambda_0}{(\lambda_0+\alpha_i+[(M+1)-(s+k)]\lambda_1)}, \quad p_{U_i(s,k),U_i(s+1,k)} = \frac{[(M+1)-(s+k)]\lambda_1}{(\lambda_0+\alpha_i+[(M+1)-(s+k)]\lambda_1)} \\ p_{U_1(s,k),U_1(s,k-1)} = \frac{\alpha_1}{(\lambda_0+\alpha_1+[(M+1)-(s+k)]\lambda_1)}, \quad p_{U_2(s,k),U_2(s-1,k)} = \frac{\alpha_2}{(\lambda_0+\alpha_2+[(M+1)-(s+k)]\lambda_1)t} \\ \text{where } i = 1 \text{ or } 2,\ 3 \le s \le N-1,\ 1 \le k \le N-s, 3 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{10}$$

$$\begin{cases} p_{U_i(2,s),U_i(2,s+1)} = \frac{\lambda_0}{(\lambda_0+\alpha_i+[M-1-s]\lambda_1)}, \quad p_{U_i(2,s),U_i(3,s+1)} = \frac{(M-1-s)\lambda_1}{(\lambda_0+\alpha_i+[M-1-s]\lambda_1)} \\ p_{U_i(2,s),U_1(2,s-1)} = \frac{\alpha_i}{(\lambda_0+\alpha_i+[M-1-s]\lambda_1)} \\ \text{where } i = 1 \text{ or } 2,\ 0 \le s \le N-2, 1 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{11}$$

From operative states to failed states;

$$\begin{cases} p_{U_i(s,N+1-s),F_i(s,N+1-s)} = \frac{\lambda_0}{(\lambda_0+\alpha_i)}, \quad \text{where } i = 1 \text{ os } 2,\ 2 \le s \le N+1 \\ p_{U_1(s,M+1-s),U_1(s,M-s)} = \frac{\alpha_1}{(\lambda_0+\alpha_1)} \quad ; 2 \le s \le N, 1 < M \le N \text{ and } M \text{ is an integer} \\ p_{U_1(N+1,0),U_2(N,0)}(t) = \frac{\alpha_1}{(\lambda_0+\alpha_1)} \end{cases} \tag{12}$$

$$\begin{cases} p_{U_2(s,M+1-s),U_2(s-1,M+1-s)} = \frac{\alpha_2}{(\lambda_0+\alpha_2)}, \quad p_{U_2(2,N-1),U_1(2,N-2)} = \frac{\alpha_2}{(\lambda_0+\alpha_2)} \\ \text{where } i = 1 \text{ or } 2,\ 3 \le s \le N+1, 1 < M \le N \text{ and } M \text{ is an integer} \end{cases} \tag{13}$$

From failed states to operative states;

$$\begin{cases} p_{F_1(s,M+1-s),U_1(s,M+1-s)} = 1, \quad p_{F_1(N+1,0),U_2(N,1)} = 1 \\ where\ 2 \le s \le N, 1 < M \le N\ and\ M\ is\ an\ integer \\ p_{F_2(s,M+1-s),U_2(s-1,M+2-s)} = 1, \quad p_{F_2(2,N-1),U_1(2,N-2)} = 1 \\ where\ 3 \le s \le N+1, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{14}$$

We may, therefore, verify the following:

$$p_{U(1),U_1(2,0)} + p_{U(1),U_2(2,0)} = 1 \tag{15}$$

$$\left\{ p_{U_i(s,0),U_i(s,1)} + p_{U_i(s,0),U_i(s+1,0)} + p_{U_i(s,0),U_2(s-1,0)} = 1, where\ i = 1\ or\ 2,\ 3 \le s \le N \right. \tag{16}$$

$$\begin{cases} p_{U_1(s,k),U_i(s,k+1)} + p_{U_1(s,k),U_i(s+1,k)} + p_{U_1(s,k),U_1(s,k-1)} = 1 \\ p_{U_2(s,k),U_2(s,k+1)} + p_{U_2(s,k),U_i(s+1,k)} + p_{U_2(s,k),U_2(s-1,k)} = 1 \\ where\ 3 \le s \le N-1,\ 1 \le k \le N-s, N > 3 \end{cases} \tag{17}$$

$$\begin{cases} p_{U_i(2,s),U_i(2,s+1)} + p_{U_i(2,s),U_i(3,s+1)} + p_{U_i(2,s),U_1(2,s-1)} = 1 \\ where\ i = 1\ or\ 2,\ 0 \le s \le N-2, N > 1 \end{cases} \tag{18}$$

$$\begin{cases} p_{U_1(s,M+1-s),F_1(s,M+1-s)} + p_{U_1(s,M+1-s),U_1(s,M-s)} = 1 \\ where\ 2 \le s \le N, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{19}$$

$$\begin{cases} p_{U_2(s,M+1-s),F_2(s,M+1-s)} + p_{U_2(s,M+1-s),U_2(s-1,M+1-s)} = 1 \\ where\ 3 \le s \le N+1, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{20}$$

$$\begin{cases} p_{U_1(N+1,0),F_1(N+1,0)} + p_{U_1(N+1,0),U_2(N,0)} = 1 \\ p_{U_2(2,N-1),F_2(2,N-1)} + p_{U_2(2,N-1),U_1(2,N-2)} = 1 \end{cases} \tag{21}$$

Thus Mean Sojourn Time (μ_i) are:

$\mu_{U(1)} = \frac{1}{(\lambda_0+N\lambda_1)}$

$\mu_{U_i(s,0)} = \frac{1}{(\lambda_0+\alpha_i+[(N+1)-s]\lambda_1)}$ *where* $i = 1\ or\ 2,\ 3 \le s \le N, N > 2$

$\mu_{U_i(s,k)} = \frac{1}{(\lambda_0+\alpha_i+[(N+1)-(s+k)]\lambda_1)}$ *where* $i = 1\ or\ 2,\ 3 \le s \le N-1,\ 1 \le k \le N-s, N > 3$

$\mu_{U_i(2,s)} = \frac{1}{(\lambda_0+\alpha_i+[N-1-s]\lambda_1)}$ *where* $i = 1\ or\ 2,\ 0 \le s \le N-2, N > 1$

$\mu_{U_i(s,N+1-s)} = \frac{1}{(\lambda_0+\alpha_i)}$ *where* $i = 1\ or\ 2,\ 2 \le s \le N+1$

$\mu_{F_i(s,N+1-s)} = \frac{1}{\alpha_i}$ *where* $i = 1\ or\ 2,\ 2 \le s \le N+1$

Using state i as the starting point, the unconditional mean times (m_{ij}) are computed as

$$m_{U(1),U_1(2,0)} = \frac{\lambda_0}{(\lambda_0+N\lambda_1)^2}, \quad m_{U(1),U_2(2,0)} = \frac{N\lambda_1}{(\lambda_0+N\lambda_1)^2} \tag{22}$$

$$\begin{cases} m_{U_i(s,0),U_i(s,1)} = \frac{\lambda_0}{(\lambda_0+\alpha_i+[(M+1)-s]\lambda_1)^2}, \quad m_{U_i(s,0),U_i(s+1,0)} = \frac{(M+1-s)\lambda_1}{(\lambda+\alpha_i+[(M+1)-s]\lambda_1)^2} \\ m_{U_i(s,0),U_2(s-1,0)} = \frac{\alpha_i}{(\lambda+\alpha_i+[(M+1)-s]\lambda_1)^2} \\ where\ i = 1\ or\ 2,\ 3 \le s \le N, 2 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{23}$$

$$\begin{cases} m_{U_i(s,k),U_i(s,k+1)} = \frac{\lambda_0}{(\lambda_0+\alpha_i+[(M+1)-(s+k)]\lambda_1)^2},\ m_{U_i(s,k),U_i(s+1,k)} = \frac{[(M+1)-(s+k)]\lambda_1}{(\lambda_0+\alpha_i+[(M+1)-(s+k)]\lambda_1)^2} \\ m_{U_1(s,k),U_1(s,k-1)} = \frac{\alpha_1}{(\lambda_0+\alpha_1+[(M+1)-(s+k)]\lambda_1)^2},\ m_{U_2(s,k),U_2(s-1,k)} = \frac{\alpha_2}{(\lambda_0+\alpha_2+[(M+1)-(s+k)]\lambda_1)^2} \\ where\ i = 1\ or\ 2,\ 3 \le s \le N-1,\ 1 \le k \le N-s, 3 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{24}$$

$$\begin{cases} m_{U_i(2,s),U_i(2,s+1)} = \frac{\lambda_0}{(\lambda_0+\alpha_i+[M-1-s]\lambda_1)^2}, \quad m_{U_i(2,s),U_i(3,s+1)}(t) = \frac{[M-1-s)]\lambda_1}{(\lambda_0+\alpha_i+[M-1-s]\lambda_1)^2} \\ m_{U_i(2,s),U_1(2,s-1)} = \frac{\alpha_i}{(\lambda_0+\alpha_i+[M-1-s]\lambda_1)^2} \\ where\ i = 1\ or\ 2,\ 0 \le s \le N-2, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{25}$$

From operative states to failed states

$$\begin{cases} m_{U_i(s,N+1-s),F_i(s,N+1-s)} = \frac{\lambda_0}{(\lambda_0+\alpha_i)^2}, \quad where\ i = 1\ or\ 2,\ 2 \le s \le N+1 \\ m_{U_1(s,M+1-s),U_1(s,M-s)} = \frac{\alpha_1}{(\lambda_0+\alpha_1)^2} \quad ;2 \le s \le N, 1 < M \le N\ and\ M\ is\ an\ integer \\ m_{U_1(N+1,0),U_2(N,0)}(t) = \frac{\alpha_1}{(\lambda_0+\alpha_1)^2} \end{cases} \tag{26}$$

$$\begin{cases} m_{U_2(s,M+1-s),U_2(s-1,M+1-s)} = \frac{\alpha_2}{(\lambda_0+\alpha_2)^2}, \quad m_{U_2(2,N-1),U_1(2,N-2)} = \frac{\alpha_2}{(\lambda_0+\alpha_2)^2} \\ where\ 3 \le s \le N+1, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{27}$$

From failed states to operative state;

$$\begin{cases} m_{F_1(s,M+1-s),U_1(s,M+1-s)} = \frac{1}{\alpha_1}, \quad m_{F_1(N+1,0),U_2(N,1)} = \frac{1}{\alpha_1} \\ where\ i = 1\ or\ 2,\ 2 \le s \le N, 1 < M \le N\ and\ M\ is\ an\ integer \\ m_{F_2(s,M+1-s),U_2(s-1,M+2-s)} = \frac{1}{\alpha_2}, \quad m_{F_2(2,N-1),U_1(2,N-2)} = \frac{1}{\alpha_2} \\ where\ 3 \le s \le N+1, 2 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{28}$$

It may be verified that

$$m_{U(1),U_1(2,0)} + m_{U(1),U_2(2,0)} = \mu_{U(1)} \tag{29}$$

$$\begin{cases} m_{U_i(s,0),U_i(s,1)} + m_{U_i(s,0),U_i(s+1,0)} + m_{U_i(s,0),U_2(s-1,0)} = \mu_{U_i(s,0)} \\ where\ i = 1\ or\ 2,\ 3 \le s \le N, N > 2 \end{cases} \tag{30}$$

$$\begin{cases} m_{U_1(s,k),U_i(s,k+1)} + m_{U_1(s,k),U_i(s+1,k)} + m_{U_1(s,k),U_1(s,k-1)} = \mu_{U_1(s,k)} \\ m_{U_2(s,k),U_2(s,k+1)} + m_{U_2(s,k),U_i(s+1,k)} + m_{U_2(s,k),U_2(s-1,k)} = \mu_{U_2(s,k)} \\ where\ 3 \le s \le N-1,\ 1 \le k \le N-s, N > 3 \end{cases} \tag{31}$$

$$\begin{cases} m_{U_i(2,s),U_i(2,s+1)} + m_{U_i(2,s),U_i(3,s+1)} + m_{U_i(2,s),U_1(2,s-1)} = \mu_{U_i(2,s)} \\ where\ i = 1\ or\ 2,\ 0 \le s \le N-2, N > 1 \end{cases} \tag{32}$$

$$\begin{cases} m_{U_1(s,M+1-s),F_1(s,M+1-s)} + m_{U_1(s,M+1-s),U_1(s,N-s)} = \mu_{U_1(s,M+1-s)} \\ where\ 2 \le s \le N, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{33}$$

$$\begin{cases} m_{U_2(s,M+1-s),F_2(s,M+1-s)} + m_{U_2(s,M+1-s),U_2(s-1,M+1-s)} = \mu_{U_2(s,M+1-s)} \\ where\ 3 \le s \le N+1, 1 < M \le N\ and\ M\ is\ an\ integer \end{cases} \tag{34}$$

$$\begin{cases} m_{U_1(N+1,0),F_1(N+1,0)} + m_{U_1(N+1,0),U_2(N,0)} = \mu_{U_1(N+1,0)} \\ m_{U_2(2,N-1),F_2(2,N-1)} + m_{U_2(2,N-1),U_1(2,N-2)} = \mu_{U_2(2,N-1)} \end{cases} \tag{35}$$

6. Measures of System Effectiveness

6.1. Mean Time to System Failure (MTSF)

The following may result from viewing the failed state as an absorbing state:

$$\phi_{U(0)}(t) = Q_{U(1),U_1(2,0)}(t) \textcircled{S} \phi_{U_1(2,0)}(t) + Q_{U(1),U_2(2,0)}(t) \textcircled{S} \phi_{U_2(2,0)}(t) \tag{36}$$

$$\begin{aligned} \phi_{U_i(s,0)}(t) = {} & Q_{U_i(s,0),U_i(s,1)}(t) \textcircled{S} \phi_{U_i(s,1)}(t) + Q_{U_i(s,0),U_i(s+1,0)}(t) \textcircled{S} \phi_{U_i(s+1,0)}(t) \\ & + Q_{U_i(s,0),U_2(s-1,0)}(t) \textcircled{S} \phi_{U_2(s-1,0)}(t) where\ i = 1\ or\ 2,\ 3 \le s \le N, N > 2 \end{aligned} \tag{37}$$

$$\begin{aligned} \phi_{U_1(s,k)}(t) = {} & Q_{U_1(s,k),U_1(s,k+1)}(t) \textcircled{S} \phi_{U_1(s,k+1)}(t) + Q_{U_1(s,k),U_1(s+1,k)}(t) \textcircled{S} \phi_{U_1(s+1,k)}(t) \\ & + Q_{U_1(s,k),U_1(s,k-1)}(t) \textcircled{S} \phi_{U_1(s,k-1)}(t) \\ & where\ 3 \le s \le N-1,\ 1 \le k \le N-s, N > 3 \end{aligned} \tag{38}$$

$$\phi_{U_2(s,k)}(t) = Q_{U_2(s,k),U_2(s,k+1)}(t)\textcircled{S}\phi_{U_2(s,k+1)}(t) + Q_{U_2(s,k),U_2(s+1,k)}(t)\textcircled{S}\phi_{U_2(s+1,k)}(t) + Q_{U_2(s,k),U_2(s-1,k)}(t)\textcircled{S}\phi_{U_2(s-1,k)}(t) \quad (39)$$
where $3 \le s \le N-1,\ 1 \le k \le N-s, N > 3$

$$\phi_{U_i(2,s)}(t) = Q_{U_i(2,s),U_i(2,s+1)}(t)\textcircled{S}\phi_{U_i(2,s+1)}(t) + Q_{U_i(2,s),U_i(3,s+1)}(t)\textcircled{S}\phi_{U_i(3,s+1)}(t) \quad (40)$$
where $i = 1$ *or* $2,\ 0 \le s \le N-2, N > 1$

$$\phi_{U_1(N+1,0)}(t) = Q_{U_i(N+1,0),F_1(N+1,0)}(t) + Q_{U_1(N+1,0),U_2(N,0)}(t)\textcircled{S}\phi_{U_2(N,0)}(t) \quad (41)$$

$$\phi_{U_1(s,N+1-s)}(t) = Q_{U_1(s,N+1-s),F_1(s,N+1-s)}(t) + Q_{U_1(s,N+1-s),U_1(s,N-s)}(t)\textcircled{S}\phi_{U_1(s,N-s)}(t) \text{where } 2 \le s \le N, N > 1 \quad (42)$$

$$\phi_{U_2(s,N+1-s)}(t) = Q_{U_2(s,N+1-s),U_2(s-1,N+1-s)}(t)\textcircled{S}\phi_{U_2(s-1,N+1-s)}(t) + Q_{U_2(s,N+1-s),F_2(s,N+1-s)}(t) \text{where } 3 \le s \le N+1, N > 1 \quad (43)$$

$$\phi_{U_2(2,N-1)}(t) = Q_{U_2(2,N-1),F_2(2,N-1)}(t) + Q_{U_2(2,N-1),U_1(2,N-2)}(t)\textcircled{S}\phi_{U_1(2,N-2)}(t) \quad (44)$$

Laplace transform can solve these equations, and the function $\phi_0^{(N)}(t)$ can be obtained after inversion. So the MTSF is determined by

$$\phi_0^{(N)} = \lim_{s\to 0} \frac{K^{(N)}(s) - L^{(N)}(s)}{sK^{(N)}(s)} = \frac{L^{(N)}}{K^{(N)}} \quad (45)$$

where the value of $L^{(N)}$ and $K^{(N)}$ in determinant form can be evaluated using MATLAB or MATHEMATICA software.

6.2. **Availability of the System**

By definition of $AT_m(t)$ and the transitions that occur, we have:

$$A_{U(0)}(t) = M_{U(0)}(t) + q_{U(1),U_1(2,0)}(t)\textcircled{C}A_{U_1(2,0)}(t) + q_{U(1),U_2(2,0)}(t)\textcircled{C}A_{U_2(2,0)}(t) \quad (46)$$

$$A_{U_i(s,0)}(t) = M_{U_i(s,0)}(t) + q_{U_i(s,0),U_i(s,1)}(t)\textcircled{C}A_{U_i(s,1)}(t) + q_{U_i(s,0),U_i(s+1,0)}(t)\textcircled{C}A_{U_i(s+1,0)}(t) + q_{U_i(s,0),U_2(s-1,0)}(t)\textcircled{C}A_{U_2(s-1,0)}(t) \quad (47)$$
where $i = 1$ *or* $2,\ 3 \le s \le N, N > 2$

$$A_{U_1(s,k)}(t) = M_{U_1(s,k)}(t) + q_{U_1(s,k),U_1(s,k+1)}(t)\textcircled{C}A_{U_1(s,k+1)}(t) + q_{U_1(s,k),U_1(s+1,k)}(t)\textcircled{C}A_{U_1(s+1,k)}(t) + q_{U_1(s,k),U_1(s,k-1)}(t)\textcircled{C}A_{U_1(s,k-1)}(t) \quad (48)$$
where $3 \le s \le N-1,\ 1 \le k \le N-s, N > 3$

$$A_{U_2(s,k)}(t) = M_{U_2(s,k)}(t) + q_{U_2(s,k),U_2(s,k+1)}(t)\textcircled{C}A_{U_2(s,k+1)}(t) + q_{U_2(s,k),U_2(s+1,k)}(t)\textcircled{C}A_{U_2(s+1,k)}(t) + q_{U_2(s,k),U_2(s-1,k)}(t)\textcircled{C}A_{U_2(s-1,k)}(t) \quad (49)$$
where $3 \le s \le N-1,\ 1 \le k \le N-s, N > 3$

$$A_{U_i(2,s)}(t) = M_{U_i(2,s)}(t) + q_{U_i(2,s),U_i(2,s+1)}(t)\textcircled{C}A_{U_i(2,s+1)}(t) + q_{U_i(2,s),U_i(3,s+1)}(t)\textcircled{C}A_{U_i(3,s+1)}(t) \text{ where } i = 1 \text{ or } 2,\ 0 \le s \le N-2, N > 1 \quad (50)$$

$$A_{U_1(s,N+1-s)}(t) = M_{U_1(s,N+1-s)}(t) + q_{U_1(s,N+1-s),F_1(s,N+1-s)}(t)\textcircled{C}A_{F_1(s,N+1-s)}(t) + q_{U_1(s,N+1-s),U_1(s,N-s)}(t)\textcircled{C}A_{U_1(s,N-s)}(t) \text{ where } 2 \le s \le N, N > 1 \quad (51)$$

$$A_{U_1(N+1,0)}(t) = M_{U_1(N+1,0)}(t) + q_{U_i(N+1,0),F_1(N+1,0)}(t)\textcircled{C}A_{F_1(N+1,0)}(t) + q_{U_1(N+1,0),U_2(N,0)}(t)\textcircled{C}A_{U_2(N,0)}(t) \quad (52)$$

$$A_{U_2(2,N-1)}(t) = M_{U_2(2,N-1)}(t) + q_{U_2(2,N-1),F_2(2,N-1)}(t)©A_{F_2(2,N-1)}(t) + q_{U_2(2,N-1),U_1(2,N-2)}(t)©A_{U_1(2,N-2)}(t) \tag{53}$$

$$A_{U_2(s,N+1-s)}(t) = M_{U_2(s,N+1-s)}(t) + q_{U_2(s,N+1-s),F_2(s,N+1-s)}(t)©A_{F_2(s,N+1-s)}(t) + q_{U_2(s,N+1-s),U_2(s-1,N+1-s)}(t)©A_{U_2(s-1,N+1-s)}(t) \quad \textit{where } 3 \le s \le N+1, N > 1 \tag{54}$$

$$A_{F_1(s,N+1-s)}(t) = q_{F_1(s,N+1-s),U_1(s,N+1-s)}(t)©A_{U_1(s,N+1-s)}(t) \textit{ where } 2 \le s \le N, N > 1 \tag{55}$$

$$A_{F_2(s,N+1-s)}(t) = q_{F_2(s,N+1-s),U_2(s-1,N+2-s)}(t)©A_{U_2(s-1,N+2-s)}(t) \quad \textit{where } 3 \le s \le N+1, N > 1 \tag{56}$$

$$A_{F_1(N+1,0)}(t) = q_{F_1(N+1,0),U_2(N,1)}(t)©A_{U_2(N,1)}(t); 2 \le s \le N \tag{57}$$

$$A_{F_2(2,N-1)}(t) = q_{F_2(2,N-1),U_1(2,N-2)}(t)©A_{U_1(2,N-2)}(t) \tag{58}$$

where

$$\begin{cases} M_{U(0)}(t) = e^{-(\lambda_0+N\lambda_1)t} \\ M_{U_i(s,0)}(t) = e^{-(\lambda_0+\alpha_i+[(N+1)-s]\lambda_1)t} \textit{ where } i = 1 \textit{ or } 21',\ 3 \le s \le N \\ M_{U_i(s,k)}(t) = e^{-(\lambda_0+\alpha_i+[(N+1)-(s+k)]\lambda_1)t} \textit{ where } i = 1 \textit{ or } 2, 3 \le s \le N-1,\ 1 \le k \le N-s \\ M_{U_i(2,s)}(t) = e^{-(\lambda_0+\alpha_i+[N-1-s]\lambda_1)t} \textit{ where } i = 1 \textit{ or } 2,\ 0 \le s \le N-2, N > 1 \\ M_{U_1(s,N+1-s)}(t) = e^{-(\lambda_0+\alpha_i)t};\ 2 \le s \le N, N > 1 \\ M_{U_2(s,N+1-s)}(t) = e^{-(\lambda_0+\alpha_i)t};\ 3 \le s \le N+1, N > 1 \\ M_{U_1(N+1,0)}(t) = e^{-(\lambda_0+\alpha_1)t} \\ M_{U_2(2,N-1)}(t) = e^{-(\lambda_0+\alpha_2)t} \end{cases} \tag{59}$$

Laplace transform can solve these equations, and the function $A_0^{(N)}(t)$ can be obtained after inversion. As a result, the availability of the system is determined by

$$A_0^{(N)} = \lim_{s\to 0} sAT_0^*(s) = \lim_{s\to 0} \frac{sN_1^{(N)}(s)}{D_1^{(N)}(s)} = \frac{N_1^{(N)}}{D_1^{(N)}} \tag{60}$$

where the value of $N_1^{(N)}$ and $D_1^{(N)}$ in determinant form can be evaluated using MATLAB or MATHEMATICA software.

6.3. Other Measures

Using the definitions of B_0 and V_0 (specified in section 3) and the same procedures described in the section preceding it, The Expected Busy Period, the expected number of Visits for Repair, is given as:

6.3.1 Expected Busy Period

$$B_0^{(N)} = \lim_{s\to 0} \frac{sN_2^{(N)}(s)}{D_1^{(N)}(s)} = \frac{N_2^{(N)}}{D_1^{(N)}} \tag{61}$$

6.3.2 Expected Number Visits for Repair

$$V_0^{(N)} = \lim_{s\to 0} sV_0^{**}(s) = \lim_{s\to 0} \frac{sN_3^{N}(s)}{D_1^{(N)}(s)} = \frac{N_3^{(N)}}{D_1^{(N)}} \tag{62}$$

where the value of $N_2^{(N)}$, $N_3^{(N)}$ and $D_1^{(N)}$ in determinant form can be evaluated using MATLAB or MATHEMATICA software.

7. Profit Analysis

The profit equation, therefore, is

$$Profit(P_N) = C_0 A_0^{(N)} - C_1 B_0^{(N)} - C_2 V_0^{(N)} \tag{63}$$

C_0, C_1, C_2 are revenue, repair cost, and repairman visit cost respectively and all the costs are per unit time.

8. Graphical Interpretation and Numerical Results

Here, reliability metrics are determined for a system of N+1 units using arbitrary parameter values. For parameters with fixed values, the reliability measure's trend has been graphically displayed.

(i) For $\lambda_1 = 0.02, \alpha_1 = 0.35, \alpha_2 = 0.45$ the numerical values and of MTSF for various values of λ are provided in Table (1)

Table 1: *MTSF w.r.t. Failure Rate (λ) for N=1,2 and 3*

	MTSF		
λ	$N=1$	$N=2$	$N=3$
0.1	15.9994	38.298	81.9643
0.2	9.118	18.4515	32.6082
0.3	6.2559	11.6205	18.735
0.4	4.7245	8.3449	12.7916
0.5	3.7813	6.4651	9.609
0.6	3.1456	5.2588	7.6582
0.7	2.6897	4.4239	6.3505
0.8	2.3475	3.8139	5.4171
0.9	2.132	3.534	5.381

From Table (1) It can be noted that decreasing the MTSF by increasing the failure rate of operational units is in contrast to an almost linear trend observed when standby units are increasing.

(ii)For $\lambda = 0.1, \alpha_1 = 0.35, \alpha_2 = 0.45$ the numerical values of MTSF for various values of λ_1 are provided in Table (2)

Table 2: *MTSF w.r.t. Failure Rate of standby units(λ_1) for N=1,2 and 3*

	MTSF		
λ_1	$N=1$	$N=2$	$N=3$
0.1	33.5714	67.5989	98.7371
0.2	25.9677	40.602	48.1599
0.3	22.0732	30.3347	33.501
0.4	19.7059	25.0817	26.8425
0.5	18.1148	21.932	23.0889
0.6	16.9718	19.8471	20.6911
0.7	16.1111	18.3709	19.0302
0.8	15.4396	17.2734	17.813
0.9	14.901	16.4268	16.88

From Table (2) It can be seen that as the failure rate of hot standby units increases, the MTSF reduces as the trend for the number of standby units increases almost linearly.

(iii) For $\lambda_1 = 0.001, \alpha_1 = 0.2, \alpha_2 = 0.3$ Availability when N=1,2 and 3 w.r.t λ

Table 3: *Availability of the system (N=1,2,3) w.r.t. (λ)*

	Availability		
λ	$N = 1$	$N = 2$	$N = 3$
0.1	0.8268	0.8814	0.9151
0.2	0.6262	0.6826	0.7143
0.3	0.4871	0.5248	0.5505
0.4	0.3932	0.4173	0.4543
0.5	0.3276	0.3434	0.3606
0.6	0.2799	0.2906	0.3155
0.7	0.2438	0.2513	0.272
0.8	0.2157	0.2212	0.2361
0.9	0.1933	0.1974	0.2154

Table (3) shows that the availability is decreasing while the failure rate is increasing. However, it increases as the number of standby units increases.

(iv) For $\lambda_1 = 0.001, \alpha_1 = 0.2, \alpha_2 = 0.3, C_0 = 3000, C_1 = 1000, C_2 = 500$ the graph below depicts the profit when N=1,2, and 3 w.r.t *lambda.* , as seen in Figure (3)

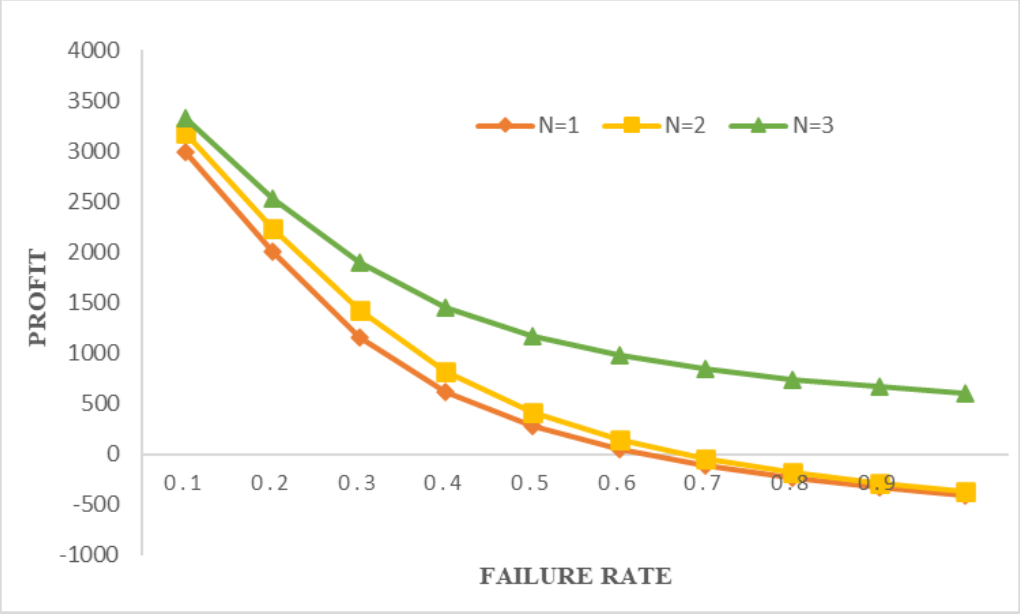


Figure 3: *Impact of Numbers of Standby Units on Profit w. r. t. λ*

From Figure (3), it is found that the Profit decreases with an increase in failure rate and the increase in the number of standby units.

(v) For $\lambda = 0.1, \lambda_1 = 0.1, \alpha_1 = 0.2, \alpha_2 = 0.3, C_0 = 4000, C_1 = 1000$, As indicated in Figure (4), the following graph has been generated.

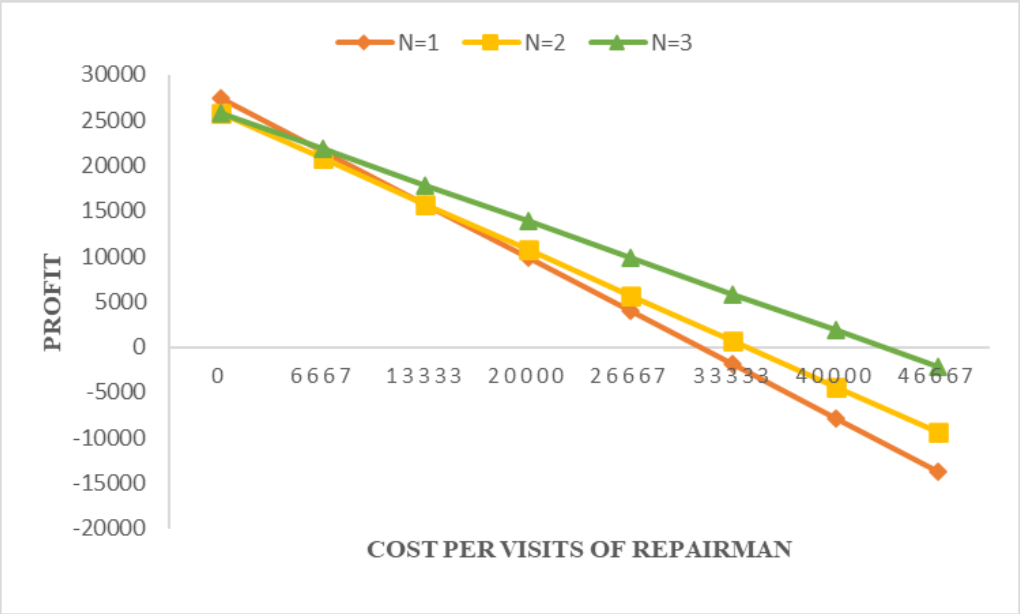


Figure 4: *Profit versus the repairman's visit costs* (C_2) *for N=1,2,3*

One might suggest that based on the above graph

- It is suggested to bring one unit along for standby if $C_2 < 6623$.
- Three or fewer units should be kept on standby if $C_2 > 6623$ for this value of C_2 since using more than three standby units causes revenue to decrease.
- Further C_2 should not exceed 32245.76 for $N = 1$, 35351.23 for $N = 2$, 44267.5 for $N = 3$ the system to be profitable. However, the above two conditions may also be kept in view while deciding about the number of standby units that will be needed.

(vi) Figure (5) and Figure (6) shows the change in profit for varied Revenue and N=1,2,3 and with fixed values for $\lambda = 0.02, \lambda_1 = 0.01, \alpha_1 = 0.1, \alpha_2 = 0.1, C_1 = 2000, C_2 = 1000$

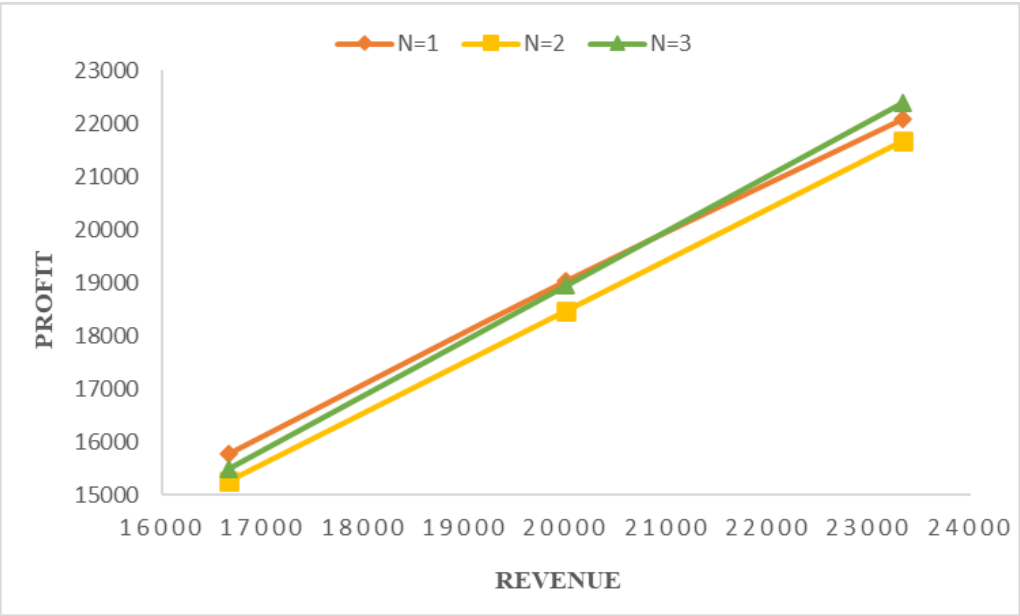


Figure 5: *Profit versus Revenue* (C_0) *for N=1,2,3*

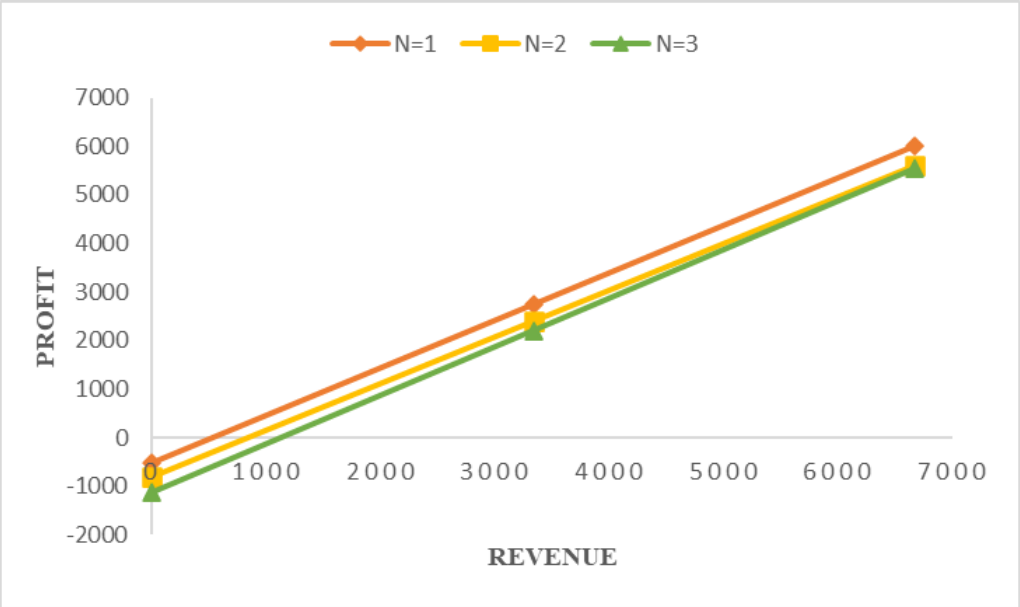


Figure 6: *Profit versus Revenue* (C_0) *for N=1,2,3*

From the graphs, the following can be determined

- From Figure (5) We notice that when C_0 is less than 20842.53, It would be profitable to keep one standby unit, and if C_0 is greater than 20842.53, keeping three standby units would be profitable.
- From Figure (6) we come to the decision that the cost of the products that the system will produce should be fixed in such a way that the value of C_0 is not less than 674.22 for N=1, 896.87 for N=2, 1144.12 for N=3, so as to make the system always profitable.

9. Conclusion

This study evaluates a crucial repairable system with a operational unit and N hot standby units. Arbitrary distributions have been used to get general findings for reliability metrics. However, the model has been validated taking specific parametric values for N=1,2,3. It has been observed taking a number of standby units is situational and may use 1 or 2 or 3 standby depending upon the cut-off points obtained in different situations.

References

[1] Gopalan, M. N. (1975). Probabilistic analysis of a single-server n-unit system with (n-1) warm standbys. *Operations Research*, 23(3): 591-598.

[2] Venkatachalam, P. (1978). Stochastic behavior of a 1-server 2-unit hot standby system. *Microelectronics Reliability*, 17(6):603-604.

[3] Goel, L. R. Gupta, R. and Gupta, P. (1983). Analysis of a two-unit hot standby system with three modes. *Microelectronics Reliability*,23(6):1029-1033.

[4] Fujii, S. and Sandoh, H. (1984). Bayes reliability assessment of a 2-Unit hot-standby redundant system. *IEEE Transactions on Reliability*,33(4):297-300.

[5] Zhang, T. Xie, M. Horigome, M. (2006). Availability and reliability of k-out-of-(M+N):G warm standby systems. *Reliability Engineering and System Safety* , 91(4):381-387.

[6] Papageorgiou, E., and Kokolakis, G. (2007). A two-unit general parallel system with (n-2) cold standbys Analytic and simulation approach. *European Journal of Operational Research*, 176(2):1016-1032.

[7] Tavakkoli-Moghaddam, R. Safari, J. and Sassani, F. (2008). Reliability optimization of series-parallel systems with a choice of redundancy strategies using a genetic algorithm. *Reliability Engineering and System Safety*,93(4):550-556.

[8] Ren, S., and Zhang, C. (2009). Study on the reliability of hot standby repairable supply system based on Markov model. *6th International Conference on Service Systems and Service Management IEEE*, 318-322.

[9] Ebrahimipour, V. Sheikhalishahi, M. Shoja, B. M. and Goldansaz, M. (2010). A universal generating function approach for redundancy optimization for hot-standby multi-state series-parallel k-out-of-n systems. *Fourth UKSim European Symposium on Computer Modeling and Simulation IEEE*, 235-239.

[10] Papageorgiou, E., and Kokolakis, G. (2010).. Reliability analysis of a two-unit general parallel system with (n-2) warm standbys. *European Journal of Operational Research* , 201(3):821-827.

[11] Rizwan, S. M. Khurana, V. and Taneja, G. (2010). Reliability analysis of a hot standby industrial system. *International Journal of Modelling and Simulation*,30(3):315-322.

[12] Jain, M., and Gupta, R. (2014). Availability analysis of a repairable redundant system with three types of failures subject to common cause failure.*International Journal of Mathematics in Operational Research*, 6(3):271-296.

[13] Malhotra, R. and Taneja, G. (2014). Stochastic analysis of a two-unit cold standby system wherein both units may become operative depending upon the demand. *Journal of Quality and Reliability Engineering*.

[14] Manocha, A. Taneja, G. and Singh S. (2017). Stochastic and cost-benefit analysis of two units hot standby database system. *International Journal of Performability Engineering*, 13(1):63.

[15] Batra, S. and Taneja G. (2018). Reliability and optimum analysis for number of standby units in a system working with one operative Unit. *International Journal of Applied Engineering Research*, 13(5): 2791-2797.

[16] Wang, W. Wu, Z. Xiong, J. and Xu, Y. (2018). Redundancy optimization of cold-standby systems under periodic inspection and maintenance. *Reliability Engineering and System Safety* , 180:394-402.

[17] Batra, S. and Taneja G. (2019). Reliability modeling and optimization of the number of hot standby units in a system working with two operative units. *Compusoft* ,8(2):3059-3068.

[18] Patowary M., Panda G. and Deka B. C. (2019). Reliability modeling of microgrid system using hybrid methods in hot standby mode. *IEEE Systems Journal*,13(3):3111-3119.

[19] Houankpo, H. G. K. Kozyrev, D. V. Nibasumba, E. and Mouale, M. N. D. B. (2020). Reliability analysis of a homogeneous hot standby data transmission system. *Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference (ESREL2020 PSAM15)*, 1-8.

[20] Wen, J. Cao, Y. Ma, L. and Du, S. (2020). Design and analysis of double one out of two with a hot standby safety redundant structure. *Chinese Journal of Electronics* 29(3):586-594.

[21] Behboudi, Z. Mohtashami Borzadaran, G.R. Asadi, M. (2021). Reliability modeling of two-unit cold standby systems: A periodic switching approach. *Applied Mathematical Modelling*,92:176-195.

[22] Zhang, Y. Wang, X. Yu, Y. Chu, P. and Xu, Y. (2021). Design of system-level hot standby redundancy-based signal system. *Journal of Physics: Conference Series*, 2010(1).

[23] Bai, S. Jia, X. Cheng, Z. Guo, B. Zhao, Q. and Zhang, X. (2022). Operation optimization model for warm standby system based on nonperiodic and imperfect multiple active switching policy. *Computers and Industrial Engineering*, 167:108001.

Fuzzy Control Charts based on Ranking of Pentagonal Fuzzy Numbers

[1,*]Mohammad Ahmad, [2]Weihu Cheng, [3]Zhao Xu, [4]Abdul Kalam, [5]Ahteshamul Haq

•

[1,2,3,4] Faculty of Science, Beijing University of Technology, Beijing, China
[5]Department of Mathematics and Statistics, Integral University Lucknow, Uttar Pradesh India
[1,*]mahmador@gmail.com, [2]chengweihu@bjut.edu.cn, [3]zhaox@bjut.edu.cn,
[4]faisal.stats@gmail.com,[5]a.haq@myamu.ac.in
[*]Corresponding author

Abstract

A Control Chart is a fundamental approach in Statistical Process Control. When uncommon causes of variability are present, sample averages will plot beyond the control boundaries, making the control chart a particularly effective process monitoring approach. Uncertainties are caused by the measuring system, including the gauges operators and ambient circumstances. In this paper, the concept of fuzzy set theory is used for dealing with uncertainty. The control limits are to converted into fuzzy control limits using the membership function. The fuzzy $\bar{X}$-R and $\bar{X}$-S control chart is developed by using the ranking of the pentagonal fuzzy number system. An illustrative example is done with the discussed technique to make fuzzy $\bar{X}$-R and $\bar{X}$-R control charts and increase the flexibility of the control limit.

Keywords: Statistical Process Control, Rank Membership Function, Fuzzy Pentagonal Number

I. Introduction

Quality has long been recognized as a significant influencing element in the performance and competitiveness of manufacturing and service organizations in both domestic and global markets. The return on capital is the consequence of well-executed strategies. Appropriate quality techniques provide productive outcomes. Fuzzy number ranking is a component of the quality control planning system. The fuzzy mathematical model for transportation of vegetable diet plan based on the ranking function of fuzzy pentagonal number (FPN) and solved by Vogel's approximation method to minimize the cost is discussed by Venkatesh and Manoj [1]. In constructing an FPN and related arithmetic operations, A. Panda and M. Pal [2] established the logical definition. The construction and fundamental features of pentagonal fuzzy matrices (PFMs) are investigated using FPN. The algebraic natures of several particular types of PFMs (trace of PFM, adjoint of PFM, determinant of PFM, etc.) are addressed. Pathinathan and Ponnivalavan [3] discussed FPN in continuation with the other defined fuzzy numbers and addressed some basic arithmetic operations. A. Chakraborty et al. [4] discuss different measures of interval-valued pentagonal fuzzy numbers (IVFPN) associated with assorted membership functions, and the ranking function is the main feature. The ranking functions of FPN develop real application and comprehend the uncertainty of the parameters more precisely in the evaluation process. A.

Chakraborty et al. [5] dealt with the idea of pentagonal neutrosophic number (PNN) from a different frame of reference and discussed some properties of PNN with real-life operational research applications, which is more reliable than the other method. A. Shafqat et al. [6] used the lower record values for developing the $\bar{X}$ control chart for the Inverse Rayleigh Distribution (IRD) is designed under repetitive group sampling. The mean and standard deviation of the Inverse Rayleigh Distribution based on lower record values are used to determine the width and power of the $\bar{X}$ control limits. Lim S. A. H. [7] evolved $\bar{X}$-R and $\bar{X}$-S chart for the food industry in the UK. The control charts developed using triangular and trapezoidal fuzzy control for balanced and unbalanced [8, 9, 10]. Ozdemir [11] developed the fuzzy control chart with a triangular fuzzy system into three phases for $\bar{X}$-S chart and process capability indices using unbalanced data, converted the data for each sample into the fuzzy form and then decided the fuzzy limits and illustrated it for uncertain data. Yeh [12] Shows an example of weighted triangular approximation of fuzzy numbers, which Zheng and Li propose. Senturk et al. [13] researched the most popular control chart for univariate data, the exponential weighted moving average control chart under fuzzy environment and applied this work into real case applications in Turkey. Senturk et al. [14] consider the control chart for fuzzy nonconformities per unit by using alpha cut and applied this technique in real case applications for truck engine manufacture. Alipour and Noorossana [15] created a control chart using a fuzzy multivariate exponentially weighted moving average (F-MEWMA). The proposed technique is developed using a combination of multivariate statistical quality control and fuzzy set theory in this study. Erginel and Şentürk [16] derived the fuzzy exponential weighted moving average and cumulative sum control chart (CUSUM) with a suitable example. Erginel [17] developed a fuzzy P control chart by using the rules that introduces the fuzzy np chart based on the constant sample size and variable sample size. In addition, the decision is taken whether it is under control or out of control. In the uncertainty theory for modelling, fuzzy sets theory plays numerous important roles. An essential consideration is that if somebody wants to take an FPN, what should its graphical representations (uncertainty quantification area) look like? How should the membership functions be defined? From this perspective, we developed the phases of an FPN control chart that may be a good choice for a decision-maker in a real-world scenario. In this study we proposed a Fuzzy control chart by using rank membership function of Pentagonal fuzzy number and the example is done with this proposed technique.

II. Development of Proposed Methodology

There are many articles published related to FPN. Venkatesh and Manoj [1] developed the pentagonal fuzzy model for transportation problems using the ranking membership function. Some definitions of FPN are as follows:

Definition 1: (Zadeh [19]) Let X be a fixed set. A fuzzy set $\tilde{A}$ of X is an object having the form $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in X\}$ where $\mu_{\tilde{A}}(x) \in [0,1]$ represents the degree of membership of the element $x \in X$ being in $\tilde{A}$, and $\mu_{\tilde{A}} : X \to [0,1]$ is called the membership function.

Definition 2: The α -cut of the fuzzy set $\tilde{A}$ is defined as:

$$\bar{A}_{\alpha} = \{x \in x/\mu_{\tilde{A}}(x) \geq \alpha, \text{ where } \alpha \in (0, 1).$$

Definition 3: A set $\tilde{A}$ is defined on the real numbers $\Re$ is said to be a fuzzy number if its membership function $\mu_{\tilde{A}} : X \to [0,1]$ follows:

(i) $\tilde{A}$ is continuous.

(ii) $\tilde{A}$ is normal such that $\mu_{\tilde{A}}(x) = 1$ there exists an $x \in$ R

Definition 4: A fuzzy number $\tilde{A}_P$ is an FPN denoted by $\tilde{A}_P = (a_1, a_2, a_3, a_4, a_5)$, where a_1, a_2, a_3, a_4, a_5 are real numbers, and its membership function will be:

$$\mu_{\tilde{A}_P}(x) = \begin{cases} 0 & , \ x < a_1 \\ \frac{(x-a_1)}{(a_2-a_1)} & , \ a_1 \le x \le a_2 \\ \frac{1}{2}\frac{(x-a_2)}{(a_3-a_2)} & , \ a_2 \le x \le a_3 \\ 1 & , \ x = a_3 \\ \frac{(a_4-x)}{2(a_4-a_3)} & , \ a_3 \le x \le a_4 \\ \frac{(a_5-x)}{(a_5-a_4)} & , \ a_4 \le x \le a_5 \\ 0 & , \ x > a_5 \end{cases} \quad (1)$$

Ranking of FPN: Ranking a fuzzy number entails comparing up to two fuzzy numbers, and defuzzification is a technique for converting a fuzzy number to an estimated crisp number. Just as the decision-maker compares two ideas that are the same, we must convert the fuzzy number to a comparable crisp number and compare the numbers based on crisp values in this problem.

Fuzzy numbers become the real line directly by using the ranking method [1]. Let $\tilde{A}_P$ be a generalized FPN. The ranking of $\tilde{A}_P$ is symbolised by $R(\tilde{A}_P)$ and it is calculated as follows:

$$R(\tilde{A}_P) = \left[\frac{a_1+2a_2+3a_3+2a_4+a_5}{9}\right] \quad (2)$$

Statistical Process Control: Many quality attributes may be stated numerically. A bearing's diameter, for example, might be measured using a micrometre and represented in mm. A variable is a single quantifiable qualitative attribute, such as a dimension, weight, or volume. Control charts for variables are widely used. When dealing with a variable quality characteristic, it is frequently required to monitor both the mean value and the variability of the quality characteristic. The control chart for means, or the control chart, is typically used to regulate the process average or mean quality level. Process variability may be tracked using either a control chart for the standard deviation known as an S control chart or a control chart for the range, known as an R control chart. The $\bar{X}$, R-chart and $\bar{X}$, S-chart is the most widely used control chart for the production process [16, 18].

$\bar{X}$ and R control charts

Table 1: *Formula for $\bar{X}$ and R Control Charts*

Chart	Lower Control Limit (LCL)	Central Line	Upper Control Limit (UCL)
$\bar{X}$	$\bar{\bar{X}} - A_2\bar{R}$	$\bar{\bar{X}}$	$\bar{\bar{X}} + A_2\bar{R}$
R	$\bar{R}D_3$	$\bar{R}$	$\bar{R}D_4$

where, $\bar{\bar{X}} = \frac{\sum \bar{x}}{m}$, $\bar{x} = \frac{\sum_{i=1}^{5} R(\tilde{A}_P(x_i))}{n}$, A_2 is constant tabulated value for n.

where, $\bar{R} = \frac{\sum Range(R(\tilde{A}_P(x_i)))}{m}$, and D_3, D_4 are constant tabulated values for n.

$\bar{X}$ and S Control Charts

Table 2: *Formula for $\bar{X}$ and S Control Charts*

Charts	LCL	Central Line	UCL
$\bar{X}$	$\bar{\bar{X}} - A_3\bar{S}$	$\bar{\bar{X}}$	$\bar{\bar{X}} + A_3\bar{S}$
S	$B_3\bar{S}$	$\bar{S}$	$B_4\bar{S}$

where, $\bar{\bar{X}} = \frac{\Sigma \bar{x}}{m}$, $\bar{x} = \frac{\sum_{i=1}^{5} R(\tilde{A}_P(x_i))}{n}$, $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}}$, $\bar{S} = \frac{\sum_{i=1}^{m} S}{m}$ and A_3, B_3, B_4 are constant tabulated values for n.

The next step will be complete with three steps for the proposed fuzzy control chart by using the following procedure.

Step 1: ***Normality Assumption***- Check the normality assumption using Anderson Darling Test (Anderson and Darling, 1954).

Step 2: ***Use of FPN Control Chart for $\bar{X}$-R* and *$\bar{X}$-S***

A summary of the work on the FPN Control Chart is as follows:

(i) The development of FPN
(ii) The representation of the FPNs in parametric form.
(iii) Apply ranking and defuzzification of FPN for the data.
(iv) Put the calculated crisp value into the control chart formula and set up the FPN control limit.

Step 3: ***Interpretation of FPN Control Chart for $\bar{X}$-R and $\bar{X}$-S***. The fuzzy CLs of the recommended $\bar{X}$-R and $\bar{X}$-S control charts are used to assess the fuzzy sample mean and standard deviation. If the fuzzy sample mean and standard deviation are inside the control bounds, the process is under control for the sample. Otherwise, the process will spiral out of control.

III. Illustrative Example

In this article, we choose the simulation data shown in Table 3 are the deviations from milling a slot in an aircraft terminal block. A high rate of rejections for many of the components manufactured in an aviation company's machine shop highlighted the necessity for an investigation into the causes of the problems. Because the majority of the rejections were for failing to satisfy dimensional limits, it was decided to utilize $\bar{X}$-R and $\bar{X}$-S charts to try to pinpoint the source of the problem. These charts, which needed real dimension measurement, were to be utilized just for the dimensions that were creating a high number of rejections. Among many such dimensions, the ones chosen for control charts were those with significant spoilage costs and reworked for those where rejections caused delays in assembly processes. Although the primary objective of all of the $\bar{X}$-R and $\bar{X}$-S charts was to diagnose problems, it was expected that some of the charts would be kept for routine process control and potentially for acceptance inspection.

Table 3: *Data for width of slot in an aircraft terminal block.*

Sample	X1	X2	X3	X4	X5
1	773	803	780	720	776
2	757	786	734	740	735
3	755	774	720	761	746
4	745	779	755	775	772
5	800	728	747	759	745

6	784	806	787	763	758
7	745	765	754	759	768
8	789	751	782	769	763
9	757	747	741	746	747
10	746	731	762	781	744
11	742	731	754	736	750
12	748	726	764	735	733
13	748	763	779	786	770
14	770	768	784	771	767
15	772	759	770	772	772
16	765	768	773	791	787
17	780	777	742	761	761
18	775	765	746	782	745
19	759	748	781	764	756
20	765	782	739	754	768
21	770	784	730	759	781
22	773	765	777	760	750
23	758	780	761	746	757
24	761	771	756	772	758
25	752	785	764	758	773
26	745	774	786	739	796
27	759	766	790	730	781
28	739	795	750	780	776
29	747	781	763	768	756
30	767	752	774	746	769

The first step is to test the normality assumption. It is shown that Figure 1 holds the normality assumption for the above data.

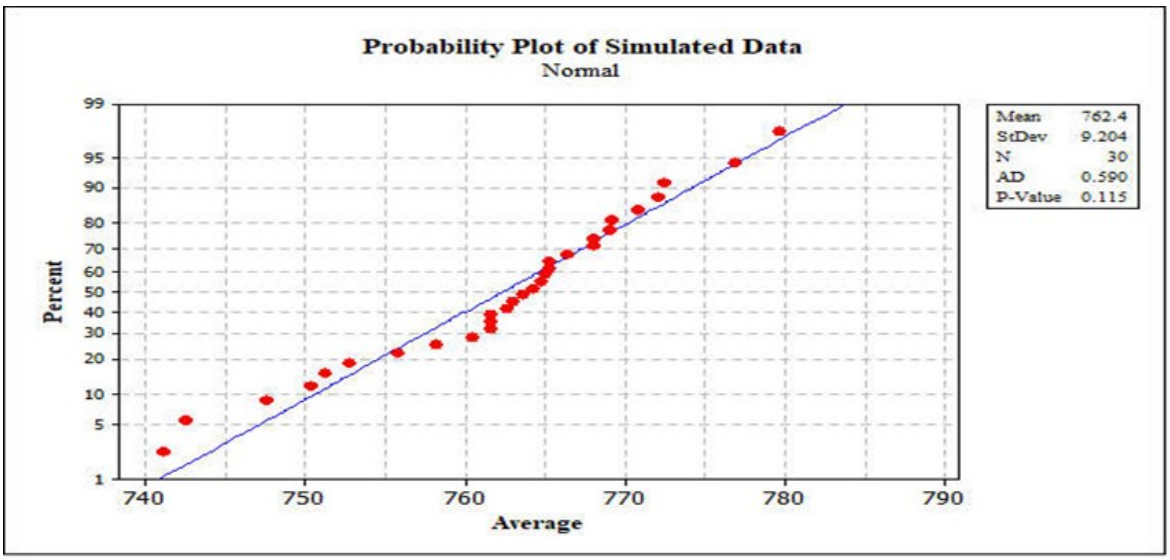


Figure 1: *Normal Probability Plot*

Now the data shows normal, and we convert the data into FPN form. Table [4-8] shows the FPN form for the variables X1, X2, X3, X4 and X5.

Table 4: *FPN for X1*

Xa1	Xb1	Xc1	Xd1	Xe1	$R(\widetilde{A}_P(x_1))$
768	770	773	777	785	774
752	754	757	761	771	758.2222
750	752	755	759	768	756.1111
740	742	745	749	757	746
795	797	800	804	812	801
777	781	784	788	796	784.7778
738	741	745	749	757	745.5556
781	786	789	793	801	789.6667
748	754	757	760	769	757.3333
740	743	746	749	752	746
721	727	730	733	741	730.2222
739	745	748	751	758	748.1111
740	744	748	751	759	748.1111
764	767	770	773	780	770.4444
766	769	772	775	778	772
757	761	765	769	786	766.4444
775	777	780	784	792	781
770	772	775	779	788	776.1111
749	756	759	763	771	759.4444
758	763	765	769	778	766.1111
763	767	770	774	783	770.8889
768	770	773	777	785	774
750	755	758	762	770	758.6667
756	758	761	765	774	762.1111
745	749	752	756	764	752.7778
738	743	745	750	759	746.4444
754	756	759	763	772	760.1111
734	736	739	743	751	740
740	745	747	751	760	748.1111
760	764	767	771	779	767.7778

Table 5: *FPN for X2*

Xa2	Xb2	Xc2	Xd2	Xe2	$R(\widetilde{A}_P(x_2))$
796	799	803	808	816	803.8889
778	781	786	790	798	786.2222
768	771	774	779	785	775
771	776	779	784	791	779.8889
719	724	728	735	742	729.2222
799	802	806	811	819	806.8889
758	762	765	770	776	765.8889
746	749	751	758	765	753.1111

741	745	747	752	759	748.3333
725	729	731	736	744	732.4444
724	728	731	735	745	732
719	723	726	730	737	726.6667
757	760	763	769	775	764.3333
760	765	768	773	778	768.6667
752	755	759	764	772	759.8889
760	765	768	774	780	769.1111
770	774	777	781	789	777.7778
760	762	765	769	778	766.1111
743	745	748	752	760	749
774	779	782	786	794	782.6667
777	782	784	788	797	785.1111
758	762	765	769	777	765.7778
772	777	780	784	792	780.6667
764	769	771	775	780	771.6667
780	782	785	789	796	785.8889
768	771	774	779	789	775.4444
759	761	766	771	782	767
790	793	795	799	808	796.3333
776	779	781	786	794	782.5556
747	750	752	757	765	753.5556

Table 6: *FPN for X3*

Xa3	**Xb3**	**Xc3**	**Xd3**	**Xe3**	$\boldsymbol{R(\widetilde{A}_P(x_3))}$
773	777	780	783	788	780.1111
727	731	734	740	744	735
715	717	720	725	728	720.7778
750	753	755	760	764	756.1111
741	745	747	753	755	748.1111
782	785	787	793	796	788.3333
749	752	754	760	763	755.3333
777	779	782	788	791	783.1111
737	740	741	747	750	742.6667
756	759	762	765	769	762.1111
748	751	754	760	763	755
758	761	764	770	773	765
772	775	779	785	787	779.5556
778	781	784	791	793	785.2222
764	767	770	776	779	771
768	771	773	780	782	774.5556
737	740	742	748	756	743.8889
740	744	746	750	758	747.1111
775	779	781	786	794	782.4444
734	737	739	746	754	741.2222
725	728	730	737	746	732.3333
771	774	777	783	791	778.5556

757	759	761	768	776	763.3333
751	754	756	760	768	757.2222
759	762	764	769	777	765.5556
781	784	786	790	798	787.2222
785	788	790	795	805	791.7778
745	747	750	758	768	752.5556
758	760	763	768	775	764.2222
768	771	774	779	787	775.2222

Table 7: *FPN for X4*

Xa4	**Xb4**	**Xc4**	**Xd4**	**Xe4**	$\boldsymbol{R(\widetilde{A}_P(x_4))}$
715	717	720	727	729	721.3333
736	738	740	749	751	742.3333
756	759	761	768	771	762.6667
770	773	775	781	784	776.3333
753	756	759	766	768	760.2222
758	761	763	769	771	764.2222
753	756	759	766	768	760.2222
764	767	769	775	778	770.3333
740	743	746	753	756	747.3333
776	779	781	787	789	782.2222
728	733	736	742	743	736.5556
729	732	735	741	743	735.8889
781	783	786	793	795	787.3333
766	769	771	778	781	772.6667
768	770	772	778	782	773.5556
787	789	791	797	799	792.3333
755	758	761	768	775	762.7778
777	780	782	787	794	783.4444
757	760	764	769	778	765
748	752	754	759	767	755.4444
754	757	759	765	773	760.8889
755	758	760	765	774	761.6667
741	744	746	750	758	747.2222
767	770	772	777	784	773.4444
753	755	758	764	773	759.7778
734	736	739	745	755	740.8889
725	728	730	735	743	731.5556
775	778	780	785	794	781.6667
763	765	768	772	780	769
741	744	746	751	760	747.6667

Table 8: *FPN for X5*

Xa5	**Xb5**	**Xc5**	**Xd5**	**Xe5**	$R(\widetilde{A}_P(x_5))$
770	773	776	782	784	776.8889
729	733	735	742	745	736.5556
741	744	746	753	755	747.5556
767	769	772	778	780	773
739	743	745	751	753	746.1111
752	754	758	764	767	758.7778
762	766	768	774	776	769.1111
758	760	763	769	772	764.1111
741	745	747	754	758	748.6667
738	741	744	750	752	744.8889
744	747	750	756	758	750.8889
727	730	733	739	743	734.1111
764	768	770	776	779	771.2222
761	765	767	774	778	768.6667
767	770	772	778	781	773.3333
782	785	787	793	795	788.2222
756	758	761	766	772	762.1111
740	743	745	750	759	746.6667
751	754	756	760	768	757.2222
763	766	768	771	779	768.8889
776	779	781	785	793	782.2222
745	748	750	755	762	751.4444
752	754	757	761	769	758
753	755	758	763	770	759.2222
768	770	773	778	784	774.1111
790	793	796	800	808	796.8889
776	779	781	786	794	782.5556
771	774	776	780	788	777.2222
751	753	756	760	767	756.8889
765	767	769	773	780	770.2222

After the FPN form, we use the rank membership function (equation (1)) for the crisp value.

Table 9: *The crisp value for X1, X2, X3, X4, X5*

$R(\widetilde{A}_P(x_1))$	$R(\widetilde{A}_P(x_2))$	$R(\widetilde{A}_P(x_3))$	$R(\widetilde{A}_P(x_4))$	$R(\widetilde{A}_P(x_5))$
774	803.8889	780.1111	721.3333	776.8889
758.2222	786.2222	735	742.3333	736.5556
756.1111	775	720.7778	762.6667	747.5556
746	779.8889	756.1111	776.3333	773
801	729.2222	748.1111	760.2222	746.1111
784.7778	806.8889	788.3333	764.2222	758.7778
745.5556	765.8889	755.3333	760.2222	769.1111
789.6667	753.1111	783.1111	770.3333	764.1111
757.3333	748.3333	742.6667	747.3333	748.6667
746	732.4444	762.1111	782.2222	744.8889

730.2222	732	755	736.5556	750.8889
748.1111	726.6667	765	735.8889	734.1111
748.1111	764.3333	779.5556	787.3333	771.2222
770.4444	768.6667	785.2222	772.6667	768.6667
772	759.8889	771	773.5556	773.3333
766.4444	769.1111	774.5556	792.3333	788.2222
781	777.7778	743.8889	762.7778	762.1111
776.1111	766.1111	747.1111	783.4444	746.6667
759.4444	749	782.4444	765	757.2222
766.1111	782.6667	741.2222	755.4444	768.8889
770.8889	785.1111	732.3333	760.8889	782.2222
774	765.7778	778.5556	761.6667	751.4444
758.6667	780.6667	763.3333	747.2222	758
762.1111	771.6667	757.2222	773.4444	759.2222
752.7778	785.8889	765.5556	759.7778	774.1111
746.4444	775.4444	787.2222	740.8889	796.8889
760.1111	767	791.7778	731.5556	782.5556
740	796.3333	752.5556	781.6667	777.2222
748.1111	782.5556	764.2222	769	756.8889
767.7778	753.5556	775.2222	747.6667	770.2222

All the crisp calculated value of the simulated data is in Table 9. Now we calculate the mean of crisps value and then put it into the formula of $\bar{X}$- R and $\bar{X}$-S, which is given in Table [1&2]. Now we found that the Control limits for the data are as follows:
UCL=785.65, LCL=741.39 and CL=763.52 for $\bar{X}$ chart,
UCL=81.11, LCL=0 and CL=38.36 for R chart and
UCL=32.38, LCL=0 and CL=15.50 for S chart.

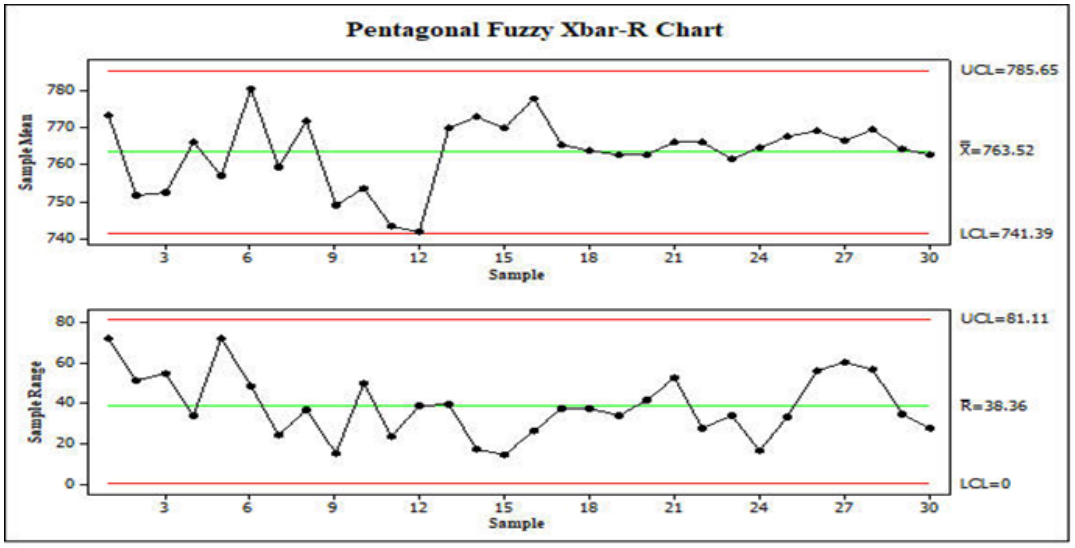


Figure 2: *Pentagonal Fuzzy $\bar{X} - R$ Chart*

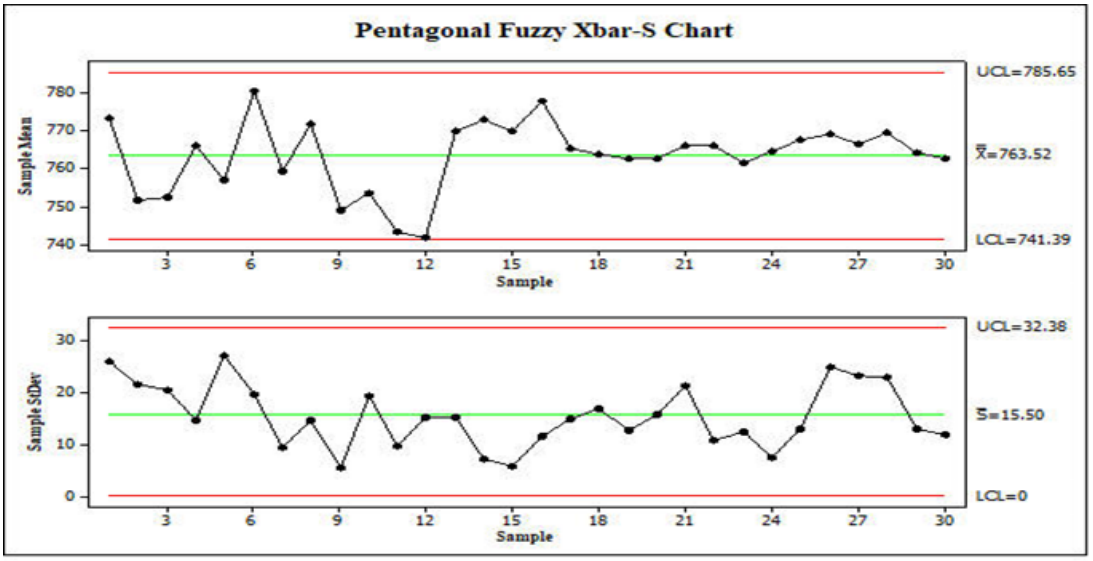


Figure 3: *Pentagonal Fuzzy* $\bar{X} - S$ *Chart*

It is shown that there is no point out of control after plotting the $\bar{X}$-R and $\bar{X}$-S control charts.

IV. Conclusion

In this paper, it is shown that FPN is suitable for traditional variable control charts. If uncertainty is presented in the data, the FPN control chart theory should control the process. In this study, the population parameter (μ *and* σ) is unknown, and we develop the theory for fuzzy $\bar{X}$-R and $\bar{X}$-S chart by using the rank membership function of FPN. More fuzzy control charts for variable and attribute data were done by $\alpha - cut$, triangular and trapezoidal fuzzy numbers in several published articles. The result of the illustrative example is done with this proposed technique, and it is shown that the fuzzy process is under control and capable. The proposed FPN control chart effectively increases the process flexibility. For further studies, the process capability indices, p-np chart, fuzzy CUSUM and fuzzy EWMA chart can be used to detect the slight shifting in the FPN process control with fuzzy data.

Reference

[1] Venkatesh, A., & Manoj, A. B. (2020). A fuzzy mathematical model for vegetable diet plan using ranking of pentagonal fuzzy number. *Malaya Journal of Matematik (MJM)*, *8*(4, 2020), 1995-1999.

[2] Panda, A., & Pal, M. (2015). A study on pentagonal fuzzy number and its corresponding matrices. *Pacific Science Review B: Humanities and Social Sciences*, *1*(3), 131-139.

[3] Pathinathan, T., & Ponnivalavan, K. (2014). Pentagonal fuzzy number. *International journal of computing algorithm*, *3*, 1003-1005.

[4] Chakraborty, A., Mondal, S. P., Alam, S., Ahmadian, A., Senu, N., De, D., & Salahshour, S. (2019). The pentagonal fuzzy number: its different representations, properties, ranking, defuzzification and application in game problems. *Symmetry*, *11*(2), 248.

[5] Chakraborty, A., Broumi, S., & Singh, P. K. (2019). *Some properties of pentagonal neutrosophic numbers and its applications in transportation problem environment*. Infinite Study.

[6] Shafqat, A., Huang, Z., & Aslam, M. (2021). Design of X-bar control chart based on Inverse Rayleigh Distribution under repetitive group sampling. *Ain Shams Engineering Journal*, *12*(1), 943-953.

[7] Lim, S. A. H., Antony, J., He, Z., & Arshed, N. (2017). Critical observations on the statistical process control implementation in the UK food industry: A survey. *International Journal of*

Quality & Reliability Management.

[8] Ahmad, M., & Cheng, W. (2022). A novel approach of fuzzy control chart with fuzzy process capability indices using alpha cut triangular fuzzy number. Mathematics, 10(19), 3572.

[9] Ahmad, M., Cheng, W., Haq, A., & Shah, S. K. (2023). Construction of fuzzy $\bar{X}-$ S control chart using trapezoidal fuzzy number with unbalanced data. Journal of Statistical Computation and Simulation, 93(4), 634-645.

[10] Jabbar, R. R., & Alkhafaji, A. A. A. (2023). Analysis of Traditional and Fuzzy Quality Control Charts to Improve Short-Run Production in the Manufacturing Industry. *Journal of Engineering*, *29*(6), 159-176.

[11] Özdemir, A. (2021). Development of fuzzy $$\bar {X}-S $$ $\bar{X}$-S control charts with unbalanced fuzzy data. *Soft Computing*, *25*(5), 4015-4025.

[12] Yeh, C. T. (2008). On improving trapezoidal and triangular approximations of fuzzy numbers. *International Journal of Approximate Reasoning*, *48*(1), 297-313.

[13] Şentürk, S., Erginel, N., Kaya, İ., & Kahraman, C. (2014). Fuzzy exponentially weighted moving average control chart for univariate data with a real case application. *Applied Soft Computing*, *22*, 1-10.

[14] Şentürk, S., Erginel, N., Kaya, I., & Kahraman, C. (2011). Design of fuzzy ũ control charts. *Journal of Multiple-Valued Logic & Soft Computing*, *17*(5-6), 459-473.

[15] Alipour, H., & Noorossana, R. (2010). Fuzzy multivariate exponentially weighted moving average control chart. *The International Journal of Advanced Manufacturing Technology*, *48*(9), 1001-1007.

[16] Erginel, N., & Şentürk, S. (2016). Fuzzy EWMA and fuzzy CUSUM control charts. In *Fuzzy Statistical Decision-Making* (pp. 281-295). Springer, Cham.

[17] Erginel, N. (2014). Fuzzy rule-based $\tilde p $ and $ n\tilde p $ control charts. *Journal of intelligent & fuzzy systems*, *27*(1), 159-171.

[18] Montgomery, D. C. (2009). *Introduction to statistical quality control*. John Wiley & Sons.

[19] Zadeh, L. A., "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338-353, 1965.

π-Power Half Logistic-G Family: Applications to Medical and Traffic Data

Pankaj Kumar[1], Laxmi Prasad Sapkota*[2] and Vijay Kumar[3]

•

[1,3]Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur, India
[2]Department of Statistics, Tribhuvan University, Tribhuvan Multiple Campus, Palpa, Nepal
[1]pankajagadish@gmail.com, [2]laxmisapkota75@gmail.com, [3]vkgkp@rediffmail.com

Abstract

This research article introduces a novel family of distributions achieved through the methodology of the π-power transformation technique. The study focuses on one specific member that is inverse Weibull distribution within this family, which showcases a hazard function exhibiting distinct J, reverse-J, bathtub, or monotonically increasing shapes. The article explores the essential characteristics of this distribution and employs the maximum likelihood estimation (MLE) method to estimate its associated parameters. To evaluate the accuracy of the estimation procedure, a simulation experiment is conducted, revealing a decrease in biases and mean square errors as sample sizes increase, even when working with small samples.

Furthermore, the practical application of the proposed distribution is demonstrated by analyzing two real medical and traffic datasets. By employing model selection criteria and conducting goodness-of-fit test statistics, the article establishes that the proposed model surpasses existing models in performance. The application of this research work can be significant in various fields where modeling and analyzing hazard functions or survival data are essential, while also making contributions to probability theory and statistical inferences.

Keywords: π-Power transformation, Half logistic, Reliability, Pivotal quantity

1. Introduction

Statistical models play a crucial role in representing and analyzing datasets in practical applications. While traditional distributions such as Weibull, Lomax, gamma, log-normal, exponential, beta, etc. have been widely used, they may not always provide a satisfactory fit for complex datasets. To address this limitation, researchers have been actively working on developing new models that offer greater adaptability and generality. These advancements often involve techniques such as exponentiation and the T-X approach to generate more flexible distributions. In this research article, we concentrate on an alternative approach called the π-power transformation (PPT) family, which was introduced by Lone and Jan [14]. The PPT family provides a distinct blend of high skewness and flexibility to the base distribution. The authors specifically examined the Pie-exponentiated Weibull as a member of this family. Prior to its introduction, the alpha power transformation (APT) approach had gained significant popularity among researchers in the fields of probability theory and survival analysis. Using the APT technique, numerous authors have put forward new generalized models and distribution families. For instance, Nassar et al. [20] employed the APT technique to define the new family of distributions using log transformation. Mead et al. [18] have further studied the APT family by providing some mathematical properties which were not provided in Mahdavi and Kundu [15]. Further, Lomax distribution was transformed using APT by Maruthan and Venkatachalam [17]. Ihtisham et al. [9] and Ihtisham et al. [10] studied the Pareto and inverse Pareto distributions using the APT approach, with the inverse Pareto distribution being applied to model real data related to extreme values. Similarly, Hozaien et al. [8] and Klakattawi and Aljuhani [12] introduced new models using the APT family of distributions. Alotaibi et al. [1] have introduced a new distribution as a weighted form of the

APT method while Gomma et al. [6] introduced the Alpha power of the power Ailamujia distribution, which offers a flexible hazard function. They utilized this distribution to model COVID-19 datasets from Italy and the UK. Furthermore, Nassar et al. [22] defined a new family utilizing the quantile function of the APT family whose cumulative distribution function (CDF) is

$$F(t) = \frac{\log\left[1 + (\alpha - 1)\, G\left(t;\phi\right)\right]}{\log\left(\alpha\right)}; \quad t > 0, \alpha > 0, \alpha \neq 1,$$

where $G\left(t;\phi\right)$ is the CDF and $\phi$is the parameter space of base distribution. Similarly, Elbatal et al. [5] introduced another new APT family whose CDF is

$$F\left(x\right) = \frac{\alpha^{G(x)} G\left(x\right)}{\alpha}; \alpha > 0, x \in \Re.$$

Similarly, Kyurkchiev [13] has introduced a family of distribution based on the Verhulst logistic function and its CDF is

$$F\left(x\right) = \frac{2G\left(x\right)}{1 - G\left(x\right)}; x \in \Re.$$

Another new method for transformation can be found in Kavya and Manoharan [11] and the CDF of this transformation is

$$F\left(x\right) = \frac{e}{e-1}\left\{1 - e^{-G(x)}\right\}; x \in \Re.$$

Also using the APT method Mandouch et al. [16] have reported a new two-parameter family of distributions whose CDF is

$$F\left(x\right) = \frac{\alpha^{kW\{G(x)\}} - 1}{\alpha - 1}; \alpha > 0, \alpha \neq 1, x \in \Re.$$

Lone and Jan [14] have introduced another new family using the concept of the APT family and named it the Pie-Exponentiated transformed (PET) family whose CDF is

$$F\left(x\right) = \frac{\pi^{\{G(x)\}^{\alpha}} - 1}{\pi - 1}; \alpha > 0, x \in \Re.$$

Hence, researchers are continuously developing and exploring new models and families of distributions to better capture the characteristics of complex datasets. The PET family has emerged as a popular approach, offering increased skewness and flexibility to the base distribution. These advancements have led to the proposal of various generalized models and distributions, which have been successfully applied to a range of datasets, including those related to COVID-19, reliability engineering, and extreme values. Building upon the concept of the PET, we have introduced a novel method to enhance existing distributions by incorporating a logistic form of CDF of any continuous distribution, which we refer to as the π-power half logistic-G (π-PHL-G) family of distributions. This new family offers increased robustness compared to other compound probability distributions and demonstrates great potential for modeling real-life datasets. The suggested family possesses two parameters that enable it to capture a broader range of characteristics exhibited by a dataset, including skewness, kurtosis, failure rate, and mathematically tractable. This enhanced flexibility allows for a more accurate representation of complex data patterns and distributional properties. By considering the π-PHL-G family, researchers, and practitioners can better account for the intricate nature of real-world datasets, leading to improved modeling outcomes. Among the members of the π-PHL-G family, one distribution stands out as particularly noteworthy-the inverse Weibull distribution. The inverse Weibull distribution has long been employed in reliability theory and life testing due to its ability to capture failure rates and survival probabilities effectively see [19, 24, 25]. With the integration of the π-PHL-G framework, the Weibull distribution can be further adapted and refined to better align with the unique characteristics observed in various applications. We have organized the remaining sections of this paper as follows; π-PHL-G family is introduced in Section 2, while its particular member, the π-PHL-Weibull distribution, is presented in Section 3. Some statistical properties are discussed in Section 4, and in Section 5, we discuss statistical inferences of the π-PHLIW distribution. The simulation experiment, application, and conclusion of the suggested model are presented in Sections 6, 7, and 8 respectively.

2. π-PHL-G Family and Some Important Functions

Let $Y \sim$ π-PHL-G family, then the CDF and PDF of π-PHL-G family $U(y;\Psi)$ and $u(y;\Psi)$ for $y \in \Re$, and $\Psi > 0$ is vector of parameters are defined as

$$U(y;\Psi) = \frac{\pi^{\left(\frac{2T(y;\Psi)}{1+T(y;\Psi)}\right)} - 1}{\pi - 1}; \quad y \in \Re. \tag{1}$$

$$u(y;\Psi) = \frac{(\log \pi)}{\pi - 1}\pi^{\left(\frac{2T(y;\Psi)}{1+T(y;\Psi)}\right)}\frac{2t(y;\Psi)}{\left[1+T(y;\Psi)\right]^2}; \quad y \in \Re. \tag{2}$$

where $T(y;\Psi)$ and $t(y;\Psi)$ are the CDF and PDF of any continuous distribution and $\bar{T}(y;\Psi)$ is the reliability function. Further reliability and hazard functions of π-PHL-G family can be expressed as

$$R(y;\Psi) = 1 - \left\{\frac{\pi^{\left(\frac{2T(y;\Psi)}{1+T(y;\Psi)}\right)} - 1}{\pi - 1}\right\}; \quad y \in \Re.$$

$$h(y;\Psi) = \frac{(\log \pi)}{\pi - 1}\pi^{\left(\frac{2T(y;\Psi)}{1+T(y;\Psi)}\right)}\frac{2t(y;\Psi)}{\left[1+T(y;\Psi)\right]^2}\left[1 - \left\{\frac{\pi^{\left(\frac{2T(y;\Psi)}{1+T(y;\Psi)}\right)} - 1}{\pi - 1}\right\}\right]^{-1}.$$

2.1. Qunatile function and Random deviation

$$Q_Y(p) = T^{-1}\left\{\frac{\log\left((\pi-1)p+1\right)}{\left(2\log\pi - \log\left((\pi-1)p+1\right)\right)}\right\}. \tag{3}$$

and

$$y = T^{-1}\left\{\frac{\log\left((\pi-1)u+1\right)}{\left(2\log\pi - \log\left((\pi-1)u+1\right)\right)}\right\}.$$

2.2. Linear form of π-PHL-G distribution

The CDF defined in Equation 6 can be expressed in the linear form as

$$U(y;\Psi) = \frac{1}{\pi-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{2^i(\log\pi)^i}{i!}\binom{i}{j}T^{i+j}(y;\Psi) - \frac{1}{\pi-1}. \tag{4}$$

Now differentiating Equation 5 with respect to y we get the linear form of PDF as

$$u(y;\Psi) = \frac{1}{\pi-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{2^i(i+j)(\log\pi)^i}{i!}\binom{i}{j}T^{i+j-1}(y;\Psi)t(y;\Psi).$$

$$u(y;\Psi) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}T^{i+j-1}(y;\Psi)t(y;\Psi), \tag{5}$$

where $\Delta_{ij} = \left(\frac{1}{\pi-1}\right)\frac{2^i(i+j)(\log\pi)^i}{i!}\binom{i}{j}$.

3. π-Power Half Logistic Inverse Weibull (π-PHLIW) Distribution

Let Y be a continuous random variable following the inverse Weibull distribution, then the CDF and PDF are

$$T(y;\Psi) = e^{-\beta y^{-\delta}}.$$

$$t(y;\Psi) = \beta\delta y^{-(\delta+1)}e^{-\beta y^{-\delta}}.$$

Now using Equation 3 as a base distribution, we introduce a new distribution π-PHLIW distribution as a special member having CDF

$$U(y;\beta,\delta)=\frac{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)}-1}{\pi-1} \quad ;\beta,\delta>0, y>0. \tag{6}$$

The PDF of the π-PHLIW distribution can be expressed as

$$u(y;\beta,\delta)=\frac{2\beta\delta(\log\pi)}{\pi-1}\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)}\frac{y^{-(\delta+1)}e^{-\beta y^{-\delta}}}{\left[1+e^{-\beta y^{-\delta}}\right]^2}; \quad y>0. \tag{7}$$

Now some key functions like reliability and hazard function of the π-PHLIW distribution can be presented as

$$R(y;\beta,\delta)=1-\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)}-1\right\}; \quad y>0.$$

$$h(y;\beta,\delta)=\frac{2\beta\delta(\log\pi)}{\pi-1}\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)}\frac{y^{-(\delta+1)}e^{-\beta y^{-\delta}}}{\left[1+e^{-\beta y^{-\delta}}\right]^2}\left[1-\left\{\frac{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)}-1}{\pi-1}\right\}\right]^{-1}.$$

3.1. Quantile function and Random deviation

The quantile function for the suggested distribution can be obtained by inverting the CDF defined in Equation 6 as

$$Q_Y(p)=\left[\log\left\{\frac{\log\left((\pi-1)\,p+1\right)}{\left(2\log\pi-\log\left((\pi-1)\,p+1\right)\right)}\right\}^{-1/\beta}\right]^{-1/\delta}. \tag{8}$$

also, random number deviate can be expressed as

$$y=\left[\log\left\{\frac{\log\left((\pi-1)\,u+1\right)}{\left(2\log\pi-\log\left((\pi-1)\,u+1\right)\right)}\right\}^{-1/\beta}\right]^{-1/\delta}.$$

The π-PHLIW distribution has a density plot that can take on a diversity of shapes, including symmetrical, left-skewed, right-skewed, or decreasing, and Figure 1 (left) shows some examples of these shapes. The HRF, on the other hand, can take on the shapes of an increasing, a j, or a reverse-j, and Figure 1 (right) shows some examples of these shapes.

4. Statistical properties of π-PHLIW distribution

4.1. Linear form of PDF of π-PHLIW distribution

After some mathematics using Equation 7, the PDF of π-PHLIW distribution can be obtained in a linear form as

$$u(y;\beta,\delta)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}y^{-(\delta+1)}e^{-(i+j)\beta y^{-\delta}}, \tag{9}$$

where $\Delta_{ij}^{*}=\beta\delta\Delta_{ij}$.

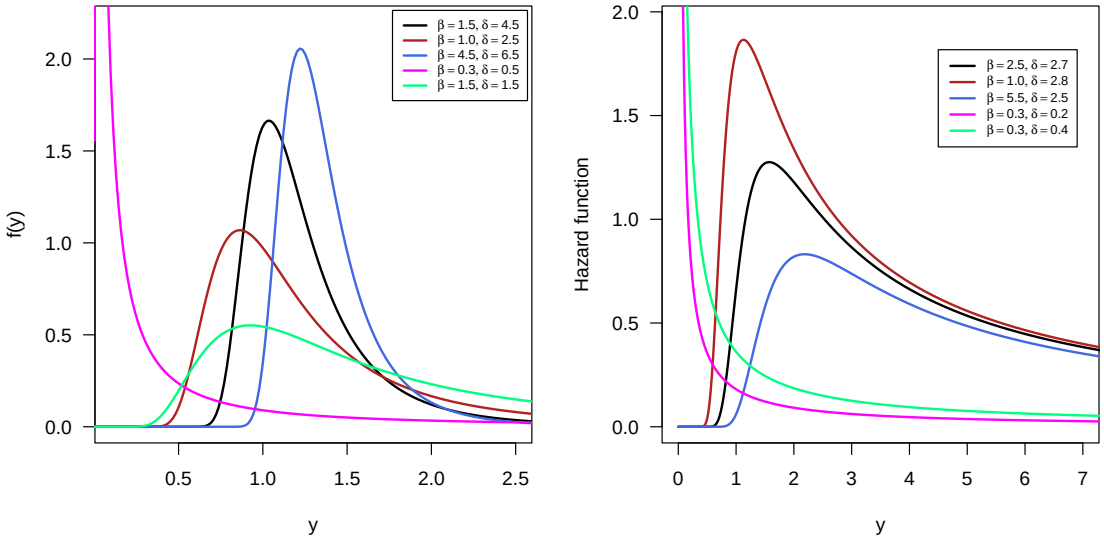


Figure 1: *Shapes of PDF and HRF of π-PHLIW distribution.*

4.2. Moments

The r^{th} moment of π-PHLIW distribution is

$$\begin{aligned} E[Y^r] &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\int_0^{\infty} y^{r-\delta-1}e^{-(i+j)\beta y^{-\delta}}dy \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\int_0^{\infty} t^{-\frac{r}{\delta}+1-1}e^{-(i+j)\beta t}dt \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\Gamma\left(-\frac{r}{\delta}+1\right)}{\{(i+j)\,\beta\}^{-\frac{r}{\delta}+1}} \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\Gamma\left(\frac{\delta-r}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-r}{\delta}}}; \quad \delta > r. \end{aligned} \tag{10}$$

Now mean and variance of π-PHLIW distribution can be expressed as

$$E[Y] = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\Gamma\left(\frac{\delta-1}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-1}{\delta}}}; \quad \delta > 1.$$

and

$$E\left[Y^2\right] = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\Gamma\left(\frac{\delta-2}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-2}{\delta}}}; \quad \delta > 2.$$

$$\begin{aligned} V[Y] &= E\left[Y^2\right] - [E(Y)]^2 \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\Gamma\left(\frac{\delta-2}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-2}{\delta}}} - \left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\Gamma\left(\frac{\delta-1}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-1}{\delta}}}\right]^2; \delta > 2. \end{aligned}$$

4.3. Moment Generating Function (MGF)

For any real number t, the MGF of π-PHLIW distribution can be defined as

$$\begin{aligned}
M_Y(t) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{t^k}{k!}\int_0^{\infty} y^{r-(\delta+1)}e^{-(i+j)\beta y^{-\delta}}dy \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{t^k}{k!}\int_0^{\infty} t^{-\frac{r}{\delta}+1-1}e^{-(i+j)\beta t}dt \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{t^k}{k!}\frac{\delta^{-1}\Gamma\left(-\frac{r}{\delta}+1\right)}{\{(i+j)\,\beta\}^{-\frac{r}{\delta}+1}} \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{t^k}{k!}\frac{\delta^{-1}\Gamma\left(\frac{\delta-r}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-r}{\delta}}};\delta>r.
\end{aligned}$$

4.4. Characteristic Function

For any real number t, the characteristic function of π-PHLIW distribution can be defined as

$$\begin{aligned}
\Phi_Y(t) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{(vt)^k}{k!}\int_0^{\infty} y^{r-(\delta+1)}e^{-(i+j)\beta y^{-\delta}}dy \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{(vt)^k}{k!}\int_0^{\infty} t^{-\frac{r}{\delta}+1-1}e^{-(i+j)\beta t}dt \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{(vt)^k}{k!}\frac{\delta^{-1}\Gamma\left(-\frac{r}{\delta}+1\right)}{\{(i+j)\,\beta\}^{-\frac{r}{\delta}+1}} \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\Delta_{ij}^{*}\frac{(vt)^k}{k!}\frac{\delta^{-1}\Gamma\left(\frac{\delta-r}{\delta}\right)}{\{(i+j)\,\beta\}^{\frac{\delta-r}{\delta}}};\delta>r.
\end{aligned}$$

where $v=\sqrt{-1}$.

4.5. Incomplete moment

The incomplete moment for π-PHLIW distribution is given by

$$\begin{aligned}
M_r(z) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\int_0^{z} y^{r-\delta-1}e^{-(i+j)\beta y^{-\delta}}dy \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\left[\gamma\left(\frac{\delta-r}{\delta},(i+j)\,\beta z^{-\delta}\right)\right]}{\{(i+j)\,\beta\}^{\frac{\delta-r}{\delta}}};\delta>r.
\end{aligned}$$

where $\gamma(.)$ incomplete gamma function.

4.6. Mean Residual Life

The Mean residual life for π-PHLIW distribution is given by

$$\begin{aligned}
\overline{M(z)} &= \frac{1}{F(z)}\left[\mu-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\int_0^{z} y^{-\delta}e^{-(i+j)\beta y^{-\delta}}dy\right]-z \\
&= \frac{1}{F(z)}\left[\mu-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Delta_{ij}^{*}\frac{\delta^{-1}\left[\gamma\left(\frac{\delta+1}{\delta},(i+j)\,\beta z^{-\delta}\right)\right]}{\{(i+j)\,\beta\}^{\frac{\delta+1}{\delta}}}\right]-z
\end{aligned}$$

where $\gamma(.)$ incomplete gamma function.

4.7. Order Statistics

Let $y_i(i = 1, 2, ..., n) \sim \pi - PHLIW(y; \beta, \delta)$ with CDF $U(y_i; \beta, \delta)$ and PDF $u(y_i; \beta, \delta)$. If $u_r(y)$ denote the PDF of rth order statistic $Y_{(r)}$, then their CDF and PDF are given by $U_r(y) = I_{U(y)}(r, n - r + 1)$

$$u_r(y) = \frac{d}{dy}\left[U_r(y)\right] = \frac{d}{dy}\left[I_{U(y)}(r, n-r+1)\right] = \frac{1}{B(r, n-r+1)} U^{r-1}(y)u(y)\left[1 - U(y)\right]^{n-r}.$$

$$u_r(y) = \frac{1}{B(r, n-r+1)} \frac{2\beta\delta(\log\pi)}{\pi - 1} \pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} \frac{y^{-(\delta+1)} e^{-\beta y^{-\delta}}}{\left[1 + e^{-\beta y^{-\delta}}\right]^2} \left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right]^{r-1}$$
$$\left[1 - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right\}\right]^{n-r}.$$

The CDF and PDF of first order statistic $Y_{(1)}$ are given by

$$U_1(y) = 1 - \left[1 - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right\}\right]^n; y > 0.$$

$$u_1(y) = \frac{2n\beta\delta(\log\pi)}{\pi - 1} \pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} \frac{y^{-(\delta+1)} e^{-\beta y^{-\delta}}}{\left[1 + e^{-\beta y^{-\delta}}\right]^2} \left[1 - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right\}\right]^{n-1}; y > 0.$$

The CDF and PDF of first order statistic $Y_{(n)}$ are given by

$$U_n(y) = \left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right]^n; y > 0.$$

$$u_n(y) = \frac{2n\beta\delta(\log\pi)}{\pi - 1} \pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} \frac{y^{-(\delta+1)} e^{-\beta y^{-\delta}}}{\left[1 + e^{-\beta y^{-\delta}}\right]^2} \left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right]^{n-1}; y > 0.$$

The Joint PDF of r^{th} and sth order statistics is given by

$$u_{rs}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} U^{r-1}(x).u(x)\left[U(y) - U(x)\right]^{s-r-1} u(y).\left[1 - U(y)\right]^{n-s}$$

$$u_{rs}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \pi^{\left\{\left(\frac{2e^{-\beta x^{-\delta}}}{1+e^{-\beta x^{-\delta}}}\right) + \left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)\right\}} \frac{e^{-\beta(x^{-\delta} + y^{-\delta})} (xy)^{-(\delta+1)}}{\left[1 + e^{-\beta y^{-\delta}}\right]^2 \left[1 + e^{-\beta x^{-\delta}}\right]^2}$$
$$\left[\frac{2\beta\delta(\log\pi)}{\pi - 1}\right]^2 \left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta x^{-\delta}}}{1+e^{-\beta x^{-\delta}}}\right)} - 1\right\}\right]^{r-1}$$
$$\left[\left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right\} - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta x^{-\delta}}}{1+e^{-\beta x^{-\delta}}}\right)} - 1\right\}\right\}\right]^{s-r-1}$$
$$\left[1 - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right\}\right]^{n-s}; x > 0, y > 0.$$

The Joint PDF of the 1^{st} and n^{th} order statistics are given by

$$u_{1n}(x, y) = n(n-1)\left[U(y) - U(x)\right]^{n-2} u(x).u(y)$$

$$u_{1n}(x, y) = n(n-1)\left(\frac{2\beta\delta(\log\pi)}{\pi - 1}\right)^2 \left[\left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)} - 1\right\}\right\} - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta x^{-\delta}}}{1+e^{-\beta x^{-\delta}}}\right)} - 1\right\}\right\}\right]^{n-2}$$
$$\pi^{\left\{\left(\frac{2e^{-\beta x^{-\delta}}}{1+e^{-\beta x^{-\delta}}}\right) + \left(\frac{2e^{-\beta y^{-\delta}}}{1+e^{-\beta y^{-\delta}}}\right)\right\}} \frac{e^{-\beta(x^{-\delta} + y^{-\delta})} (xy)^{-(\delta+1)}}{\left[1 + e^{-\beta y^{-\delta}}\right]^2 \left[1 + e^{-\beta x^{-\delta}}\right]^2}; x > 0, y > 0$$

4.8. System Reliability Function

4.8.1 Series System:

Consider a system with n independent components, each component follows $\pi - PHLIW(y;\beta,\delta)$ distribution. Let's assume $T_i(i=1,2,...,n) \sim \pi - PHLIW(y;\beta,\delta)$ with CDF $U(t_i;\beta,\delta)$ and PDF $u(t_i;\beta,\delta)$, then the system reliability for linear consecutive (series system) is given by

$$R_S(t) = \prod_{i=1}^{n} R_i(t) = \left[1 - \left\{\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right\}\right]^n \quad ;\beta,\delta > 0, t > 0.$$

The CDF system reliability for linear consecutive (series system) is given by

$$F_S(t) = 1 - \prod_{i=1}^{n} R_i(t) = 1 - \left[1 - \left\{\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right\}\right]^n \quad ;\beta,\delta > 0, t > 0. \qquad (11)$$

Differentiating the Equation 11, the PDF system reliability for linear consecutive (series system) is given by

$$f_S(t) = \frac{dF_S(t)}{dt} = \frac{2n\beta\delta(\log\pi)}{\pi-1}\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)}\frac{t^{-(\delta+1)}e^{-\beta t^{-\delta}}}{\left[1+e^{-\beta t^{-\delta}}\right]^2}\left[1 - \left\{\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right\}\right]^{n-1}$$

The Hazard function system reliability for linear consecutive (series system) is obtained by the ratio of PDF and Reliability function of linear consecutive (series system).

$$h_S(t) = \frac{f_S(t)}{R_S(t)} = \frac{2n\beta\delta(\log\pi)}{\pi-1}\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)}\frac{t^{-(\delta+1)}e^{-\beta t^{-\delta}}}{\left[1+e^{-\beta t^{-\delta}}\right]^2}\left[1 - \left\{\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right\}\right]^{-1}.$$

4.8.2 Parallel System:

Consider a system with n independent components, each component follows $\pi - PHLIW(y;\beta,\delta)$ distribution. Lets assume $T_i(i=1,2,...,n) \sim \pi - PHLIW(y;\beta,\delta)$ with CDF $U(t_i;\beta,\delta)$ and PDF $u(t_i;\beta,\delta)$, then the reliability function for parallel system is given by

$$R_P(t) = 1 - \prod_{i=1}^{n}(1 - R_i(t)) = 1 - \left[\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right]^n \quad ;\beta,\delta > 0, t > 0.$$

The CDF for the parallel system is given by

$$F_P(t) = \left[\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right]^n \quad ;\beta,\delta > 0, t > 0.$$

Differentiating the above equation, the PDF formula for the parallel system is given by

$$f_P(t) = \frac{2n\beta\delta(\log\pi)}{\pi-1}\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)}\frac{t^{-(\delta+1)}e^{-\beta t^{-\delta}}}{\left[1+e^{-\beta t^{-\delta}}\right]^2}\left[\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right]^{n-1}$$

The Hazard function for a parallel system is obtained by the ratio of PDF and the Reliability function of a parallel system.

$$h_P(t) = \frac{\frac{2n\beta\delta(\log\pi)}{\pi-1}\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)}\frac{t^{-(\delta+1)}e^{-\beta t^{-\delta}}}{\left[1+e^{-\beta t^{-\delta}}\right]^2}\left[\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right]^{n-1}}{1 - \left[\frac{1}{\pi-1}\left\{\pi^{\left(\frac{2e^{-\beta t^{-\delta}}}{1+e^{-\beta t^{-\delta}}}\right)} - 1\right\}\right]^n}.$$

5. Statistical Inference

5.1. Estimation

The parameters of the suggested model are estimated using the maximum likelihood method (MLE). Let $y_i(i = 1, 2, ..., m) \sim \pi - PHLIW(y; \beta, \delta)$ with PDF $u(y_i; \beta, \delta)$then the log-likelihood function can be calculated as

$$\ell\left(\underline{y}; \beta, \delta\right) = n\log\left(2\beta\delta(\log\pi)\right) - n\log\left(\pi - 1\right) + \left(\log\pi\right)\sum_{i=1}^{n}\left(\frac{2e^{-\beta y_i^{-\delta}}}{1 + e^{-\beta y_i^{-\delta}}}\right) - \left(\delta + 1\right)\sum_{i=1}^{n}\log y_i - \beta\sum_{i=1}^{n} y_i^{-\delta} - 2\sum_{i=1}^{n}\log\left(1 + e^{-\beta y_i^{-\delta}}\right) \tag{12}$$

Differentiating Equation 12 with respect to associated parameters, we get the first-order derivatives as

$$\frac{\partial\ell}{\partial\beta} = \frac{n}{\beta} - \left(\log\pi\right)\sum_{i=1}^{n}\frac{2y_i^{-\delta}e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2} - \sum_{i=1}^{n} y_i^{-\delta} + 2\sum_{i=1}^{n}\frac{y_i^{-\delta}e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}}.$$

$$\frac{\partial\ell}{\partial\delta} = \frac{n}{\delta} + \left(\log\pi\right)\sum_{i=1}^{n}\frac{2\beta y_i^{-\delta}\left(\log y_i\right)e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2} - \sum_{i=1}^{n}\log y_i + \beta\sum_{i=1}^{n} y_i^{-\delta}\left(\log y_i\right) - 2\sum_{i=1}^{n}\frac{y_i^{-\delta}\left(\log y_i\right)\beta e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}}$$

By solving the above two non-linear equations using suitable software one can obtain the estimates under the MLE method.

5.2. Cramer-Rao (CR) Inequality

If $T(y_1, ..., y_n)$ is an unbiased estimator for $k(\omega)$, a function of parameter ω, then

$$Var[T(y_1, ..., y_n)] \geqslant \frac{\left\{\frac{d}{d\omega}k(\omega)\right\}^2}{E\left(\frac{\partial}{\partial\omega}logL\right)} = \frac{\left\{k'(\omega)\right\}^2}{I(\omega)},$$

where $I(\omega)$ is the information on ω, supplied by the sample. To define CR lower bound (CRLB) for β when δ is supposed to be known, then CRLB for an unbiased estimator $T_1(y_1, ..., y_n)$ of a parameter β is given by $\frac{1}{I(\beta)}$, where

$$I\left(\beta\right) = -E\left[\frac{\partial^2\ell}{\partial\beta^2}\right] = \frac{n}{\beta^2} + \left(\log\pi\right)\sum_{i=1}^{n}E\left\{\frac{\partial}{\partial\beta}\left\{\frac{2y_i^{-\delta}e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2}\right\}\right\} - 2\sum_{i=1}^{n}E\left\{\frac{\partial}{\partial\beta}\left\{\frac{y_i^{-\delta}e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}}\right\}\right\}$$

and

$$\frac{\partial^2\ell}{\partial\beta^2} = -\frac{n}{\beta^2} - \left(\log\pi\right)\sum_{i=1}^{n}\frac{\partial}{\partial\beta}\left\{\frac{2y_i^{-\delta}e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2}\right\} + 2\sum_{i=1}^{n}\frac{\partial}{\partial\beta}\left\{\frac{y_i^{-\delta}e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}}\right\}$$

Again CRLB for δ when β is supposed to be known, then CRLB for an unbiased estimator $T_2(y_1, ..., y_n)$ of a parameter δ is given by $\frac{1}{I(\delta)}$, where

$$I\left(\delta\right) = -E\left[\frac{\partial^2\ell}{\partial\delta^2}\right] = \frac{n}{\delta^2} - \left(\log\pi\right)\sum_{i=1}^{n}E\left\{\frac{\partial}{\partial\delta}\left\{\frac{2\beta y_i^{-\delta}\left(\log y_i\right)e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2}\right\}\right\} + \beta\sum_{i=1}^{n}E\left\{y_i^{-\delta}\left(\log y_i\right)^2\right\} + 2\sum_{i=1}^{n}E\left\{\frac{\partial}{\partial\delta}\left\{\frac{y_i^{-\delta}\left(\log y_i\right)\beta e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}}\right\}\right\}$$

And

$$\frac{\partial^2 \ell}{\partial \delta^2} = -\frac{n}{\delta^2} + (\log \pi) \sum_{i=1}^{n} \frac{\partial}{\partial \delta} \left\{ \frac{2\beta y_i^{-\delta} (\log y_i) e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2} \right\} - \beta \sum_{i=1}^{n} y_i^{-\delta} (\log y_i)^2 - 2 \sum_{i=1}^{n} \frac{\partial}{\partial \delta} \left\{ \frac{y_i^{-\delta} (\log y_i) \beta e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}} \right\}.$$

5.3. Asymptotical Properties

A consistent solution of the likelihood equation is asymptotically normally distributed about true value θ_0. Thus, $\hat{\theta}$ is asymptotically $N\left(\theta_0, \frac{1}{I(\theta_0)}\right)$ as $n \to \infty$, Now $\hat{\beta}$ is asymptotically $N\left(\beta, \frac{1}{I(\beta)}\right)$ as $n \to \infty$ where

$$I(\beta) = -E\left[\frac{\partial^2 \ell}{\partial \beta^2}\right] = \frac{n}{\beta^2} + (\log \pi) \sum_{i=1}^{n} E\left\{ \frac{\partial}{\partial \beta} \left\{ \frac{2 y_i^{-\delta} e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2} \right\} \right\} - 2 \sum_{i=1}^{n} E\left\{ \frac{\partial}{\partial \beta} \left\{ \frac{y_i^{-\delta} e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}} \right\} \right\}.$$

$\hat{\delta}$ is asymptotically $N\left(\delta, \frac{1}{I(\delta)}\right)$ as $n \to \infty$ where

$$I(\delta) = -E\left[\frac{\partial^2 \ell}{\partial \delta^2}\right] = \frac{n}{\delta^2} - (\log \pi) \sum_{i=1}^{n} E\left\{ \frac{\partial}{\partial \delta} \left\{ \frac{2\beta y_i^{-\delta} (\log y_i) e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}^2} \right\} \right\} + \beta \sum_{i=1}^{n} E\left\{ y_i^{-\delta} (\log y_i)^2 \right\}$$
$$+ 2 \sum_{i=1}^{n} E\left\{ \frac{\partial}{\partial \delta} \left\{ \frac{y_i^{-\delta} (\log y_i) \beta e^{-\beta y_i^{-\delta}}}{\left\{1 + e^{-\beta y_i^{-\delta}}\right\}} \right\} \right\}$$

Pivotal Quantity (PQ): Let $y_i (i = 1, ..., m) \sim \pi - PHLIW(y; \beta, \delta)$ with CDF $U(y_i; \beta, \delta)$ then pivotal quantity is defined as

$$-2 \sum_{i=1}^{n} \ln\left[U(y_i; \beta, \delta)\right] \sim \chi^2_{2n} \quad \Rightarrow -2 \sum_{i=1}^{n} \ln\left[1 - U(y_i; \beta, \delta)\right] \sim \chi^2_{2n}$$

$$PQ = -2 \sum_{i=1}^{n} \ln\left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y_i^{-\delta}}}{1+e^{-\beta y_i^{-\delta}}}\right)} - 1\right\}\right] \sim \chi^2_{2n}$$

$$\Rightarrow PQ = -2 \sum_{i=1}^{n} \ln\left[1 - \left\{\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y_i^{-\delta}}}{1+e^{-\beta y_i^{-\delta}}}\right)} - 1\right\}\right\}\right] \sim \chi^2_{2n}.$$

Let $x_i (i = 1, ..., m) \sim \pi - PHLIW(x; \beta, \delta)$ and $y_i (i = 1, ..., m) \sim \pi - PHLIW(y; \beta, \delta)$ are two independent random variable with CDF $U(x_i; \beta, \delta)$ and $U(y_i; \beta, \delta)$ respectively, then $\frac{PQ_1}{PQ_2} \sim Beta_2(m, n)$ and $\frac{PQ_1}{PQ_1 + PQ_2} \sim Beta_1(m, n)$ and $\frac{n}{m}\frac{PQ_1}{PQ_2} \sim F(m, n)$, where

$$PQ_1 = -2 \sum_{i=1}^{n} \ln\left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta x_i^{-\delta}}}{1+e^{-\beta x_i^{-\delta}}}\right)} - 1\right\}\right].$$

and

$$PQ_2 = -2 \sum_{i=1}^{n} \ln\left[\frac{1}{\pi - 1}\left\{\pi^{\left(\frac{2e^{-\beta y_i^{-\delta}}}{1+e^{-\beta y_i^{-\delta}}}\right)} - 1\right\}\right].$$

5.4. Confidence interval for Large Sample

Under certain conditions, the first derivative of the logarithm of the likelihood function w. r. to parameter θ viz., $\frac{\partial \log L}{\partial \theta}$, is asymptotically normal with mean zero and variance given by:

$$Var\left(\frac{\partial \log L}{\partial \theta}\right) = E\left(\frac{\partial \log L}{\partial \theta}\right)^2 = -E\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)$$

Hence for large n, $Z = \frac{\frac{\partial}{\partial\theta}\log L}{\sqrt{Var\left(\frac{\partial}{\partial\theta}\log L\right)}} \sim N(0,1)$ The result enables us to obtain a confidence interval for the parameter θ in a large sample. Thus for a large sample, the confidence interval for θ with confidence coefficient $(1-c)\%$ is obtained by converting the inequalities in $P(|Z| \leqslant \gamma_c) = 1-c$ where γ_c is given by $\frac{1}{2\pi}\int_{-\gamma_c}^{\gamma_c} \exp\left(-t^2/2\right)dt = 1-c$. Thus confidence interval for β and δ are given by $\widehat{\beta} \pm SE\left(\widehat{\beta}\right)$ and $\widehat{\delta} \pm SE\left(\widehat{\delta}\right)$ at the confidence coefficient $(1-c)\%$.

6. Simulation

In our research study, we employed the maxLik R package, developed by Henningsen and Toomet [7], to generate samples from the quantile function described in Equation 8 for various combinations of parameters of the $\pi - PHLIW$ distribution. The MLEs were then computed for each sample using the `maxLik()` function and the BFGS algorithm. This analysis allowed us to investigate issues related to parameter estimation and determine the direction and magnitude of bias in the MLEs, whether it be overestimation or underestimation.

In our simulation, we utilized sample sizes ranging from 100 to 400 with increments of 100. The entire process was repeated 1000 times in order to obtain estimates of the mean square error (MSE). The Biases and MSEs for the four different parameter combinations are presented in Tables 1 to 4. The results demonstrate that as the sample size increases, the corresponding Biases and MSEs decrease toward zero. This finding suggests that the MLE method exhibits asymptotic efficiency, and consistency, and follows the invariance property.

Table 1: *Bias and MSE for* $(\beta = 1.25, \delta = 0.5)$

Sample size	Bias		MSE	
	β	δ	β	δ
100	0.008	0.0085	0.0153	0.0018
200	0.001	0.0042	0.0072	9.00E-04
300	0.0035	0.0015	0.005	5.00E-04
400	0.0016	0.0026	0.0041	4.00E-04

Table 2: *Bias and MSE for* $(\beta = 0.75, \delta = 0.75)$

Sample size	Bias		MSE	
	β	δ	β	δ
100	0.0015	0.0102	0.0077	0.0036
200	-9.00E-04	0.0057	0.0035	0.0018
300	0.0032	0.0019	0.0025	0.0012
400	2.00E-04	0.0031	0.0018	9.00E-04

Table 3: *Bias and MSE for* $(\beta = 0.5, \delta = 1.25)$

Sample size	Bias		MSE	
	β	δ	β	δ
100	0.0035	0.0164	0.0046	0.0109
200	-0.0011	0.0083	0.0022	0.0055
300	0.0005	0.0031	0.0015	0.0036
400	-0.0018	0.0051	0.0011	0.0024

Table 4: *Bias and MSE for* $(\beta = 0.25, \delta = 2.5)$

Sample size	Bias		MSE	
	β	δ	β	δ
100	-6.00E-04	0.0207	0.0019	0.0149
200	0.0012	0.0056	9.00E-04	0.0068
300	7.00E-04	0.005	7.00E-04	0.005
400	-8.00E-04	0.0055	5.00E-04	0.0038

7. Application

In this section, we demonstrate the application of the π-PHLIW distribution using two real datasets. The datasets utilized for applying the suggested distribution are presented below.

Data set I:

A real data set of the relief time of 20 patients taking an analgesic is provided in this section and can be found in Chaudhary et al. [3]. Data are as follows: 1.4, 1.1, 1.7, 1.3, 1.8, 1.9, 2.2, 1.6, 2.7, 1.7, 1.8, 4.1, 1.2, 1.5, 3, 1.4, 2.3, 1.7, 2.0, 1.6

Data set II:

We consider the TRAFFIC data set given by Bain and Engelhardt [2] which represents 128 observations on times, in seconds, between the arrival of vehicles at a particular location on a road: "0.2, 0.5, 0.8, 0.8, 0.8, 1.0, 1.1, 1.2, 1.2, 1.2, 1.2, 1.2, 1.3, 1.4, 1.5, 1.5, 1.6, 1.6, 1.6, 1.7, 1.8, 1.8, 1.8, 1.8, 1.8, 1.9, 1.9, 1.9, 1.9, 1.9, 1.9, 1.9, 2.0, 2.1, 2.1, 2.2, 2.3, 2.3, 2.4, 2.4, 2.5, 2.5, 2.5, 2.6, 2.6, 2.7, 2.8, 2.8, 2.9, 3.0, 3.0, 3.1, 3.2, 3.4, 3.7, 3.9, 3.9, 3.9, 4.6, 4.7, 5.0, 5.1, 5.6, 5.7, 6.0, 6.0, 6.1, 6.6, 6.9, 6.9, 7.3, 7.6, 7.9, 8.0, 8.3, 8.8, 8.8, 9.3, 9.4, 9.5, 10.1, 11.0, 11.3, 11.9, 11.9, 12.3, 12.9, 12.9, 13.0, 13.8, 14.5, 14.9, 15.3, 15.4, 15.9, 16.2, 17.6, 20.1, 20.3, 20.6, 21.4, 22.8, 23.7, 24.7, 29.7, 30.6, 31.0, 33.7, 34.1, 34.7, 36.8, 40.1, 40.2, 41.3, 42.0, 44.8, 49.8, 51.7, 55.7, 56.5, 58.1, 70.5, 72.6, 87.1, 88.6, 91.7, 119.8, 125.3"

7.1. Model Analysis

We computed several well-known goodness-of-fit statistics to analyze both data sets I and II. The fitted models were evaluated using various metrics, including the log-likelihood value (-2logL), Akaike information criterion (AIC), Hannan-Quinn information criterion (HQIC), Anderson-Darling (AD), Kolmogrov-Smirnov (KS), and Cramer-von Mises (CVM) with corresponding p-values. All the essential computations and graphical plots were performed using the R software Wickham and Grolemund [26] and the R Core Team [23]. To compare the fitting capability of the $\pi - PHLIW$ model, we have selected several models such as the Lindley Weibull (LW) Cordeiro et al. [4], alpha power exponential (APE) Mahdavi and Kundu [15], Weibull, APT-Weibull (APTW) Nassar et al. [21], and new APT-Weibull (NAPTW) Elbatal et al. [5]. We have presented the KS and PP plots which provide an estimate of the CDF based on both data sets under study in Figures 2 and 3 (left for dataset-I, right for dataset-II). The estimated values of the parameters and their associated standard errors (SE) for both datasets were presented in Tables 5 and 7, which were obtained using the MLE method. Additionally, Tables 5 and 7 showcase model selection criteria such as log-likelihood, HQIC, and AIC, and goodness of fit statistics such as KS along with p-value for both datasets. Our observations show that the π-PHLIW model has the least statistics compared to the LW, APE, Weibull, APTW, and NAPTW distributions, along with the corresponding highest p-values. This indicates that the PiPHLIW model is more flexible and provides a good fit. Furthermore, we have provided graphical illustrations of the fitted models in Figures 6 and 8, which support our findings that the π-PHLIW model outperforms the other candidate models.

8. Conclusion

In this study, we have introduced an innovative distribution family called the π-power half logistic-G family. Drawing inspiration from the PPT methodology, we selected the Inverse Weibull distribu-

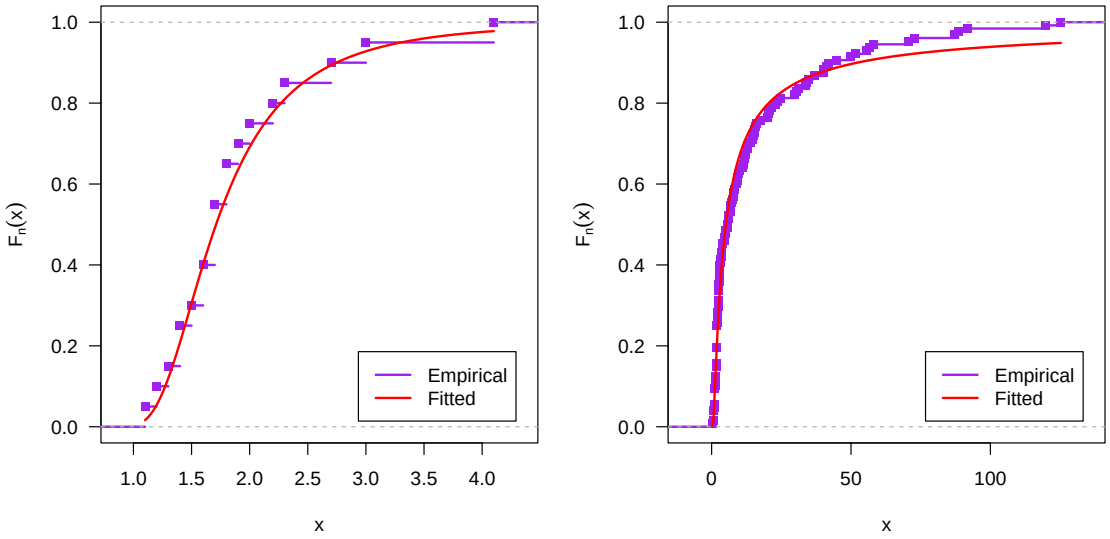


Figure 2: *Graphs of KS plot of $\pi - PHLIW$ distribution (left for dataset-I, right for dataset-II)*

Table 5: *Estimated parameters using the MLE method for the data set-I*

Model	parameter	SE	parameter	SE	parameter	SE
π-PHLIW(β, δ)	6.0338	1.9673	3.8496	0.6916	–	–
LW(α, β, λ)	9.2825	21.6417	2.0201	0.3020	0.0053	0.0241
APE(α, λ)	229.1815	16.7718	1.2110	0.1415	–	–
Weibull(α, β)	0.1216	0.0563	2.7869	0.4274	–	–
APTW$(\alpha, \delta, \lambda)$	93.1808	5.0649	1.6946	0.2693	0.6935	0.1681
NAPTW(α, β, λ)	105.5443	172.8892	1.5803	0.2237	0.8243	0.0895

tion as the foundation for this new family. The π-PHLIW distribution offers a wide range of hazard function shapes, including increasing, bathtub, J-shaped, and reverse-J-shaped. By employing the maximum likelihood estimation technique, we explored the statistical properties of this distribution and estimated its parameters. To assess the accuracy of our estimation method, we conducted a Monte Carlo simulation. The results revealed that the mean square errors decrease as the sample size increases, even when dealing with small samples. To demonstrate the practical utility of the π-PHLIW distribution, we applied it to two real engineering datasets. Through model selection criteria and goodness-of-fit tests, we compared its performance against seven existing models. Our findings strongly support the superiority of the π-PHLIW distribution over the alternative models, suggesting its potential application in various fields, such as medical and life sciences, reliability engineering, actuarial science, and survival analysis. Additionally, the π-power transformation family of distributions holds promise as a foundation for developing novel models in the future.

Conflict of interest: The authors declare that there is no conflict of interest.

References

[1] Alotaibi, R., Okasha, H., Rezk, H., and Nassar, M. (2021). A new weighted version of alpha power transformation method: Properties and applications to COVID-19 and software reliability data. *Physica Scripta*, 96(12), 125221.

[2] Bain, L. J., and Engelhardt, M. (1980). Probability of correct selection of Weibull versus gamma based on livelihood ratio. *Communications in statistics-theory and methods*, 9(4),

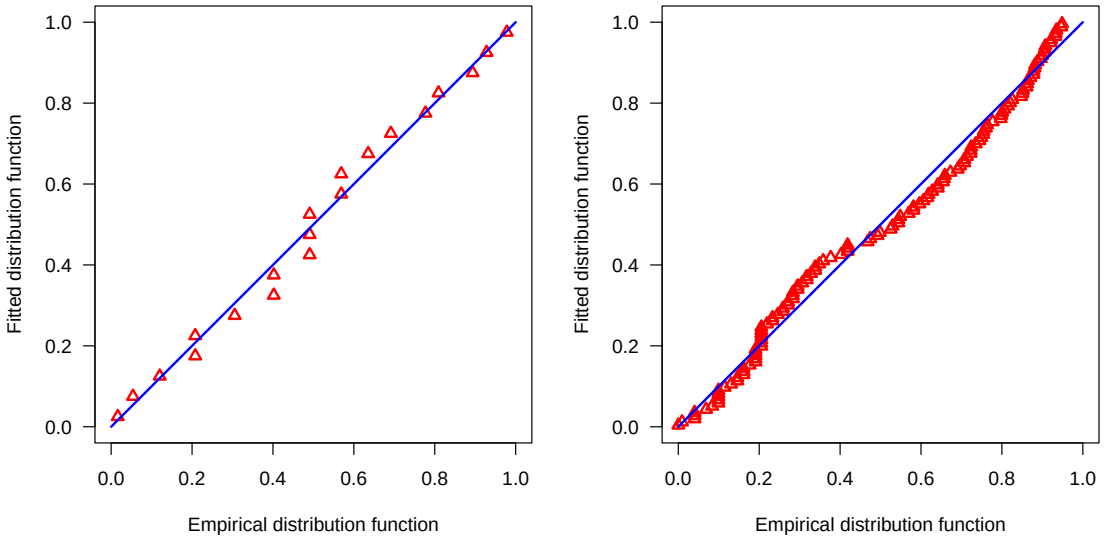


Figure 3: *Graphs of PP plot of $\pi - PHLIW$ distribution (left for dataset-I, right for dataset-II)*

Table 6: *Fitted statistics for the data set-I*

Model	-2logL	AIC	HQIC	KS	p(KS)	CVM	p(CVM)	AD	p(AD)
π-PHLIW	30.8149	34.8149	35.2036	0.1020	0.9854	0.0259	0.9895	0.1515	0.9986
LW	38.6683	44.6683	45.2514	0.1811	0.5282	0.1456	0.4060	0.8604	0.4380
APE	43.9186	47.9186	48.3074	0.2350	0.2195	0.2907	0.1433	1.6494	0.1449
Weibull	41.1728	45.1728	45.5616	0.1849	0.5009	0.1834	0.3037	1.0834	0.3155
APTW	38.0798	44.0798	44.6629	0.1557	0.7176	0.1082	0.5505	0.7223	0.5386
NAPTW	36.6835	42.6835	43.2666	0.1586	0.6955	0.0957	0.6109	0.6204	0.6268

375-381.

[3] Chaudhary, A. K., Sapkota, L. P., and Kumar, V. (2023). Inverse exponential power distribution: theory and applications. *International Journal of Mathematics, Statistics and Operations Research*. 3(1), 175-185.

[4] Cordeiro, G. M., Afify, A. Z., Yousof, H. M., Cakmakyapan, S., and Ozel, G. (2018). The Lindley Weibull distribution: properties and applications. *Anais da Academia Brasileira de Ciencias*, 90, 2579-2598.

[5] Elbatal, I., Ahmad, Z., Elgarhy, M., and Almarashi, A. M. (2018). A new alpha power transformed family of distributions: properties and applications to the Weibull model. *Journal of Nonlinear Science and Applications*, 12(1), 1-20.

[6] Gomaa, R. S., Hebeshy, E. A., El Genidy, M. M., and El-Desouky, B. S. (2023). Alpha-Power of the Power Ailamujia Distribution: Properties and Applications.

[7] Henningsen, A., and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26, 443-458.

[8] Hozaien, H. E., Dayian, G. A., and EL-Helbawy, A. A. (2021). Kumaraswamy Distribution Based on Alpha Power Transformation Methods. *Asian Journal of Probability and Statistics*, 11(1), 14-29.

[9] Ihtisham, S., Khalil, A., Manzoor, S., Khan, S. A., and Ali, A. (2019). Alpha-Power Pareto distribution: Its properties and applications. *PloS one*, 14(6), e0218027.

[10] Ihtisham, S., Manzoor, S., Khalil, A., Badshah, S., Ijaz, M., and Atta, H. (2023). Modeling Extreme Values with Alpha Power Inverse Pareto Distribution. *Measurement Science Review*, 23(2), 55-62.

[11] Kavya, P. and Manoharan, M. (2021). Some parsimonious models for lifetimes and applications. *Journal of Statistical Computation and Simulation*, 91(18), 3693-3708.

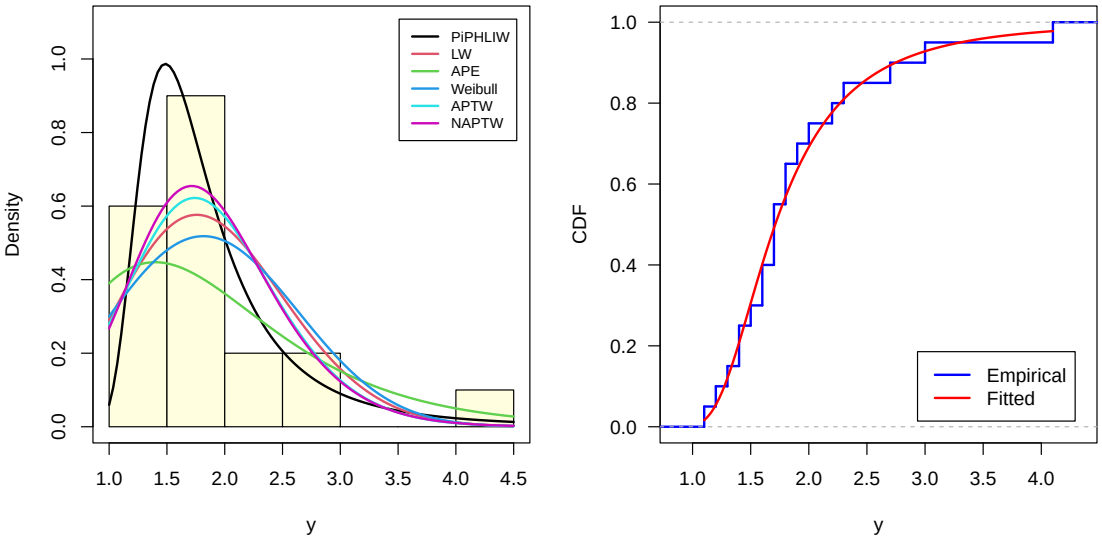


Figure 4: *Fitted PDF (left) and fitted CDF vs empirical CDF (right) (dataset-I)*

Table 7: *Estimated parameters using MLE method for the data set-II*

Model	**parameter**	**SE**	**parameter**	**SE**	**parameter**	**SE**
$\boldsymbol{\pi}$-PHLIW$(\boldsymbol{\beta}, \boldsymbol{\delta})$	2.7544	0.2511	0.7849	0.0530	–	–
LW$(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$	90.4554	4.5499	0.5342	0.0312	0.0498	0.0117
APE$(\boldsymbol{\alpha}, \boldsymbol{\lambda})$	0.0544	0.0426	0.0322	0.0075	–	–
Weibull$(\boldsymbol{\alpha}, \boldsymbol{\beta})$	0.1488	0.0267	0.7463	0.0489	–	–
APTW$(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda})$	1.0000	0.6757	0.7404	0.0303	0.1501	0.0444
NAPTW$(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$	307.9351	272.1076	0.3584	0.0184	1.1521	0.0184

[12] Klakattawi, H. S., and Aljuhani, W. H. (2021). A new technique for generating distributions based on a combination of two techniques: alpha power transformation and exponentiated TX distributions family. *Symmetry*, 13(3), 412.

[13] Kyurkchiev, N. (2017). A new transmuted cumulative distribution function based on the Verhulst logistic function with application in population dynamics. *Biomath Communications*, 4(1).

[14] Lone, M. A., and Jan, T. R. (2023). A New Pi-Exponentiated Method for Constructing Distributions with an Application to Weibull Distribution. *Reliability: Theory & Applications*, 18(1 (72)), 94-109.

[15] Mahdavi, A., and Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13), 6543-6557.

[16] Mandouh, R. M., Sewilam, E. M., Abdel-Zaher, M. M., and Mahmoud, M. R. (2022). A New Family of Distributions in the Class of the Alpha Power Transformation with Applications to Income. *Asian Journal of Probability and Statistics*, 19(1), 41-55.

[17] Maruthan, S. K., and Venkatachalam, N. (2022). Alpha Power Transformation of Lomax Distribution: Properties and Applications. *Thailand Statistician*, 20(3), 669-685.

[18] Mead, M. E., Cordeiro, G. M., Afify, A. Z., and Al Mofleh, H. (2019). The alpha power transformation family: properties and applications. *Pakistan Journal of Statistics and Operation Research*, 525-545.

[19] Musleh, R. M., and Helu, A. (2014). Estimation of the inverse Weibull distribution based on progressively censored data: Comparative study. *Reliability Engineering & System Safety*, 131, 216-227.

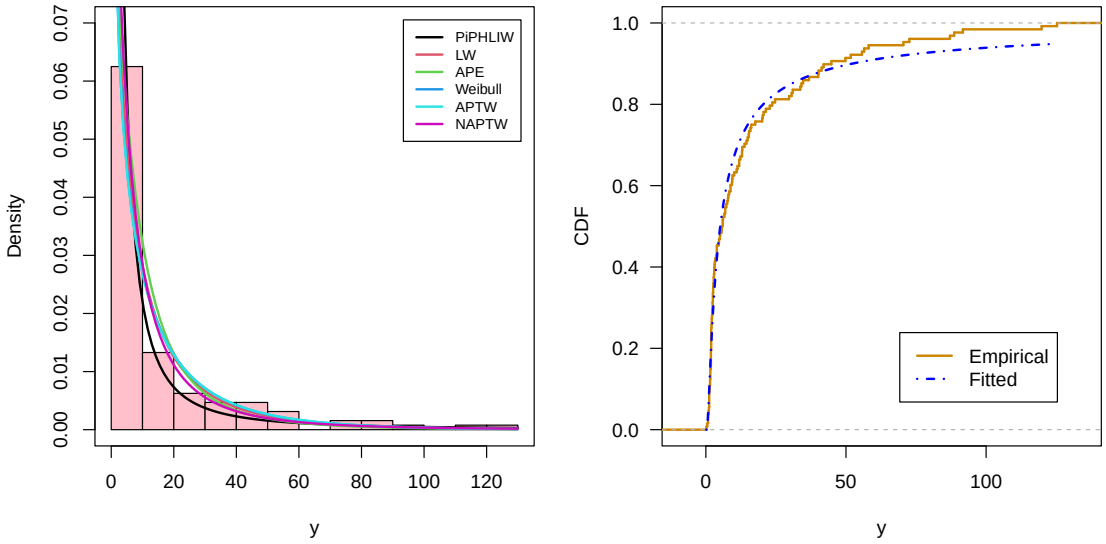


Figure 5: *Fitted PDF (left) and fitted CDF vs empirical CDF (right) (dataset-II)*

Table 8: *Fitted statistics for the data set-II*

Model	-2logL	AIC	HQIC	KS	p(KS)	CVM	p(CVM)	AD	p(AD)
π-PHLIW	921.5701	925.5701	927.8877	0.0599	0.7479	0.1416	0.4165	1.0183	0.3477
LW	933.4466	939.4466	942.923	0.1219	0.0445	0.4260	0.0616	2.6348	0.0422
APE	935.1745	939.1745	941.4921	0.1518	0.0055	0.6127	0.0207	3.6848	0.0125
Weibull	939.3848	943.3848	945.7024	0.1163	0.0628	0.4712	0.0470	2.9577	0.0288
APTW	939.4115	945.4115	948.8879	0.1160	0.0637	0.4767	0.0455	2.9670	0.0285
NAPTW	926.9602	932.9602	936.4366	0.1096	0.0926	0.2956	0.1390	1.9828	0.0940

[20] Nassar, M., Alzaatreh, A., Abo-Kasem, O., Mead, M., and Mansoor, M. (2018a). A new family of generalized distributions based on alpha power transformation with application to cancer data. *Annals of Data Science*, 5, 421-436.

[21] Nassar, M., Alzaatreh, A., Mead, M., and Abo-Kasem, O. (2017). Alpha power Weibull distribution: Properties and applications. *Communications in Statistics - Theory and Methods*, 46(20):10236-10252.

[22] Nassar, M., Dey, S., and Kumar, D. (2018). Logarithm transformed Lomax distribution with applications. *Calcutta Statistical Association Bulletin*, 70(2), 122-135.

[23] R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[24] Sapkota, L. P., and Kumar, V. (2022). Odd Lomax Generalized Exponential Distribution: Application to Engineering and COVID-19 data. *Pakistan Journal of Statistics and Operation Research*, 18(4), 883-900. https://doi.org/10.18187/pjsor.v18i4.4149

[25] Sapkota, L. P. and Kumar, V. (2023). Half Cauchy Generalized Rayleigh Distribution: Theory and Application. *South East Asian J. of Mathematics and Mathematical Sciences*, 19(1), 335-360.

[26] Wickham, H., and Grolemund, G. (2017). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc.

POWER KOMAL DISTRIBUTION WITH PROPERTIES AND APPLICATION IN RELIABILITY ENGINEERING

Rama Shanker, Mousumi Ray, Hosenur Rahman Prodhani*

•

Department of Statistics, Assam University, Silchar, India
e-mail: shankerrama2009@gmail.com, mousumiray616@gmail.com,
*hosenur72@gmail.com

Abstract

The statistical analysis and modeling of reliability data from engineering is really a challenge for statistician because the reliability data from engineering are stochastic in nature. Recently one parameter Komal distribution was introduced in statistics literature for the analysis and modeling of failure time data from engineering. Komal distribution, being one parameter distribution, does not provide good fit to some engineering data due to its theoretical or applied point of view. In this article we propose a two-parameter power Komal distribution, which includes Komal distribution as particular case, for the analysis and modeling of data from reliability engineering. Its statistical properties including behavior of probability density function and cumulative distribution function for varying values of parameters have been presented. The first four raw moments and the variance of the proposed distribution has been derived and given. The expressions for hazard rate function and mean residual life function have been obtained and their behaviors for varying values of parameters have been presented. The stochastic ordering which is very much useful comparing the stochastic nature has also been discussed. Method of maximum likelihood has been discussed for estimating the parameters. Application of the distribution has been investigated using a real lifetime dataset from engineering. The goodness of fit of power Komal distribution has been tested using Akaike Information criterion and Kolmogorov-Smirnov statistic. The goodness of fit of power Komal distribution shows that it gives much closure fit over two-parameter power Garima distribution, Power Shanker distribution and Weibull distribution and one parameter exponential distribution, Shanker distribution, Garima distribution and Komal distribution. As the power Komal distribution gives much better fit over Weibull distribution, which is very much useful for modeling and analysis of data from reliability engineering, the final recommendation is that the power Komal distribution should be preferred over the considered distributions including Weibull for modeling data from reliability engineering.

Keywords: *Komal distribution, Descriptive measures, Reliability properties, Maximum likelihood estimation, Application.*

1. Introduction

Shanker [1] introduced one parameter lifetime distribution named Komal distribution defined by probability density function (pdf) and cumulative distribution function (cdf) as

$$f(y;\theta)=\frac{\theta^2}{\theta^2+\theta+1}(1+\theta+y)e^{-\theta y};\ y>0,\theta>0 \tag{1}$$

$$F(y,\theta)=1-\left[1+\frac{\theta y}{\theta^2+\theta+1}\right]e^{-\theta y};\ y>0,\theta>0 \tag{2}$$

Komal distribution is a mixture of exponential (θ) and gamma $(2,\theta)$ distributions with mixing proportion $\frac{\theta(\theta+1)}{\theta^2+\theta+1}$. We have

$$f(y,\theta)=p\,g_1(y)+(1-p)\,g_2(y)$$

where $p=\frac{\theta+1}{\theta+2}$, $g_1(y)=\theta e^{-\theta y}$, and $g_2(y)=\theta^2 y e^{-\theta y}$.

Shanker [1] has discussed some of its statistical properties including shapes for varying values of parameter, moments, skewness, kurtosis, hazard rate function, mean residual life function and stochastic ordering. Estimation of parameter and applications of Komal distribution has also been discussed by Shanker [1].

The pdf and the cdf of Power Shanker distribution (PSD) obtained by Shanker and Shukla [2] are given by

$$f(x;\theta,\alpha)=\frac{\alpha\theta^2}{\theta^2+1}x^{\alpha-1}(\theta+x^{\alpha})e^{-\theta x^{\alpha}};\ x>0,\theta>0,\alpha>0 \tag{3}$$

$$F(x;\theta,\alpha)=1-\left[1+\frac{\theta x^{\alpha}}{\theta^2+1}\right]e^{-\theta x^{\alpha}};\ x>0,\theta>0,\alpha>0 \tag{4}$$

It can be easily shown that at $\alpha=1$, PSD reduces to one parameter Shanker distribution, introduced by Shanker [3] having pdf and cdf

$$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x};\ x>0,\theta>0 \tag{5}$$

$$F(x;\theta)=1-\left[1+\frac{\theta x}{\theta+1}\right]e^{-\theta x};\ x>0,\theta>0 \tag{6}$$

The pdf and the cdf of Power Garima distribution (PGD) obtained by Berhane et al [4] are given by

$$f(x;\theta,\alpha)=\frac{\alpha\theta}{\theta+2}(1+\theta+\theta x^{\alpha})x^{\alpha-1}e^{-\theta x^{\alpha}},\ x>0,\theta>0,\ \alpha>0 \tag{7}$$

$$F(x;\theta,\alpha)=1-\frac{(\theta x^{\alpha}+\theta+2)}{\theta+2},\ x>0,\theta>0,\alpha>0 \tag{8}$$

It can be easily shown that at $\alpha=1$, PGD reduces to one parameter Garima distribution, introduced by Shanker [5] having pdf and cdf

$$f\left(x;\theta\right)=\frac{\theta}{\theta+2}\left(1+\theta+\theta x\right)e^{-\theta x};x>0,\theta>0 \tag{9}$$

$$F\left(x;\theta\right)=1-\left(1+\frac{\theta x}{\theta+2}\right)e^{-\theta x};x>0,\theta>0 \tag{10}$$

Komal distribution, being one parameter lifetime distribution, has less flexibility to model data of various natures. The main motivation of considering power Komal distribution lies in the fact that as the Komal distribution found to be better than exponential distribution, Shanker distribution of Shanker [3] and Garima distribution of Shanker [5], it is expected that power Komal distribution would provide better fit than power Shanker distribution (PSD) by Shanker and Shukla [2], Power Garima distribution (PGD) by Berhane et al [4] and Weibull distribution of Weibull [6]. In the present paper an attempt has been made to obtain two-parameter power Komal distribution using power transformation of Komal distribution. The statistical properties of the distribution including shapes of density for varying values of parameters, the moments, survival function, hazard rate function and mean residual life function have been discussed. The maximum likelihood estimation has been discussed. The goodness of fit of the proposed distribution has been discussed with a real lifetime dataset from engineering and fit shows quite satisfactory over other one parameter and two-parameter lifetime distributions.

2. Power Komal Distribution

Assuming the power transformation $X=Y^{1/\alpha}$ in (1.1), the pdf of the random variable X can be obtained as

$$\begin{aligned}f\left(x;\theta,\alpha\right)&=\frac{\alpha\theta^{2}}{\theta^{2}+\theta+1}x^{\alpha-1}\left(1+\theta+x^{\alpha}\right)e^{-\theta x^{\alpha}},\ x>0,\theta>0,\ \alpha>0\\&=p\,g_{1}\left(x;\theta,\alpha\right)+\left(1-p\right)g_{2}\left(x;\theta,\alpha,2\right)\end{aligned} \tag{11}$$

where $p=\dfrac{\theta\left(\theta+1\right)}{\theta^{2}+\theta+1}$

$$g_{1}\left(x;\theta,\alpha\right)=\alpha\theta e^{-\theta x^{\alpha}}x^{\alpha-1};x>0,\theta>0,\alpha>0$$

$$g_{2}\left(x;\theta,\alpha,2\right)=\alpha\theta^{2}e^{-\theta x^{\alpha}}x^{2\alpha-1};x>0,\theta>0,\alpha>0\,.$$

We would call the density in (2.1) as power Komal distribution (PKD) with parameters θ and α, and it is denoted by PKD$\left(\theta,\alpha\right)$. Like Komal distribution, PKD is also a convex combination of Weibull$\left(\theta,\alpha\right)$ distribution, a generalized gamma$\left(2,\alpha,\theta\right)$ distribution.

The corresponding cdf of PKD can be obtained as

$$F\left(x;\theta,\alpha\right)=1-\left[1+\frac{\theta x^{\alpha}}{\theta^{2}+\theta+1}\right]e^{-\theta x^{\alpha}},\ \ x>0,\theta>0\,,\alpha>0 \tag{12}$$

The nature of the pdf and the cdf of PKD has been studied with the help of the graphs for varying values of parameters and presented in figures 1 and 2 respectively.

On careful examination of the graphs of pdf of PKD it is obvious that the shapes of PKD are decreasing, symmetric, negatively skewed, positively skewed, platykurtic and leptokurtic for varying values of parameters.

For example, for fixed $\alpha = 1$ and $\theta \geq 1$, pdf of PKD is decreasing ,for fixed value of α i.e. $\alpha = 0.01$ and $\alpha = 0.05$ at $\theta \geq 2$ pdf of PKD becomes positively Skewed and symmetric. For $\alpha \geq 1$ and $\theta \geq 1$ the pdf graph becomes platykurtic and leptokurtic. This means that PKD is applicable for modeling lifetime data of various natures.

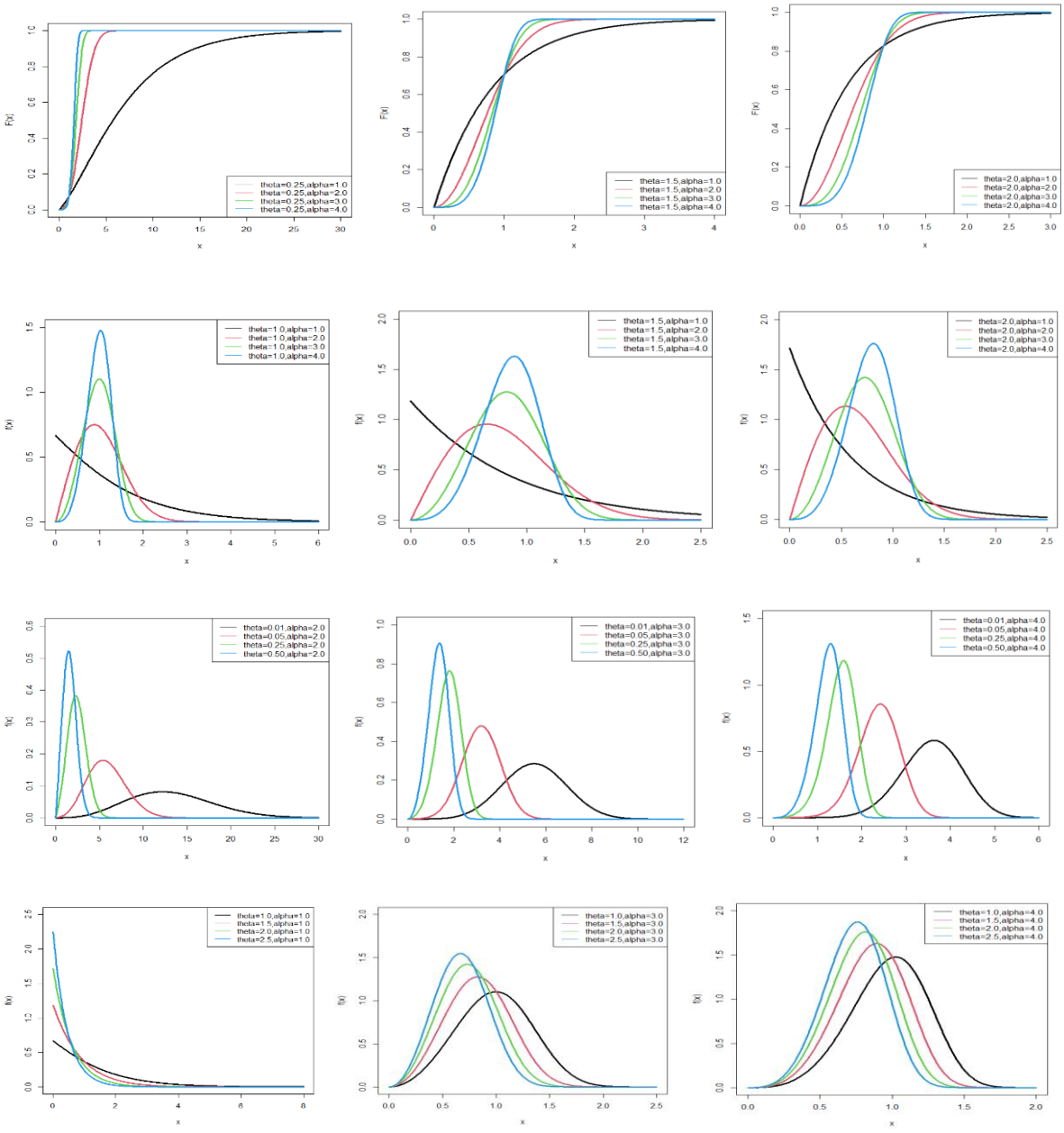


Figure 1: *Graphs of pdf of PKD for varying values of parameters* (θ, α)

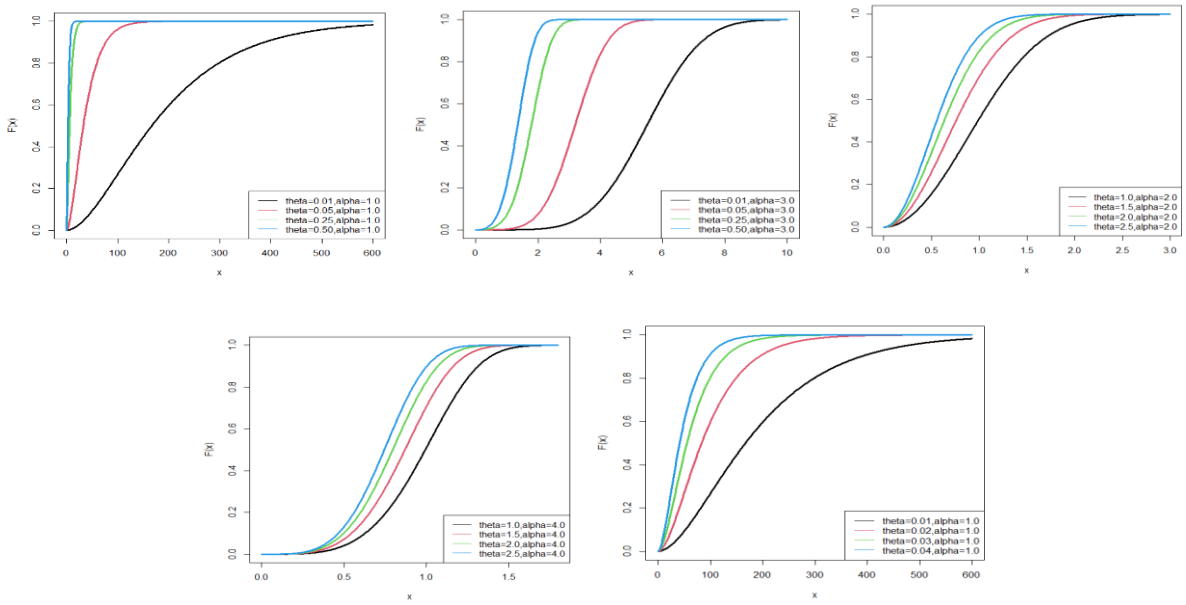


Figure 2: *Graphs of the cdf of PKD for varying values of the parameters* (θ, α)

3. Moments of Power Komal Distribution

Using the mixture representation in (2.2), the r th moment about the origin, μ_r' of PKD can be obtained as

$$\mu_r' = E\left(X^r\right) = p\int_0^\infty x^r g_1\left(x;\theta,\alpha\right)dx + \left(1-p\right)\int_0^\infty x^r g_2\left(x;\theta,\alpha,2\right)dx$$

$$= \frac{r\left(\alpha\theta^2+\alpha\theta+\alpha+r\right)\Gamma\left(\frac{r}{\alpha}\right)}{\alpha^2\left(\theta^2+\theta+1\right)\theta^{\frac{r}{\alpha}}}; r = 1,2,3,\ldots$$

It should be noted that at $\alpha = 1$, the above expression reduces to the r th moment about origin of Komal distribution given by

$$\mu_r' = \frac{r!\left(\theta^2+\theta+r+1\right)}{\theta^r\left(\theta^2+\theta+1\right)}; r = 1,2,3,\ldots$$

Thus, the first four moments about origin of the PKD are thus obtained as

$$\mu_1' = \frac{\left(\alpha\theta^2+\alpha\theta+\alpha+1\right)\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2\left(\theta^2+\theta+1\right)\theta^{\frac{1}{\alpha}}}, \qquad \mu_2' = \frac{2\left(\alpha\theta^2+\alpha\theta+\alpha+2\right)\Gamma\left(\frac{2}{\alpha}\right)}{\alpha^2\left(\theta^2+\theta+1\right)\theta^{\frac{2}{\alpha}}}$$

$$\mu_3' = \frac{3\left(\alpha\theta^2+\alpha\theta+\alpha+3\right)\Gamma\left(\frac{3}{\alpha}\right)}{\alpha^2\left(\theta^2+\theta+1\right)\theta^{\frac{3}{\alpha}}}, \qquad \mu_4' = \frac{4\left(\alpha\theta^2+\alpha\theta+\alpha+4\right)\Gamma\left(\frac{4}{\alpha}\right)}{\alpha^2\left(\theta^2+\theta+1\right)\theta^{\frac{4}{\alpha}}}$$

Therefore, the variance of PKD can be obtained as

$$\mu_2 = \mu_2' - \left(\mu_1'\right)^2 = \frac{2\left(\alpha\theta^2+\alpha\theta+\alpha+2\right)\left(\theta^2+\theta+1\right)\Gamma\left(\frac{2}{\alpha}\right) - \left(\alpha\theta^2+\alpha\theta+\alpha+1\right)^2\left(\Gamma\left(\frac{1}{\alpha}\right)\right)^2}{\alpha^2\left(\theta^2+\theta+1\right)^2\theta^{\frac{2}{\alpha}}}$$

The higher order central moments are not being given here because their expressions are big. However, higher order central moments, if required, can be easily obtained using relationship between moments about mean and moments about origin. Finally, coefficient of variation, skewness, kurtosis and index of dispersion, if needed, can be obtained using their formulae in terms of central moments.

4. Reliability Properties of Power Komal Distribution

The survival function of PKD can be obtained as

$$S\left(x;\theta,\alpha\right) = 1 - F\left(x;\theta,\alpha\right) = \left(1+\frac{\theta x^{\alpha}}{\theta^2+\theta+1}\right)e^{-\theta x^{\alpha}};\ x>0,\theta>0,\alpha>0$$

The hazard rate function $h(x;\theta,\alpha)$ and the mean residual function $m(x;\theta,\alpha)$ of PKD are given respectively as:

$$h\left(x;\theta,\alpha\right) = \frac{f\left(x;\theta,\alpha\right)}{S\left(x;\theta,\alpha\right)} = \frac{\alpha\theta^2\left(1+\theta+x^{\alpha}\right)x^{\alpha-1}}{\theta^2+\theta+1+\theta x^{\alpha}}.$$

$$m\left(x;\theta,\alpha\right) = E\left(X-x \mid X \geq x\right) = \frac{1}{S\left(x;\theta,\alpha\right)}\int_x^{\infty} t\, f\left(t;\theta,\alpha\right)dt - x$$

$$= \frac{\left(\theta^2+\theta\right)\Gamma\left(\frac{1}{\alpha}+1,\theta x^{\alpha}\right)+\Gamma\left(\frac{1}{\alpha}+2,\theta x^{\alpha}\right)}{\theta^{\frac{1}{\alpha}}\left(\theta^2+\theta+1+\theta x^{\alpha}\right)e^{-\theta x^{\alpha}}} - x$$

It can be easily verified that at $x=0$ it reduces to the expression for the mean of PKD. The nature of hazard rate function of PKD for varying values of parameters has been shown graphically in figure 3. Depending upon the values of the parameter the shapes of hazard rate function of PKD is increasing very quickly and slowly in nature, for example, for $\alpha \geq 1$ and $\theta < 1$, the hazard rate is increasing very quickly and for $\alpha \geq 2$ and $\theta > 1$ it is increasing slowly. Again, the nature of mean residual life function of PKD for varying values of parameters has been shown in figure 4. The graphs of mean residual life function of PKD are monotonically decreasing for varying values of parameters.

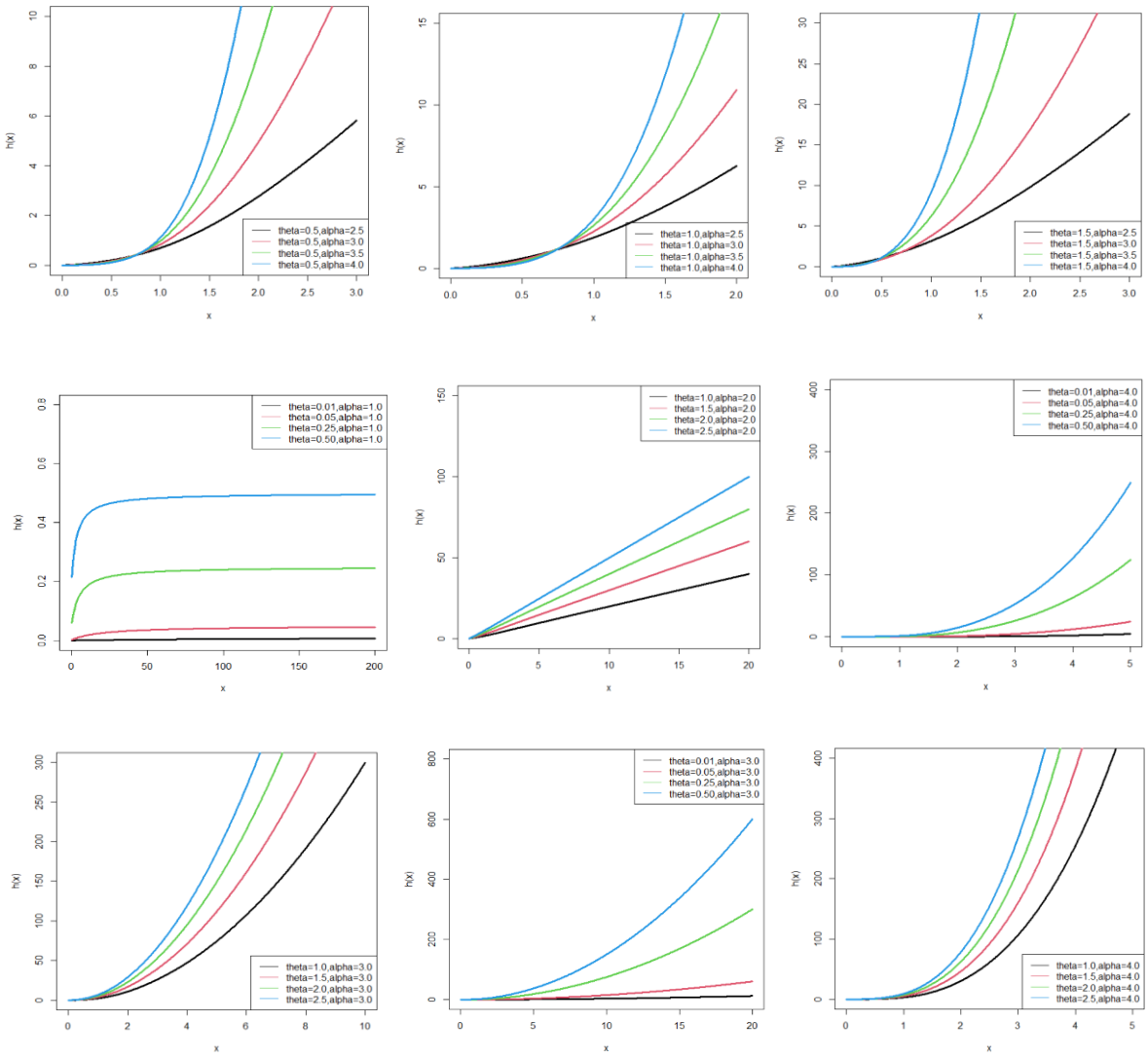


Figure 3: *Hazard rate function of PKD for varying values of the parameters* (α, θ)

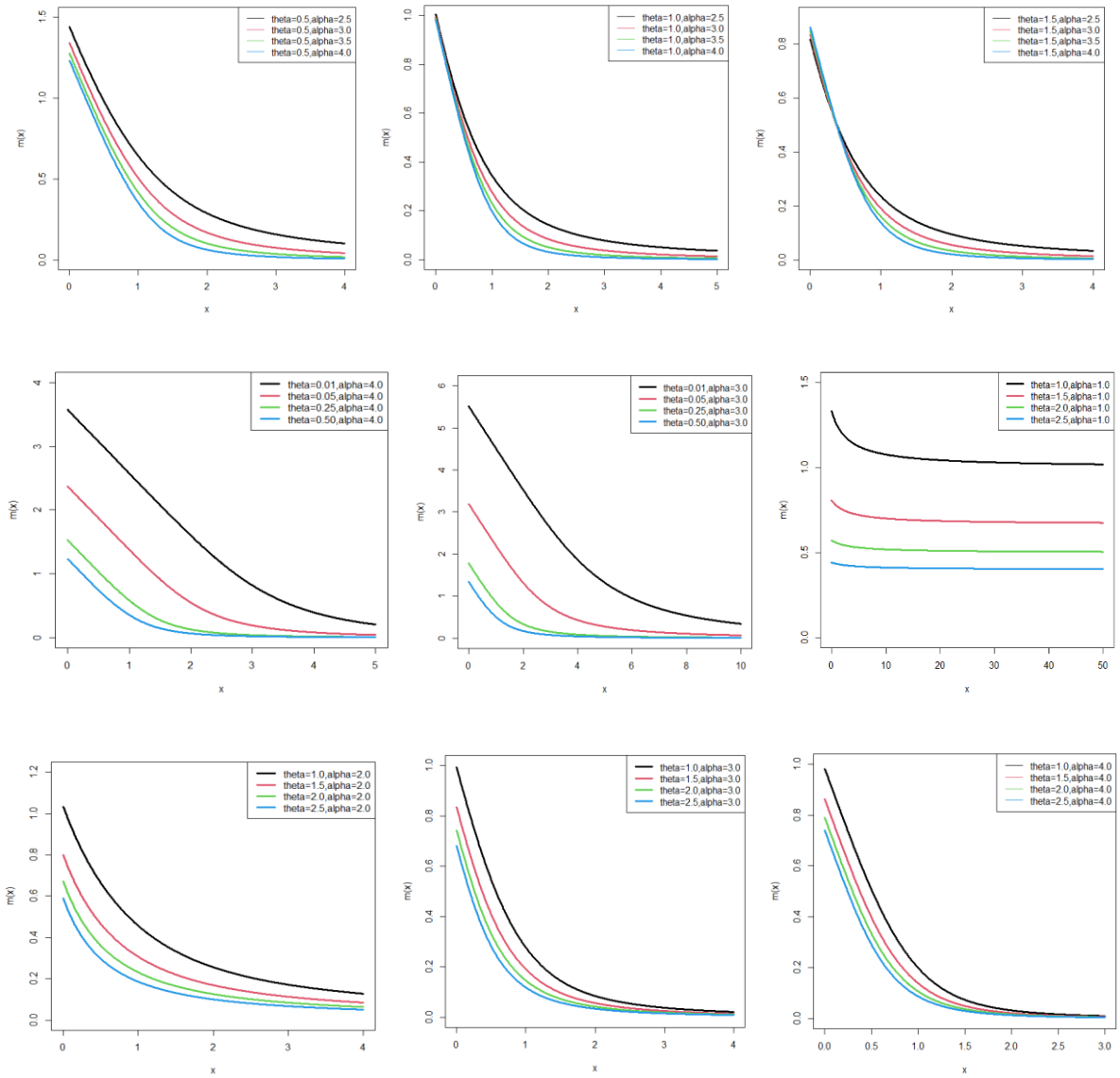


Figure 4: *Mean residual life function of PKD for varying values of the parameters* (α, θ)

5. Stochastic Ordering

The stochastic ordering of positive continuous random variables is an important tool for judging their comparative behavior. A random variable X is said to be smaller than a random variable Y in the

(i) stochastic order $\left(X \leq_{st} Y\right)$ if $F_X(x) \geq F_Y(x)$ for all x

(ii) hazard rate order $\left(X \leq_{hr} Y\right)$ if $h_X(x) \geq h_Y(x)$ for all x

(iii) mean residual life order $\left(X \leq_{mrl} Y\right)$ if $m_X(x) \leq m_Y(x)$ for all x

(iv) likelihood ratio order $\left(X \leq_{lr} Y\right)$ if $\dfrac{f_X(x)}{f_Y(x)}$ decreases in x.

The following interrelationships due to Shaked and Shanthikumar [7] are well known for establishing stochastic ordering of distributions

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y$$
$$\Downarrow$$
$$X \leq_{st} Y$$

It can be easily shown that PKD is ordered with respect to the strongest 'likelihood ratio' ordering. The stochastic ordering of PKD has been explained in the following theorem:

Theorem: Suppose $X \sim \text{PKD}\left(\theta_1, \alpha_1\right)$ and $Y \sim \text{PKD}\left(\theta_2, \alpha_2\right)$. If $\alpha_1 \leq \alpha_2$ and $\theta_1 > \theta_2$ (or $\alpha_1 < \alpha_2$ and $\theta_1 \geq \theta_2$), then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Proof: We have

$$\frac{f_X\left(x;\theta_1,\alpha_1\right)}{f_Y\left(x;\theta_2,\alpha_2\right)} = \left[\frac{\alpha_1\theta_1^{\,2}\left(\theta_2^{\,2}+\theta_2+1\right)}{\alpha_2\theta_2^{\,2}\left(\theta_1^{\,2}+\theta_1+1\right)}\right]\left(\frac{1+\theta_1+x^{\alpha_1}}{1+\theta_2+x^{\alpha_2}}\right)x^{\alpha_1-\alpha_2}\, e^{-\left(\theta_1 x^{x^{\alpha_1}}-\theta_2 x^{x^{\alpha_2}}\right)};\, x>0$$

Now, taking logarithm both sides, we get

$$\log\frac{f_X\left(x;\theta_1,\alpha_1\right)}{f_Y\left(x;\theta_2,\alpha_2\right)} = \log\left[\frac{\alpha_1\theta_1^{\,2}\left(\theta_2^{\,2}+\theta_2+1\right)}{\alpha_2\theta_2^{\,2}\left(\theta_1^{\,2}+\theta_1+1\right)}\right]+\log\left(\frac{1+\theta_1+x^{\alpha_1}}{1+\theta_2+x^{\alpha_2}}\right)+\left(\alpha_1-\alpha_2\right)\log x-\left(\theta_1 x^{x^{\alpha_1}}-\theta_2 x^{x^{\alpha_2}}\right)$$

This gives

$$\frac{d}{dx}\ln\frac{f_X\left(x;\theta_1,\alpha_1\right)}{f_Y\left(x;\theta_2,\alpha_2\right)} = \frac{\alpha_1-\alpha_2}{x}+\frac{\left(\alpha_1 x^{\alpha_1-1}-\alpha_2 x^{\alpha_2-1}\right)+\left(\alpha_1\theta_2 x^{\alpha_1-1}-\alpha_2\theta_1 x^{\alpha_2-1}\right)+\left(\alpha_1-\alpha_2\right)x^{\alpha_1+\alpha_2-1}}{\left(1+\theta_1+x^{\alpha_1}\right)\left(1+\theta_2+x^{\alpha_2}\right)}$$
$$-\left(\theta_1\alpha_1 x^{\alpha_1-1}-\theta_2\alpha_2 x^{\alpha_2-1}\right)$$

.

Thus for $\alpha_1 \leq \alpha_2$ and $\theta_1 > \theta_2$ (or $\alpha_1 < \alpha_2$ and $\theta_1 \geq \theta_2$), $\dfrac{d}{dx}\ln\dfrac{f_X\left(x;\theta_1,\alpha_1\right)}{f_Y\left(x;\theta_2,\alpha_2\right)} < 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

6. Estimation of Parameters of Power Komal Distribution

Suppose $(x_1, x_2, ..., x_n)$ be a random sample of size n from PKD(θ, α) The log-likelihood function of PKD can be expressed as

$$\log L = \sum_{i=1}^{n} \log f(x;\theta,\alpha)$$

$$= 2n\log\theta - n\log\left(\theta^2+\theta+1\right) + (\alpha-1)\sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log\left(1+\theta+x_i^{\alpha}\right) - \theta\sum_{i=1}^{n}x_i^{\alpha}$$

The maximum likelihood estimates $(\hat{\theta}, \hat{\alpha})$ of parameters (θ, α) of PKD are the solution of the following log-likelihood equations

$$\frac{\partial \log L}{\partial \theta} = \frac{2n}{\theta} - \frac{n(2\theta+1)}{\theta^2+\theta+1} + \sum_{i=1}^{n}\frac{1}{1+\theta+x_i^{\alpha}} - \sum_{i=1}^{n}x_i^{\alpha} = 0$$

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\frac{x_i^{\alpha}\log x_i}{1+\theta+x_i^{\alpha}} - \theta\sum_{i=1}^{n}x_i^{\alpha}\log(x_i)$$

These two natural log-likelihood equations do not seem to be solved directly, because they cannot be expressed in closed forms. The (MLE's) $(\hat{\theta}, \hat{\alpha})$ of (θ, α) can be computed directly by solving the natural log-likelihood equation using Newton-Raphson iteration available in R-software till sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

For Fisher's scoring method, we have

$$\frac{\partial^2 \log L}{\partial \theta^2} = \frac{-2n}{\theta^2} - \frac{n(2\theta^2+2\theta-1)}{\left(\theta^2+\theta+1\right)^2} - \sum_{i=1}^{n}\frac{1}{\left(1+\theta+x_i^{\alpha}\right)^2}$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = -\sum_{i=1}^{n}\frac{x_i^{\alpha}\log x_i}{\left(1+\theta+x_i^{\alpha}\right)^2} - \sum_{i=1}^{n}x_i^{\alpha}\log x_i = \frac{\partial^2 \log L}{\partial \alpha \partial \theta}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \sum_{i=1}^{n}\frac{(\log x_i)^2\left\{\left(1+\theta+x_i^{\alpha}\right)x_i^{\alpha} - x_i^{2\alpha}\right\}}{\left(1+\theta+x_i^{\alpha}\right)^2} - \theta\sum_{i=1}^{n}(\log x_i)^2 x_i^{\alpha}$$

For finding the MLEs $(\hat{\theta}, \hat{\alpha})$ of parameters (θ, α) of PKD, following equations can be solved

$$\begin{bmatrix} \frac{\partial^2 \ln L}{\partial \theta^2} & \frac{\partial^2 \ln L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \ln L}{\partial \theta \partial \alpha} & \frac{\partial^2 \ln L}{\partial \alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta}-\theta_0 \\ \hat{\alpha}-\alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L}{\partial \theta} \\ \frac{\partial \ln L}{\partial \alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}}$$

where θ_0 and α_0 are the initial values of θ and α, as given by the method of moments. These equations are solved iteratively till close estimates of parameters are obtained.

7. Goodness of Fit

In this section, the goodness of fit of PKD using maximum likelihood estimates of parameters has been discussed with one real dataset. The goodness of fit has been compared with other one parameter and two-parameter lifetime distributions. Following dataset from engineering has been considered.

Dataset 1: The following symmetric dataset, discussed by Murthy et al [8], is the failure times of windshields and the values are:
0.04, 0.3, 0.31, 0.557, 0.943, 1.07, 1.124, 1.248, 1.281, 1.281, 1.303, 1.432, 1.48, 1.51, 1.51, 1.568, 1.615, 1.619, 1.652, 1.652, 1.757, 1.795, 1.866, 1.876, 1.899, 1.911, 1.912, 1.9141, 0.981, 2.010, 2.038, 2.085, 2.089, 2.097, 2.135, 2.154, 2.190, 2.194, 2.223, 2.224, 2.23, 2.3, 2.324, 2.349, 2.385, 2.481, 2.610, 2.625, 2.632, 2.646, 2.661, 2.688, 2.823, 2.89, 2.9, 2.934, 2.962, 2.964, 3, 3.1, 3.114, 3.117, 3.166, 3.344, 3.376, 3.385, 3.443, 3.467, 3.478, 3.578, 3.595, 3.699, 3.779, 3.924, 4.035, 4.121, 4.167, 4.240, 4.255, 4.278, 4.305, 4.376, 4.449, 4.485, 4.570, 4.602, 4.663, 4.694

In order to compare the considered distributions, values of $\text{MLE}\left(\hat{\theta},\hat{\alpha}\right)$ along with their standard errors, $-2\log L$, AIC (Akaike Information Criterion), K-S (Kolmogorov-Smirnov Statistics) and p-values for the real lifetime dataset have been computed and presented in table1. The AIC and K-S Statistics are computed using the following formulae: $AIC = -2\log L + 2k$ and $\text{K-S} = \sup_x \left|F_n(x) - F_0(x)\right|$, where $k =$ the number of parameters, $n =$ the sample size, $F_n(x)$ is the empirical (sample) cumulative distribution function, and $F_0(x)$ is the theoretical cumulative distribution function. The best distribution is the distribution corresponding to lower values of $-2\log L$, AIC, and K-S statistic.

Table 1: MLE's, $-2\log L$, $S.E\left(\hat{\theta},\hat{\alpha}\right)$, AIC, K-S and P-value of the fitted distributions of dataset 1.

Distributions	ML estimates and Standard Error		$-2\log L$	AIC	K-S	P-Value
	$\hat{\theta}$ $S.E\left(\hat{\theta}\right)$	$\hat{\alpha}$ $S.E\left(\hat{\alpha}\right)$				
PKD	0.2572 (0.0401)	1.8786 (0.1329)	**271.8212**	**275.8212**	**0.0540**	**0.9826**
PGD	0.1563 (0.0373)	2.1489 (0.1823)	272.7053	276.7053	0.1009	0.4340
PSD	0.3422 (0.0483)	1.7015 (0.1203)	272.7581	276.7581	0.0912	0.5729
WD	0.0829 (0.0223)	2.3563 (0.2031)	274.1806	278.1806	0.2990	0.0000
KD	0.5826 (0.0445)	…	325.0655	327.0655	0.2189	0.0015
GD	0.5572 (0.0506)	…	331.4615	333.4615	0.5131	0.0000
SD	0.6430 (0.0451)	…	314.4891	316.4891	0.2075	0.0031
ED	0.3893 (0.0415)	…	342.0450	344.045	0.5256	0.0000

In order to see the closeness of the fit given by one parameter exponential, Shanker and Garima distributions and two-parameter Weibull distribution, PGD and PSD, the fitted plot of pdfs of these distributions for the given dataset have been shown in figure 5. It is also obvious from the goodness of fit given in the table 1 and the fitted plots of the distributions in figure 5 that the PKD gives much closer fit as compared to other considered distributions.

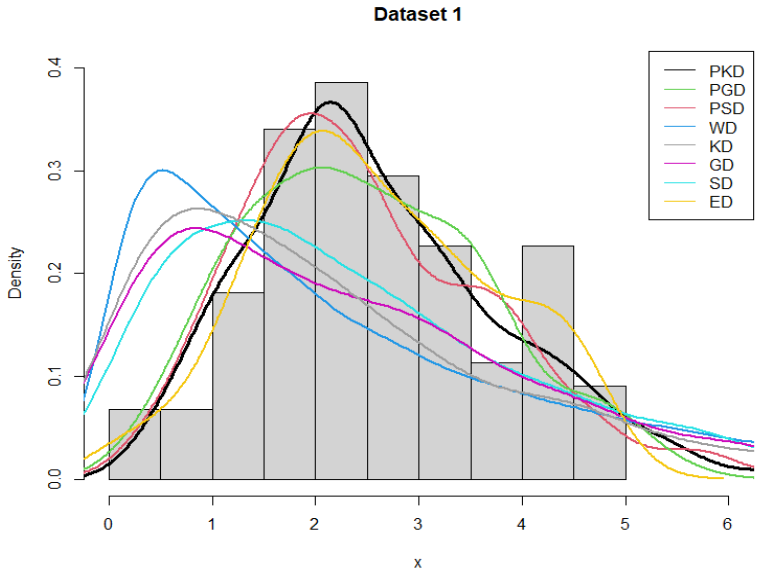


Figure 5: *Fitted pdf plots of distributions for dataset 1*

8. Conclusions

A two-parameter Power Komal distribution (PKD) has been introduced which includes Komal distribution, introduced by Shanker (2023), as a particular case. The statistical and reliability properties including shapes of density for varying values of parameters, the moments about the origin, the variance, survival function, hazard rate function, mean residual function of PKD have been discussed. The method of maximum likelihood for estimating the parameters has been discussed. Finally, the goodness of fit of PKD has been discussed with a real lifetime dataset and the fit has been found quite satisfactory as compared to one parameter exponential Shanker and Garima distributions and two-parameter Power Shanker distribution (PSD), power Garima distribution (PGD) exponential, and Weibull distribution. Therefore, PKD can be considered as an important lifetime distribution for modeling lifetime data from engineering.

Conflict of Interest

The Authors declare that there is no conflict of interest.

References

[1] Shanker, R. (2023). Komal Distribution with properties and Application in Survival Analysis, *Biometrics & Biostatistics International Journal*, 12(2): 40 – 44.

[2] Shanker, R. and Shukla, K.K. (2017). Power Shanker Distribution and Application, *Turliye Kliniklery Journal of Biostatistics* ,9(3):175-87.

[3] Shanker, R. (2015). Shanker distribution and its Applications, *International Journal of Statistics and Applications*, 5(6): 338 – 348.

[4] Berhane, A., Tesfay, M., Tesfalem, E., Shanker, R. (2019). A Two Parameter Power Garima Distribution with Properties and Applications, Annals of Biostatistics and Biometric applications ,1(3):1-7.

[5] Shanker, R., (2016). Garima distribution and Its Application to model behavioral science data, Biometrics and Biostatistics International Journal, 4(7): 1-9.

[6] Weibull, W. (1951). A statistical distribution of wide applicability, Journal of Applied Mathematics, 18: 293 – 297.

[7] Shaked, M. and Shanthikumar, J.G. (1994): Stochastic Orders and Their Applications, Academic Press, New York.

[8] Murthy, D.N.P., Xie, M., Jiang, R. (2004). Weibull Models, John Wiley & Sons Inc., Hoboken.

RT&A, No 4 (76)
Volume 18, December 2023

PROFIT ANALYSIS OF REPAIRABLE COLD STANDBY SYSTEM UNDER REFRESHMENTS

[1]Ajay Kumar and [2]Ashish Sharma

•

[1]SOET, Raffles University, Neemrana, Rajasthan
[2]Department of Pharmacy, Sushant University, Gurugram
ajaykumar.soet@rafflesuniversity.edu.in, ashishlpu10@gmail.com

Abstract

In the generation of science and technology, every company wants to increase the reliability of their products. So, they used the concept of cold standby redundancy, timely repair of the failed unit and providing limited refreshments to the available technician when required. This paper aims to explore the system of two identical units where the primary unit is operative and the secondary unit is in cold standby mode. When the primary unit fails due to any fault then secondary unit starts working immediately. Here, times of failure of unit and technician refreshment request follow the general distribution whereas times of repair of unit and refreshment follow the exponential distributions. Such types of systems are used in industries and education systems to prevent losses. The system's performance is calculated by using concepts of mean time to system failure, availability, busy period of the server, expected number of visits made by the server and profit function using the semi Markov process and regenerative point technique. Tables are used to explore the performance of the system.

Keywords: Cold standby, refreshment, regenerative point, semi Markov process

I. Introduction

Reliability and maintainability are the essential parameters of items and products that satisfy customers' requirements. In today's era, several approaches for performance improvement of maintainable systems have been adopted by scientists and engineers during designing them. A large amount of research work has been done on repairable systems such that Subramanian [17] explored the idea of preventive maintenance in two distinct units system under repair. Bao and Mays [3] analyzed the hydraulic reliability of water distribution systems under demand, pipe roughness and pressure head. Gnedenko and Igor [8] explored reliability and probability studies for engineering purposes. Diaz et al. [6] threw light on the warranty cost management system. Jack and Murthy [10] discovered the role of limited warranty and extended warranty for the product. Wang and Zhang [19] examined the repairable system of two non identical components under repair facility using geometric distributions. Mahmoud and Moshref [14] described the cold standby system under hardware failure and preventive maintenance using the semi Markov process. Deswal and Malik [5] evaluate the non identical units system under different working conditions by using the semi Markov process. Kumar et al. [11] examined the stochastic behavior of two unit system where one unit in cold standby mode and subject to maximum repair time using the regenerative point technique. Kumar and Goel [12] analyzed the preventive maintenance in two unit cold standby system under general distributions. Malik and Rathee [15] threw light on the two parallel units system under preventive maintenance and maximum operation time.

Temraz [18] analyzed the performance of two parallel components system under load sharing and degradation facility. Levitin et al. [13] explored the results of optimal preventive replacement of failed units in a cold standby system by using the poisson process. Barak et al. [2] threw light on the availability and profit values of milk plant under repair facility. Agarwal et al. [1] described the reliability and availability of water reservoir system under repair facility. Chaudhary and Sharma [4] explored the parallel non identical units system that gives priority to repair over preventive maintenance. Garg and Garg [7] analyzed the reliability and profit values of briquette machine under neglected faults like sound and overheating. Jia et al. [9] explored the two unit system under demand and energy storage techniques. Sengar and Mangey [16] examined the performance of complicated systems under inspection using copula methodology.

II. System Assumptions

There are following system assumptions:

- Initially, the system has two units such that one is an operative (primary) unit and the other is a cold standby (secondary) unit.
- When the operative unit fails then the cold standby unit starts working.
- An expert repairman is always available to repair the failed unit.
- The failed unit behaves like a new one after repair.
- Repair and refreshment times are exponentially distributed whereas times for failure of unit and server refreshment request are general.

III. System Notations

There are following system notations:

R	Collection of regenerative states S_i $(i = 0,1,2,3)$
O/Cs	Operative unit / cold standby unit of the system
a/b	The probability that the cold standby unit is working/ not working
λ/μ	Failure rate of the unit/ rate by which the server needs refreshment
$g(t)/G(t)$	PDF/ CDF of the repair time of the unit
$f(t)/F(t)$	PDF/ CDF of refreshments time that restores freshness to the server
$q_{r,s}(t)/Q_{r,s}(t)$	PDF/ CDF of first passage time from rth to sth regenerative state or sth failed state without halting in any other $S_i \in R$ in (0,1]
$M_r(t)$	Represents the probability of the system that it initially works $S_r \in R$ at a time (t) without moving through another state $S_i \in R$
$W_r(t)$	Probability that up to time (t) the server is busy at the state S_r without transit to another state $S_i \in R$ or before return to the same state through one or more non regenerative states
$\oplus/\otimes$	Laplace convolution / Laplace Stieltjes Convolution
$*/**/'$	Symbol for Laplace Transform/ Laplace Stieltjes Transform/ Function's derivative
○ / ● / ▭	Upstate/ regenerative state/ failed state

IV. State Descriptions

The individual state description is given by the table 1:

Table 1: *State Descriptions*

States	Descriptions
S_0	It is a regenerative upstate with two units such that one is operative (O) and other is cold standby (Cs).
S_1	This regenerative upstate has two units such that one is failed under repair (*Fur*) and the other is in operative mode (O).
S_2	It is a regenerative upstate under refreshment facility (*sut*) where one unit is failed & waiting for repair (*Fwr*) and the other is in operative mode (O).
S_3	It is a regenerative down state and the system has two units such that one is failed under repair (*Fur*) and the other is failed and waiting for repair (*Fwr*).
S_4	It is a down state where one unit fails under repair (*FUR*) continuously from the prior state and the other unit is failed & waiting for repair (*Fwr*).
S_5	It is a down state that has two units under refreshment facility (*sut*) such that one is failed and waiting for repair (*FWR*) continuously from the previous state and the other is failed and waiting for repair (*Fwr*).
S_6	At this down state, the system has two units such that one is failed under repair (*FUR*) continuously from the previous state and the other unit is failed and waiting for repair (*FWR*) continually from the prior state.
S_7	This down state has two units under continuous refreshment facility (*SUT*) such that one is failed and waiting for repair (*Fwr*) and the other is failed and waiting for repair (*FWR*) continuously from the previous state.
S_8	This down state has two units under refreshment facility (*sut*) such that one is failed and waiting for repair (*Fwr*) and the other is failed and waiting for repair (*FWR*) continuously from the previous state.

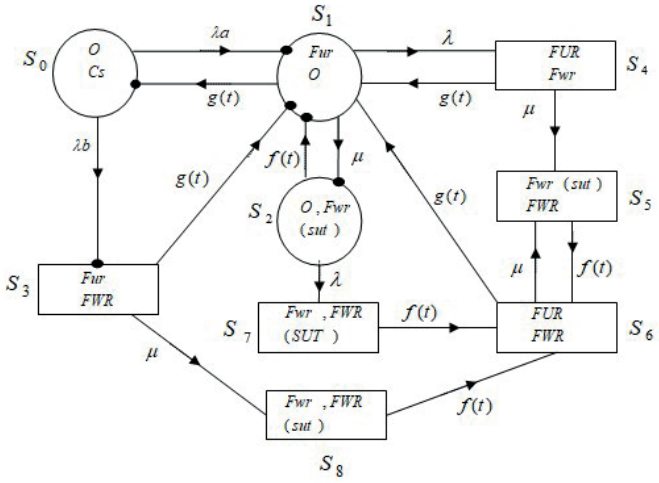


Figure 1: *State Transition Diagram*

V. Transition Probabilities

The transition probabilities are calculated using $f(t) = \theta e^{-\theta t}$, $g(t) = \phi e^{-\phi t}$ and get

$$p_{01} = a,\ p_{03} = b,\ p_{10} = \frac{\phi}{\phi+\mu+\lambda},\ p_{12} = \frac{\mu}{\phi+\mu+\lambda},\ p_{14} = \frac{\lambda}{\phi+\mu+\lambda}$$

$$p_{21} = \frac{\theta}{\theta+\lambda},\ p_{27} = \frac{\lambda}{\theta+\lambda},\ p_{31} = \frac{\phi}{\phi+\mu},\ p_{38} = \frac{\mu}{\phi+\mu},\ p_{41} = \frac{\phi}{\phi+\mu},\ p_{45} = \frac{\mu}{\phi+\mu}$$

$$p_{61} = \frac{\phi}{\phi+\mu},\ p_{65} = \frac{\mu}{\phi+\mu},\ p_{56} = p_{76} = p_{86} = 1 \tag{1}$$

It has been conclusively established that

$$p_{01} + p_{03} = 1,\ p_{10} + p_{12} + p_{14} = 1,\ p_{21} + p_{27} = 1,\ p_{31} + p_{38} = 1$$

$$p_{41} + p_{45} = 1,\ p_{56} = p_{76} = p_{86} = 1,\ p_{31} + p_{31.8(65)^n} = 1$$

$$p_{10} + p_{12} + p_{11.4} + p_{11.4(56)^n} = 1,\ p_{21} + p_{21.7(65)^n} = 1 \tag{2}$$

VI. Mean Sojourn Time

In the cold standby redundant system, μ_i represents the mean sojourn time. Mathematically, time consumed by a system in a particular state is, $\mu_i = \sum_j m_{i,j} = \int_0^\infty P(T > t)\,dt$. Then

$$\mu_0 = m_{01} + m_{03} = \frac{1}{\lambda},\ \mu_1 = m_{10} + m_{12} + m_{14} = \frac{1}{\phi+\mu+\lambda},\ \mu_2 = m_{21} + m_{27} = \frac{1}{\theta+\lambda}$$

$$\mu_3 = m_{31} + m_{38} = \frac{1}{\phi+\mu},\ \mu_4 = m_{41} + m_{45} = \frac{1}{\phi+\mu},\ \mu_6 = m_{61} + m_{65} = \frac{1}{\phi+\mu}$$

$$\mu_5 = \mu_7 = \mu_8 = \frac{1}{\theta},\ \mu'_1 = m_{10} + m_{12} + m_{11.4} + m_{11.4(56)^n} = \frac{\theta\phi + \lambda(\theta+\mu)}{\theta\phi(\phi+\mu+\lambda)}$$

$$\mu'_2 = m_{21} + m_{21.7(65)^n} = \frac{\theta\phi + \lambda(\theta+\phi+\mu)}{\theta\phi(\theta+\lambda)},\ \mu'_3 = m_{31} + m_{31.8(65)^n} = \frac{(\theta+\mu)}{\theta\phi} \tag{3}$$

VII. Reliability Measures Evaluations

I. Mean Time to System Failure (MTSF)

Let the cumulative distribution function of the first elapsed time be $\varphi_i(t)$ from the regenerative state S_i to the failed state of the system. Treating the failed states as an absorbing state then the repetitive interface for $\varphi_i(t)$ being

$$\varphi_0(t) = Q_{0,3}(t) + Q_{0,1}(t) \otimes \varphi_1(t)$$

$$\varphi_1(t) = Q_{1,4}(t) + Q_{1,2}(t) \otimes \varphi_2(t) + Q_{1,0}(t) \otimes \varphi_0(t)$$

$$\varphi_2(t) = Q_{2,7}(t) + Q_{2,1}(t) \otimes \varphi_1(t) \tag{4}$$

Taking LST of the relation (4) and solving for $\varphi_0^{**}(s)$ then

$$MTSF = \lim_{s\to 0} \frac{1 - \varphi_0^{**}(s)}{s} \tag{5}$$

System reliability can be obtained by using the inverse LT of equation (5). We have

$$MTSF = \frac{[1-p_{12}p_{21}]\mu_0 + p_{01}\mu_1 + p_{01}p_{12}\mu_2}{1-p_{01}p_{10}-p_{12}p_{21}} \tag{6}$$

$$MTSF = \frac{\theta\phi+\lambda\theta+\phi\lambda+\mu\lambda+\lambda^2+a\lambda\theta+a\lambda^2+a\mu\lambda}{\theta\phi+\lambda\theta+\phi\lambda+\mu\lambda+\lambda^2-a\theta\phi-a\lambda\phi} \tag{7}$$

II. Availability of the system

From the transition diagram, the system is available at the regenerative up states S_0, S_1 and S_2. Let $A_i(t)$ is the probability that the system is in upstate at time (t) specified that the system arrives at the regenerative state S_i at t = 0. Then the repetitive interface for $A_i(t)$ is

$$\begin{aligned}
A_0(t) &= M_0(t) + q_{01}(t) \oplus A_1(t) + q_{03}(t) \oplus A_3(t) \\
A_1(t) &= M_1(t) + q_{10}(t) \oplus A_0(t) + q_{12}(t) \oplus A_2(t) \\
&\quad + q_{11.4}(t) \oplus A_1(t) + q_{11.4(56)^n}(t) \oplus A_1(t) \\
A_2(t) &= M_2(t) + q_{21}(t) \oplus A_1(t) + q_{21.7(65)^n}(t) \oplus A_1(t) \\
A_3(t) &= q_{31}(t) \oplus A_1(t) + q_{31.8(65)^n}(t) \oplus A_1(t)
\end{aligned} \tag{8}$$

where, $M_0(t) = e^{-\lambda t}$, $M_1(t) = e^{-(\mu+\lambda_2)t}\overline{G(t)}$, $M_2(t) = e^{-\lambda t}\overline{G(t)}$ (9)

Using LT of the above relation (8), there exist

$$\therefore A_0 = \lim_{s\to 0}\frac{N_A}{D_1'} = \frac{[\mu_0 p_{1,0} + \mu_1 + \mu_2 p_{1,2}]}{[\mu_0 p_{1,0} + \mu_1' + \mu_2' p_{1,2} + \mu_3' p_{0,3} p_{1,0}]} \tag{10}$$

$$\therefore A_0 = \frac{\theta\phi[\phi(\theta+\lambda)+(a-b)(\theta+\lambda)\lambda+\lambda\mu]}{\left[\begin{array}{l}\theta\phi^2(\theta+\lambda)+(\theta\phi+\lambda(\theta+\mu))(a-b)\lambda(\theta+\lambda) \\ +(\theta\phi+\lambda(\theta+\phi+\mu))\mu(a-b)\lambda-(\theta+\mu)b\lambda\phi(\theta+\lambda)\end{array}\right]} \tag{11}$$

III. Busy Period of the Server

From the transition diagram, it is clear that the technician is busy at states S_1, S_2 and S_3. Let $B_i(t)$ is the probability that the repairman is busy due to the repair of the failed unit at time 't' specified that the system arrives at the regenerative state S_i at t = 0. Then the repetitive interface for $B_i(t)$ is

$$\begin{aligned}
B_0(t) &= q_{01}(t) \oplus B_1(t) + q_{03}(t) \oplus B_3(t) \\
B_1(t) &= W_1(t) + q_{10}(t) \oplus B_0(t) + q_{12}(t) \oplus B_2(t) \\
&\quad + q_{11.4}(t) \oplus B_1(t) + q_{11.4(56)^n}(t) \oplus B_1(t) \\
B_3(t) &= W_2 + q_{21}(t) \oplus B_1(t) + q_{21.7(65)^n}(t) \oplus B_1(t) \\
B_3(t) &= W_3(t) + q_{31}(t) \oplus B_1(t) + q_{31.8(65)^n}(t) \oplus B_1(t)
\end{aligned} \tag{12}$$

where, $W_1(t) = \overline{G_1(t)}e^{-(\lambda+\mu)t} + \overline{G_1(t)}\lambda e^{-(\lambda+\mu)t} \oplus \overline{G_1(t)}e^{-\mu t} + ...$

$$\begin{aligned}
W_2(t) &= \overline{G_1(t)} \oplus \lambda e^{-\lambda t}\overline{F_1(t)} \oplus f_1(t) \\
W_3(t) &= \overline{G_1(t)}e^{-\mu t} + \overline{G_1(t)}\mu e^{-\mu t} \oplus f_1(t) \oplus \overline{G_1(t)}e^{-\mu t} + ...
\end{aligned} \tag{13}$$

Using LT on relations (12) then we get

$$B_0 = \lim_{s\to 0}\frac{N_B}{D_1'} = \frac{W_1^*(0) + W_2^*(0)p_{1,2} + W_3^*(0)p_{0,3}p_{1,0}}{[\mu_0 p_{1,0} + \mu_1' + \mu_2' p_{1,2} + \mu_3' p_{0,3} p_{1,0}]} \tag{14}$$

$$B_0 = \frac{\lambda\mu[\theta\phi + \lambda(\phi+\theta)]}{\left[\begin{array}{l}\theta\phi^2(\theta+\lambda) + (\theta\phi + \lambda(\theta+\mu))(a-b)\lambda(\theta+\lambda) \\ + (\theta\phi + \lambda(\theta+\phi+\mu))\mu(a-b)\lambda - (\theta+\mu)b\lambda\phi(\theta+\lambda)\end{array}\right]} \tag{15}$$

IV. Estimated number of visits made by the server

The transition diagram explores that the technician visits at states S_1 and S_2. Let $N_i(t)$ is the estimated number of visits made by the repairman for repair in (0, t] specified that the system arrives at the regenerative state S_i at t = 0. Then the repetitive interface for $N_i(t)$ is

$$\begin{aligned}
N_0(t) &= Q_{01}(t) \otimes [1 + N_1(t)] + Q_{03}(t) \otimes [1 + N_3(t)] \\
N_1(t) &= Q_{10}(t) \otimes N_0(t) + Q_{12}(t) \otimes N_2(t) \\
&\quad + Q_{11.4}(t) \otimes N_1(t) + Q_{11.4(56)^n}(t) \otimes N_1(t) \\
N_2(t) &= Q_{21}(t) \otimes N_1(t) + Q_{21.7(65)^n}(t) \otimes N_1(t) \\
N_3(t) &= Q_{31}(t) \otimes N_1(t) + Q_{31.8(65)^n}(t) \otimes N_1(t)
\end{aligned} \tag{16}$$

Using LST of the above relations (16) then we get

$$N_0 = \lim_{s\to 0} \frac{N_v}{D_1'} = \frac{p_{1,0}}{[\mu_0 p_{1,0} + \mu_1' + \mu_2' p_{1,2} + \mu_3' p_{0,3} p_{10}]} \tag{17}$$

$$N_0 = \frac{\lambda\theta\phi^2(\theta+\lambda)}{\left[\begin{array}{l}\theta\phi^2(\theta+\lambda) + (\theta\phi + \lambda(\theta+\mu))(a-b)\lambda(\theta+\lambda) \\ + (\theta\phi + \lambda(\theta+\phi+\mu))\mu(a-b)\lambda - (\theta+\mu)b\lambda\phi(\theta+\lambda)\end{array}\right]} \tag{18}$$

V. Profit Analysis

It is an integral part of reliability measures that tell customers and system developers whether the system is beneficial or not. The profit values depend upon the MTSF, availability of the system, busy period of server and extended number of visits. Then the profit function of the system is defined by

$$P = T_0 A_0 - T_1 B_0 - T_2 N_0 \tag{19}$$

where, T_0 =1500 (Revenue per unit up-time)

T_1 =500 (Charge per unit for server busy period)

T_2 =200 (Charge per visit made by the server)

4. Discussion

The transition diagram is used to calculate the system reliability measures like MTSF, availability of the system, busy period of the server, expected number of visits made by the server and profit values. It can be seen from the table 2 that the tendency of MTSF increases smoothly with respect to increments in refreshment rate (θ); however, other parameters such as failure rate of unit (λ=0.55), server refreshment request rate (μ=0.4), repair rate of unit (ϕ=0.5), cold stand by unit working probability (*a*=0.8) and not working probability (*b*=0.2) have fixed values. It is clear that when failure rate (λ) increases then MTSF declines. When technician refreshment rate (μ) enhances then MTSF also declines but when repair rate (ϕ) increases then MTSF enhances. Thus, the concept of refreshment is beneficial for the owner and technician. When MTSF enhances then system reliability also enhances.

Table 2: *MTSF vs. Refreshment Rate*

θ ↓	λ=0.55, μ=0.4 ϕ=0.5, *a*=0.8 *b*=0.2	λ=0.65,	μ=0.6	ϕ=0.7
0.1	3.8613371	3.218144	3.775294	4.0647311
0.2	3.8894957	3.235483	3.806505	4.1010786
0.3	3.9129156	3.250148	3.833119	4.1312067
0.4	3.9327001	3.262713	3.856081	4.1565858
0.5	3.9496349	3.273599	3.876094	4.1782569
0.6	3.9642941	3.283122	3.893693	4.1969772
0.7	3.9771076	3.291522	3.90929	4.2133109
0.8	3.9884034	3.298987	3.923207	4.2276872
0.9	3.998436	3.305664	3.935703	4.240438
1	4.007406	3.311673	3.946984	4.2518242

The availability of the redundant system is affected by the refreshment rate (θ), repair rate (ϕ), unit failure rate (λ) and server refreshment request rate (μ). Table 3 explores the availability of the system and its value increase corresponding to increments in refreshment rate (θ) when the system's other parameters λ=0.55, μ=0.4, ϕ=0.5, *a*=0.8, *b*=0.2 possess constant values. When the failure rate of unit changes (λ=0.55 to 0.65) then the availability of system declines. Also, when the technician request rate changes (μ=0.4 to 0.6) then the system's availability declines but when the repair rate of unit changes (ϕ=0.5 to 0.7) then the availability of the system enhances.

Table 3: *Availability vs. Refreshment Rate*

θ ↓	λ=0.55, μ=0.4 ϕ=0.5, *a*=0.8 *b*=0.2	λ=0.65,	μ=0.6	ϕ=0.7
0.1	0.214076	0.185918	0.190417	0.25028
0.2	0.319498	0.281347	0.293845	0.364639
0.3	0.381465	0.338725	0.358019	0.427923
0.4	0.421888	0.376681	0.401314	0.467056
0.5	0.450148	0.403461	0.432266	0.493101
0.6	0.470912	0.423258	0.455361	0.511375
0.7	0.48675	0.438421	0.473168	0.524719
0.8	0.499191	0.450366	0.487264	0.534774
0.9	0.509198	0.45999	0.498663	0.542545
0.1	0.517404	0.46789	0.508046	0.548679

It is an important part of the system that tells the customers about the performance of the product that it is beneficial or not. So, the cold standby redundant system is used to enhance the system's profit. It is evident from table 4 that the system uses constant parameters such that λ=0.55, μ=0.4, ϕ=0.5, *a*=0.8, *b*=0.2 and the trend of profit values enhanced with respect to increments in refreshment rate (θ). When the failure rate of unit (λ) changes from 0.55 to 0.65 then the profit of system declines. Also, when the technician request rate (μ) changes from 0.4 to 0.6 then profit values decline but when the repair rate of unit (ϕ) changes from 0.5 to 0.7 then the profit of the system enhances.

Table 4: *Profit vs. Refreshment Rate*

θ ↓	λ=0.55, μ=0.4 ϕ=0.5, a=0.8 b=0.2	λ=0.65,	μ=0.6	ϕ=0.7
0.1	2046.417	1588.805	1690.491	2692.9
0.2	3640.771	3027.839	3256.171	4428.411
0.3	4576.626	3891.418	4226.372	5389.1
0.4	5186.276	4461.578	4880.064	5983.344
0.5	5611.935	4863.085	5346.797	6378.977
0.6	5924.286	5159.353	5694.6	6656.653
0.7	6162.259	5385.877	5962.46	6859.482
0.8	6348.978	5563.998	6174.235	7012.365
0.9	6498.993	5707.276	6345.293	7130.568
1	6621.895	5824.711	6485.948	7223.9

7. Conclusion

The results of the study show that providing refreshments to the server during the job generally enhances his efficiency which is crucial for any repairable system. From the above discussion, the MTSF, availability and profit values of the system increase with respect to increments in refreshment rate as well as repair rate but the reliability values decline when server refreshment request rate and failure rate of unit are enhanced. It is clear from tables that the server has to override his emotions and try to satisfy the customers. The idea of refreshment is used by corporate sectors, industries, cybercafés, education, university systems, etc.

8. Future Scope

Refreshment to the server plays an essential role in the water-boosting station system where one unit is operative and another is kept on cold standby.

References

[1] Agrawal, A., Garg, D., Kumar, A. and Kumar, R. (2021). Performance analysis of the water treatment reverses osmosis plant. *Reliability: Theory & Applications*, *16*(3): 16-25.
[2] Barak, M. S., Garg, R. and Kumar, A. (2021). Reliability measures analysis of a milk plant using RPGT. *Life Cycle Reliability and Safety Engineering*, 10: 295-302.
[3] Bao, Y. and Mays, L. W. (1990). Model for water distribution system reliability. *Journal of Hydraulic Engineering*, 116: 1119-1137.
[4] Chaudhary, P. and Sharma, A. (2022). A two non-identical unit parallel system with priority in repair and correlated lifetimes. *Reliability Theory and Application*,17(67): 113-122.
[5] Deswal, S. and Malik, S. C. (2015). Reliability measures of a system of two non-identical units with priority subject to weather conditions. *Journal of Reliability and Statistical Studies*, 181-190.
[6] Diaz, V. G., Gomez, J. F., Lopez, M., Crespo, A., and de Leon, P. M. (2009). Warranty cost models State-of-Art: A practical review to the framework of warranty cost management. *ESREL*, 2051–2059.
[7] Garg, D. and Garg, R. (2022). Reliability analysis and profit optimization of briquette machine by considering neglected faults. *Reliability Theory and Application,* 17(67): 210-222.
[8] Gnedenko B. and Igor A.U. (1995). Probabilistic Reliability Engineering. *John Wiley and Sons.*

[9] Jia, H., Peng, R., Yang, L., Wu, T., Liu, D., and Li, Y. (2022). Reliability evaluation of demand based warm standby systems with capacity storage, *Reliability Engineering& System Safety,* 218: 108–132.

[10] Jack, N., and Murthy, D. P. (2007). A flexible extended warranty and related optimal strategies. *Journal of the Operational Research Society*, 58(12): 1612–1620.

[11] Kumar, A., Baweja, S. and Barak, M. (2015). Stochastic behavior of a cold standby system with maximum repair time. *Decision Science Letters*, *4*(4): 569-578.

[12] Kumar, J. and Goel, M. (2016). Availability and profit analysis of a two-unit cold standby system for general distribution. *Cogent Mathematics*, *3*(1): 1262937.

[13] Levitin, G., Finkelstein, M., and Xiang, Y. (2020). Optimal preventive replacement for cold standby systems with elements exposed to shocks during operation and task transfers. *IEEE Transactions on Systems, Man and Cybernetics: Systems*, 10: 2168–2216.

[14] Mahmoud, M. A. W. and Moshref, M. E. (2009). On a two-unit cold standby system considering hardware, human error failures and preventive maintenance. *Math. Comput. Model,* 51: 736–745.

[15] Malik, S. C. and Rathee, R. (2016). Reliability modelling of a parallel system with maximum operation and repair times. *International Journal of Operational Research*, 25(1): 131-142.

[16] Sengar S. and Mangey R. (2022). Reliability and performance analysis of a complex manufacturing system with inspection facility using copula methodology. *Reliability: Theory & Applications,* 4(71): 494-508.

[17] Subramanian, R. (1978). Availability of two-unit system with preventive maintenance and one repair facility. *IEEE Transactions on Reliability*, 27(2): 171-172.

[18] Temraz, N. S. Y. (2018). Availability and reliability analysis for dependent system with load sharing and degradation facility. *International Journal of Systems Science and Applied Mathematics*, 3(1): 10-15.

[19] Wang, G. J. and Zhang, Y. L. (2007). An optimal replacement policy for repairable cold standby system with priority in use. *International Journal of Systems Science*, *38*(12): 1021-1027.

RT&A, No 4 (76)
Volume 18, December 2023

POWER WEIBULL QUANTILE FUNCTION AND IT'S RELIABILITY ANALYSIS

JEENA JOSEPH AND SONITTA TONY

•

Department of Statistics
St. Thomas' college (Autonomous)
Thrissur, India
sony.jeena@gmail.com, sonittachakkery2014@gmail.com

Abstract

In this article, we propose a new class of distributions defined by a quantile function, which is the sum of the quantile functions of the Power and Weibull distributions. Various distributional properties and reliability characteristics of the class are discussed. To examine the usefulness of the model, the model is applied to a real life datasets. Parameters are estimated using maximum likelihood estimation technique.

Keywords: Power distribution; Weibull distribution; L-moments; Hazard quantile function; Mean residual quantile function; Residual variance quantile function; Reversed hazard quantile function.

1. Introduction

In modelling and analysis of statistical data, probability distribution can be specified either in terms of distribution function or by the quantile function. Quantile functions have several interesting properties that are not shared by distributions, which makes it more convenient for analysis. For example, the sum of two quantile functions is again a quantile function. For a nonnegative random variable X with distribution function F(x), the quantile function $Q(u)$ is defined by Nair and Sankaran [7]

$$Q(u) = F^{-1}(u) = inf\{x : F(x) \geq u\}, \quad 0 \leq u \leq 1 \tag{1}$$

For every -$\infty < x < \infty$ and $0 < u < 1$, we have

$$F(x) \geq u \text{ if and only if } Q(u) \leq x.$$

Thus, if there exists an x such that $F(x) = u$, then $F(Q(u)) = u$ and $Q(u)$ is the smallest value of x satisfying $F(x) = u$. Further, if $F(x)$ is continuous and strictly increasing, $Q(u)$ is the unique value x such that $F(x) = u$, and so by solving the equation $F(x) = u$, we can find x in terms of u which is the quantile function of X.

If $f(x)$ is the probability function of X, then $f(Q(u))$ is called the density quantile function. The derivative of $Q(u)$,

$$q(u) = Q^{'}(u),$$

is known as the quantile density function of X. If $F(x)$ is right continuous and strictly increasing, we have

$$F(Q(u)) = u \tag{2}$$

so that F(x) = u implies x = Q(u). When f(x) is the probability density function (PDF) of X; we have from (2)

$$q(u)f(Q(u)) = 1 \tag{3}$$

Quantile function has several properties that are not shared by distribution function. See Nair and Sankaran [7] for details. For example, the sum of two quantile functions is again a quantile function. Further, the product of two positive quantile functions is again a quantile function in the nonnegative setup. There are explicit general distribution forms for the quantile function of order statistics. It is easier to generate random numbers from the quantile function. A major development in portraying quantile functions to model statistical data is given by Hastings *et al.* [5], who introduced a family of distributions by a quantile function. This was refined later by Tukey [13] to form a symmetric distribution, called the Tukey lambda distribution.

This model was generalized in different ways, referred as lambda distributions. These include various forms of quantile functions discussed in Ramberg and Schmeiser [10], Ramberg [8], Ramberg *et al.* [9] and Freimer *et al.* [1]. Govindarajulu [3] introduced a new quantile function by taking the weighted sum of quantile functions of two power distributions. Hankin and Lee [4] presented a new power - Pareto distribution by taking the product of power and Pareto quantile functions. Van Staden and Loots [14] developed a four-parameter distribution, using a weighted sum of the generalized Pareto and its reflection quantile functions. Sankaran *et al.* [12] developed a new quantile function based on the sum of quantile functions of generalized Pareto and Weibull quantile functions. Sankaran and Dileep [11] developed a new quantile function based on the sum of quantile functions of half logistic and exponential geometric distributions.

The aim of the present work is to introduce a new quantile function that is useful in reliability analysis. The proposed quantile function is derived by taking the sum of quantile functions of power and Weibull distributions. The survival function and quantile function of power distribution are respectively given by

$$\bar{F}(x) = 1 - \left(\frac{x}{\alpha}\right)^{\beta} \qquad 0 \le x \le \alpha; \alpha, \beta > 0$$

and

$$Q_1(u) = \alpha u^{\frac{1}{\beta}} \qquad 0 \le u \le 1, \qquad \alpha, \beta > 0. \tag{4}$$

The survival function and quantile function of Weibull distribution are respectively given by

$$\bar{F}(x) = exp\left[-\left(\frac{x}{\sigma}\right)^{\lambda}\right] \qquad x > 0; \lambda, \sigma > 0$$

and

$$Q_2(u) = \sigma(-log(1-u))^{\frac{1}{\lambda}} \qquad 0 \le u \le 1, \quad \alpha, \lambda > 0 \tag{5}$$

We now propose a new class of distributions defined by a quantile function, which is the sum of quantile functions of power and Weibull distributions.

2. Power–Weibull (PW) quantile function

Let X and Y be two nonnegative random variables with distribution functions F(x) and G(x) with quantile functions $Q_1(u)$ and $Q_2(u)$, respectively. Then

$$Q(u) = Q_1(u) + Q_2(u), \tag{6}$$

is also a quantile function. We now introduce a class of distributions given by the quantile function,

$$Q(u) = \alpha u^{\frac{1}{\beta}} + \sigma(-log(1-u))^{\frac{1}{\lambda}} \qquad 0 \le u \le 1, \quad \alpha, \beta, \sigma, \lambda > 0 \tag{7}$$

Thus Q(u) is the sum of (4) and (5). It is named as Power–Weibull (PW) quantile function. The quantile density function is obtained as

$$q(u) = \frac{\alpha u^{\frac{1}{\beta}-1}}{\beta} + \frac{\sigma(-log(1-u))^{\frac{1}{\lambda}-1}}{\lambda(1-u)}. \tag{8}$$

For the proposed class of distribution, the density function $f(x)$ can be written in terms of the distribution function as

$$f(x) = \frac{\beta\lambda(1-F(x))}{\alpha\lambda(1-F(x))F(x)^{\frac{1}{\beta}-1} + \sigma\beta(-log^{\frac{1}{\lambda}-1}(1-F(x)))}. \tag{9}$$

For all values of the parameters, the density is strictly decreasing in x and it tends to zero as x $\rightarrow$ ∞.
The quantile function (7) represents a family of distributions with a variety of shapes for its probability density function. Plots of the density function for different combinations of parameters are shown in figure 1, 2 and 3.

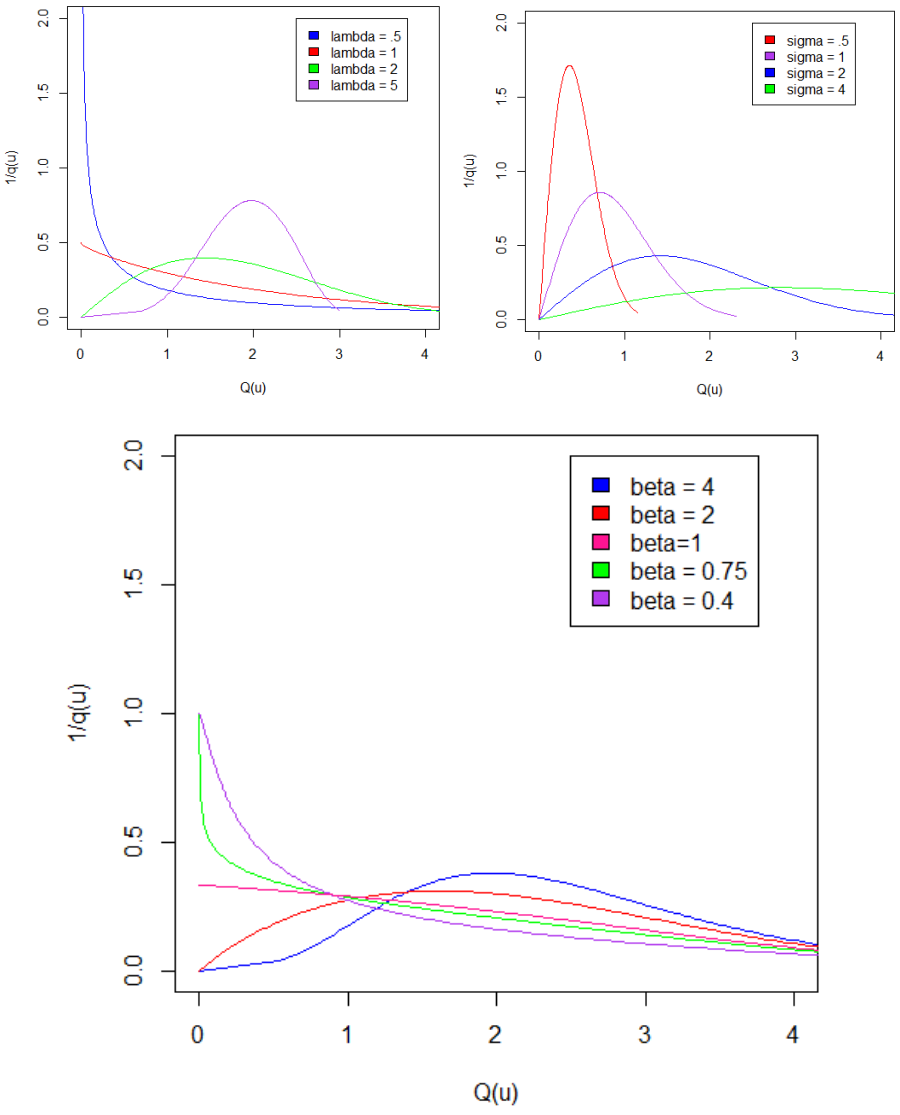


Figure 1: *Plot of the density function for various values of λ , σ and β*

3. Members of the family

The proposed family of distributions (7) includes several well-known distributions for various values of the parameters.

Case 1. $\alpha = 0$, $\sigma > 0$.

$$Q(u) = \sigma(-log(1-u))^{\frac{1}{\lambda}} \tag{10}$$

is the quantile function of the Weibull distribution, which contains the exponential distribution with mean σ for $\lambda = 1$ and the Rayleigh distribution when $\lambda = 2$.

Case 2. $\sigma = 0$, $\alpha > 0$, $\beta > 0$.

$$Q(u) = \alpha u^{\frac{1}{\beta}} \tag{11}$$

is the quantile function of the power distribution.

Case 3. $\alpha = 0$, $\lambda = 1$, $\sigma > 0$.

$$Q(u) = \alpha log\left(\frac{1-pu}{1-u}\right) \tag{12}$$

which is quantile function of exponential geometric distribution with $p = 0$ and $\alpha = -\sigma$.

We can derive some well-known distributions from the proposed model by making use of various transformations described in Gilchrist [2].

Case 4. Consider the power u-transformation $T(u) = u^{\frac{1}{\theta}}$ with $\alpha = 0$. Then,

$$Q(u) = \sigma(-log(1-u^{\frac{1}{\theta}}))^{\frac{1}{\lambda}} \tag{13}$$

is the quantile function of exponentiated Weibull distribution. If $\lambda = 1$,

$$Q(u) = \sigma(-log(1-u^{\frac{1}{\theta}})) \tag{14}$$

is the quantile function of generalized exponential distribution.

Case 5. By reciprocal transformation with σ=0 then,

$$Q(u) = \frac{1}{Q(1-u)} = \sigma(1-u)^{-\frac{1}{\alpha}} \tag{15}$$

which is the quantile function of Pareto distribution with $\sigma = \frac{1}{\alpha}$ and $\alpha = \beta$.

4. Distributional characteristics

The quantile based measures of the distributional characteristics like location, dispersion, skewness, and kurtosis are popular in statistical analysis. These measures are also useful for estimating parameters of the model by matching population characteristics with corresponding sample characteristics. For the model (7), we have,

$$Median = Q(\frac{1}{2}) = \alpha(0.5)^{\frac{1}{\beta}} + \sigma(log(2))^{\frac{1}{\lambda}}. \tag{16}$$

The inter-quartile-range, IQR is obtained as,

$$\begin{aligned} IQR &= Q(\frac{3}{4}) - Q(\frac{1}{4}) \\ &= \alpha[0.75^{\frac{1}{\beta}} - 0.25^{\frac{1}{\beta}}] + \sigma[(log(4))^{\frac{1}{\lambda}} - (-log(0.75))^{\frac{1}{\lambda}}]. \end{aligned} \tag{17}$$

The Galton's coefficient of skewness, S is given by,

$$\begin{aligned} S &= \frac{Q(\frac{3}{4}) + Q(\frac{1}{4}) - 2Median}{IQR} \\ &= \frac{\alpha[0.75^{\frac{1}{\beta}} + 0.25^{\frac{1}{\beta}} - 2(0.5)^{\frac{1}{\beta}}] + \sigma[(log4)^{\frac{1}{\lambda}} + (-log0.75)^{\frac{1}{\lambda}} - 2(log2)^{\frac{1}{\lambda}}]}{\alpha[0.75^{\frac{1}{\beta}} - 0.25^{\frac{1}{\beta}}] + \sigma[(log4)^{\frac{1}{\lambda}} - (-log0.75)^{\frac{1}{\lambda}}]} \end{aligned} \tag{18}$$

and the Moor's coefficient of kurtosis,

$$\begin{aligned} T &= \frac{Q(\frac{7}{8}) - Q(\frac{5}{8}) + Q(\frac{3}{8}) - Q(\frac{1}{8})}{IQR} \\ &= \frac{\alpha(7^{\frac{1}{\beta}} - 5^{\frac{1}{\beta}} + 3^{\frac{1}{\beta}} - 1)8^{\frac{-1}{\beta}} + \sigma[log^{\frac{1}{\lambda}}8 - log^{\frac{1}{\lambda}}(8/3) + log^{\frac{1}{\lambda}}(8/5) - log^{\frac{1}{\lambda}}(8/7)]}{\alpha[0.75^{\frac{1}{\beta}} - 0.25^{\frac{1}{\beta}}] + \sigma[(log4)^{\frac{1}{\lambda}} - (-log0.75)^{\frac{1}{\lambda}}]}. \end{aligned} \tag{19}$$

5. L-moments

The L-moments are often found to be more desirable than the conventional moments in describing the characteristics of the distributions as well as for inference. A unified theory and a systematic study on L-moments have been presented by Hosking [6]. The L-moments have generally lower sampling variances and are robust against outliers.
The rth L moment is given by

$$L_r = \int_0^1 \sum_{k=0}^{r-1} (-1)^{r-1-k} \binom{r-1}{k} \binom{r-1+k}{k} u^k Q(u) du \tag{20}$$

For the model (7), the first L moment L_1 is the mean of the distribution.

$$L_1 = \int_0^1 Q(u) du = \frac{\alpha\beta}{1+\beta} + \sigma\Gamma(\frac{1}{\lambda} + 1) \tag{21}$$

The second L-moment for the family is obtained as

$$L_2 = \int_0^1 (2u-1)Q(u) du = \frac{\alpha\beta}{1+3\beta+2\beta^2} + \sigma\Gamma(\frac{1}{\lambda} + 1)(1 - 2^{-\frac{1}{\lambda}}) \tag{22}$$

which is twice the mean differences of the population.
The third and fourth L-moments are obtained as

$$\begin{aligned} L_3 &= \int_0^1 (6u^2 - 6u + 1)Q(u) du \\ &= \frac{\alpha\beta}{\beta+1} - \frac{6\alpha\beta^2}{1+5\beta+6\beta^2} + \sigma\Gamma(\frac{1}{\lambda} + 1)(1 - 32^{-\frac{1}{\lambda}} + 23^{-\frac{1}{\lambda}}) \end{aligned} \tag{23}$$

and

$$\begin{aligned} L_4 &= \int_0^1 (20u^3 - 30u^2 + 12u - 1)Q(u) du \\ &= \frac{20\alpha\beta}{1+4\beta} - \frac{30\alpha\beta}{1+3\beta} + \frac{12\alpha\beta}{1+2\beta} - \frac{\alpha\beta}{1+\beta} + \sigma\Gamma(\frac{1}{\lambda} + 1)(1 - 32^{1-\frac{1}{\lambda}} + 103^{-\frac{1}{\lambda}} - 54^{-\frac{1}{\lambda}}) \end{aligned} \tag{24}$$

The L-coefficient of variation (τ_2), analogous to the coefficient of variation based on ordinary moments is given by,

$$\tau_2 = \frac{L_2}{L_1} = \frac{\frac{\alpha\beta}{1+3\beta+2\beta^2} + \sigma\Gamma(\frac{1}{\lambda} + 1)(1 - 2^{-\frac{1}{\lambda}})}{\frac{\alpha\beta}{1+\beta} + \sigma\Gamma(\frac{1}{\lambda} + 1)}. \tag{25}$$

L-coefficient of skewness (τ_3) for the PW quantile function is obtained as

$$\begin{aligned}\tau_3 &= \frac{L_3}{L_2} \\ &= \frac{\frac{\alpha\beta}{\beta+1} - \frac{6\alpha\beta^2}{1+5\beta+6\beta^2} + \sigma\Gamma(\frac{1}{\lambda}+1)(1-32^{-\frac{1}{\lambda}}+23^{-\frac{1}{\lambda}})}{\frac{\alpha\beta}{1+3\beta+2\beta^2} + \sigma\Gamma(\frac{1}{\lambda}+1)(1-2^{-\frac{1}{\lambda}})}.\end{aligned} \tag{26}$$

L-coefficient of kurtosis (τ_4) for the PW quantile function is obtained as

$$\begin{aligned}\tau_4 &= \frac{L_4}{L_3} \\ &= \frac{\frac{20\alpha\beta}{1+4\beta} - \frac{30\alpha\beta}{1+3\beta} + \frac{12\alpha\beta}{1+2\beta} - \frac{\alpha\beta}{1+\beta} + \sigma\Gamma(\frac{1}{\lambda}+1)(1-32^{1-\frac{1}{\lambda}}+103^{-\frac{1}{\lambda}}-54^{-\frac{1}{\lambda}})}{\frac{\alpha\beta}{\beta+1} - \frac{6\alpha\beta^2}{1+5\beta+6\beta^2} + \sigma\Gamma(\frac{1}{\lambda}+1)(1-32^{-\frac{1}{\lambda}}+23^{-\frac{1}{\lambda}})}.\end{aligned} \tag{27}$$

6. Order statistics

If $X_{r:n}$ is the rth order statistic in a random sample of size n, then the density function of $X_{r:n}$ can be written as

$$f_r(x) = \frac{1}{B(r, n-r+1)} f(x) F^{r-1}(x)(1-F(x))^{n-r} \tag{28}$$

From (9), we have

$$f_r(x) = \frac{1}{B(r, n-r+1)} \frac{\beta\lambda F^{r-1}(x)(1-F(x))^{n-r+1}}{\alpha\lambda(1-F(x))F(x)^{\frac{1}{\beta}-1} + \sigma\beta(-log^{\frac{1}{\lambda}-1}(1-F(x)))}. \tag{29}$$

Hence,

$$\begin{aligned}\mu_{r:n} = E(X_{r:n}) &= \int x f_r(x) dx \\ &= \frac{1}{B(r, n-r+1)} \int_0^\infty x \frac{\beta\lambda F^{r-1}(x)(1-F(x))^{n-r+1}}{\alpha\lambda(1-F(x))F(x)^{\frac{1}{\beta}-1} + \sigma\beta(-log^{\frac{1}{\lambda}-1}(1-F(x)))} dx.\end{aligned} \tag{30}$$

In quantile terms, we have

$$E(X_{r:n}) = \frac{1}{B(r, n-r+1)} \int_0^1 Q(u) \frac{\beta\lambda u^{r-1}(1-u)^{n-r+1}}{\alpha\lambda(1-u)u^{\frac{1}{\beta}-1} + \sigma\beta(-log^{\frac{1}{\lambda}-1}(1-u))} du. \tag{31}$$

For the class of distributions (7), the first-order statistic $X_{1:n}$ has the quantile function

$$\begin{aligned}Q_1(u) &= Q(1-(1-u)^{\frac{1}{n}}) \\ &= \alpha[1-(1-u)^{\frac{1}{n}}]^{\frac{1}{\beta}} + \sigma[-log(1-u)^{\frac{1}{n}}]^{\frac{1}{\lambda}},\end{aligned} \tag{32}$$

and the nth order statistic $X_{n:n}$ has the quantile function

$$\begin{aligned}Q_n(u) &= Q(u^{\frac{1}{n}}) \\ &= \alpha u^{\frac{1}{n\beta}} + \sigma(-log(1-u^{\frac{1}{n}}))^{\frac{1}{\lambda}}.\end{aligned} \tag{33}$$

7. Hazard quantile function

One of the basic concepts employed for modeling and analysis of lifetime data is the hazard rate. In a quantile setup, Nair and Sankaran [7] defined the hazard quantile function, which is equivalent to the hazard rate. The hazard quantile function $H(u)$ is defined as

$$H(u) = h(Q(u))) = (1-u)^{-1} f Q(u) = [(1-u)q(u)]^{-1}. \tag{34}$$

Thus $H(u)$ can be interpreted as the conditional probability of failure of a unit in the next small interval of time given the survival of the unit until $100(1-u)\%$ point of the distribution. Note that $H(u)$ uniquely determines the distribution using the identity,

$$Q(u) = \int_0^u \frac{dp}{(1-p)H(p)}. \tag{35}$$

The hazard quantile functions of Power and Weibull distribution is given by

$$H_1(u) = \beta\alpha^{-1}(1-u)^{-1}u^{1-\frac{1}{\beta}} \tag{36}$$

and

$$H_2(u) = \lambda\sigma^{-1}(-log(1-u))^{1-\frac{1}{\lambda}}. \tag{37}$$

Since the proposed class of distributions is the sum of quantile functions of power and Weibull quantile functions, (34) and (35) give

$$\frac{1}{H(u)} = \frac{1}{H_1(u)} + \frac{1}{H_2(u)} \tag{38}$$

where $H(u)$,$H_1(u)$ and $H_2(u)$ are the hazard quantile functions of the proposed class of distributions, power, and Weibull quantile functions, respectively.

For the PW quantile function (7), we have

$$H(u) = \frac{1}{\beta^{-1}\alpha(1-u)u^{\frac{1}{\beta}-1} + \lambda^{-1}\sigma(-log(1-u))^{\frac{1}{\lambda}-1}} \tag{39}$$

with $H(0) = \infty$ and $H(1) = 0$. Plots of hazard quantile function for different values of parameters are given in figure (2).

The shape of the hazard function is determined by the derivative of $H(u)$, which is obtained as

$$H^{'}(u) = \frac{\beta^{-1}\alpha u^{\frac{1}{\beta}-1}[1+(\frac{1}{u}-1)(1-\frac{1}{\beta})] + \lambda^{-1}\sigma(1-\frac{1}{\lambda})(1-u)^{-1}(-log(1-u))^{\frac{1}{\lambda}-2}}{[\beta^{-1}\alpha(1-u)u^{\frac{1}{\beta}-1} + \lambda^{-1}\sigma(-log(1-u))^{\frac{1}{\lambda}-1}]^2}. \tag{40}$$

Since $[\beta^{-1}\alpha(1-u)u^{\frac{1}{\beta}-1} + \lambda^{-1}\sigma(-log(1-u))^{\frac{1}{\lambda}-1}]^2 > 0$. For all values of the parameters, the sign of $H^{'}(u)$ depends only on

$$g(u) = \beta^{-1}\alpha u^{\frac{1}{\beta}-1}[1+(\frac{1}{u}-1)(1-\frac{1}{\beta})] + \lambda^{-1}\sigma(1-\frac{1}{\lambda})(1-u)^{-1}(-log(1-u))^{\frac{1}{\lambda}-2}. \tag{41}$$

The parameters α, σ being always > 0 do not affect the sign of the two terms in g(u). Now we consider the following cases.

Case 1. $o < \beta < 1$ and $o < \lambda < 1$.
$g(u) < 0$ and distribution has an decreasing hazard rate (DHR).

Case 2. $\beta = 1$ and $\lambda = 1$.
$g(u) = \alpha$ and distribution has an increasing hazard rate (IHR).

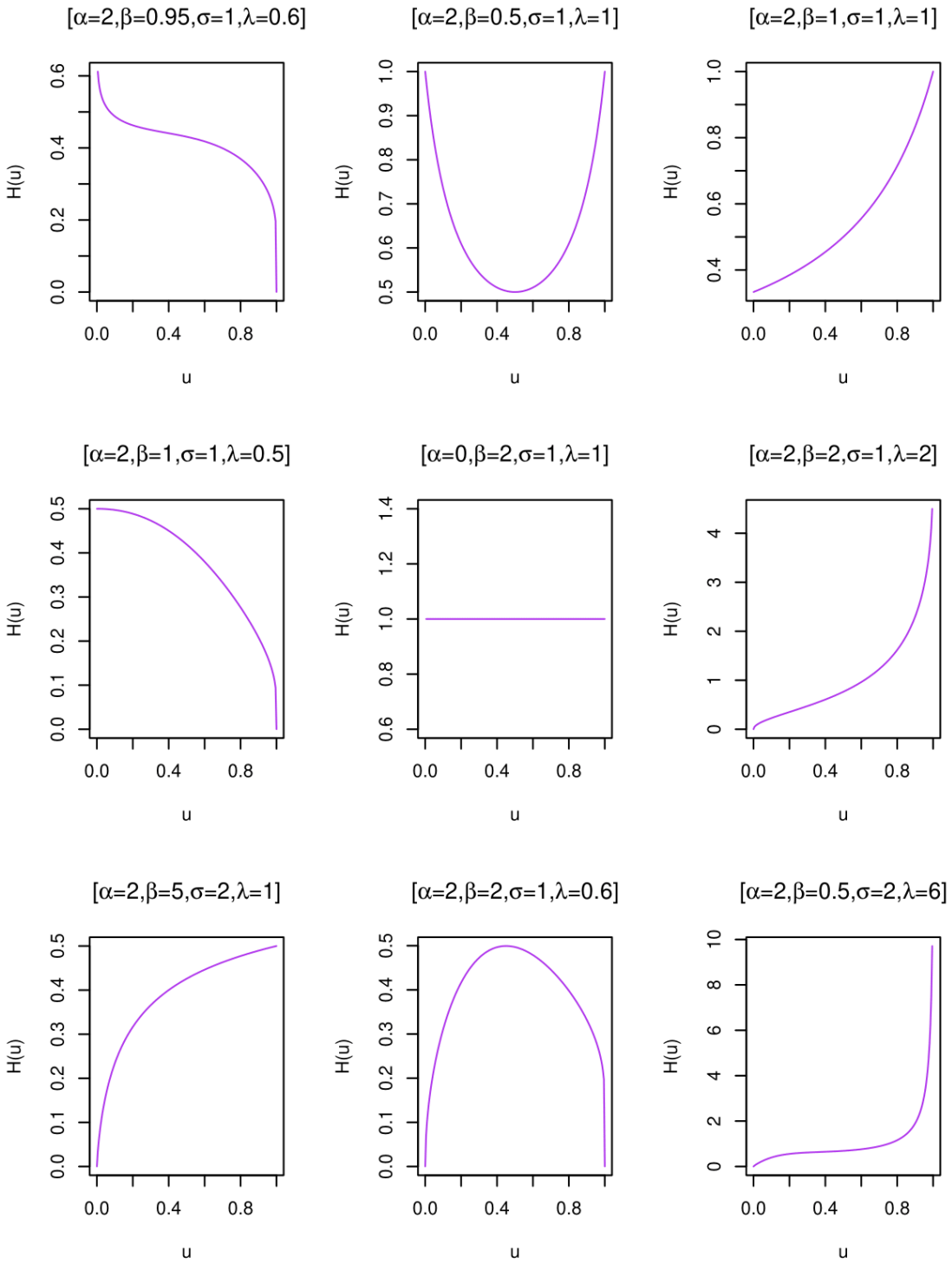


Figure 2: *Plots of hazard quantile function*

Case 3. $\beta = 1$ and $\lambda > 1$.
$g(u) > 0$ and distribution has an increasing hazard rate (IHR).

Case 4. $\beta > 1$ and $\lambda = 1$.
The first term in $g(u)$ is positive and second term is zero, so that $g(u) > 0$ and distribution has an increasing hazard rate (IHR).

Case 5. $\beta > 1$ and $\lambda > 1$.
The first and second term in $g(u)$ is positive, so that $g(u) > 0$ and distribution has an increasing hazard rate (IHR).

Case 6. $\beta = 1, 0 < \lambda < 1$.
Distribution has an decreasing hazard rate (DHR).

Case 7. $o < \beta < 1$ and $\lambda = 1$.
$H(u)$ attains a minimum at $u_0 = 1 - \beta$ and therefore $H(u)$ is bathtub shaped.

For the remaining cases, ($o < \beta < 1, \lambda > 1$) and ($\beta > 1, 0 < \lambda < 1$), one term in $g(u)$ is positive and the other is negative so that $g(u)$ can be zero. From (40) and(41) we obtain the first derivative of $H(u)$,

$$H^{'}(u) = \frac{g(u)}{[\beta^{-1}\alpha(1-u)u^{\frac{1}{\beta}-1} + \lambda^{-1}\sigma(-log(1-u))^{\frac{1}{\lambda}-1}]^2}. \tag{42}$$

For further analysis, we take the second derivative of $H(u)$, the sign of $H^{''}(u)$ depends on,

$$g^{'}(u)\Big(\beta^{-1}\alpha(1-u)u^{\frac{1}{\beta}-1} + \lambda^{-1}\sigma(-log(1-u))^{\frac{1}{\lambda}-1}\Big)^2 + 2g(u)\Big(\beta^{-1}\alpha(1-u)u^{\frac{1}{\beta}-1} + \lambda^{-1}\sigma(-log(1-u))^{\frac{1}{\lambda}-1}\Big)g(u).$$

Let u_0 be the solution of the equation $g(u) = 0$. Then the sign of $H^{''}(u)$ at u_0 depends on $g^{'}(u_0)$. Since u_0 is a solution of $g(u) = 0$.

$$g^{'}(u) = \beta^{-1}\alpha(\frac{1}{\beta} - 1)u^{\frac{1}{\beta}-2}\Big(1 + (\frac{1}{u} - 1)(1 - \frac{1}{\beta}) + u^{-1}\Big) + \lambda^{-1}\sigma(\frac{1}{\lambda} - 1)(1-u)^{-2}$$
$$(-log(1-u))^{\frac{1}{\lambda}-2}\Big(1 - (\frac{1}{\lambda} - 2)(-log(1-u))^{-1}\Big).$$

Then,

$$g^{'}(u_0) = \beta^{-1}\alpha(\frac{1}{\beta} - 1)u_0^{\frac{1}{\beta}-2}\Big(1 + (\frac{1}{u_0} - 1)(1 - \frac{1}{\beta}) + u_0^{-1}\Big) + \lambda^{-1}\sigma(\frac{1}{\lambda} - 1)(1-u_0)^{-2}$$
$$(-log(1-u_0))^{\frac{1}{\lambda}-2}\Big(1 - (\frac{1}{\lambda} - 2)(-log(1-u_0))^{-1}\Big). \tag{43}$$

Case 8. $0 < \beta < 1$ and $\lambda > 1$.
$H(u)$ has an increasing hazard rate (IHR).

Case 9. $\beta > 1$ and $0 < \lambda < 1$.
$H(u)$ has an upside-down bathtub-shaped hazard quantile function.

The patterns of $H(u)$ for various parameter values are summarized in table 1.

Table 1: *Behavior of the hazard quantile function for different regions of parameter space.*

No.	Parameter region	Shape of hazard quantile function
1	$o < \beta < 1$ and $o < \lambda < 1$	DHR
2	$\beta = 1$ and $\lambda = 1$	IHR
3	$\beta = 1$ and $\lambda > 1$	IHR
4	$\beta > 1$ and $\lambda = 1$	IHR
5	$\beta > 1$ and $\lambda > 1$	IHR
6	$\beta = 1, 0 < \lambda < 1$	DHR
7	$o < \beta < 1$ and $\lambda = 1$	Bathtub
8	$o < \beta < 1$ and $\lambda > 1$	IHR
9	$\beta > 1$ and $0 < \lambda < 1$	Upside-down Bathtub

8. Mean residual quantile function

Another concept used in reliability is that of residual life $X_t = (X - t|X > t)$ with survival function

$$\bar{F}_t(x) = \bar{F}(t + x)/\bar{F}(t), \quad x \geq 0, o < t < T.$$

The mean residual life function is then

$$m(t) = E(X_t) = [\bar{F}(t)]^{-1} \int_t^{\infty} \bar{F}(x)dx.$$

Accordingly, the mean residual quantile function is defined by Nair and Sankaran [7] as

$$M(u) = mQ(u) = (1 - u)^{-1} \int_u^1 (Q(t) - Q(u))dt \tag{44}$$

which is the average remaining life beyond the $100(1 - u)\%$ point of the distribution. For the class of distributions (7), $M(u)$ has the form

$$M(u) = \frac{\alpha(1 - u^{\frac{1}{\beta}+1})}{(1 - u)(\frac{1}{\beta} + 1)} - \alpha u^{\frac{1}{\beta}} + \sigma(1 - u)^{-1}\Gamma\big(\frac{1}{\lambda} + 1, -log(1 - u)\big) - \sigma(-log(1 - u))^{\frac{1}{\lambda}}. \tag{45}$$

9. Residual variance quantile function

The quantile form of variance residual function, the residual variance quantile function is defined as

$$V(u) = (1 - u)^{-1} \int_u^1 Q^2(p)dp - (M(u) + Q(u))^2. \tag{46}$$

In the above equation, the variance residual life function is obtained by letting $Q(u) = x$.
Nair and Sankaran [7] derived the relationship between $M(u)$ and $V(u)$ as

$$M^2(u) = V(u) - (1 - u)V^{'}(u) \tag{47}$$

or

$$V(u) = (1 - u)^{-1} \int_u^1 M^2(p)dp. \tag{48}$$

Since $M(u)$ characterizes the distribution, from above equations it follows that $V(u)$ also characterizes the distribution.

For the PW qantile function, residual variance quantile function is

$$V(u) = \frac{1}{1-u}\left\{\frac{\beta\alpha^2(1-u^{\frac{2+\beta}{\beta}})}{2+\beta} + \sigma^2\Gamma(\frac{2}{\lambda}+1, -log(1-u)) + 2\alpha\sigma\int_u^1 p^{1/\beta}\right.$$
$$\left.(-log(1-p))^{1/\lambda}dp\right\} - \left\{\frac{\alpha(1-u^{\frac{1}{\beta}+1})}{(1-u)(1/\beta+1)} + \sigma(1-u)^{-1}\Gamma(\frac{1}{\lambda}+1,\right.$$
$$\left.-log(1-u))\right\}^2 \tag{49}$$

10. Reversed hazard quantile function

The reversed hazard quantile function [8] is defined by

$$A(u) = \frac{1}{uq(u)} \tag{50}$$

and it determines the distribution through the formula

$$Q(u) = \int_0^u \frac{1}{pA(p)}dp. \tag{51}$$

For power-Weibull distribution,

$$A(u) = \Big[\frac{\alpha u^{\frac{1}{\beta}}}{\beta} + \frac{\sigma u(-log(1-u))^{\frac{1}{\lambda}-1}}{\lambda(1-u)}\Big]^{-1}. \tag{52}$$

11. Data Analysis

There are different methods for the estimation of parameters of the quantile function. The method of percentiles, method of L-moments, method of minimum absolute deviation, method of least squares, and method of maximum likelihood are commonly used techniques. To estimate the parameters of (7), we use the method of maximum likelihood estimation procedure.

To illustrate the application of the proposed class of distributions we consider a real data set reported in Zimmer *et al.* [15]. The data consist of times to first failure of 20 electric carts used for internal transportation and delivery in a manufacturing company. The estimates of the parameters are obtained using R software as,

$$\hat{\alpha} = 0.421, \hat{\beta} = 2.088, \hat{\sigma} = 0.027 \quad \text{and} \quad \hat{\lambda} = 0.312.$$

To examine the adequacy of the model, we use chi-squared goodness of fit. The test gives the p-value 0.082. This indicates the adequacy of proposed model for the given data set.

12. Summary and Conclusion

In this paper, we introduced a class of distributions (7), which is the sum of the quantile function of the power and Weibull distributions known as Power-Weibull (PW) quantile function and its graphical representation of density function is included. We have identified several well-known distributions which are either the members of the proposed class of distributions and also through suitable transformations such as Weibull distribution, power distribution, generalized exponential distribution, etc. Various distributional characteristics and L-moments are discussed. The hazard quantile function and its shape in various parameter region are analysed. Increasing, decreasing,

bathtub and upside-down bathtub hazard quantile function are obtained. Mean residual quantile function of PW quantile function was studied. Finally PW model is applied to a real life data set, and parameters are estimated by using maximum likelihood estimation procedure and model adequacy is checked by chi-squared goodness of fit test using the R software.

There are several properties and extensions for the PW quantile function not considered in this article, such as parameter estimation using L-moments, stochastic orderings and generalization of PW quantile function.

References

[1] Freimer, M.G., Kollia, G. S. Mudholkar, and C. T. Lin. (1988). A study of the generalized Tukey lambda family.*Communications in Statistics—Theory and Methods 17 (10)*,3547–67.

[2] Gilchrist, W. (2000). Statistical Modelling with Quantile Functions,*(1st ed.). Chapman and Hall/CRC, Abingdon*.

[3] Govindarajulu, Z. (1977). A class of distributions useful in life testing and reliability. *IEEE Transactions on Reliability, 26(1)*, 67–69.

[4] Hankin, R. K., and A. Lee. (2006). A new family of non-negative distributions. *Australian and New Zealand Journal of Statistics, 48(1)*, 67–78.

[5] Hastings, C., F. Mosteller, J. W. Tukey, and C. P. Winsor. (1947). Low moments for small samples: A comparative study of order statistics.*Annals of Mathematical Statistics, 18*, 413–426.

[6] Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society Series B, 52*, 105–24.

[7] Nair, N. U., and P. G. Sankaran. (2009). Quantile-based reliability analysis. *Communications in Statistics—Theory and Methods, 38(2)*, 222–32.

[8] Ramberg, J. S. (1975). A probability distribution with applications to Monte Carlo simulation studies.*In Statistical distributions in scientific work: Model building and model selection*, eds. G.P. Patil, S. Kotz, J. K. Ord, vol. 2. Dordrecht: D. Reidel.

[9] Ramberg, J. S., E. J. Dudewicz, P. R. Tadikamalla, and E. F. Mykytka. 1979. A probability distribution and its uses in fitting data.*Technometrics 21 (2)*, 201–14.

[10] Ramberg, J. S., and B. W. Schmeiser. 1972. An approximate method for generating symmetric random variables. *Communications of the ACM 15 (11)*, 987–90.

[11] Sankaran, P. G. and M, Dileep Kumar. (2018). A new class of quantile functions useful in reliability analysis. *Journal of Statistical Theory and Practice, 12(3)*, 615-634.

[12] Sankaran, P. G., N. Unnikrishnan Nair and N. N. Midhu. (2016). A New Quantile Function with Applications to Reliability Analysis. *Communications in Statistics - Simulation and Computation, 45(2)*, 566-582.

[13] Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics, 33(1)*, 1–67.

[14] Van Staden, P. J., and M. T. Loots. (2009). Method of L-moment estimation for the generalized lambda distribution. *Proceedings of the Third Annual ASEARC Conference, New Castle, Australia, December 7–8*.

[15] Zimmer, W.J., J. B. Keats, and F. K. Wang.(1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology 30(4)*, 386-94.

GENERALISED EXPONENTIAL RATIO-CUM-PRODUCT ESTIMATOR FOR ESTIMATING POPULATION VARIANCE IN SIMPLE RANDOM SAMPLING

Rafia Jan[1], T. R. Jan[2] and Faizan Danish[3*]

•

[1]Department of Statistics, Government Degree College Bejbehara Anantnag, J&K, India
[2] Department of Statistics, University of Kashmir, J&K, India
[3]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amravati, Andhra Pradesh-522237, India.
Email: 1:rafiajan836@gmail.com
2:drtrjan@gmail.com
3.danishstat@gmail.com

Abstract

This study presents a comprehensive investigation into the estimation of population variance for a study variable using Simple Random Sampling, with the incorporation of two auxiliary variables. To address the challenges of variance estimation in complex scenarios, a novel approach termed the "Proposed Generalized Exponential Ratio-cum-Product Estimator" is introduced. This innovative estimator belongs to a class of estimators that rely on exponential functions of the auxiliary variables, providing enhanced precision and efficiency in variance estimation. To thoroughly assess the performance of the proposed estimators, the research develops equations for both Mean Square Errors and Biases, unveiling their statistical properties. The study systematically explores the conditions under which these estimators demonstrate superior efficiency compared to traditional alternative estimators, thereby enabling researchers to identify contexts where their utilization is most beneficial. The empirical aspect of the research constitutes a significant contribution to the study's validity. Through empirical analysis, the proposed estimators are directly compared against the conventional Unbiased Sample Variance Estimator, showcasing their clear superiority in terms of efficiency. Furthermore, Mean Square Errors and Percent Relative Efficiency are calculated for all estimators and subjected to theoretical and empirical comparisons with existing estimation methods. These findings corroborate the advantageous attributes of the proposed estimator in real-world scenarios, reinforcing its practicality and reliability in various research domains. Beyond methodological developments, this study also delves into the real-world implications and applications of the proposed estimators. It highlights the potential benefits of utilizing these estimators in situations where study variables exhibit intricate relationships with auxiliary variables, offering valuable insights into multifaceted data sets and multidimensional factors. Additionally, a comprehensive sensitivity analysis is undertaken to assess the robustness of the proposed estimators under varying assumptions and sampling schemes. The researchers' meticulous evaluation enhances the credibility of the proposed estimators and ensures their adaptability across diverse practical scenarios. Overall, this study's significance extends beyond statistical theory, presenting valuable practical implications for researchers and practitioners across different fields. Improved population variance estimation leads to enhanced decision-making, optimal resource allocation, and deeper insights into underlying phenomena. By introducing the proposed estimator and thoroughly examining its performance through rigorous theoretical and empirical analyses, this research lays a solid foundation for more robust and efficient variance estimation techniques. The insights gained from this study can reshape statistical practices, paving the way for advancements in diverse scientific disciplines and inspiring further knowledge exploration.

Keywords: Exponential estimator, Auxiliary variable, Mean Square Error and Percent Relative Efficiency

I. Introduction

In recent years, the use of sample surveys has gained popularity due to the practicality of overcoming logistical challenges associated with conducting comprehensive census surveys. This trend has led to the widespread adoption of estimators like the ratio, product, and regression estimators for efficiently estimating population parameters, particularly the mean of the variable of interest. These estimators capitalize on the inherent correlation between the study variable and auxiliary variables, either during the survey design or at the estimation stage, to yield accurate results while optimizing resources. The central focus of this research is to develop a novel modified exponential ratio estimator for the population mean. This estimator aims to address potential limitations of existing estimators and enhance the precision of estimates, as evaluated through mean squared error comparisons. By exploring alternative approaches and incorporating adjustments, the researchers anticipate achieving more reliable and efficient estimates of the population mean.

Over the years, several scholars have made significant contributions to the field of survey estimation. Various authors have made numerous work for the estimation of population variance from time to time including [14],[9] , [13], [8] [1], [5], [11],[12],[15] and [10] have made important studies on this topic in the literature. Notably, [17] made pioneering strides by explicitly utilizing auxiliary information for estimation purposes, laying the foundation for the ratio estimator. Subsequently, [18] further advanced this concept by employing auxiliary information to refine estimations.

When dealing with scenarios where the coefficient of correlation is negative between the study variable and auxiliary variables, [19] introduced the product-type estimator, which has proven to be valuable in specific contexts. Additionally, [20] proposed an innovative approach by combining multiple ratio estimators based on individual auxiliary variables positively correlated with the study variable. This technique allowed for greater accuracy in estimation. The product estimator was formalized by [21], providing a well-defined framework for its application. Furthermore, [22] delved into the complexities of ratio estimators involving two or more correlated variables, shedding light on new possibilities for refining estimation methods. The exponential type estimators of population mean were thoroughly investigated by [23] using auxiliary data, resulting in a comprehensive analysis of their performance and potential improvements. [24] took a unique approach by incorporating transformed auxiliary variables, which led to promising results in estimating the mean of the study character. The literature offers an array of other contributions in this area, including the works of [25], [26], [27], and [28], who introduced their respective estimators and demonstrated their efficacy in diverse sampling scenarios. Moreover, [29] and [30] took on the challenge of developing superior exponential type estimators by considering information from two altered auxiliary variables, further expanding the range of available estimation techniques. To gain a more comprehensive understanding of this topic, interested readers can refer to [31], which offers an in-depth exploration of various aspects of survey estimation. In recent times, [32], [33], and [34] have made notable contributions to this area of study, introducing novel ideas and methodologies that hold promise for advancing the field of survey estimation even further.

In conclusion, this research endeavors to create a Generalized Ratio-cum-product estimator of population variance that builds upon the knowledge and advancements made by previous scholars. By harnessing the power of auxiliary information and exploring innovative avenues, the researchers aim to provide an enhanced and efficient approach to estimating the population mean and contributing to the growing body of knowledge in survey estimation techniques.

.

Consider a population of size N. Let y and x be the variable of interest and the Auxiliary variables respectively.

The usual unbiased variance estimator is,

$$t_0 = s_y^2, \tag{1}$$

With variance

$$V(t_0) = \gamma S_y^4 [\lambda_{40} - 1], \tag{2}$$

where, $S_y{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$ and $S_x{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$ are the population variances of *y* and *x* respectively.

[6] proposed the ratio type estimator of S_y^2

$$t_1 = s_y{}^2 \frac{S_x{}^2}{s_x{}^2}, \tag{3}$$

and the MSE is

$$MSE(t_1) = \gamma S_y{}^4 [(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)], \tag{4}$$

where,

$\lambda_{22} = \frac{\mu_{11}}{\sqrt{\mu_{02}\,\mu_{20}}}$ = covariance between $S_y{}^2$ and $S_x{}^2$.

$\lambda_{40} = \beta_2(y) = \frac{\mu_{40}}{\mu_{20}{}^2}$ = kurtosis for population of Y.

$\lambda_{04} = \beta_2(x) = \frac{\mu_{04}}{\mu_{02}{}^2}$ = kurtosis for population of X.

$\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^r (x_i - \bar{X})^s$ and $\gamma = \frac{1}{n}$.

The exponential ratio type and product type estimators for S_y^2

$$t_2 = s_y^2 \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right), \tag{5}$$

$$t_3 = s_y^2 \exp\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right), \tag{6}$$

And the MSE is

$$MSE(t_2) = \gamma S_y{}^4 \left[(\lambda_{40} - 1) + \frac{(\lambda_{04} - 1)}{4} - (\lambda_{22} - 1)\right], \tag{7}$$

$$MSE(t_3) = \gamma S_y{}^4 \left[(\lambda_{40} - 1) + \frac{(\lambda_{04} - 1)}{4} + (\lambda_{22} - 1)\right]. \tag{8}$$

II. Proposed Estimator

We propose the generalized exponential ratio-cum product estimator following [16],

$$t^* = s_y^2 \exp\left[\frac{\left(\frac{S_x^2}{s_x^2}\right)^{\alpha} - 1}{\left(\frac{S_x^2}{s_x^2}\right)^{\alpha} + 1}\right], \tag{9}$$

where α is a suitable constant to minimize the $MSE(t^*)$.

Special cases:

i. If $\alpha = 1$, we get exponential ratio type estimator.

ii. If $\alpha = -1$, we get exponential product type estimator.

To get the bias and MSE

$s_y^{\,2} = S_y^{\,2}(1+e_0),\ s_x^{\,2} = S_x^{\,2}(1+e_1),$

such that

$$E(e_0) = E(e_1) = 0,$$

and

$E(e_0^{\,2}) = \gamma(\beta_2(y)-1) = \gamma(\lambda_{40}-1),$

$E(e_1^{\,2}) = \gamma(\beta_2(x)-1) = \gamma(\lambda_{04}-1),$

$E(e_0 e_1) = \gamma(\lambda_{22}-1).$

Expressing (9) in $e's$ we have

$$t^* = S_y^2(1+e_0)\exp\left[\frac{\left(\frac{S_x^2}{S_x^2(1+e_1)}\right)^{\alpha}-1}{\left(\frac{S_x^2}{S_x^2(1+e_1)}\right)^{\alpha}+1}\right]$$

After simplification we have,

$$t^* = S_y^2(1+e_0)\exp\left[-\frac{\alpha e_1}{2}\left(1-\frac{\alpha e_1}{2}\right)^{-1}\right]$$

After solving we get the MSE given as

$$MSE(t_2) = \gamma S_y^{\,4}\left[(\lambda_{40}-1)+\frac{\alpha^2(\lambda_{04}-1)}{4}-\alpha(\lambda_{22}-1)\right] \tag{10}$$

To get minimum $MSE(t^*)$ differentiating above equation with respect to α,

$$\alpha_{opt} = 2\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)}. \tag{11}$$

Using the value of α_{opt}, we get the $MSE_{\min}(t^*)$ given as

$$MSE_{\min}(t^*) = \gamma S_y^4\left[(\lambda_{04}-1)-\frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)}\right] \tag{12}$$

Which coincides with the MSE of the regression estimator.

III. Efficiency Comparison

$$MSE(t_0) - MSE_{\min}(t^*) = \frac{S_y^4}{n}\left[\frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)}\right] \geq 0$$

$$MSE(t_1) - MSE_{\min}(t^*) = \frac{S_y^4}{n}\left[\frac{(\lambda_{04}-\lambda_{22})^2}{(\lambda_{04}-1)}\right] \geq 0$$

$$MSE(t_2) - MSE_{\min}(t^*) = \frac{S_y^4}{n}\left[\frac{(\lambda_{04}-2\lambda_{22}+1)^2}{(\lambda_{04}-1)}\right] \geq 0$$

All the above conditions are always satisfied.

IV. Numerical Illustration

For empirical study three data sets are given as:
The Population I has been taken from [7] , Population II from [3] and Population III from [4].

Table 1: *Descriptive Statistics of the populations*

Population I	Population II	Population III
$N=106$	$N=100$	$N=278$
$n=20$	$n=10$	$n=30$
$\rho=0.82$	$\rho=0.6500$	$\rho=0.7213$
$S_y=64.25$	$S_y=14.6595$	$S_y=56.4571$
$S_x=491.89$	$S_x=7.5323$	$S_x=40.6747$
$\beta_{2(x)}=25.71$	$\beta_{2(x)}=2.2387$	$\beta_{2(x)}=38.8898$
$\beta_{2(y)}=80.13$	$\beta_{2(y)}=2.3523$	$\beta_{2(y)}=25.8969$
$\lambda_{22}=33.30$	$\lambda_{22}=1.5432$	$\lambda_{22}=26.8142$

Table 2: *Results obtained for different estimators using above populations*

Population	Estimators	MSE	PRE
Population I	t_0	67,423,045.30	100
	t_1	33,431,595.28	201.66
	t_2	45164775.33	149.28
	t^*	314489.80	214.28
Population II	t_0	6245.40	100
	t_1	6948.36	89.87
	t_2	5166.55	120.88
	t^*	5145.31	121
Population III	t_0	8431372.53	100
	t_1	82660.52	102
	t_2	68547.74	123
	t^*	35575.41	237

IV. Conclusion

In the realm of survey sampling, a novel approach called the Generalized Exponential Ratio-Cum-Product Estimator (GERPE) for sampling variance in Simple Random Sampling (SRS) has been introduced. This method offers a more robust and efficient means of estimating sampling variance as compared to existing techniques discussed in the literature. The crux of GERPE lies in its ability to strike a balance between ratio and product estimators, leading to enhanced precision in variance estimation. Remarkably, when GERPE is operating at its optimal settings, its Mean Squared Error (MSE) attains the same value as that of the conventional regression estimator. This is a significant advantage because the usual regression estimator has been proven to exhibit lower MSE compared to other existing estimators. Consequently, the application of GERPE not only outperforms the methods previously documented but also aligns itself with the performance level of the superior regression estimator, thus promising more accurate and reliable variance estimates in SRS

scenarios. This innovative approach opens up new avenues for improving the precision and efficiency of sampling variance estimation in the field of survey sampling and can be a valuable addition to the statistical toolbox of researchers and practitioners alike.

Acknowledgement: We are thankful to Editor-in-chief for his suggestions that improved the manuscript.

References

[1] Agrawal, M. C., Sthapit, A. B. (1995). Unbiased ratio-type variance estimation. *Stat. Prob. Lett*., 25:361–364.

[2] Bahl, S. & Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Journal of Information and Optimization Sciences*, 12(1): 159 163.

[3] Cochran, W. G. Sampling techniques (3rd ed.). Wiley Eastern Limited,1977.

[4] Das, A. K. (1988). Contribution to the theory of sampling strategies based on auxiliary information. Ph.D. thesis submitted to BCKV, Mohanpur, Nadia, West-Bengal, India.

[5] Gupta, S., Shabbir, J. (2008). Variance estimation in simple random sampling using auxiliary information. *Hacett. J. Math. Stat*., 371:57–67.

[6] Isaki, C.T. (1983). Variance estimation using auxiliary information. J. *Am. Stat. Assoc*., 78381:117–123.

[7] Kadilar, C., Cingi, H. (2007). Improvement in variance estimation in simple random sampling. *Commun. Stat. Theo. Meth*., 36:2075–2081.

[8] Prasad, B., Singh, H. P. (1990). Some improved ratio-type estimators of finite population variance in sample surveys. *Commun. Stat. Theo. Meth*., 193:1127–1139.

[9] Singh, H.P. (1986). A note on the estimation of variance of sample mean using the knowledge of coefficients of variation in normal population. *Commun. Stat. Theo. Meth.,* 1512:3737–3746.

[10] Singh, H. P., Pal, S. K. (2016). An efficient class of estimators of finite population variance using quartiles. *J. Appl. Stat*., 10.1080/02664763.2015.1125865

[11] Singh, H. P., Solanki, R. S. (2013a). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical Papers,* 542:479–497.

[12] Singh, H. P., Solanki, R. S. (2013b). Improved estimation of finite population variance using auxiliary information. *Commun. Stat. Theo. Meth*., 215:2718–2730.

[13] Singh, H. P., Upadhyaya, L. N., Namjoshi, U. D. (1988). Estimation of finite population variance. *Curr. Sci.,* 5724:1331–1334.

[14] Srivastava, S. K., Jhajj, H. S. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhya, C*, 421(2):87–96.

[15] Upadhyaya, L. N., Singh, H. P. (1986). On a dual to ratio estimator for estimating finite population variance. *Nepal Math. Sci. Rep*., 111:37–42.

[16] Vishwakarma, G. K., Zeeshan, S. M. and Bouza, C. N. (2017). Ratio and product type exponential estimators for population mean using ranked set sampling. *Revista Investigacion Operacional*, 38(3): 266-271.

[18] J. Neyman (1934). On two different aspects of the representative methods, the method of stratified sampling and the method of purposive selection, *Journal of. Royal Statistical. Society*, 97: 558-606.

[19] Cochran, W. G. Ratio Estimators, In Sampling Techniques, 3rd ed., India: Wiley Eastern Limited.1940.

[20] Murthy, M. N. (1964). Product Method of Estimation, *Sankhya*., 26(A): 69–74.

[21] Olkin , I. (1958). Multivariate ratio estimation for finite populations, *Biometrika*., 45(1-2), 154–165.

[22] Singh, M. P. (1967) . Ratio-cum-product method of estimation, *Metrika*., 12: 34–72, .

[23] Chand, L. (1975). Some ratio type estimators based on two or more auxiliary variables,

Unpublished Ph.D. dissertation, Iowa State University, Iowa.

[24] Bahl, S. and Tuteja, R. K. (1991). Ratio and product type exponential estimator, *Journal of Information and Optimization Sciences*., 12: 159–163.

[25] Upadhyaya, L. N. and Singh, H. P. (1999). Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean, *Biometrical Journal*., 41: 627–636.

[26] Abu-Dayeh, W. A., Ahmed, M. S. , Ahmed, R. A. and Muttlak, H. A. (2003). Some Estimators of a Finite Population Mean Using Auxiliary Information, *Applied Mathematics and Computation*., 139: 287–298.

[27] Kadilar, C. and Cingi, H. (2005). A new estimator using two auxiliary variables, *Applied Mathematics and Computation*., 162: 901– 908.

[28] Diana, G. and Perri, P. F. (2007). Estimation of finite population mean using multi-auxiliary information, *Metron*., 65(1): 99–112.

[29] Shabbir, J. and Gupta, S. (2011). On estimating finite population mean in simple and stratified random sampling, *Communications in Statistics- Theory & Methods*., 40: 199–212.

[30] Yasmeen, U. , Noor ul Amin, M. and Hanif , M. (2016). Exponential ratio and product type estimators of finite population mean, *Journal of Statistics and Management Systems*., 19: 55–71.

[31] Singh, S. Advanced Sampling Theory with Applications-How Michael 'selected' Amy, Vol-II. London: Kluwer Academic Publishers. 2003

[32] Gupta, R. K. and Yadav, S. K. (2017). New Efficient Estimators of Population Mean Using Non-Traditional Measures of Dispersion, *Open Journal of Statistics*., 7: 394-404.

[33] Muili, J. O. , Agwamba, E. N. , Erinola, Y. A. , Yunusa, M. A . , Audu, A. and Hamza, M. A. (2020). Modified Ratio-Cum-Product Estimators of Population Mean Using Two Auxiliary Variables, *Asian Journal of Research in Computer Science*., 6(1): 55-65.

[34] Singh, H. P. and Yadav, A. (2020). A New Exponential Approach for Reducing the Mean Squared Errors of the Estimators of Population Mean Using Conventional and Non-Conventional Location Parameters, *Journal of Modern Applied Statistical Methods*., 18(1).

RT&A, No 4 (76)
Volume 18, December 2023

EXPONENTIAL-PARETO MIXTURE DISTRIBUTION

Irina Peshkova

•

Petrozavodsk State university,
Institute of Applied Mathematical Research of the Karelian Research Centre of RAS
iaminova@petrsu.ru

Abstract

In this paper we introduce the Exponential-Pareto mixture distribution. This distribution is associated as mixture of light and heavy-tailed data which arise in a wide class applications including risk analysis. Characteristic function, failure rate function, mean excess, conditional excess distribution are derived. It is proved that the limiting distribution of maxima among n values of rv's with Exponential-Pareto distribution has Frechet-type form. The maximal likelihood estimation of parameters is discussed. The upper bound of uniform distance between Exponential-Pareto mixture and Pareto distributions is derived.

Keywords: finite mixture, Exponential-Pareto distribution

1. Introduction

Finite mixture models are often used in the modelling time to failure of the systems in the competing risk situations. The exponential, Weibull, lognormal and Pareto distributions occupy a central role because of their demonstrated usefulness in the analysis of lifetime data and in problems related to the modelling of ageing or failure processes. Finite mixtures are also useful in medical and biology research, in artificial neural networks and robustness studies, income analysis [1].

In lifetime data analysis, the population of lifetimes usually can be decomposed into sub-populations. Moreover, The data may be actually generated by quite different distributions, in particular, distributions with the so-called *light or heavy tails*. For example, an insurance portfolio may include both many small (light) claims and also a few large (heavy) claims. Hence the claims distribution can be modeled as a finite mixture distributions with different tail behavior. [2] In this paper we propose to consider a mixture of exponential distribution and the Pareto distribution to model such situations since the Pareto distribution relates to the class of long-tailed distributions, while the exponential distribution is light-tailed.

The exponential distribution appeared suitable for modelling the lifetimes of various types of manufactured items. The Pareto model is applied in many fields, for example, in income analysis, in signal processing for simulation of X-band maritime surveillance radar clutter [3].

The mixtures of light and heavy-tailed distribution can be useful in the simulation of IoT traffic, because, as it is known, the smart-home and smart-city environments can generate both short and large-sized packets. These environments involve several sensors dedicated to specific tasks, such as monitoring systems or collecting cyber-physical values (temperature, humidity, etc.). Smart cameras generate continuous data flows with large-sized packets, while smart plugs generate small-sized packets at a slow pace [4].

Various issues of the Exponential-Pareto mixture were considered by authors in the papers [5, 6, 7] in relation to queueing systems.

The main contribution of the paper is to introduce the Exponential-Pareto mixture distribution. It is the two-component mixture of exponential and Pareto distributions, which is used to approximate mixed data with light and heavy tails.

The paper is organized as follows. The moments and characterization function are derived in Section 2. The subexponentiality of the distribution is proved in Section 2. In Section 4 we discuss the properties of the conditional excess distribution. In Section 5 the extreme behavior of the Exponential-Pareto mixture distribution is discussed. In Section 6 We apply the known upper bound for Kullback-Leibler divergence between exponential and Pareto distributions [3, 8] to derive the convergence of uniform distance between Exponential-Pareto mixture and exponential distributions to zero as shape parameter goes to infinity. The estimation of parameters by log-likelihood maximization is discussed in Section 7..

2. Definition and moments of Exponential-Pareto mixture distribution

Definition 1. We say that r.v. Z has an *Exponential-Pareto mixture distribution* if its distribution function (df) has the following form:

$$F_Z(x) = 1 - pe^{-\lambda x} - (1-p)\left(\frac{x_0}{x_0+x}\right)^{\alpha},\ \lambda > 0, \alpha > 0, x_0 > 0, x \geq 0, \tag{1}$$

where $0 < p < 1$ is *mixing proportion*.

Equation (1) shows that r.v. Z coincides with exponential distribution with the probability p, and with Pareto distribution with the probability $1-p$.The above can be reformulated as follows. Suppose that the random variables X, Y with distribution functions F_X, F_Y, respectively, are independent, and let I be indicator function independent of X, Y, taking value 1 with probability p (value 0 with probability $1-p$). Then it is said that the variable

$$Z = I\,X + (1-I)\,Y \tag{2}$$

has *two-component mixture distribution*. If X has exponential distribution $Exp(\lambda)$ and Y has Pareto distribution $Pareto(\alpha, x_0)$, then df of Z is given by (1).

The density function of Exponential-Pareto mixture distribution has form

$$f_Z(x) = p\lambda e^{-\lambda x} + (1-p)\frac{\alpha x_o^{\alpha}}{(x_0+x)^{\alpha+1}}.$$

and has the following limits

$$\begin{aligned}\lim_{x\to\infty} f_Z(x) &= 0;\\ \lim_{x\to 0} f_Z(x) &= p\lambda + \frac{(1-p)\alpha}{x_0}.\end{aligned}$$

Figure 1 depicts cdf of Exponential-Pareto mixture distribution with parameters $x_0 = 0.5, \lambda = 2, p = 0.5$ and varied $\alpha = 0.5; 1.5; 2.5$.

Lemma 1. Let rv Z follow the Exponential-Pareto mixture distribution, then its characteristic function is given by

$$\phi_Z(t) = \frac{\lambda\alpha}{\lambda - ipt} e^{-i(1-p)tx_0} \sum_{k=0}^{\infty} \frac{(i(1-p)tx_0)^k}{k!(\alpha-k)} \tag{3}$$

Proof. Substitute df (1) into the formula for characteristic function $\phi_X(t)$ of rv Z

$$\phi_Z(t) = e^{it(pX+(1-p)Y)} = e^{pitX} e(1-p)itY = \phi_X(pt)\phi_Y((1-p)t).$$

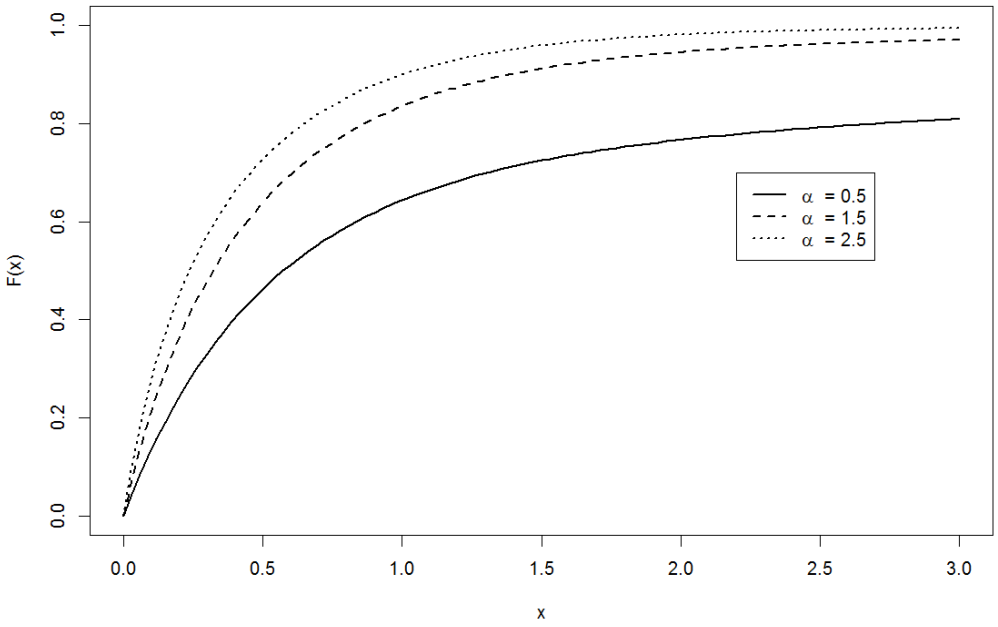


Figure 1: *Cumulative distribution function (cdf) of Exponential-Pareto distribution.*

where the characteristic function of first component X is given by

$$\phi_X(t) = \int_0^\infty e^{itx}\lambda e^{-\lambda x}dx = \frac{\lambda}{\lambda - it}.$$

The characteristic function of Pareto distribution can be calculated using the expansion of the exponential function in a Taylor series in the vicinity of zero

$$\begin{aligned}
\phi_Y(t) &= \int_0^\infty e^{itx}\alpha x_0^\alpha (x_0+x)^{-\alpha-1}dx \\
&= e^{-itx_0}\int_{x_0}^\infty e^{itx}x_0^\alpha \alpha x^{-\alpha-1}dx \\
&= e^{-itx_0}x_0^\alpha\alpha\int_{x_0}^\infty \sum_{k=0}^\infty \frac{(itx)^k}{k!}x^{-\alpha-1}dx \\
&= e^{-itx_0}x_0^\alpha\alpha\sum_{k=0}^\infty \frac{(it)^k}{k!}\int_{x_0}^\infty x^{k-\alpha-1}dx \\
&= e^{-itx_0}\alpha\sum_{k=0}^\infty \frac{(itx_0)^k}{k!(\alpha-k)}.
\end{aligned}$$

Hence, the characteristic function of Exponential-Pareto mixture distribution takes form (3). ■

Lemma 2. Let rv Z follow the Exponential-Pareto mixture distribution, then its moments are given by

$$\mathbf{E}Z^k = \Gamma(1+k)\left(\frac{p}{\lambda}\right)^k + \frac{\Gamma(1+k)x_0^k(1-p)^k}{\Gamma(\alpha)}\sum_{j=0}^{k-1}\left(\frac{p}{x_0\lambda(1-p)}\right)^j\Gamma(\alpha-k+j),\ \alpha > k,$$

where $\Gamma(z) = \int\limits_0^\infty t^{z-1}e^{-t}dt$ is Gamma function.

The proof obviously follows from the well-known relation between moments and characterization function

$$i^{-k}\frac{d^k\phi_Z(t)}{dt^k}|_{t=0} = \mathbf{E}Z^k,$$

In particular, the mean is

$$\mathbf{E}Z = \frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha-1},\ \alpha > 1,$$

the second moment is

$$\mathbf{E}Z^2 = \frac{2p^2}{\lambda^2} + \frac{2p(1-p)x_0}{\lambda(\alpha-1)} + \frac{2(1-p)^2x_0^2}{(\alpha-1)(\alpha-2)},\ \alpha > 2,$$

and variance is

$$\mathbf{Var}Z = \frac{p^2}{\lambda^2} + \frac{(1-p)^2x_0^2\alpha}{(\alpha-1)^2(\alpha-2)},\ \alpha > 2.$$

Note that moments of order k exist only for $\alpha > k$.

3. FAILURE RATE FUNCTION AND SUBEXPONENTIALITY

Consider the equilibrium df F_e of rv Z with df (1)

$$F_e(x) \quad = \quad \frac{1}{\mathbf{E}Z}\int\limits_0^x \overline{F}_Z(t)dt = 1 - \frac{1}{\mathbf{E}Z}\left(\frac{pe^{-\lambda x}}{\lambda} + \frac{(1-p)x_0^\alpha}{(\alpha-1)(x_0+x)^{\alpha-1}}\right), \tag{4}$$

where $\overline{F}(x) = 1 - F(x)$ is the tail of df F. We emphasize that expression (4) exists if parameter $\alpha > 1$.

We calculate the failure rate function of the Exponential-Pareto mixture distribution

$$r(x) := \frac{f_Z(x)}{\overline{F}_Z(x)} = \frac{p\,\lambda a(x) + (1-p)\,\alpha/(x_0+x)}{p\,a(x) + (1-p)}, \tag{5}$$

where

$$a(x) = e^{-\lambda x}\left(1 + \frac{x}{x_0}\right)^\alpha. \tag{6}$$

is an auxiliary function.

Note that

$$r(x) \longrightarrow 0 \quad \text{as} \quad x \to \infty,$$

which is typical for long-tailed distributions. Recall that df F_Z is *long-tailed*, if for each fixed $x > 0$ the following relation holds

$$\lim_{u\to\infty} \mathbf{P}(X > u + x | X > u) = 1.$$

This asymptotic property means, that for each fixed $x > 0$, the random variable Z exceeds the threshold $x + u$ with probability approaching 1 as u increasing. Long-tailed distributions have asymptotically decaying to zero failure rate functions (not necessary monotone) [9]. Below we prove that Exponential-Pareto distribution belongs to a subclass of long-tailed distribution – so-called *subexponential* distributions.

Note, that the failure rate function has the same limit as density function as $x \to 0$,

$$r_Z(x) \longrightarrow p\lambda + (1-p)\frac{\alpha}{x_0} \quad \text{as} \quad x \to 0.$$

Provided that the following relation between the parameters of the mixture is satisfied $\lambda \geq \alpha/x_0$, failure rate function is bounded from above, $r_Z(x) \leq \lambda$.

To verify the monotonicity of the failure rate function we calculate its derivative

$$\frac{dr_Z(x)}{dx} = -(1-p)\frac{pa(x)(r_X(x)-r_Y(x))^2 + r_Y^2(x)/\alpha(pa(x)+(1-p))}{(pa(x)+(1-p))^2} < 0,$$

Since $\frac{dr_Z(x)}{dx}$ is negative for all x, then $r_Z(x)$ decreases monotonically and, therefore, the Exponential-Pareto distribution belongs to the class of distributions with decreasing failure rate functions.

The following lemma states that the distribution (1) and the equilibrium distribution (4) are both belong to the class of so-called subexponential distributions. Recall that df F_Z is called *subexponential* [9] if

$$\lim_{x\to\infty} \frac{\overline{F_Z^{*n}}(x)}{n\overline{F}_Z(x)} = 1 \text{ for all } n \geq 2,$$

where $\overline{F_Z^{*n}}(x)$ is the tail of n-convolution of the distribution $F_Z(x)$ with itself.

To verify that the Exponential-Pareto mixture distribution F_Z and the corresponding equilibrium distribution F_e (with parameter $\alpha > 1$) both belong to the class of the subexponential distributions, it is enough [9] to verify that F_Z belongs to a special subclass $\mathcal{S}^*$ of the subexponential distributions.

Lemma 3. The Exponential-Pareto mixture distribution with df defined by expression (1) (with parameter $\alpha > 1$) belongs to a special subclass $\mathcal{S}^*$ of the subexponential distributions.

Proof. One of the following criteria for df to belong $\mathcal{S}^*$ can be applied [9].

1. If
$$\limsup_{x\to\infty} xr(x) < \infty, \tag{7}$$
then $F_Z \in \mathcal{S}^*$.

2. Suppose that
$$\lim_{x\to\infty} r(x) = 0. \tag{8}$$
Then
$$F_Z \in \mathcal{S}^* \iff \lim_{x\to\infty} \int_0^x e^{yr(x)}\overline{F}_Z(y)dy = \mathbf{E}Z. \tag{9}$$

It is easy to check that for Exponential-Pareto mixture distribution

$$r(x) \longrightarrow 0 \quad \text{as} \quad x \to \infty \quad \text{and} \quad xr(x) \longrightarrow \alpha \quad \text{as} \quad x \to \infty.$$

Moreover,

$$\begin{aligned}\int_0^x e^{yr(x)}\overline{B}(y)dy &= -\frac{pe^{-y(\lambda-r(x))}}{\lambda-r(x)}\Big|_0^x + (1-p)\frac{(x_0r(x))^\alpha}{r(x)}e^{-x_0r(x)}\cdot \\ &\cdot \quad [\gamma(-\alpha+1, x_0r(x)+xr(x)) - \gamma(-\alpha+1, x_0r(x))] \longrightarrow \\ &\longrightarrow \quad \frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha-1} = \mathbf{E}Z \quad \text{as} \quad x\to\infty\end{aligned}$$

since

$$\gamma(-\alpha+1, x_0r(x)+xr(x)) \longrightarrow \gamma(-\alpha,\alpha) \quad \text{as} \quad x\to\infty$$

and

$$\gamma(-\alpha+1, x_0r(x)) \sim \frac{(xr(x))^{-\alpha+1}}{-\alpha+1} \quad \text{as} \quad x\to\infty,$$

where $\gamma(\alpha, x) = \int_0^x y^{\alpha-1}e^{-y}dy$ is the lower incomplete gamma function.

Hence conditions (7)-(9) are satisfied and Exponential-Pareto distribution belongs to subclass $\mathcal{S}^*$. ■

4. Conditional Excess Distribution

In this section we consider conditional excess distribution of Exponential-Pareto mixture distribution.

On the event $\{Z > u\}$, define the excess $Z_u := Z - u$, which has the *conditional excess distribution over the threshold* u [10]

$$P(Z_u \le x) = P(Z - u \le x | Z > u),\ u \ge 0,\ \ x \ge 0. \tag{10}$$

Let $X_u,\ Y_u$ be conditional excesses of rv's X and Y with conditional excess distributions defined by (10), relatively. In the case of Exponential-Pareto mixture distribution expression for the tail of conditional excess distribution becomes

$$\overline{F}_{Z_u}(x) = \frac{\overline{F}_Z(x+u)}{\overline{F}_Z(u)} = \frac{pe^{-\lambda(x+u)} + (1-p)\left(\frac{x_0}{x_0+x+u}\right)^{\alpha}}{pe^{-\lambda u} + (1-p)\left(\frac{x_0}{x_0+u}\right)^{\alpha}}.$$

Note that rv Z_u is not a mixture of rv's X_u and Y_u, namely $Z_u \ne IX_u + (1-I)Y_u$.

The expected value of conditional excess is given by *the mean excess function*, defined as [10]

$$\begin{aligned} e_Z(u) = \mathbf{E}Z_u &= \int_0^{\infty} \frac{\overline{F}_Z(x+u)}{\overline{F}_Z(u)}dx = \frac{\int_u^{\infty} \overline{F}_Z(x)dx}{\overline{F}_Z(u)} \\ &= \frac{\int_u^{\infty} (pe^{-\lambda x} + (1-p)x_0^{\alpha}(x_0+x)^{-\alpha})dx}{pe^{-\lambda x} + (1-p)x_0^{\alpha}(x_0+x)^{-\alpha}} \\ &= \frac{pe^{-\lambda u}(\alpha-1)(x_0+u)^{\alpha} + \lambda(1-p)x_0^{\alpha}(x_0+u)}{\lambda p(\alpha-1)(x_0+u)^{\alpha}e^{-\lambda u} + \lambda(1-p)(\alpha-1)x_0^{\alpha}}.\ \ \alpha > 1. \end{aligned}$$

Note that the mean excess function is an increasing function, $e_Z(u) \to \infty$ as $u \to \infty$. The Failure rate function of Z_u is given by

$$r_{Z_u}(x) = \frac{p\, r_{X_u}(x) a(x+u) + (1-p)\, r_{Y_u}(x+u)}{p\, a(x+u) + (1-p)} = r_Z(x+u), \tag{11}$$

where $a(x)$ is defined by expression (6).

Now we derive the condition on the parameters of the Exponential-Pareto mixture distribution that guarantees the mixture to be bounded by its components in terms of the failure rate ordering. To prove it we need the following lemma.

Lemma 4. Let $r_X,\ r_Y$ be the failure rates of F_X, F_Y, respectively. Then if

$$\sup_{x \ge 0} r_X(x) < \infty \ \text{ and } \ \inf_{x \ge 0} r_Y(x) > 0 \tag{12}$$

and

$$\sup_{x \ge 0} r_X(x) \le \inf_{x \ge 0} r_Y(x), \tag{13}$$

then

$$X \underset{r}{\ge} Y,$$

where notation $X \underset{r}{\ge} Y$ means that X is *more than Y in failure rate*, i.e. $r_X(x) \le r_Y(x)$ for all x.

Theorem 1. Let rv Z have Exponential-Pareto mixture distribution with df defined by (1). If the following inequality holds

$$\frac{\alpha}{x_0} \le \lambda, \tag{14}$$

then

$$X \underset{r}{\le} Z \underset{r}{\le} Y, \tag{15}$$

$$X_u \underset{r}{\le} Z_u \underset{r}{\le} Y_u, \tag{16}$$

$$e_X(u) \le e_Z(u) \le e_Y(u) \tag{17}$$

Proof. To prove the theorem, it suffices to find conditions under which the relations

$$r_Y(x) \le r_Z(x) \le r_X(x), \tag{18}$$

are satisfied. From lemma 4 for this it is sufficient that the following relation be satisfied:

$$r_Y(x) \le \sup_{x\ge 0} r_Y(x) = r_Y(0) = \frac{\alpha}{x_0} \le r_X(x) = \lambda. \tag{19}$$

Inequality (19) can be rewrite in form (14) whence it follows (15).

From statement 2.1 in [11] it follows that if rv's are ordered in failure rate, then their conditional excesses are ordered in failure rate too, that proves the statement (16).

Now we calculate the mean excess functions for rv's X and Y. Easy to check, that for exponential distribution mean excess function is equal to mathematical expectation, $e_X(u) = \mathbf{E}X = 1/\lambda$. For Pareto distribution we have

$$e_Y(u) = \frac{x_0 + u}{\alpha - 1}.$$

Obviously, if the condition of the theorem (14) is fulfilled, then inequality

$$\lambda \ge \frac{\alpha}{x_0} \ge \frac{\alpha - 1}{x_0 + u}$$

holds and

$$\frac{1}{\lambda} \le \frac{pe^{-\lambda u}(\alpha - 1)(x_0 + u)^\alpha + \lambda(1-p)x_0^\alpha(x_0 + u)}{\lambda p(\alpha - 1)(x_0 + u)^\alpha e^{-\lambda u} + \lambda(1-p)(\alpha - 1)x_0^\alpha} \le \frac{x_0 + u}{\alpha - 1},$$

that proves the relation (17) of the theorem. ■

Figure 2 demonstrates the failure rate functions of rv's X, Y, Z, where $\lambda = 1, \alpha = 0.5, x_0 = 0.6$. It can be seen from the graph that, for the given parameters, the failure rate functions are ordered in accordance with the relation (18), hence it follows from Theorem 1 that rv's are ordered in accordance with (17).

Figure 3 shows mean excess functions of rv's X, Y, Z with $\lambda = 5, \alpha = 2.1, x_0 = 0.5, p = 0.5$. It can be seen from the graph that, for the given parameters, these functions are ordered as in (17).

5. Extreme behavior of the Exponential-Pareto distribution

Let $\{X_n,\ n \ge 1\}$ be a family of the independent and identically distributed (iid) rv's with a distribution function F. Then the distribution of $M_n = \max(X_1, \dots, X_n)$ satisfies $\mathbf{P}(M_n \le x) = F^n(x)$.

Suppose there exists a sequence of real constants $b_n,\ a_n > 0,\ n \ge 1$ such that

$$\lim_{n\to\infty} \mathbf{P}((M_n - b_n)/a_n \le x) = G(x), \quad n \to \infty, \tag{20}$$

for every continuity point x of G, and G a nondegenerate distribution function. Then $G(x)$ is one of the three types of *extreme value distributions*: Gumbel, Frechet or reversed Weibull [12].

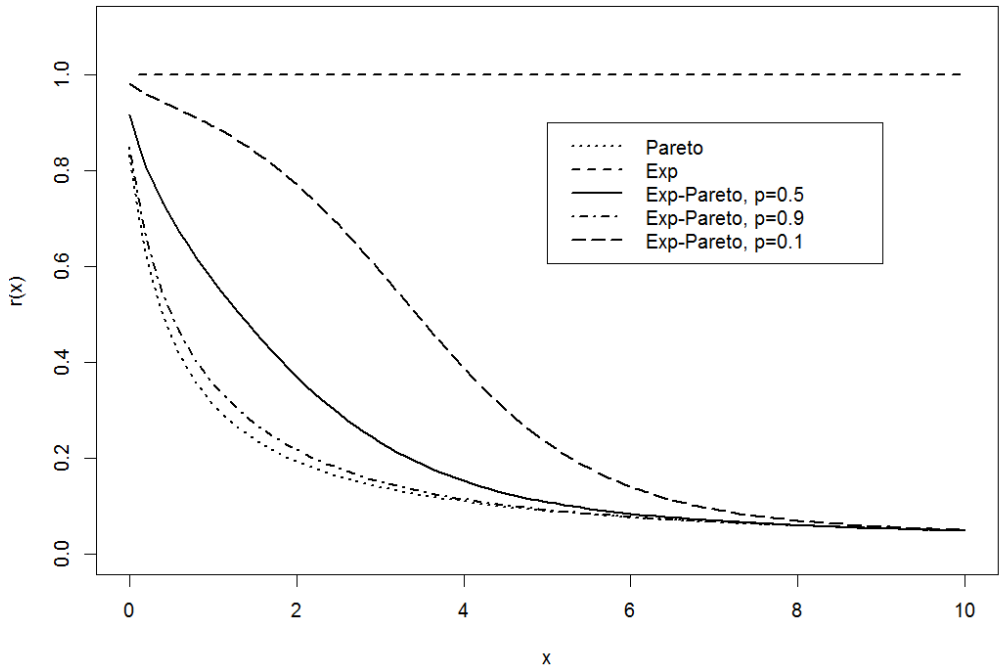


Figure 2: *The failure rate function of distributions* $Exp(1)$, $Pareto(0.5; 0.6)$ *and mixture* $Exp - Pareto(1; 0.5; 0.6; p)$ *with different mixing proportions* p.

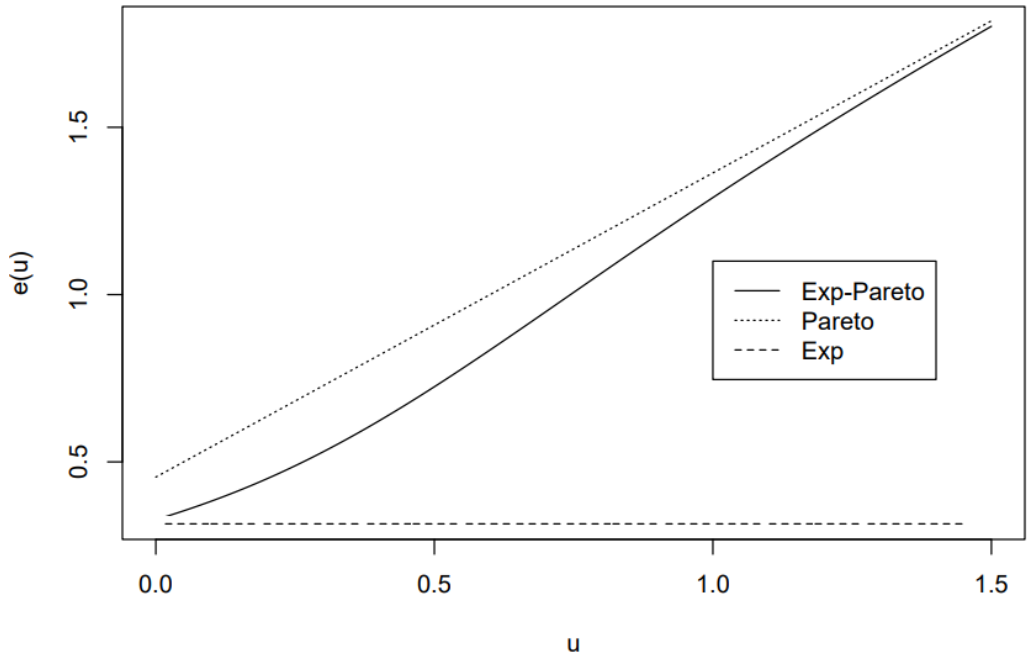


Figure 3: *Mean excess functions of distributions* $Exp(5)$, $Pareto(2.1; 0.5)$ *and mixture* $Exp - Pareto(5; 2.1; 0.5; 0.5)$.

The class of extreme value distributions (which combines all three types) is $G_\eta(cx + d)$ with real

$c > 0,\ d$, where

$$G_\eta(x) = \begin{cases} e^{-(1+\eta x)^{-1/\eta}}, & \eta \neq 0,\ 1+\eta x > 0; \\ e^{-e^{-x}}, & \eta = 0. \end{cases}$$

It is easy to check that the first component of Exponential-Pareto mixture distribution is in the maximum domain of attraction (MDA) of the Gumbel law, while the second is in MDA of the Frechet law.

For given $0 \leq \tau \leq \infty$ and a sequence $\{u_n,\ n \geq 1\}$ of real numbers the following are equivalent [10]

$$n\overline{F}(u_n) \to \tau \text{ as } n \to \infty \tag{21}$$

and

$$\mathbf{P}(M_n \leq u_n) \to e^{-\tau} \text{ as } n \to \infty. \tag{22}$$

If condition (20) is satisfied, then convergence (22) is preserved for any linear normalizing sequence $u_n(x) = a_n x + b_n,\ n \geq 1$ and expression (22) becomes

$$\mathbf{P}(M_n \ \leq u_n(x)) \to \tau(x),$$

where a concrete form of the function $\tau(x)$ depends on the type of the limiting distribution.

Theorem 2. Let the sequence of independent rv's $X_1, \ldots, X_n$ have Exponential-Pareto mixture distribution with df defined by (1). Define $M_n = \max(X_1, \ldots, X_n)$ is maxima among n values of sequence. Then $(M_n - b_n)/a_n \in MDA(\Phi_\alpha)$, where

$$a_n = x_0 n^{1/\alpha},\ b_n = -x_0. \tag{23}$$

Proof. First we find $\tau(x)$ substituting $u_n(x) = x_0 n^{1/\alpha}\, x - x_0$ into the relation (21):

$$\begin{aligned} n\overline{F}(u_n(x)) &= n\left[pe^{-\lambda u_n(x)} + \frac{(1-p)x_0^\alpha}{(x_0 + u_n(x))^\alpha}\right] \\ &= n\left[pe^{-\lambda u_n(x)} + \frac{(1-p)x_0^\alpha}{(x_0\, n^{1/\alpha}\, x))^\alpha}\right] \\ &= \frac{n(1-p)x^{-\alpha}}{n}\left[\frac{p\, n\, e^{-\lambda u_n(x)} x^\alpha}{1-p} + 1\right] \\ &\longrightarrow (1-p)x^{-\alpha} \end{aligned}$$

since $u_n(x) \to \infty$ and $n\, e^{-\lambda u_n(x)} \to 0$ as $n \to \infty$. In accordance with (22) we get the following asymptotic distribution:

$$P(M_n \leq u_n(x)) \to e^{-(1-p)x^{-\alpha}} \text{ as } n \to \infty, \tag{24}$$

that is Frechet distribution. ■

6. Uniform distance and Kullback-Leibler divergence

The uniform distance between two distributions F_X and F_Y [13],

$$\Delta(F_X, F_Y) = \sup_x |F_X(x) - F_Y(x)|, \tag{25}$$

is a widely used probability metric in sensitivity analysis [14].

It follows from the Pinsker-Csiszar Inequality [3] that uniform distance is bounded by Kullback-Leibler divergence, namely

$$\Delta(F_X, F_Y) \leq \sqrt{2D_{KL}(X||Y)}, \tag{26}$$

where

$$D_{KL}(X||Y) = \int_0^\infty f_X(x) \log\left(\frac{f_X(x)}{f_Y(x)}\right) dx \tag{27}$$

is Kullback-Leibler divergence.

It is shown in [3] that the minimum divergence between Exponential and Pareto distribution is reached at $\lambda = \frac{\alpha-1}{x_0}$ and

$$D_{KL}(Y||X)_{min} \leq \frac{1}{\alpha(\alpha-1).} \tag{28}$$

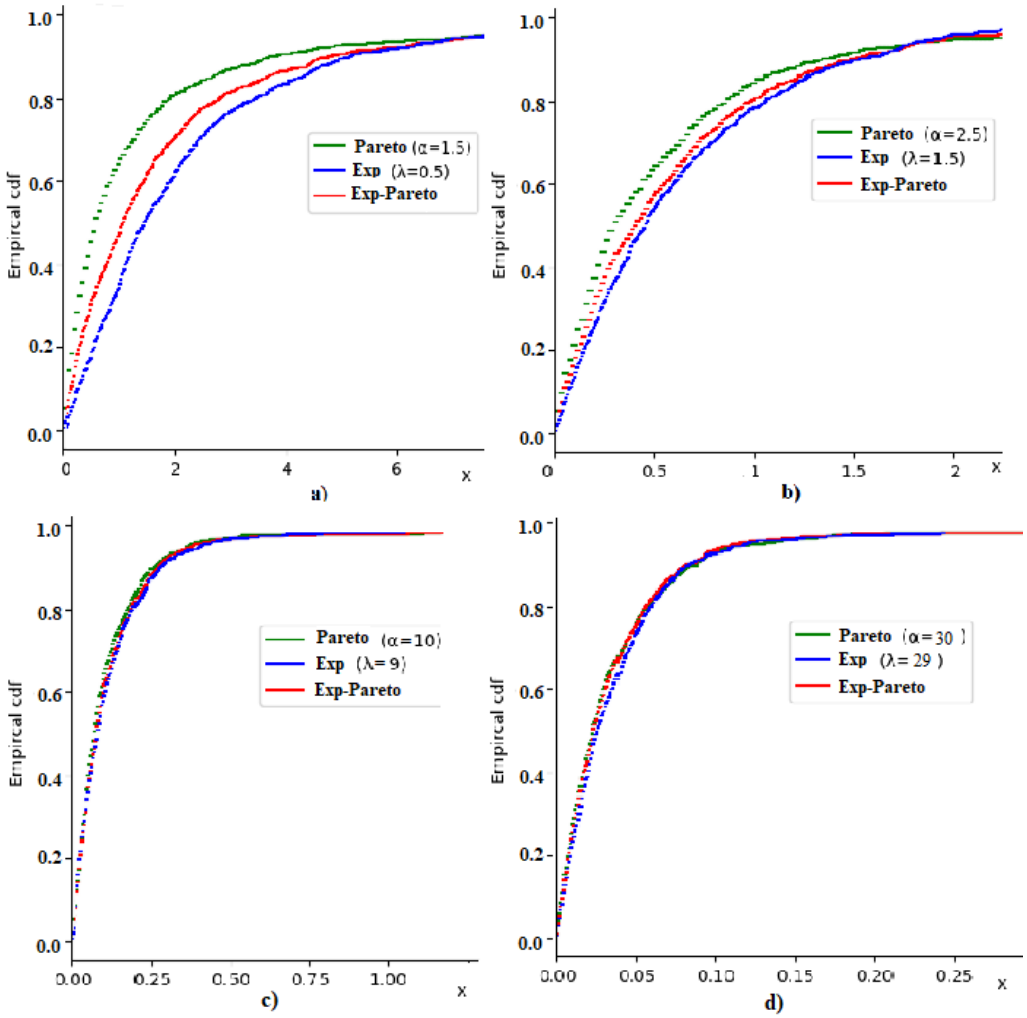


Figure 4: *Cumulative distribution functions of Exponential, Pareto and Exponential-Pareto distributions with different* α, λ, *and* $x_0 = 1, p = 0.5$

Clearly if the Kullback-Leibler divergence is close to zero, the uniform distance inherits this and thus implies that Exponential and Pareto distribution are close. We apply this to estimate the uniform distance between the Exponential-Pareto mixture and the Pareto distributions (for case $\alpha > 1$).

$$\begin{aligned}\Delta(F_Z, F_Y) &= \sup_x |pF_X(x) + (1-p)F_Y(x) - F_Y(x)| = \sup_x |pF_X(x) - pF_Y(x)| \\ &= p\Delta(F_X, F_Y) \le p\sqrt{2D_{KL}(Y||X)} \le \frac{\sqrt{2}p}{\sqrt{\alpha(\alpha-1)}}. \end{aligned} \tag{29}$$

The last inequality demonstrates the convergence rate of $\Delta(F_Z, F_Y)$ to zero as $\alpha \to \infty$. We will get the same effect by approximating Pareto distribution via Exponential distribution as $\alpha \to \infty$ [3]. Let

$$\lambda = \frac{\alpha + o_\alpha(1)}{x_0}, \tag{30}$$

where $o_\alpha(1) \to 0$ as $\alpha \to \infty$, then

$$\overline{F}_Y(x) = \lim_{\alpha\to\infty} \left(1 + \frac{\lambda x}{\alpha(1 + o_\alpha(1))}\right)^{-\alpha} = e^{-\lambda x} = \overline{F}_X(x). \tag{31}$$

The discussions above allow us to formulate the following lemma about approximation mixture by Exponential distribution for large α.

Lemma 5. Let df $F_Z(x)$ has form (1), then

$$F_Z(x) \to F_X(x) \text{ as } \alpha \to \infty, \text{ for all } x \ge 0,$$

where $F_X(x) = 1 - e^{-\lambda x},\ x \ge 0$.

Figure 4 demonstrates results of Exponential, Pareto and Exponential-Pareto mixture distributions with $x_0 = 1, p = 0.5$, a) $\alpha = 1.5, \lambda = 2.5$, b) $\alpha = 2.5, \lambda = 3.5$, c) $\alpha = 10, \lambda = 11$, d) $\alpha = 30, \lambda = 31$ for $n = 1000$ sample size. As expected, Kolmogorov-Smirnov test confirms that data sets from Exponential-Pareto mixture and Pareto distributions are homogeneous only for the case d) $\alpha = 30$. The uniform distance $\Delta(F_Z, F_Y)$ is $0.79, 0.34, 0.074, 0.012$, for cases a)-d), relatively. The upper bound for uniform distance $\frac{\sqrt{2}p}{\sqrt{\alpha(\alpha-1)}}$ defined by relation (29) is $0.81, 0.36, 0.074, 0.023$, relatively.

7. Parameters estimation

In this section we discuss the estimating the parameters of Exponential-Pareto mixture distribution by the method of moments and via maximization of log-likelihood function.

The method of moments gives the following estimates of the parameters α and λ, expressed in terms of the parameter x_0:

$$\begin{aligned}\alpha &= \frac{2(1-p)^2x_0^2 - 2(1-p)x_0\overline{Z}}{S_Z^2 - (\overline{Z})^2} + 1; \\ \lambda &= \frac{2p(1-p)x_0 - 2p\overline{Z}}{2(1-p)x_0 - \overline{Z^2}}, \end{aligned}$$

where $\overline{Z}$ – sample mean, $\overline{Z^2}$ – sample second moment, S_Z^2 – sample variance of random sample $x_!, \ldots, x_n$ from Exponential-Pareto mixture distribution. As x_0, it is possible to choose the first order statistic of random sample $x_!, \ldots, x_n$.

Let $x = (x_1, \ldots, x_k)$ be a realization of rv with Exponential-Pareto mixture distribution. Then likelihood function can be written as

$$L(x, \lambda, x_0, \alpha) = \prod_{k=1}^{n} (p\lambda e^{-\lambda x_k} + (1-p)\alpha x_0^\alpha (x_0 + x_k)^{-\alpha-1}).$$

The log likelihood function is given by

$$l(x,\lambda,x_0,\alpha)=\sum_{k=1}^{n}\log(p\lambda e^{-\lambda x_k}+(1-p)\alpha x_0^{\alpha}(x_0+x_k)^{-\alpha-1}),$$

hence, the derivatives satisfy the following equations

$$\begin{aligned}
\frac{\partial l}{\partial\lambda} &= -p\lambda\sum_{k=1}^{n}\frac{x_k}{y_k}e^{-\lambda x_k};\\
\frac{\partial l}{\partial x_0} &= (1-p)\alpha x_0^{\alpha-1}\sum_{k=1}^{n}\frac{x_0+\alpha x_k}{y_k(x_0+x_k)^{\alpha+2}};\\
\frac{\partial l}{\partial\alpha} &= (1-p)x_0^{\alpha}(1-\alpha\log(1+1/\alpha))\sum_{k=1}^{n}\frac{1}{y_k(x_0+x_k)^{1+\alpha}};\\
\frac{\partial l}{\partial p} &= \lambda\sum_{k=1}^{n}\frac{e^{-\lambda x_k}}{y_k}-\alpha x_0^{\alpha}\sum_{k=1}^{n}\frac{1}{y_k(x_0+x_k)^{1+\alpha}},
\end{aligned}$$

where $y_k=p\lambda e^{-\lambda x_k}+(1-p)\alpha x_0^{\alpha}(x_0+x_k)^{-\alpha-1}$.

Setting the last equations equal to zero, the numerical maximum likelihood estimates of α, x_0, λ can be obtained by standard numerical methods like Newton Raphson method. The EM algorithm can be applied for iterative calculation of maximum likelihood estimates. Denote

$$g_{k1}=\frac{p_1 f_X(x_k|\lambda)}{p_1 f_X(x_k|\lambda)+p_2 f_Y(x_k|x_0,\alpha)},\quad g_{k2}=\frac{p_2 f_Y(x_k|x_0,\alpha)}{p_1 f_X(x_k|\lambda)+p_2 f_Y(x_k|x_0,\alpha)},$$

where $p_1=p,\ p_2=1-p$. Then maximization by parameters of

$$\sum_{k=1}^{n}(g_{k1}(\log p_1+\log f_X(x_k|\lambda))+g_{k2}(\log p_2+\log f_Y(x_k|x_0,\alpha))\to\max$$

leads to the following relations

$$\begin{aligned}
p_j &= \sum_{k=1}^{n}g_{kj}/n,\quad j=1,2,\\
\lambda &= \frac{\sum\limits_{k=1}^{n}g_{k1}}{\sum\limits_{k=1}^{n}g_{k1}x_k},\\
\alpha &= \frac{\sum\limits_{k=1}^{n}g_{k2}}{\sum\limits_{k=1}^{n}g_{k2}\log(z_k/x_0)},
\end{aligned}$$

where $z_k=x_k+x_0$ and parameter x_0 can be obtained from the equality

$$\frac{\sum\limits_{k=1}^{n}g_{k2}}{\sum\limits_{k=1}^{n}g_{k2}/z_k}=\frac{\sum\limits_{k=1}^{n}g_{k2}}{\sum\limits_{k=1}^{n}g_{k2}\log(z_k/x_0)}.$$

Table 1 demonstrates the results of identification of the distribution's parameters of the request processing time for web server 'dots.center' for different sample size of data. The web server processes industrial Internet data related to fuel consumption and operation of vessel equipment.

	Parameters of Exponential-Pareto distribution				
sample size	10^2	10^3	10^4	10^5	10^6
p	0.500	0.450	0.425	0.245	0.207
α	3.37	2.43	2.62	2.48	2.28
x_0	0.41	0.40	0.36	0.38	0.36
λ	6.45	6.71	6.67	6.70	6.47
MSE	0.070	0.037	0.017	0.240	0.370

Table 1: *The parameters of Exponential-Pareto distribution for the of the request processing time.*

8. Conclusion

In this paper we introduced the Exponential-Pareto mixture distribution. Characteristic function, failure rate function, mean excess, conditional excess distribution are derived. It is proved that this distribution belong to the subexponential distributions. Under condition on parameters $\lambda \geq \alpha/x_0$ Exponential, Exponential-Pareto mixture and Pareto distributions are ordered in failure rate, as well as conditional excesses and mean excesses. It is proved that the limiting distribution of maxima among n values of rv's with Exponential-Pareto distribution has Frechet-type form. The upper bound of uniform distance between Exponential-Pareto mixture and Pareto distributions is derived. The maximal likelihood estimates for parameters are given.

9. Funding

The publication has been prepared with the support of Russian Science Foundation according to the research project No. 21-71-10135, https://rscf.ru/en/project/21-71-10135/..

References

[1] Nanuwong N., Bodhisuwan W., Pudprommarat Ch. (2015). A New Mixture Pareto Distribution and Its Application. *Thailand Statistician* 13(2): 191-207.

[2] Marshall A., Olkin I. Life distributions. Structure of nonparametric, semiparametric and parametric families. Springer Science+Business Media, LLC., 2007.

[3] Weinberg G. V. (2016) Kullback-Leibler divergence and the Pareto-Exponential approximation. *SpringerPlus* 5(604)

[4] Nguyen-An H., Silverston T., Yamazaki T., Miyoshi T. (2021) IoT Traffic: Modeling and Measurement Experiments. *IoT* 2: 140–162.

[5] Peshkova, I., Morozov, E., Maltseva, M. (2020). On Comparison of Multiserver Systems with Exponential-Pareto Mixture Distribution. *Gaj, P., Guminski, W., Kwiecien, A. (eds) Computer Networks. CN 2020. Communications in Computer and Information Science, Springer, Cham* 1231.

[6] Peshkova I., Morozov E. (2022). On comparison of multiserver systems with multicomponent mixture distributions. *J. Math. Sci.* 267(2): 260–272.

[7] Peshkova, I., Golovin, A., Maltseva, M. (2023). On Waiting Time Maxima in Queues with Exponential-Pareto Service Times. *Vishnevskiy, V.M., Samouylov, K.E., Kozyrev, D.V. (eds) Distributed Computer and Communication Networks. DCCN 2022. Communications in Computer and Information Science Springer, Cham* 1748.

[8] Bulinski A., Slepov N. (2022). Sharp Estimates for Proximity of Geometric and Related Sums Distributions to Limit Laws. *Mathematics* 10(4747). https://doi.org/10.3390/math10244747

[9] Adler R. J., Feldman R. E. and Taqqu M. S. A Practical Guide to Heavy Tails: Statistical Techniques and Applications. Birkhauser Boston Inc., 1998.
[10] Embrechts P., Kluppelberg C., Mikosch T. Modelling Extremal Events for Insurance and Finance. Springer Heidelberg New York Dordrecht London, 1997.
[11] Navarro J. (2016). Stochastic comparisons of generalized mixtures and coherent systems *TEST*, 25: 150–169.
[12] Leadbetter M,, Lindgren G., Rootzen H. Extremes and related properties of random sequences and processes. Springer-Verlag New-York Inc. Springer series in statistics, 1983.
[13] Zolotarev V. M. Modern theory of summarizing of independent random variables. Nauka, Moscow, 1986. (in Russian)
[14] Korolev V. Yu, Krylov V. A. and Kuz'min V. Yu. (2011) Stability of finite mixtures of generalized Gamma-distributions with respect to disturbance of parameters *Inform. Primen.* 5(1): 31–38.

RELIABILITY ASSESSMENTS USING STOCHASTIC DEGRADATION PROCESS FOR CURRENT TIME ANALYSIS CUMULATIVE DAMAGE MODELS

G. Sathya Priyanka[1*], S. Rita[2], M. Iyappan[3]

[1*] Ph.D Research Scholar, Department of Statistics, Periyar University, Salem-11, India
[2]Associate Professor& Head, Department of Statistics, Periyar University, Salem-11, India
[3]Assistant Professor, Department of Statistics St. Francis College, Bengaluru-34, India
[1*]sathyapriyankastat@gmail.com, [2]ritasamikannu@gmail.com, [3]iyappastat@gmail.com

Abstract

The reliability study of such incredibly reliable items is inappropriate for the use of failure time data analysis and testing methodologies. More trustworthy information can be obtained from degradation data than from standard censored failure-time data, especially in cases where few or no failures are anticipated. The market for lighting has given a lot of attention to high-power white light emitting diodes (HPWLEDs). But as one of the more dependable electronic goods, it may not be expected to fail in either a traditional or even an accelerated life test. DDDM, or data-driven degradation methodology, is used in this research. Using data on lumen maintenance gathered from the IES LM-80-08 lumen maintenance test standard and based on the general degradation path model, the dependability of HPWLED was predicted. Testing such devices in typical working situations, and occasionally even under worse conditions, is difficult enough without trying to collect an adequate amount of time-to-failure data. Modern items are made with superb quality and high reliability in mind. Some safety-critical parts and systems are even made to last for an incredibly long time in order to prevent the disastrous effects of probable breakdowns. A cumulative damage model based on stochastic degradation processes has been developed in this paper. A suitable numerical representation is used to support the analytical findings. As a result, the degradation analysis approach has been developed to address dependability modeling issues using data on product degradation gleaned from historical records or degradation testing.

Keywords: Reliability, Lumen Maintenance Data, Degradation Data, High-Power White Light, Failure-Time, Highly Reliable Products

I. Introduction

The present makers face solid strain to foster new, higher innovation items in record time, while further developing efficiency, item field unwavering quality what's more, by and large quality. This has spurred the advancement of techniques like simultaneous designing and energized more extensive utilization of planned tests for item and interaction improvement. The

necessities for higher unwavering quality have expanded the requirement for more forthright testing of materials, parts and frameworks. Engineers in the assembling enterprises have utilized accelerated test (AT) tests for a long time. The reason for AT tests is to secure dependability data rapidly. Test unit of a material part of subsystem or whole frameworks are exposed to higher than normal degrees of one or then again additional speeding up factors like temperature or stress. Then, at that point, the AT results are utilized to foresee life of the units at use conditions. The extrapolation is normally legitimate (accurately or erroneously) based on genuinely roused models or a blend of experimental model fitting with an adequate measure of past involvement with testing comparative units. The need to extrapolate in both time and the speeding up factors by and large requires the utilization of completely parametric models [16]. Analysts have made significant commitments in the improvement of suitable stochastic models for AT information [typically a dispersion for the reaction and relapse connections between the boundaries of this appropriation and the speeding up variable(s)], measurable strategies for AT arranging (decision of speeding up factor levels and allotment of accessible test units to those levels) and strategies for assessment of reasonable dependability measurements. This paper gives a survey of a considerable lot of the AT models that have been utilized effectively around here.

Sped up life tests are usually utilized in item configuration processes. Since there is restricted opportunity to send off new items, engineers utilize sped up tests to acquire required data on the unwavering quality by raising the levels of specific speed increase factors like temperature, voltage, dampness, stress, and strain. For exceptionally solid present day items, it frequently requires substantially more investment to acquire lifetime and debasement information under common use conditions, and this expects one to utilize sped up tests [4]. Sped up tests open the items to more noteworthy ecological feelings of anxiety so we can get lifetime and corruption estimations in an additional convenient Procedures for playing out an Accelerated Life Testing Plans(ALT) incorporate steady pressure, step pressure, and slope pressure, among others. Assessment of the fluctuation of an assessor of a log area scale dispersion quantile with shifting pressure has numerous pragmatic applications [20]. It is important to foster helpful, precise likelihood models for derivations on the lifetime of the gadgets or frameworks under study. Such models ought to sensibly consolidate the speed increase factors and estimations of debasement as well as any real disappointments noticed. In this manner, in many designing dependability tests, proportions of debasement or wear toward disappointment can frequently be seen throughout some stretch of time before disappointment happens. Since the debasement values give extra data past that given by the disappointment perceptions, the two arrangements of perceptions should be thought about while doing derivation on the factual boundaries of the item or framework lifetime circulations as examined [24]. The target of the current paper is to expand existing outcomes by creating general disappointment models in view of Stochastic cycles for corruption which consolidate a few speeding up factors, and utilize both debasement estimations, and various decisions of test-feelings of anxiety and test length can bring about various accuracy of the gauge of the dependability of the item at typical use conditions. We want to find a test plan that gives least fluctuation of the most extreme probability gauges (MLEs) of the obscure area and scale boundaries of the log-area scale group of disseminations at indicated feelings of anxiety by reasonably deciding the test length [8].

II. Methods

I. Degradation Models in Reliability Analysis

The major idea under the overall debasement way models is to restrict the example space of the corruption interaction and accept all example capabilities concede a similar practical structure yet with various boundaries allude to Lio (2004). The overall corruption way model fits the debasement perceptions by a relapse model with irregular coefficients [16]. It declares that both basic straight relapse and nonlinear relapse models are by and large utilized in debasement way displaying [24]. Straight corruption is used in some basic wear cycles, for example, auto tire wear. Nonetheless, corruption ways are in many cases nonlinear elements of time and now and again linearization is infeasible. It presents an overall nonlinear blended impacts model and a two-stage way to deal with gauge model boundaries, that are multivariate regularly dispersed [18]. What's more, fosters a Monte Carlo reproduction strategy to work out a gauge of the dissemination capability of the opportunity to-disappointment [18]. They propose a parametric bootstrap strategy to set certainty stretches as recommended [18] with the accompanying presumption.

Sample assets are randomly selected from a population or production process and random measurement errors are independent across time and assets

- Sample assets are tested in a particular homogenous environment such as the same constant temperature
- Measurement (or inspection) times are pre-specified, the same across all the test assets, and may or may not be equally spaced in time. This assumption is used for constructing confidence intervals for time to failure distribution via the bootstrap simulation technique

A general degradation path model can be expressed as:

$$y_{ij} = \eta_{ij} + \varepsilon_{ij}\eta(t_j ; \Phi, \Theta_i) + \varepsilon_{ij} ; i = 1,2,3 \ldots\ldots n \quad (1)$$

$$\varepsilon_{ij} \approx N(0, \sigma_\varepsilon^2) \; j = 1,2, \ldots\ldots m \; \Theta_i < m \quad (2)$$

Where

t_j is time of the j^{th}measurement or inspection.

ε_{ij} is the measurement error with constant variance σ_ε^2

η_{ij} is the actual path of the i^{th}asset at time t_j with unknown parameters as listed later.

Φ is the vector of fixed-effect parameters, common for all assets.

Θ_i is the vector of the i^{th} asset random-effect parameters, representing resenting individual asset characteristics.

Θ_i and ε_{ij} are independent of each other ($i = 1,2, \ldots\ldots n$) and ($j = 1,2, \ldots\ldots m \; \Theta_i$)

m is the total number of possible inspections in the experiment.

$m\,\Theta_i$ is the total number of inspections on the i^{τ}asset, a function of Θ_i

It is assumed that Θ_i, $i = 1,2, \ldots\ldots n$ follows a multivariate distribution function (.) which may depend on some unknown parameters that must be estimated from the data. The distribution function of$T$, the failure time, can be written as:

$$Pr\{T \le t\} = F_T t = F_T(t ; \Phi, G_\varepsilon(\cdot), D, \eta) \quad (3)$$

II. Stochastic Models for Degradation Process

The aleatory vulnerabilities of a corruption cycle can be portrayed utilizing different sorts of probabilistic models. Customarily, the existence time appropriation models are utilized, in which the vulnerability of the corruption is portrayed according to the point of view of the unsure disappointment season of the part. The existence time dissemination model is generally applied in age based upkeep methodologies, where a part is supplanted when its activity time arrives at specific edge [16]. At the point when the review and substitution cost is restrictively high, for example, on account of a thermal energy station, age-based support techniques are normally wasteful as investigation and substitution of part are regardless of its genuine state of corruption. In such cases, condition based support procedures are frequently utilized, which require direct displaying of debasement progress.

Stochastic models are overall more adaptable in demonstrating these mind boggling designs of corruption process. Consequently, the utilization of stochastic models in corruption appraisal and forecast has become progressively famous as of late. Unwavering quality creation in light of corruption demonstrating can be a proficient technique for assess dependability of frameworks when perception s of disappointment rate [29]. Flow research shows that there has been a rising interest in use of stochastic debasement models in dependability expectation and endurance examination.

Models that depict the course of weakening or debasement in units or frameworks are of interest by their own doing, and are likewise key fixings in processes that decide "disappointment" occasions. Comparative with disappointment based dependability, debasement based unwavering quality has gotten a humble measure of consideration in the open writing. Corruption is an action to survey part life time was tended to in the early work [2]. All the more as of late give Valuable synopsis of corruption models, accentuating the utilization of straight models with accepted log-ordinary paces of debasement [19]. In such case, the full life time conveyance can be processed scientifically [14]. Other late models experienced in the writing manage corruption of materials, for example, those because of Celina (2001). If examined general ways to deal with assessing life time disseminations in sped up life test for exceptionally factor conditions. The models introduced there in center the particular of the corruption way as it relies unequivocally upon the working climate [10]. For profoundly dependable gadgets or costly gadgets, nonetheless, lifetime information might be challenging to acquire because of the timeframe required, or the expense of perception. Sped up life testing can frequently be utilized to speed up the disappointment for exceptionally dependable gadgets and subsequently it is proposed to broaden existing outcomes by creating general disappointment models in light of stochastic cycles for debasement which consolidate a few speeding up factors, and utilize both corruption estimations, and genuine disappointments in derivation strategies [22]. Corruption models and sped up test models for induction on dependability have been concentrated on by a few creators.

Sped-up tests decline the strength or time to disappointment and the expense of testing by uncovering the test examples to more significant levels of pressure conditions (expanded sizes or levels of ecological factors) which cause prior breakdowns and more limited lifetimes than the typical use condition [24]. These ecological factors and levels of pressure conditions are alluded to as the "speeding up factors" in the measurements and unwavering quality writing. Sped up life testing (ALT) is a speedy method for getting data about the existence dissemination of a material, part or item. In Sped up life testing (ALT) things are exposed to conditions that are more serious than the typical ones, which yield more limited life yet, ideally, don't change the disappointment systems [24]. A few suppositions are required to relate the life at high feelings of anxiety to life at typical feelings of anxiety being used. In view of these presumptions, the existence conveyance

under typical feelings of anxiety can be assessed. Such approach to testing lessens both time and cost.

Three kinds of pressure loadings are typically applied in sped up life tests: steady pressure, step pressure and Moderate pressure. Consistent pressure is the most widely recognized kind of pressure stacking. Each thing is tried under a consistent level of the pressure, which is higher than typical level. In this sort of testing, we might have a few feelings of anxiety, which are applied for various gatherings of the tried things. This implies that each thing is exposed to just a single feeling of anxiety until the thing falls flat or the test is halted for different reasons. In Sync pressure stacking, the test things are exposed to progressively more significant levels of pressure at pre allotted test times. All things are first exposed to a predetermined consistent pressure for a predefined timeframe [7]. Things that don't bomb will be exposed to a more significant level of pressure for one more indicated time. The degree of stress is expanded bit by bit until all things have fizzled or the test stops for different reasons. Moderate pressure stacking is very similar to the step pressure testing with the distinction that the feeling of anxiety increments ceaselessly. Step-stress testing is an extremely normal kind of sped up testing in view of speeding up factors. It is an effective method for getting disappointments in a moderately short measure of time. There are numerous varieties of step-stress testing. A typical sort is one in which the units are tried at a given anxiety for a specific measure of time. Toward the finish of that time, assuming there are units making due, the anxiety is expanded and held for one more measure of time [21]. The information that outcome from such tests can be dissected utilizing the total harm model. For a nitty gritty concentrate on combined harm model [3o].

For exceptionally dependable items, it's anything but a simple assignment to evaluate the lifetime dispersion of the items by utilizing the conventional life-testing methods which record just opportunity to disappointment information. In any event, utilizing the systems consolidating editing and speeding up strategies, the data about the lifetime appropriation is still exceptionally restricted. Under this present circumstance, an elective methodology is to gather the "debasement" information at more elevated levels of pressure for foreseeing an item's lifetime at a specific use-feeling of anxiety. Such an investigation is called an ADT [3].

For fruitful use of ADT, many ways to deal with model debasement of items are given. Especially, Markov cycles, for example, the Brownian movement with float, the compound Poisson process, and the gamma interaction are generally utilized inferable from their autonomous augmentations property. For the stochastic displaying of monotonic and progressive corruption over the long run in a grouping of minuscule augmentations, the gamma cycle has found application as a debasement model in many examinations [17]. The fundamental goal of paper is to expand existing outcomes and creating general disappointments models based a stochastic cycle for debasement which consolidate a few speeding up factors, and utilize both corruption estimations, and genuine disappointments in inferential methodology.

III. Model Description

At stress level $x_i, (i = 1, 2 \ldots k,)$ the lifetime Y of a test unit is assumed to follow a log-location scale distribution with a cumulative distribution function (CDF)

$$F_i(y;\ \mu_i, \sigma) = \Phi\left(\frac{\ln y - \mu_i}{\sigma}\right), y \geq 0 \quad (4)$$

Where $\Phi(\cdot)$ is the standard log-location-scale CDF, and the location parameter is $\mu_i = \mu(x_i) = \beta_0 + \beta_1 x_i$, and σ is the unknown scale parameter. Here, the regression parameters β_0 and $\beta_1 (< 0)$ are unknown and need to be estimated, and the scale parameter σ is assumed to be free of stress levels. The CDF of the lifetime of a test unit under the k-level step-stress ALT is given by

$$G(y) = F_i(y + s_{i-1} - \tau_{i-1}; \mu_i, \sigma) \text{ for } \begin{cases} \tau_{i-1} \le y \le \tau_i, & \text{for } i = 1,2,\cdots k-1 \\ \tau_{i-1} \le y \le \infty, & \text{for } i = k, \end{cases} \quad (5)$$

Where s_0=0 and $s_{i-1} = \tau_{i-1} + s_{i-2} - \tau_{i-2} \exp(\mu_i - \mu_{i-1})$ is the solution of the equation $F_i(s_{i-1}; \mu_i, \sigma) = F_{i-1}(\tau_{i-1} + s_{i-2} - \tau_{i-2}; \mu_{i-1}, \sigma), i = 2,3,\cdots,k$

$$G(y) = \begin{cases} G_1(y) = \Phi\left(\frac{\ln y - \mu_i}{\sigma}\right), & \text{for } 0 < y \le \tau_1 \\ G_1(y) = \Phi\left[\frac{\ln(y + s_{i-1} - \tau_{i-1}) - \mu_i}{\sigma}\right], & \text{for } \tau_{i-1} \le y \le \tau_i \end{cases} \quad (6)$$

where $i = 2,\cdots k-1$, and $\tau_{i-1} \le y \le \infty$ where $i = k$ and the corresponding probability density function (PDF) of the lifetime of a test unit is given by

$$g(y) = \begin{cases} g_1(y) = \frac{1}{\sigma y}\Phi\left(\frac{\ln y - \mu_1}{\sigma}\right), & \text{for } 0 < y \le \tau_1 \\ g_i(y) = \frac{1}{\sigma(y + s_{i-1} - \tau_{i-1})}\Phi\left[\frac{\ln(y + s_{i-1} - \tau_{i-1}) - \mu_i}{\sigma}\right], & \text{for } \tau_{i-1} \le y \le \tau_i \end{cases} \quad (7)$$

IV. Maximum Likelihood Estimation

From Equations (6) and (7), the joint PDF of observed data is $n = (n_1, n_2, \ldots, n_k)$ and $y = (y_1, y_2, \ldots, y_k)$ with $y_i = (y_{i,1}, y_{i,2}, \ldots, y_{i,n_i})$ is given by

$$f(y, n) = \frac{n!}{(n - \sum_{i=1}^{k} n_i)!}\left\{\prod_{i=1}^{k}\left[\prod_{j=1}^{n_i} g_i(y_{i,j})\right]\right\}[1 - G_k(\tau_k)]^{n - \sum_{i=1}^{k} n_i} \quad (8)$$

Note that the MLEs of β_0, β_1, and σ exists only if $n_i > 0, i = 2,\cdots,k$, in Equation (8). By using the following expressions

$$\frac{\partial s_{i-1}}{\partial \beta_1} = \sum_{h=2}^{i}(x_h - x_{h-1}) s_{h-1} e^{\beta_1(x_i - x_h)}$$

$$\frac{\partial z_{i,j}}{\partial \beta_0} = -\frac{1}{\sigma}, \quad \frac{\partial z_{i,j}}{\partial \beta_1} = \frac{1}{\sigma}\left[\sum_{h=2}^{i}(x_h - x_{h-1})\frac{s_{h-1} e^{\beta_1(x_i - x_h)}}{y_{i,j} + s_{i-1} - \tau_{i-1}} - x_i\right], \frac{\partial z_{i,j}}{\partial \sigma} = -\frac{z_{i,j}}{\sigma}$$

$for\ i = 1,2,\cdots,k\ and\ j = 1,2,\cdots,n_i$

$$\frac{\partial \eta_k}{\partial \beta_0} = \frac{1}{\sigma}, \frac{\partial \eta_k}{\partial \beta_1} = \frac{1}{\sigma}\left[\sum_{h=2}^{k}(x_h - x_{h-1})\frac{s_{h-1} e^{\beta_1(x_k - x_h)}}{y_{i,j} + s_{k-1} - \tau_{k-1}} - x_k\right], \frac{\partial \eta_k}{\partial \sigma} = -\frac{\eta_k}{\sigma}$$

$$\frac{\partial \Phi(\eta_k)}{\partial \beta_0} = -\frac{\Phi(\eta_k)}{\sigma}, \frac{\partial \Phi(\eta_k)}{\partial \beta_1} = \frac{\Phi(\eta_k)}{\sigma}\left[\sum_{h=2}^{k}(x_h - x_{h-1})\frac{s_{h-1} e^{\beta_1(x_k - x_h)}}{\tau_k + s_{k-1} - \tau_{k-1}} - x_k\right]$$

The MLE is $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\sigma}$ can be obtained by solving the following likelihood equations:

$$\frac{\partial In\,L}{\partial_{\beta_0}} = \frac{1}{\sigma}\left[-\sum_{i=1}^{k} n_i - \sum_{i=2}^{k}\sum_{j=1}^{n_i}\frac{z_{i,j}}{\phi(z_{i,j})}\cdot\frac{d\phi(z_{i,j})}{dz} + \left(n-\sum_{i=1}^{k} n_i\right)\frac{\eta_k\phi(\eta_k)}{1-\Phi(\eta_k)}\right] = 0 \quad (9)$$

The second derivative of the log-likelihood function

$$\frac{\partial^2 In\,L}{\partial^2_{\beta_0}} = -\frac{1}{\sigma^2}\left\{\sum_{i=2}^{k}\sum_{j=1}^{n_i}\frac{1}{\phi(z_{i,j})}\left[\left(\frac{d\phi(z_{i,j})}{dz}\right)^2 - \frac{d^2\phi(z_{i,j})}{dz^2}\right] + \left(n-\sum_{i=1}^{k} n_i\right)\frac{1}{1-\Phi(\eta_k)}\left[\frac{d\phi(\eta_k)}{d\eta} + \frac{\phi^2(\eta_k)}{1-\Phi(\eta_k)}\right]\right. \quad (10)$$

$$\frac{\partial^2 InL}{\partial\sigma^2} = -\frac{1}{\sigma}\left(\frac{\partial InL}{\partial\sigma}\right) - \frac{1}{\sigma^2}\left\{\sum_{i=2}^{k}\sum_{j=1}^{n_i}\frac{z_{ij}}{\phi(z_{ij})}\left[\frac{1}{\phi(z_{ij})}\left(\frac{d\phi(z_{ij})}{dz}\right)^2 - \frac{d^2\phi(z_{ij})}{dz^2}\right]\right.$$

$$\left.\sum_{i=1}^{k} x_i\sum_{j=1}^{n_i}\frac{z_{ij}}{\phi(z_{ij})}\cdot\left(\frac{d\phi(z_{ij})}{dz}\right) + \left(n-\sum_{i=1}^{k} n_i\right)\frac{\eta k}{1-\Phi\eta_k}\phi(\eta_k) + \eta k\frac{d\phi(\eta_k)}{d\eta} + \frac{\phi^2(\eta_k)}{1-\Phi\eta_k}\right\} \quad (11)$$

Since these situations can't be addressed scientifically, mathematical techniques, for example, reenacted strengthening calculation or some other iterative system should be utilized for this assessment issue. A benefit of utilizing the reenacted strengthening calculation is that it permits us to find a worldwide ideal without relying upon the decision of the underlying qualities, which is one of the primary downsides of the regularly utilized mathematical techniques like Newton-Raphson.

IV. Discussion

I. Numerical Illustration

The ideal arrangements were acquired by the reproduced tempering calculation as proposed by Corona et al., (1987). It very well may be effectively seen that the assessed asymptotic changes in light of complete information are the littlest, trailed by those given blue-penciled information inside the looking-through range (0, 50], and afterward the ones with examination spans being picked by specific proportions. To determine the optimal unequal time points $(\tau_1 < \tau_2 < \cdots < \tau_k)$ that minimize the large-sample approximate variance of the MLE of the 200pth quantile $(0 < p < 1)$ of the log-lifetime distribution at the normal-use stressx_0. The MLE of the 200p quantile at the normal-use stress x_0 can be expressed as$\hat{y}_p = \hat{\beta}_0 + \hat{\beta}_1 x_0 + z_p\hat{\sigma}$, where z_p is the 100pth percentile of the standard log-location-scale distribution. Thus, If $x_0 = 0$, the asymptotic variance of the estimator $\hat{y}_p$ at the normal-use stressx_0 is given by

$$A\,Var(y_p) = A\,Var(\hat{\beta}_1 + z_p\hat{\sigma}) = V_{11} + z_p^2 V_{33} + 2_{zp}V_{13} \quad (12)$$

Where V_{ij}' It is verified under the C- optimality Criterion based 2-level step-stress ALT plan is preferable, whenever we optimize the general 2-level step-stress ALT plan the second K-level step–ALT plan under a censoring scheme under the considered and the results are shown in Table -1. $\tau = 2{:}\ 1$,The inspection interval for the first stage is twice as long as that for the second stage

Table 1: *Censoring scheme*

C_0	P_0	γ_0	θ_0	Parameter	MLE	RAB	MSE
1.0	1.0	1.25	0.7	C	1.10902	0.10902	0.01189
				P	0.98314	0.01686	0.00028
				γ	1.22511	0.01991	0.00062
				θ	0.88398	0.11602	0.01346
				α_1	1.38040	0.10486	0.01717
				α_2	0.69831	0.11785	0.00542
				α_3	0.46874	0.12552	0.00273
1.0	1.0	1.3	1.0	C	1.15695	0.15695	0.02463
				P	0.9846	0.0154	0.00024
				γ	1.23098	0.05309	0.00476
				θ	0.87707	0.12293	0.12446
				α_1	1.44053	0.15299	0.03654
				α_2	0.72799	0.16536	0.01067
				α_3	0.48837	0.17266	0.00517
1.0	1.0	1.5	1.0	C	1.35169	0.35169	0.12369
				P	0.99012	0.00988	0.00010
				γ	1.25504	0.16331	0.06001
				θ	0.84992	0.15008	0.02252
				α_1	1.68507	0.34872	0.18982
				α_2	0.84832	0.35799	0.05001
				α_3	0.56782	0.36344	0.02291
1.0	1.1	1.4	1.0	C	1.26473	0.26473	0.07008
				P	0.97595	0.11277	0.01539
				γ	1.37052	0.02106	0.00087
				θ	0.8474	0.1526	0.02329
				α_1	1.5717	0.23028	0.08654
				α_2	0.79906	0.34074	0.04124
				α_3	0.53793	0.4099	0.02446
1.25	1.1	1.25	1.0	C	1.14000	0.08800	0.01210
				P	0.95459	0.13219	0.02114
				γ	1.37506	0.10005	0.01564
				θ	1.01030	0.01030	0.00011
				α_1	1.40997	0.11705	0.03494
				α_2	0.72753	0.02342	0.00030
				α_3	0.49403	0.03589	0.00029
1.4	1.0	1.0	0.7	C	0.91213	0.34848	0.23802

				P	0.68352	0.31648	0.10016
				γ	1.25991	0.25991	0.06755
				θ	1.10442	0.57774	0.16356
				α_1	1.06206	0.39281	0.47207
				α_2	0.66129	0.24387	0.04549
				α_3	0.50122	0.14034	0.00670
1.4	1.2	1.0	0.9	C	0.91672	0.3452	0.23356
				P	0.84089	0.29926	0.12896
				γ	1.49253	0.49253	0.24259
				θ	1.10045	0.22272	0.04018
				α_1	1.10547	0.39552	0.52319
				α_2	0.61718	0.22467	0.03198
				α_3	0.43888	0.10314	0.00255

Table 2: *The variance optimality under the step stress setting based*

Complete data				**Censored data**			
p	X_1	X_2	τ_1^c	τ_1^c, τ_2^c	$\Delta_\tau = 2:1$	$\Delta_\tau = 2:1$	$\Delta_\tau = 1:2$
0.5	0.2	0.5	10.2895	10.0470, 20.0000	10.0312(3.48)	10.0000(3.35)	6.6667(3.99)
		1.0	11.0011	10.7887, 20,0000	10.8267(2.45)	10.8267(2.50)	6.6667(3.10)
	0.4	0.5	6.8724	6.8408, 20,0000	6.5817(15.81)	6.5817(16.02)	6.6667(15.83)
		1.0	7.3327	7.3107, 20.0000	7.1133(8.71)	7.1133(8.32)	0.6667(8.38)
0.95	0.2	0.5	13.6090	12.6071, 20,0000	13.3333(7.48)	10.0000(7.87)	6.6667(11.03)
		1.0	14.9424	14.0111, 20,0000	13.3333(5.87)	10.0000(6.56)	6.6667(9.42)
	0.4	0.5	8.1548	8.0935, 20,0000	7.8435(26.76)	8.0765(25.03)	6.6667(25.91)
		1.0	9.0548	9.0018, 20,0000	8.8341(14.17)	9.0146(14.08)	6.6667(16.19)

The stress levels x_i, $i = 1, \ldots, k$, when $\beta_0 = 2.5, \beta_1 = -1.0, \sigma = 0.5$. to identify the optimal change points leading to variance–optimality, the optimal change points, and associated asymptotic variance based on the censored data when the lengths of the inspection intervals were chosen according to a certain ratioτ.

V. Conclusion

Numerous stochastic models of equipment deterioration have been put forth based on the physics of failure and the operational environment. The essential idea of these models is that they are created by modeling the underlying mechanisms that lead to failure, like degradation and wear, using the appropriate stochastic processes. A unified theory of predictive maintenance must be developed through the creation and analysis of these stochastic deterioration models.

These models produce residual life distributions and time to failure that are highly theoretically challenging. Our study's goal is to determine whether conventional time to failure distributions, like the Weibull, can be used to approximate the time to first failure distributions

that come from stochastic deterioration models. We initially built a discrete-event simulation model that simulates the stochastic deterioration and failure of the target system, which is a single-unit system subject to a random operating environment with a variable instantaneous rate of degradation. The typical time to failure distribution is then fitted to the simulated data using a predetermined methodology. The quality of this fit is assessed using a large-scale numerical experiment with a variety of system characteristics. According to the findings of the goodness-of-fit tests, a truncated, three-parameter Weibull distribution is a fair approximation for the scenario discussed in the study.

In ALTA, the analysis is carried out in conditions of high stress, and the extrapolation from the heightened stress levels to the usage stress level is based on the association between life stress and stress at work. Any level at which the life-stress relationship holds true can be used to forecast product performance, including the usage stress level.

References

[1] Aly, H. M. and Ismail, A. A. (2008). Optimum Simple Time-Step Stress Plans for Partially Accelerated Life Testing With Censoring. *Far East Journal of Theoretical Statistics*, 24: 175-200.

[2] Abdel-Ghaly, A. A Attia, A. F, and. Aly, H. M. (1998). Estimation of the parameters of pareto distribution and the reliability function using accelerated life testing with censoring. *Communications in Statistics Part B*, 27: 469–484.

[3] Abdulla, A. Alhadeed and Shie-Shien, Yang. (2005). Optimal Simple Step-Stress Plan for Cumulative Exposure Model Using Log-Normal Distribution. *IEEE Transactions on Reliability*, 54: 64-68.

[4] Bai, D. S. and Chung S. W. (1989). An accelerated life test model with the inverse power law. *Reliability Engineering and System Safety*, 24: 223–230.

[5] Bugaighis, M. M. (1990). Properties of the MLE for parameters of a Weibull regression model under type I censoring. *IEEE Transactions on Reliability*, 39:102–105.

[6] Bai, M. S. Kim and Lee, S. H. (1989). Optimum Simple Step-Stress Accelerated Life Tests with Censoring. *IEEE Transactions on Reliability*, 38:528-532.

[7] Bhattacharyya, G. K and Zanzawi Soejoeti, (1989). A tampered failure rate model for step-stress accelerated life test. *Communications in Statistics - Theory and Methods*, 18:1627-1643.

[8] Chen-Mao Liao, and Sheng-Tsaing Tseng, (2006). Optimal Design for Step-Stress Accelerated Degradation Tests. *IEEE Transactions on Reliability*, 55: 59-66.

[9] Chengjie Xiong, and George, A. Milliken. (1999). Step-S tress Life-Testing with Random S tress-Change Times for Exponential Data. *IEEE Transactions on reliability*, 48:141-148.

[10] Chanseok Park and William, J. Padgett. (2005). New Cumulative Damage Models for Failure Using Stochastic Processes as Initial Damage. *IEEE Transactions on Reliability*, 54:530-540.

[11] El-Dessouky, E. A., On the use of bayesian approach in accelerated life testing, M.S. thesis: Institute of Statistical Studies and Research, Cairo University, Egypt, 2001.

[12] Gouno, E. (2007). Optimum step-stress for temperature accelerated life testing. *Quality and Reliability Engineering International*, 23:915–924.

[13] Jerry Lawless and Martin Crowder, (2004). Covariates and Random Effects in a Gamma Process Model with Application to Degradation and Failure. *Lifetime Data Analysis*, 10: 213–227.

[14] Johnson, N. L., Kotz, S., and Balakrishnan, N. Continuous Univariate Distributions, Probability and mathematical Statistics, 1995.

[15] Joseph .C, Lu and William. Q,Meeker. (1993). Using Degradation Measures to Estimate a Time-to- Failure Distribution, *Technometrics*, 35:161-174.

[16] Lawless, J.F. (1976).Confidence interval estimation in the inverse power law model. *Journal of the Royal Statistical Society: Series C*, 25: 128–138.

[17] Lu, C.J. and Meeker, W.Q. (1993). Using Degradation Measures to Estimate a Time-to-Failure Distribution. *Technometrics,* 35:161-174.

[18] McCool, J. I. (1980). Confidence limits for Weibull regression with censored data. *IEEE Transactions on Reliability,* 29:145–150.

[19] Meeker, W. Q, Escobar, L. A and Lu, C. J.(1998). Accelerated degradation tests: modeling and analysis. *Technometrics,* 40:89–99.

[20] Mathai, A. M. and Provost, S. B. (2004). On the distribution of order statistics from generalized logistic samples. *International Journal of Statistics*, 62:63–71.

[21] Michkle Boulanger Carey, and Reed H. Koenig. (1991). Reliability Assessment Based on Accelerated Degradation: A Case Study. *IEEE Transactions on Reliability,* 40:499-506.

[22] Nesar Ahmad, Ariful Islam and Abdus Salam, (2006). Analysis of optimal accelerated life test plans for periodic inspection. *International Journal of Quality & Reliability Management*, 23: 1019-1046.

[23] Nelson, W and Kielpinski, T. J (1976).Theory for optimum censored accelerated life tests for normal and lognormal life distributions. *Technometrics,* 18: 105–114.

[24] Park. C, and Padgett. W. J, (2005). Accelerated Degradation Models for Failure Based on Geometric Brownian motion and Gamma Processes. *Lifetime Data Analysis,* 11: 511–527.

[25] Rai, R. K. Sarkar, S. and Singh V. P. (2009). Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph. *Water Resources Management,* 23:899–929.

[26] Robert Miller and Wayne Nelson.(1983). Optimum Simple Step-Stress Plans for Accelerated Life Testing. *IEEE Transactions on Reliability,* 32:59-65.

[27] Shabri, A. Ahmad, U. N. Zakaria Z. A. (2011). TL-moments and L-moments estimation of the generalized logistic distribution. *Journal of Mathematical Research*, 10:97–106.

[28] Whitmore, G. A. and Fred Schenkelberg. (1997). Modeling Accelerated Degradation Data Using Wiener Diffusion with a Time Scale Transformation.*Lifetime Data Analysis,* 3:27–45.

[29] William Q. Meeker, Luis A. Escobar and C. Joseph Lu, (1998). Accelerated Degradation Tests: Modeling and Analysis. *Technometrics,* 40: 89-99.

[30] Watkins, A. J. (1991). On the analysis of accelerated life-testing experiments. *IEEE Transactions on Reliability*, 40:98–101.

Software Quality Analysis Based on Selective Parameters Using Enhanced Ensemble Model

Rakhi Singh
•
Department of Computer Science & Engineering, Shobhit Institute of Engineering & Technology (Deemed-to-be University), Meerut, India
research.rs2k11@gmail.com

Mamta Bansal
•
Department of Computer Science & Engineering, Shobhit Institute of Engineering & Technology (Deemed-to-be University), Meerut, India
mamta.bansal@shobhituniversity.ac.in

Surabhi Pandey
•
Department of ICT & e-Governance, Indian Institute of Public Administration, New Delhi, India
dr.pandeysurabhi@gmail.com

Abstract

Software Quality Analysis refers to the process of evaluating and assessing the quality of software products or applications. It involves analyzing various aspects of the software to determine its level of quality, identify potential issues or defects, and make informed decisions to expand the software's overall quality. There are investigated different software quality models based on machine learning algorithms. Nevertheless, the explored approaches have an inconsistent understanding of the software product quality and high complexity. This research presents an enhanced ensemble model (EEM) that involves Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Extreme Learning Machine (ELM) to assess the optimal outcomes. This model performance is computed based on multiple selective parameters namely functional suitability and maintainability of the software and compared along with several algorithms namely Decision Tree, Random Forest, AdaBoost, and Naive Bayes. The outcome of this ensemble model demonstrates that it offers highly accurate results for software validation, verification, and overall product development process to analyze the functional suitability as well as software maintainability. The measured accuracy on Decision Tree, Random Forest, Naive Bayes, AdaBoost, and EEM is found 92.08%, 93.35%, 94.50%, 95.60%, and 99.14%, respectively

Keywords: Software Quality Analysis, Enhanced Ensemble Model (EEM), Decision Tree, Random Forest, Naive Bayes, AdaBoost.

1. Introduction

The process of evaluating the quality and efficiency of software systems is known as Software Quality Analysis (SQA), commonly referred to as Software Quality Assurance or Software Testing [1]. To find flaws, errors, or variations from expected behavior in software products, systematic activities, and procedures are used. Software quality analysis' main goal is to make sure that the

entire software product works properly, complies with specifications, and offers a desired user experience. Activities like test planning, tracking of defects, test designing, execution of tests, reporting of tests, and other testing kinds including performance testing, security testing, and usability testing are all a part of this process. Organizations may reduce the chance of problems in production and produce high-quality software to satisfy user expectations by undertaking software quality analysis [2]. This allows organizations to discover and fix errors early in the development lifecycle. There are available different software quality analysis techniques such as static analysis and many others. To estimate and promise the quality of software systems, a variety of approaches and techniques are used in software quality analysis. Following are some examples of frequently used software quality analysis techniques [3]:

1.1. *Static Analysis:*
In this technique, the program is not run; instead, the source code, design documentation, and other artifacts are examined. Code concerns such as coding mistakes, coding standards violations, security flaws, and performance inefficiencies are all examined using static analysis tools.

1.2. *Dynamic Analysis:*
Dynamic analysis aids in the discovery of runtime problems such as memory leaks, performance bottlenecks, and functional flaws.

1.3. *Regression Testing:*
Regression testing is done to make sure that changes to the program do not result in the introduction of new flaws or affect current functioning.

1.4. *Performance Testing:*
It helps in the detection of performance bottlenecks, resource use problems, and response time lags. Techniques for performance testing include load, stress, and endurance testing [4].

1.5. *Security Testing:*
This method is utilized to find weaknesses and possible attack entry points. It also utilizes the strategies including penetration testing, vulnerability scanning, and security code reviews.

1.6. *Usability Testing:*
Usability testing looks at the user interface (UI) and user experience (UX) of the program to make sure that it is simple, effective, as well as satisfies user expectations.

1.7. *Code Reviews:*
Code reviews entail systematically going over source code to look for coding flaws, performance problems, maintainability difficulties, and adherence to coding standards.

1.8. *Test Automation:*
Test automation is the process of automating time-consuming and repetitive testing operations using specialized tools and frameworks.

There have been investigated multiple ways which introduce how can improve software quality analysis and according to that the study of software quality may be greatly enhanced by combining useful methods and strategies. The use of thorough testing methods at every stage of the software development lifecycle is a critical task [5]. By integrating unit tests, integration tests, regression tests, system tests, as well as acceptance testing, software developers may ensure a

thorough assessment of the quality of their product. Each testing level has a distinct function and is capable of spotting many kinds of flaws, enabling a more thorough assessment. Additionally, the use of test automation is crucial for improving software quality analysis. The formulation and execution of automated tests are made possible by automated testing tools and frameworks, which enhance test coverage, lowers human error rates, and boost the general effectiveness of testing procedures [6]. Software problems may be swiftly found and fixed by developers with the capacity to run automated tests frequently, speeding up the resolution process. The examination of software quality is also improved by code reviews and inspections. Regular code reviews enable engineers to see possible problems before they become serious. Developers can find coding errors, confirm that coding standards are being followed, and pinpoint opportunities for development by carefully studying the source code. By taking proactive measures, the software's overall quality is improved and risks are reduced [7]–[9].

J. Mona et al. in [10] explored the increasing need for software applications due to recent technological advancements. The demand for software products has significantly grown across various industries. To ensure the increase of the software industry, it is decisive to create high-quality software and maintain its reputation among users. However, implementing quality assurance standards to instill trust in consumers can be challenging. The software development team often views quality assurance as a time-consuming and documentation-intensive process that adds little value to the client. It is also stated that determining the quality of software can be difficult as it depends on individual perspectives and interests. The study also mentions the use of soft computing-rooted machine learning techniques to address quality as well as assurance in software engineering, intending to build an effective framework to predict software faults. The drawback of this method is that it does not provide specific examples of recent technological breakthroughs or discuss any particular challenges faced in quality assurance or software development. It presents a general overview of the importance of software quality but lacks specific details and concrete examples.

E. Rashid et al. [11] discussed the use of case-based reasoning (CBR), a kind of machine learning, as an expert system to forecast software quality. The major goal of this study is to reduce software expenses. The paper introduces a novel idea of building a knowledge base (KBS) for CBR, which includes new problems and their solutions. The research also aims to reduce maintenance costs by removing duplicate records from the KBS. Four similarity functions are used for error prediction. Case-based models, such as the developed case-based reasoning model, are considered effective when defining rules for a problem domain is challenging. On the other side, one of the drawbacks of using machine learning for software quality analysis is the requirement for large amounts of labeled data. Machine learning algorithms typically perform better when trained on a large and diverse dataset. However, obtaining labeled data for software quality prediction can be challenging and time-consuming.

Research Questions:

- How to improve the software quality analysis with selective parameters and an ensemble model?
- To evaluate the performance metrics using diverse datasets.

2. LITERATURE REVIEW

T. Sharma et al. in [12] discussed the importance of predicting software faults before the testing phase to improve source provision and ensure high-quality software development. Different algorithms are used to address this issue, and various analytical frameworks have been built to classify software modules as defective or non-defective. However, despite the advancements in

software development, poor prediction performance is still observed in many studies. This is attributed to challenges such as redundancy, correlation, irrelevant features, missing data, and an unbalanced distribution between classes with faults and classes without faults in the software data. To tackle these challenges, researchers globally have turned to ensemble different techniques, which have shown certain enhancements in fault estimation performance. This research critically analyzes ensemble-rooted machine-learning methods explored for software fault forecasting from 2018 to 2021. The paper aims to gain a deeper understanding of why hybrid frameworks still exhibit poor performance on accessible datasets. The drawback investigation is that it relies on meta-learning and classifiers for software defect prediction. While meta-learning can be effective in adapting to different datasets and improving generalization, it may face limitations when applied to software defect prediction.

K. Bashir et al. [13] explored that software quality analysis aims to develop forecasting models utilizing the fault data as well as software metrics to identify potentially faulty program modules. This helps in optimizing resource allocation as well as utilization. Nevertheless, the strength of the classifiers and the quality of the data both have an impact on how accurate these prediction models are. The effectiveness of classification models might be jeopardized by problems with data quality such as an excessive number of dimensional imbalances in classes, and the inclusion of diverse noise in software fault data. To improve software development procedure (SDP) models. This study presents a novel and unified framework. The system uses Data Sampling (DS), Iterative-Partition Filter (IPF), and Ranker Feature Selection (FS) approaches to handle the issues of high dimensionality, and imbalances in classes, respectively. The outcome validates the effectiveness of this suggested framework for SDP, indicating that it helps improve the accuracy as well as performance of fault prediction models. The drawback of the proposed technique is that supervised learning requires labeled training data, which may be expensive and time-consuming. Collecting and labeling a comprehensive dataset of software defects can be challenging, especially for large-scale software systems. Limited or incomplete training data can lead to biased or inaccurate predictions, reducing the reliability of the defect prediction framework.

M. Rawat et al. [14] stated that despite thorough planning, extensive documentation, and proper process control throughout software development, it is unavoidable to encounter certain defects. These defects can negatively impact the quality of the software and potentially lead to failure. In today's highly competitive environment, it is crucial to actively work towards controlling and minimizing software engineering flaws. However, these initiatives need time, money, and resources. This paper aims to identify the root causes of these defects and provide suggestions for improving software quality and productivity. Additionally, the paper demonstrates the implementation of various defect prediction models that have resulted in a decrease in the number and severity of defects. The main drawback of this is the potential for inaccurate predictions. While these models aim to identify and predict software defects, there is still a degree of uncertainty involved. The models rely on historical data and patterns to make predictions, which means they may not account for unique or unforeseen circumstances that could lead to defects. Additionally, the accuracy of these models heavily relies upon the quality as well as the relevance of the utilized datasets for training. If the dataset is incomplete, biased, or outdated, then software quality prediction may not be reliable.

S. Yamada [15] investigated that the assessment of software reliability holds significant importance in efficiently developing high-quality software products. This paper focuses on quantitative measurement as well as the evaluation of software product reliability. The methods discussed in this work are centered on nonhomogeneous Poisson processes used in the software product consistency advancement models built in Japan. These models aim to provide a more realistic depiction of the software error-detection and failure-occurrence processes in the test phase of software product expansion by refining the underlying assumptions. The paper provides an

outline of the existing software reliability growth models (SRGMs) as well as explores the application of maximum-likelihood estimations for analyzing software reliability data and assessing software reliability. Furthermore, the paper presents instances of software reliability assessment using a tool that integrates various prominent SRGMs. These examples are based on observed test data derived from actual software projects. The drawback is the complexity of measuring software quality and reliability. Software systems are intricate and involve numerous interconnected components, making it challenging to develop comprehensive measurement techniques that capture all aspects of quality and reliability accurately. It is difficult to account for all potential failure scenarios and accurately predict the behavior of software in all possible circumstances.

V. Srivastava et al. [16] discovered a more robust software product source code for effective software product quality analysis. Software product quality analysis is challenging in the case of huge software applications. The profiles of resources are necessary for appropriate staffing, finance, and the advancement of an achievable project plan throughout the lifespan of a project, even while the needs for the framework may be functional but only stated at an extreme level. The size of the program may be calculated using past data and trends for the same project, opening up the possibility of an assessment approach. Software quality is now a key component in the intended attainment of overall human and commercial safety in the modern world, where processors are used in every conceivable area. Finding the software's quality attributes and turning them into calculable measurements will be crucial to the success of the final product. The mapping of program elements to these metric values illustrates details about the behavior and complexity of the framework. Here are some of the drawbacks associated with this technique are that it has limited scope, lack of context, overemphasis on quantitative metrics, sensitivity to programming languages and paradigms, the complexity of metric selection and interpretation, and inability to capture dynamic aspects.

E. Belachew [17] discussed the significance of software metrics in software engineering and emphasizes the importance of accurate measurement in this field. As software becomes more complex, manually inspecting it becomes challenging. Software engineers are concerned about software quality and seek ways to measure and improve it. The objective of the proposed research is to assess as well as analyze software metrics used for measuring software products and processes. The researcher collected literature from electronic databases since 2008 to gain a comprehensive understanding of software metrics. The study concludes that software quality can be measured by assessing how well the software aligns with its design, considering variables such as exactness, quality of product, scalability, fullness, and absence of bugs. Since quality standards vary across organizations, using software metrics and common tools can reduce subjectivity when evaluating software quality. The study's main contributions include providing an overview of software metrics and conducting a critical study of the key metrics found in the literature. On the other side, there are some potential limitations and drawbacks associated with this approach which are interpretation ambiguity, incomplete representation, metric selection, and validity, time and effort requirements, lack of causality, etc.

I. Khan et al. in [18], discussed that the field of Software Quality Analysis (SQA) is promptly mounting in the field of software product engineering, as it is aimed to offer more robust solutions for real-world applications. SQA involves a formal process utilized to assess, document, as well as ensure software application quality at all stages of the Software Development Life Cycle (SDLC). This research focuses on identifying and understanding various factors that can impact the quality of software product development. The study investigates the relationships between these factors and the different phases of the SDLC. Additionally, the research introduces a new quality factor called testability to the existing set of quality factors in system design and analysis. The anticipated outcomes of this suggested solution emphasize the significance of testability, especially during the

systems analysis and design stage of software development. One drawback of this process is its narrow focus, which may limit the understanding of software quality improvement. While testability is undeniably crucial in software development, it represents only a single facet of the broader concept of software quality. By primarily emphasizing testability, other vital factors that contribute to software quality, such as maintainability, scalability, performance, security, and usability may be overlooked.

R. Jamwal et al. [19] explored the increasing importance of software quality in today's marketplace and how it differs for producers and users of software. Software users view software as a tool to support their specific business sector, while quality is seen as a combination of various characteristics. These characteristics are often represented in models that depict their relationships, helping to highlight what people consider important in terms of quality. Different organizations adopt different quality models based on their requirements. This study examines various concepts of software quality appearances and provides a relative analysis of different software quality models used by different organizations.

3. METHODOLOGY

3.1. *Dataset:*

The researcher utilized the PROMISE Repository dataset in this study. Specifically, the NASA Metrics Data Program (MDP) provided the NASA PROMISE datasets, which include software metrics. Before training the ensemble model, multiple preprocessing tasks such as data normalization and transformation were conducted. These preprocessing operations involved addressing missing values and outliers, as well as selecting suitable feature variables for the input model. Subsequently, the proposed model was trained to forecast software flaws. Table 1 illustrates the datasets used for the experiment.

Table 1: ***Illustrated the Datasets Used for Experiment.***

Sr. No.	Dataset Name	Total Attributes	Total Instances	Fault Percentage
1.	CM1	24	500	8.98
2.	MC1	37	9464	1.1
3.	MW1	36	400	6.95
4.	PC1	24	1111	7.1
5.	MC2	43	163	30.90

3.2. *Design:*

In this study, a new design architecture is proposed which is mentioned in Figure 1. The developed EEM model aims to empirically assess and confirm the effectiveness of the suggested methods. The performance of this EEM model is evaluated on five NASA PROMISE repository datasets. The PROMISE repository dataset is a collection of software engineering datasets that are publicly available and widely used for research purposes. PROMISE stands for "Project Repository for Object-Oriented Software Engineering," and it aims to provide a comprehensive and diverse set of datasets related to software engineering tasks.

This entire dataset is initially preprocessed for training and testing purposes. Preprocessing refers to the techniques applied to raw data before it is used for analysis tasks. It involves transforming the data into a suitable format and preparing it for further analysis or modeling. Data

preprocessing is crucial as it helps to improve the quality of the data like handling missing values outlier detection and treatment feature scaling or normalization, feature selection or extraction, handling categorical variables, addressing issues or inconsistencies, and enhancing the performance of downstream tasks. In the next stage, preprocessed data are gone under hierarchical clustering which is a data analysis technique that groups similar data points into clusters in a hierarchical manner, forming a dendrogram a kind of tree-like structure. Each data point begins as its cluster and is eventually combined with others using a bottom-up technique based on their similarities, creating a hierarchy of clusters. Furthermore, the K-Fold Cross Validation is opted for datasets splitting into separate train as well as test groups.

The K-Fold Cross Validation method divides the dataset into two groups i.e., 80% data are train data, and the rest 20% data are test data. The train data in an ensemble model pertains to the dataset utilized for training the individual base models in the ensemble. It encompasses input samples along with their associated target values or labels. Each base model is trained separately using a subset of the training data, and their predictions are amalgamated to generate a final prediction. The testing data in an ensemble model refers to a separate dataset that is used to appraise the performance and generalization of the ensemble model. It consists of input samples, similar to the training data, but without the corresponding target values or labels. The testing data is used to determine how effectively the ensemble model can categorize or forecast brand-new data. The predictions of the individual base models are combined in the ensemble model to make predictions on the testing data and evaluate its accuracy and effectiveness. Furthermore, multiple algorithms which consist of Decision Tree (DT), AdaBoost, Random Forest (RF), and Naive Bayes (NB) are used with the help of a test model and pre-trained model to compute the performance. In the other direction, the Enhanced Ensemble Model which is the combination of the Extreme Learning Machine (ELM), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM), is performed to compute the performance with the help of the test model and pre-trained model.

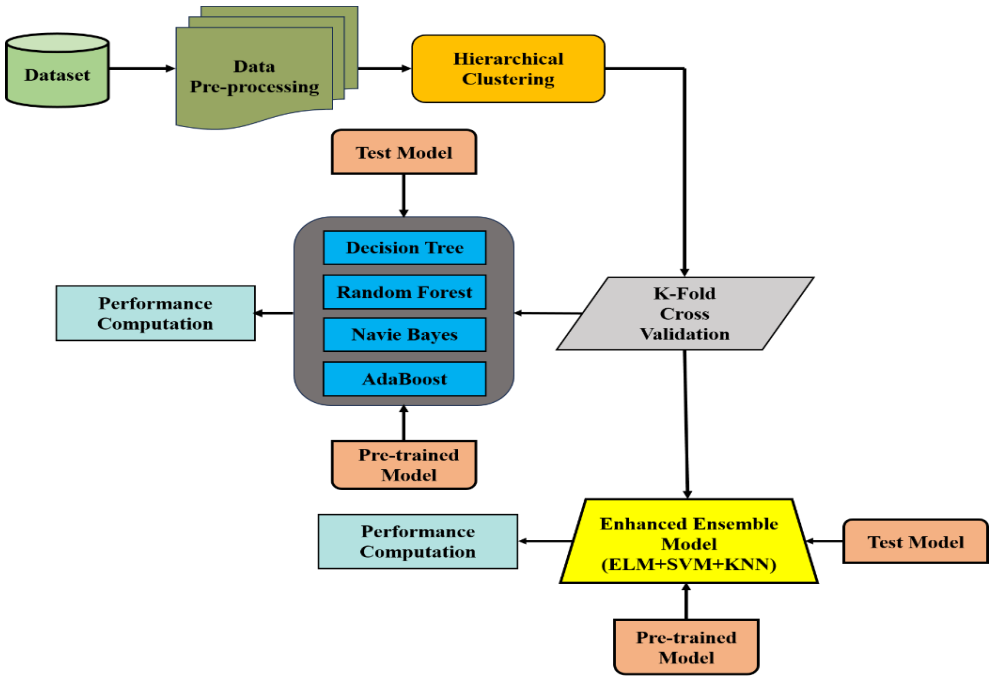


Figure 1: *Design Architecture of Proposed Model.*

3.3. *System Configuration:*

This section introduces the system configuration which is used in this research. For performing this research, there is utilized the latest system configuration. Table 2 represents the computer configuration and software details.

Table 2: ***Represents the System Configuration.***

Sr. No.	System Configuration	Details
1.	System RAM	8 GB
2.	Operating System	Windows 11
3.	Processor	13th Generation intel core i5
4.	Optimizer	Adam v1.3
5.	Programming Language	Python v3.11.0

3.4. *Performance Matrix:*

The four distinct performance measurement metrics, accuracy, recall, precision, and F1-score, are described in this section. These matrices are utilized to evaluate the effectiveness of this suggested EEM along with other algorithms on multiple NASA repository datasets.

3.4.1. *Precision*:

Precision is a metric used in ensemble learning machines to evaluate the accuracy of the ensemble model's successful predictions. It determines the proportion of true positives cases that were properly identified as positive to the total of true positives as well as false positives cases that were identified as positive. Precision aids in assessing the ensemble model's accuracy in locating positive examples. To fully comprehend the performance of the ensemble model in classification tasks must be taken into account in addition to other performance indicators.

$$\text{Precision}=TP/(FP+TP) \quad (1)$$

3.4.2. *Recall:*

The performance of an ensemble model to identify every positive case is measured by the recall metric. It aims to reduce false negatives cases in which a positive event is arbitrarily labeled as negative. The ratio of true positives (positive cases that were properly identified) to the total of true positives as well as false negatives is used to determine recall.

$$\text{Recall}=TP/(FN+TP) \quad (2)$$

3.4.3. *F1-Score:*

The harmonic mean of precision as well as recall that measures the proportion of positively predicted instances out of all positively predicted instances and measures the proportion of positively predicted instances out of all positively actualized instances, which is used to compute the F1-score. The F1-score is appropriate for unbalanced datasets where the distribution of classes is skewed since it takes the harmonic mean, which equalizes accuracy and recall.

$$\text{F1-Score}=(2\times\text{Recall}\times\text{Precision})/(\text{Recall}+\text{Precision}) \quad (3)$$

3.4.4. *Accuracy:*

Accuracy matrix refers to the measurement utilized for determining which model is better to

identify the relationships as well as patterns amongst the variables in a dataset rooted on the input, or training datasets.

$$Accuracy=(TN+TP)/(TN+TP+FN+FP) \quad (4)$$

4.RESULT AND DISCUSSION.

In this research, a novel EEM model is developed that is robust and performs the verification as well as validation in the software product development process. This model analyzes and improves selective parameters i.e., functional suitability along with software maintainability. The EEM model obtains maximum accuracy on diverse datasets of the NASA repository and performs better in comparison to other algorithms such as Decision Trees and many others for validation and verification of the software products. In addition to this, the proposed enhanced ensemble model analyzes the different selective parameter that is maintainability and functional suitability in a more effective manner. The analysis of software quality is done using an improved ensemble model based on selective these parameters. The EEM model precisely assesses the software system's quality and function suitability to improve the effectiveness of SDLC factors related to software quality, such as code complexity, maintainability, and performance. The EEM integrates ELM, SVM, and KNN to offer a thorough assessment of software quality. According to the analysis, this method provides pragmatic outcomes to identify possible problems or opportunities for software product improvement as well as improve software development processes.

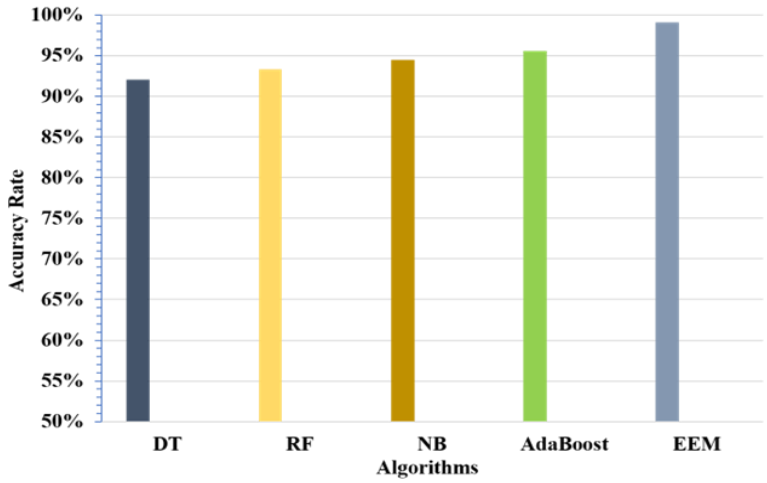


Figure 2*: Graphical Representation of Accuracy Rate.*

The graphical representation of different algorithms accuracy is shown in Figure 2. The accuracy rate has been computed and compared with the EEM approach. In this research, a performance comparison is demonstrated by implementing Decision Tree, Random Forest, Naive Bayes, and AdaBoost and compared with EEM. It is shown that the accuracy rate of Decision Tree is 92.8%, the accuracy measured on Random Forest is 93.35%, the accuracy on Naive Bayes is 94.50% and AdaBoost obtains an accuracy of 95.60%. However, the EEM obtains a maximum accuracy rate which is 99.13%. Hence, EEM is more efficient as compared to all the algorithms and offers more accurate outcomes on diverse datasets.

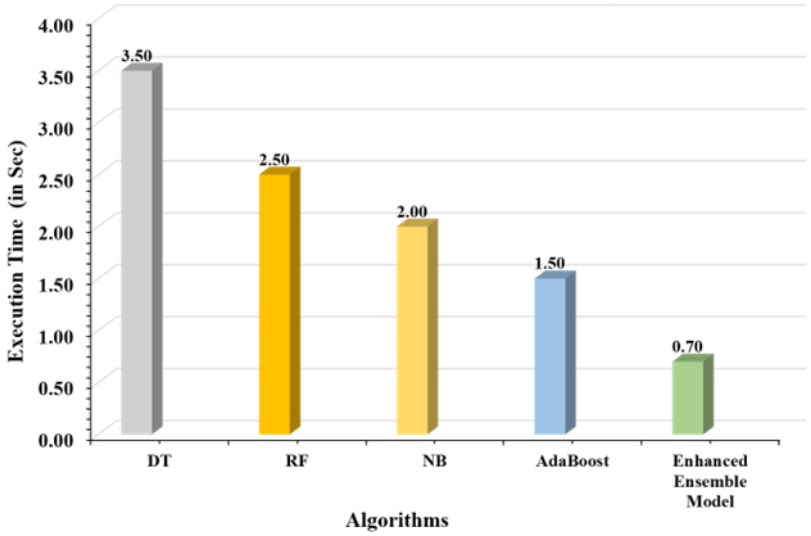


Figure 3: *Illustrates graphical representation of algorithms and their time of execution.*

Figure 3 illustrates a graphical representation of algorithms and their time of execution. While implementing the time taken by the Decision Tree, Random Forest, Navie Baye, and AdaBoost algorithm was computed at 3.50s, 2.50s, 2.00s, 1.50s, and 0.70s, respectively. The proposed EEM approach is executed with the lowest time i.e., 0.70s only. This demonstrates that this EEM is less complex in terms of time complexity and handles large data effectively in training and testing of the software product or applications.

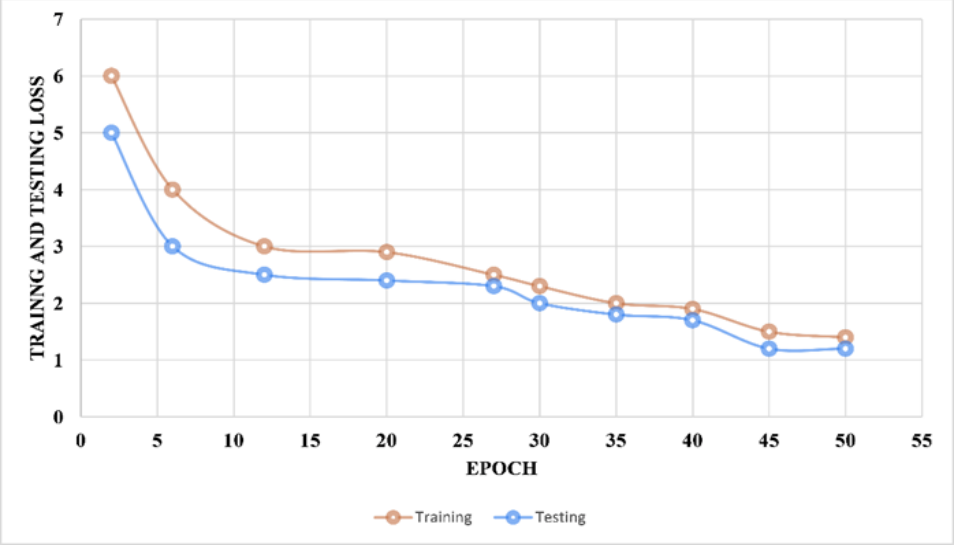


Figure 4: *Shows training as well as testing loss comparison.*

Figure 4 shows training as well as testing loss comparison of the proposed EEM approach. The computed training loss on epochs 2, 6, 12, 20, 27, 30, 35, 40, 45, and 50 is 6%, 4%, 3%, 2.9%, 2.5%, 2.3%, 2%, 1.9%, 1.5%, and 1.4%. The computed testing loss on epochs 2, 6, 12, 20, 27, 30, 35, 40, 45, and 50 is 5%, 3%, 2.5%, 2.4%, 2.3%, 2%, 1.8%, 1.7%, 1.2%, and 1.2%. Nevertheless, the proposed EEM approach performs well, and minimal train and test loss are seen.

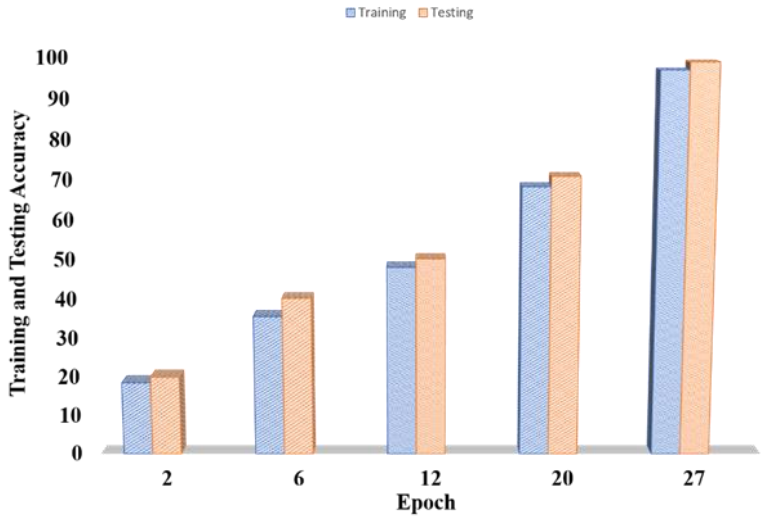


Figure 5: *Shows training as well as testing accuracy comparison.*

The training and testing accuracy of all algorithms along with the EEM approach are displayed in Figure 5. The training accuracy on epochs 2, 6, 12, 20, and 27 are 18.5%, 35.6%, 48.12%, 68.34%, and 97.01%. However, the test accuracy is found 19.98%, 40.22%, 50.25%, 70.91%, and 98.95%, respectively. It is shown that EEM achieves optimal accuracy level and is more suitable for software product quality analysis for diverse selective parameters such as functional suitability of the software product and maintainability. In addition to this, the EEM archives good accuracy for the software products and applications validation and verification.

Table 3: *Illustrates the Average Computed Accuracy of Different Algorithms.*

S. No.	Dataset Name	DT	AdaBoost	NB	RF	EEM
1.	CM1	0.89	0.88	0.90	0.91	0.98
2.	KC1	0.90	0.87	0.89	0.93	0.97
3.	MW1	0.91	0.92	0.91	0.94	0.99
4.	PC1	0.86	0.91	0.93	0.92	0.98
5.	MC2	0.88	0.89	0.89	0.95	0.99
Average		0.88	0.89	0.90	0.93	0.99

Table 3 illustrates the average computed accuracy of different algorithms along with EEM. The EEM and diverse algorithms are validated utilizing the NASA repository's five datasets namely, CM1, KC1, MW1, PC1, and MC2, respectively. The obtain accuracy using the decision tree algorithm on CM1, KC1, MW1, PC1, and MC2 are 0.89%, 0.90%, 0.91%, 0.86%, and 0.88%, respectively. The obtain accuracy using the AdaBoost algorithm on CM1, KC1, MW1, PC1, and MC2 are 0.88%, 0.87%, 0.92%, 0.91%, and 0.89%, respectively. The accuracy received using the NB algorithm on CM1, KC1, MW1, PC1, and MC2 are 0.90%, 0.89%, 0.91%, 0.93%, and 0.89%,

respectively. The accuracy received using the RF algorithm on CM1, KC1, MW1, PC1, and MC2 are 0.91%, 0.93%, 0.94%, 0.92%, and 0.95% respectively. The obtained accuracy using the EEM algorithm on CM1, KC1, MW1, PC1, and MC2 are 0.98%, 0.97%, 0.99%, 0.98%, and 0.99% respectively. The average accuracy on DT, AdaBoost, NB, RF, and EEM is found 0.88%%, 0.89%, 0.90%, 0.93%, and 0.99%, respectively.

5.CONCLUSION

Software quality analysis is becoming a very challenging and time-consuming process in modern software product and application development. There are developed multiple software testing approaches for effective analysis of the prediction of the faults in applications. However, these testing methods are inconsistent to determine the software product quality, especially in the developing phase, and consume a huge time for the developer in the unit testing process to determine the functional suitability and validation and verification of the product effectively. Therefore, this work presents an EEM approach that is based on the SVM, ELM, and KNN for software product and application quality analysis in less time. The obtained results are recorded and compared with the other algorithms i.e., DT, AdaBoost, NB, and RF for validation of the EEM. To validate the proposed EEM approach, there has been used the five diverse datasets of the PROMISE repository of NASA which includes the CM1, KC1, MW1, PC1, and MC2, respectively. It is found that the average accuracy on Decision Tree, Random Forest, Naive Bayes, AdaBoost, and EEM is 92.08%, 93.35%, 94.50%, 95.60%, and 99.14%, respectively.

References

[1] S. Alyahya, “Crowdsourced software testing: A systematic literature review,” Information and Software Technology. 2020. doi: 10.1016/j.infsof.2020.106363.

[2] D. Manikavelan and R. Ponnusamy, “Software quality analysis based on cost and error using fuzzy combined COCOMO model,” J. Ambient Intell. Humaniz. Comput., 2020, doi: 10.1007/s12652-020-01783-9.

[3] E. Ronchieri, M. Canaparo, and D. Salomoni, “A software quality model by using discriminant analysis predictive technique,” J. Integr. Des. Process Sci., 2014, doi: 10.3233/jid-2014-0016.

[4] OM Melkozerova “Software performance testing,” Bull. V.N. Karazin Kharkiv Natl. Univ. Ser. Mathematical Model. Inf. Technol. Autom. Control Syst., 2020, doi: 10.26565/2304-6201-2020-45-07.

[5] D. El-Masri, F. Petrillo, Y. G. Guéhéneuc, A. Hamou-Lhadj, and A. Bouziane, “A systematic literature review on automated log abstraction techniques,” Inf. Softw. Technol., 2020, doi: 10.1016/j.infsof.2020.106276.

[6] P. Temple, M. Acher, and J. M. Jezequel, “Empirical Assessment of Multimorphic Testing,” IEEE Trans. Softw. Eng., 2021, doi: 10.1109/TSE.2019.2926971.

[7] I. Atoum et al., “Challenges of Software Requirements Quality Assurance and Validation: A Systematic Literature Review,” IEEE Access. 2021. doi: 10.1109/ACCESS.2021.3117989.

[8] M. Varela-González, H. González-Jorge, B. Riveiro, and P. Arias, “Performance testing of LiDAR exploitation software,” Comput. Geosci., 2013, doi: 10.1016/j.cageo.2012.12.001.

[9] Z. Ding, J. Chen, and W. Shang, “Towards the use of the readily available tests from the release pipeline as performance tests,” in Proceedings - International Conference on Software Engineering, 2020. doi: 10.1145/3377811.3380351.

[10] S. M. M. A. Jammel Mona, Radhwan Hussein Abdulzhraa Al-Sagheer, “Software Quality Assurance Models and Application to Defect Prediction Techniques,” 2022, [Online]. Available:

https://ijisae.org/index.php/IJISAE/article/view/2455/1038

[11] E. A. Rashid, rikanta B. Patnaik, and V. C. Bhattacherjee, "Machine Learning and Software Quality Prediction: As an Expert System," Int. J. Inf. Eng. Electron. Bus., vol. 6, no. 2, pp. 9–27, Apr. 2014, doi: 10.5815/ijieeb.2014.02.02.

[12] T. Sharma, A. Jatain, S. Bhaskar, and K. Pabreja, "Ensemble Machine Learning Paradigms in Software Defect Prediction," Procedia Comput. Sci., vol. 218, pp. 199–209, 2023, doi: 10.1016/j.procs.2023.01.002.

[13] K. Bashir, T. Li, C. W. Yohannese, and Y. Mahama, "Enhancing software defect prediction using supervised-learning based framework," in 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Nov. 2017, pp. 1–6. doi: 10.1109/ISKE.2017.8258790.

[14] M. S. Rawat and S. K. Dubey, "Software defect prediction models for quality improvement: A literature study," Int. J. Comput. Sci. Issues, 2012.

[15] S. Yamada, "Software quality/reliability measurement and assessment: software reliability growth models and data analysis," J. Inf. Process., 1991.

[16] V. K. L. Srivastava, "An efficient Software Source Code Metrics for Implementing for Software quality analysis," Int. J. Emerg. Trends Eng. Res., pp. 216–222, Sep. 2019, doi: 10.30534/ijeter/2019/01792019.

[17] E. B. Belachew, "Analysis of Software Quality Using Software Metrics," Int. J. Comput. Sci. Appl., vol. 8, no. 4/5, pp. 11–20, Oct. 2018, doi: 10.5121/ijcsa.2018.8502.

[18] I. Khan and A. K. K. Shinwari, "Testability as a Measure for Improving Software Quality in System Analysis and Design," Kardan J. Eng. Technol., Dec. 2019, doi: 10.31841/KJET.2021.5.

[19] R. S. Jamwal, D. Jamwal, and D. Padha, "Comparative Analysis of Different Software Quality Models," Comput. Nation Dev., 2009.

OPTIMIZATION OF SYSTEM PARAMETERS OF 2: 3 GOOD SERIAL SYSTEM USING DEEP LEARNING

Shakuntla Singla[1], Shilpa Rani[2*], Umar Muhammad Modibbo[3], Irfan Ali[4]

•

[1,2] Department of Mathematics and Humanities, MMEC, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, India
Shakus25@gmail.com[1] gargshilpa46@gmail.com[2]

•

[3]Department of Operations Research,
Modibbo Adama University, P.M.B.2076, Yola Nigeria.
Email: umarmodibbo@mau.edu.ng

•

[4]Department of Statistics and Operations Research,
Aligarh Muslim University, Aligarh 202002, UP, India.
Email: irfii.st@amu.ac.in
*Corresponding author**:** gargshilpa46@gmail.com[2]

Abstract

In this paper, Optimization of System Parameters of 2:3 Good Serial System controlled with the help of a controlling unit Using Deep Learning Optimization with packing unit in series with priority in repair and single server which never fails is carried out. There are three units of different capacities working in parallel in which if three/two units are working then the system is working at full/reduced capacity. Working of these online parallel and offline units is managed by the controlling unit, which also manages the preventive maintenance of all type of units, together with a 24/7 repair facility is modeled for reliability performance measurements. Taking exponential failure and repair rates of units and facilities a steady state transition diagram (depicting transition rates and states) is drawn using Markov process. The system parameters are modelled using Regenerative Point graphical Technique (RPGT) and optimized using Deep learning methods such as Adam, SGD, RMS prop. The results of the optimization may be used to validate and challenge existing models and assumptions about the systems.

Keywords: Optimization, RPGT, Deep learning, Adam, SGD, RMS prop

I. Introduction

In this paper, Optimization of System Parameters of 2:3 Good Serial System Using Deep Learning Optimization with packing unit in series, with priority in repair managed by a controlling unit and single server who never fails is carried out. There are three main dissimilar units of different capacities performing different jobs working in parallel, together with the terminal unit (usually packing unit) controlled by a controller unit for allocation and preventive maintenance of all units in the system, together with single server who never fails is carried out. Out of three parallel units if two units are working then the system is working at reduced capacity and failed unit is maintained by the controller, with priority in repair to controlling unit performing preventive maintenance and

single server is conducted using Markov process, RPGT and Deep Learning optimization methods. Preventive maintenance is feasible only in the initial state when controlling facility is almost free. The system measures can be optimized with proper utilization of maintenance activities of any system, in most of the studies under taken so far by most of the earlier authors, it is assumed that system once installed for operation will continue to do so, but practically it is not so, sometimes, it is necessary to manage the operation of working unit and maintenance facilities with the help of some controlling unit or external supporting system which may also fail (or non-available), on failure of one or more units or controlling unit, there is need of repair facility.

In this paper two out of three units in which if three/two units are working then the system is working at full/reduced capacity. Working of these online units is managed by controlling which also manages for the preventive maintenance of all type of units, together with a repair facility modeled for reliability performance measurements. Taking exponential failure and repair rates of units, a steady state transition diagram (depicting transition rates and states) is drawn using Markov process. Various directed paths, primary, secondary, tertiary circuits, base state, simple paths w. r. t. initial and base states enumerated. Various path probabilities transition probabilities mean sojourn times and expressions for four reliability measures are modeled using RPGT, optimization of system parameters using deep learning methods such as Adam, SGD, RMS prop is provided, which gives valuable insights into the factors that affect system performance by drawing corresponding tables and graphs followed by discussion.

II. Assumptions and Notations

- The repair procedure arises soon after a unit flops.
- Repaired unit is as if a new one.
- Failure/repair rates of units are exponential
- Server facility is 24x7 hours.
- B = controlling unit
- A_i, B, D/ a_i, b, d – Working state / failed state of individual units. ($1 \le i \le 3$)
- m_1/ h_1 – represent failure /repair rates of units A_i ($1 \le i \le 3$), m_2/ h_2 = those of controlling unit, m_3/ h_3 = those of packing unit D,
- α, β- transition rates of controlling unit performing preventive maintenanceμ_i' waiting time for repair facility to arrive
- μ_i' waiting time for repair facility to arrive
- ($i \rightarrow$ j $\rightarrow$ k $\rightarrow$ i) cycle
- D = Packing unit
- $Repair\ Priority\ order\ is\ D > B >$ Ai
- ◯ : Full Capacity Working State
- ⬭ : Reduced Capacity Working State
- □ : Failed State

III. Transition Diagram Description

Taking the transition failure and repair rates, the system may be stable in the states S_i ($0 \le i \le 10$) .

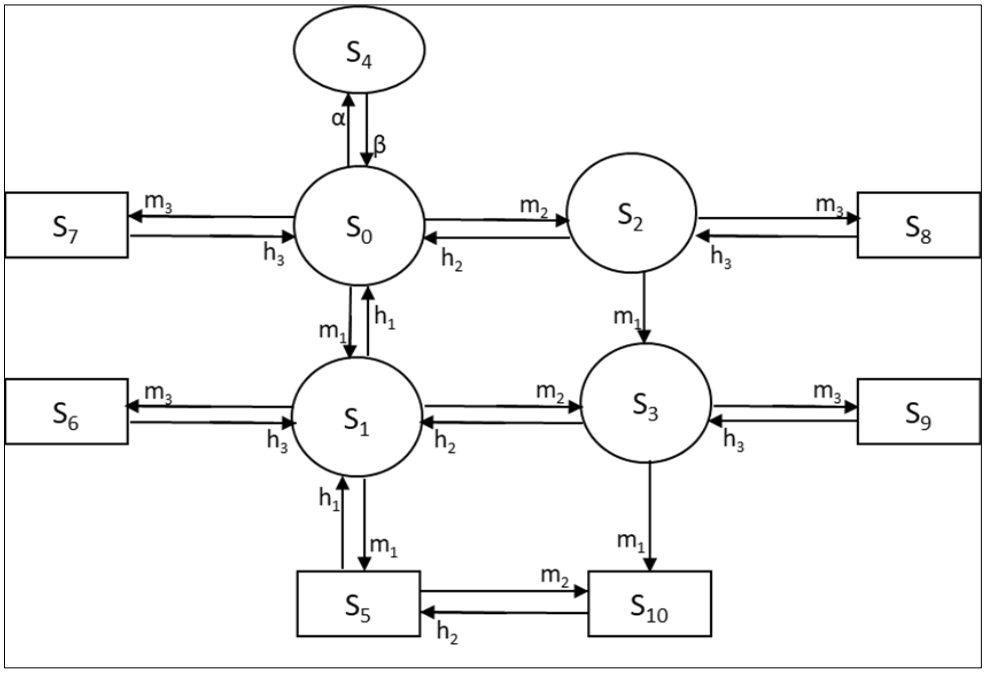


Figure 1: *Transition Diagram*

S_0 = $A_1A_2(A_3)$ BD, S_1 = $a_1A_2A_3$BD, S_2 = $A_1A_2(A_3)$ Bd, S_3 = $a_1A_2A_3$Bd,
S_4(under preventive maintenance) = $A_1A_2(A_3)$ BD,
S_5 =$a_1a_2A_3$BD, S_6 = $a_1A_2A_3$bD, S_7 = $A_1A_2A_3$bD, S_8 = $A_1A_2A_3$ bd, S_9 = $a_1A_2A_3$bd, S_{10} = $a_1a_2A_3$Bd

I. State Transition Probabilities

$q_{i\to j}(t)$
$q_{0\to1}(t) = m_1 e^{-(m_1+m_2+m_3+\alpha)t}$
$q_{0\to2}(t) = m_2 e^{-(m_1+m_2+m_3+\alpha)t}$
$q_{0\to4}(t) = \alpha e^{-(m_1+m_2+m_3+\alpha)t}$
$q_{0\to7}(t) = m_3 e^{-(m_1+m_2+m_3+\alpha)t}$
$q_{1\to0}(t) = h_1 e^{-(m_1+m_2+m_3+h_1)t}$
$q_{1\to3}(t) = m_2 e^{-(m_1+m_2+m_3+h_1)t}$
$q_{1\to5}(t) = m_1 e^{-(m_1+m_2+m_3+h_1)t}$
$q_{1\to6}(t) = m_3 e^{-(m_1+m_2+m_3+h_1)t}$
$q_{3\to1}(t) = q_{2\to0}(t) = h_2 e^{-(m_1+m_3+h_2)t}$
$q_{2\to3}(t) = m_1 e^{-(m_1+m_3+h_2)t}$
$q_{2\to8}(t) = q_{3\to9}(t) = m_3 e^{-(m_1+m_3+h_2)t}$
$q_{3\to10}(t) = m_1 e^{-(m_1+m_3+h_2)t}$
$q_{4\to0}(t) = \beta e^{-\beta t}$
$q_{5\to1}(t) = h_1 e^{-(m_2+h_1)t}$
$q_{5\to10}(t) = m_2 e^{-(m_2+h_1)t}$
$q_{6\to1} = q_{7\to0} = q_{8\to2} = h_3 e^{-h_3 t}$
$q_{10\to3} = 0$
$q_{10\to5} = h_2 e^{-h_2 t}$
$p_{i\to j} = q^*_{i\to j}(0)$
$p_{0\to1} = m_1/(m_1+m_2+m_3+\alpha)$
$p_{0\to2} = m_2/(m_1+m_2+m_3+\alpha)$
$p_{0\to4} = \alpha/(m_1+m_2+m_3+\alpha)$
$p_{0\to7} = m_3/(m_1+m_2+m_3+\alpha)$
$p_{1\to0} = h_1/(m_1+m_2+m_3+h_1)$
$p_{1\to3} = m_2/(m_1+m_2+m_3+h_1)$
$p_{1\to5} = m_1/(m_1+m_2+m_3+h_1)$
$p_{1\to6} = m_3/(m_1+m_2+m_3+h_1)$

$p_{2\to0}= h_2/ (m_1+m_3+h_2)$
$p_{3\to10} = p_{2\to3}= m_1/ (m_1+m_3+h_2)$
$p_{3\to9} = p_{2\to8}= m_3/ (m_1+m_3+h_2)$
$p_{3\to1}= h_2/ (m_1+m_3+h_2)$
$p_{2\to0}= h_2/ (m_3+h_2)$
$p_{5\to10}= m_2/ (m_2+h_1)$
$p_{4\to0} = p_{6\to1}= p_{7\to0}= p_{8\to2}=p_{9\to3}= p_{10\to8}= 1$
$p_{0\to1}+p_{0\to2}+p_{0\to4}+p_{0\to7}=p_{1\to0}+p_{1\to3}+p_{1\to5}+p_{1\to6} = p_{2\to0}+p_{2\to3}+p_{2\to8}=p_{3\to1}+p_{3\to9}+p_{3\to10} = 1$

II. Mean Sojourn Times $\mu_i=R_i^*(0)$

$R_i(t)$
$R_0(t)= e^{-(m_1+m_2+m_3+\alpha)t}$
$R_1(t)= e^{-(m_1+m_2+m_3+h_1)t}$
$R_2(t)= e^{-(m_3+h_2)t}$
$R_3(t)= e^{-(m_1+m_3+h_2)t}$
$R_4(t)= e^{-\beta t}$
$R_5(t)= e^{-(m_2+h_1)t}$
$R_6(t)=R_7(t)=R_8(t)=R_9(t)= e^{-h_3t}$
$R_{10}(t)= e^{-h_2t}$
$\mu_0 = 1/ (m_1+m_2+m_3+\alpha)$
$\mu_1= 1/ (m_1+m_2+m_3+h_1)$
$\mu_2= 1/ (m_3+h_2)$
$\mu_3= 1/ (m_1+m_3+h_2)$
$\mu_4= 1/\beta$
$\mu_5= 1/ (m_2+h_1)$
$\mu_6= \mu_7= \mu_8= \mu_9= 1/h_3$
$\mu_{10}= 1/h_2$

III. Evaluation of Transition Path Probabilities:

Applying RPGT and '0' as the initial state of the system as under: The transition likelihood factors of altogether the reachable states after the initial state 'ξ' = '0' are:
$V_{0\to0} = 1$
$V_{0\to1,} = p_{0\to1}/\{(1-p_{1\to3}p_{3\to1})/(1-p_{3\to9}p_{9\to3})\}\{(1-p_{3\to10}p_{5\to10}p_{5\to1})/(1-p_{10\to5}\ p_{5\to10})\}(1-p_{1\to6}p_{6\to1})$
$(1-p_{1\to5}p_{5\to1})$
$V_{0\to2} = p_{2\to0}=/(1-p_{2\to8}p_{8\to2})$
$V_{0\to3} =$ ....., *Continue*
The transition likelihood factors of completely the reachable states since the base state 'ξ' = '1' are Probabilities since state '1' to dissimilar vertices stand given as
$V_{1\to0} = p_{1\to0}/\{(1-p_{0\to2}p_{2\to0})/(1-p_{2\to8}p_{8\to2})\}(1-p_{0\to4}p_{4\to0})(1-p_{0\to7}p_{7\to0})$
$V_{1\to1}= 1$
$V_{1\to2}= p_{0\to2}/\{(1-p_{0\to2}p_{2\to0})/(1-p_{2\to8}p_{8\to2})\}(1-p_{0\to4}p_{4\to0})(1-p_{0\to7}p_{7\to0})(1-p_{2\to8}p_{8\to2})$
$V_{1\to3}=$ ....., *Continu*

IV. Modeling system parameters

I.MTSF (F_0):

The states to which the arrangement can transit from original state '0', before joining down state are: 'i' = 0 to 4
$F_0 = (V_{0\to0}\ \mu_0+V_{0\to3}\ \mu_3+V_{0\to2}\ \mu_2+V_{0\to1}\ \mu_1+V_{0\to4}\ \mu_4)/[\{1-(0\to1\to0)-(0\to2\to0)-(0\to4\to0)\}]$
$= m_1^2h_1\alpha/m_1+m_2+m_3+\alpha^2)(3h_1+2\alpha+m_2+m_3)$

II. Availability (Y_0):

The reformative states at which the organization is accessible are 'j' = 0 to 4 and the reformative states are 'i' = 0 to 8

$$Y_0 = \left[\sum_j V_{\xi\to j}, f_j, \mu_j\right] \div \left[\sum_i V_{\xi\to i}, f_j, \mu_i^1\right]$$
$$= (V_{1\to1}\ \mu_1 + V_{1\to2}\ \mu_2 + V_{1\to3}\ \mu_3 + V_{1\to4}\ \mu_4)/D_1$$
$$\text{Where } D_1 = V_{1\to0}\ \mu_0 + V_{1\to1}\ \mu_1 + V_{1\to4}\ \mu_4 + V_{1\to3}\ \mu_3 + V_{1\to2}\ \mu_2 + V_{1\to5}\ \mu_5 + V_{1\to8}\ \mu_8 + V_{1\to7}\ \mu_7 + V_{1\to6}\ \mu_6 + V_{1\to9}\ \mu_9 + V_{1\to10}\ \mu_{10}$$

$$Y_0 = m_1 m_2 / (m_1 + 3m_2 + \alpha + h_1)(3m_1 + m_2 + m_2^2 + \alpha h_1 + 3m_2)$$

III. Busy Period of the Server (H_0) :

The states where the attendant is busy for doing some jobs are 'i' = 1 to 8, taking 'ξ' = '0', using RPGT busy period is given as

$$H_0 = \left[\sum_j V_{\xi\to j}, n_j\right] \div \left[\sum_i V_{\xi\to i}, \mu_i^1\right]$$

$$\text{Where } D = V_{0\to0}\ \mu_0 + V_{0\to4}\ \mu_4 + V_{0\to2}\ \mu_2 + V_{0\to6}\ \mu_6 + V_{0\to1}\ \mu_1 + V_{0\to8}\ \mu_8 + V_{0\to3}\ \mu_3 + V_{0\to7}\ \mu_7 + V_{0\to5}\ \mu_5 + V_{0\to9}\ \mu_9 + V_{0\to10}\ \mu_{10}$$
$$H_0 = [m_1/(m_1 + m_2 + m_3 + \alpha)^2 + m_2(m_3 + h_2) + 1/(m_1 + m_2 + m_3 + \alpha) + (3m_1 + m_2^2)(m_1 + m_3 + 3h_1) + h_2/(m_1 + m_3 + h_1).\ h_1/(h_1 + m_2) + 3h_1/(m_1^2 + h_3) + m_2 h_2(h_2 + 3m_1 + h_1 + \alpha)$$

IV. Expected Number of Examinations by the repair man T_0:

The reformative states where the repair person appointments afresh are j = 1, 2, 4, 7 the reformative states are i = 0 to 8, Taking 'ξ' = '0',

$$T_0 = \left[\sum_j V_{\xi\to j}\right] \div \left[\sum_i V_{\xi\to i}, \mu_i^1\right]$$
$$= (V_{0\to4} + V_{0\to2} + V_{0\to1} + V_{0\to0})/D$$
$$T_0 = [(m_1 + m_2)/(m_1 + m_2 + m_3 + \alpha) + \alpha m_3(m_1 + m_2 + m_3 + \alpha)(m_1 + m_2) + m_1 m_3(m_1 + m_2 + m_3 + \alpha)(m_1 + m_2 + m_3 + h_1)(m_1 + m_3 + h_2) + m_1 h_2/(m_1 + m_3 + \alpha).\ (\alpha + h_1 + m_2) + 3\alpha^2/(m_1 + \alpha)(m_3 + h_1)$$

V. Optimization Using Deep Learning methods

Performing a optimization of a repairable using deep learning requires several steps in equation 1, 2, 3 and 4 to include for model to find different parameter [1, 2, 10]. Here is an example experiment that you could perform:

- Collection of data: Gather a dataset that contains information on the input parameters and the system's output. The input parameters could include factors such as the system's design, operating conditions, and maintenance schedule. The output could include metrics such as system availability, downtime, and failure rate in table 1 and table 2.
- Preprocess data: Clean and preprocess the dataset, splitting it into training, validation, and test sets.
- Train the model: Use a deep learning algorithm, such as a neural network, to model the connection among the input parameters and the output. Train the model using the training set and validate it using the set of values in table 1. You could use techniques such as early stopping and regularization to prevent over fitting.
- Appraise the model: After the model is proficient, appraise its performance by means of test set. Estimate metrics such as busy period.
- Perform sensitivity analysis: Using the trained model, vary the values of one parameter at a time while keeping the others constant. Record the effect on the system's output. Repeat this process for each input parameter, recording the impact of each parameter on the system's output.
- Interpret results: Analyze the consequences of the optimization examination to determine which input parameters need the most considerable influence on the system's output. You could use systems such as nose importance and fractional dependence plots to increase understandings into the mockup's behavior. Overall, performing a optimization of a repairable undertaken system using deep learning requires a combination of data collection, preprocessing, model training, and

analysis [3, 10]. It can be a powerful tool (Google Colab, Colab Notebook, Colab Python) for understanding the factors that contribute to the reliability of the system [4, 10]

.

VI. Dataset

Optimization is a way used to study how variations in the input parameters of an organization move the output. In the background of a repairable two out of three good system, optimization can help determine which parameters have the most significant impact on the system's reliability. To perform optimization using deep learning, you would need a dataset that contains information on the input parameters and the system's output [5, 6]. The input parameters could include factors such as the system's design, operating conditions, and maintenance schedule. The output could include metrics such as system availability, MTSF, and busy period

Table 1: *Table of parameter*

W (w1, w2, ---, wn)	$\lambda(\lambda 1, \lambda 2, \ldots, \lambda n)$	*S (s1, s2,---sn)*	*P*
(0-20,21-100)	*(0-30,31-100)*	*(0-100)*	*(0-80)*

Once you have a dataset, you could use a deep learning algorithm to model the relationship among the input parameters and the production. One approach could be to use a neural network, which can learn complex relationships between inputs and outputs. To perform optimization using a neural network, you could first train the network on the dataset, using a portion of the data for training and another portion for validation. Once the network trained, you could use it to make predictions on new input data, varying the values of one parameter at a time while keeping the others constant [7, 10]. By observing how changes in each parameter affect the system's output, you can determine which parameters have the most significant impact on the system's reliability to included dataset Table.1. Overall, optimization using deep learning can be an influential tool for understanding the issues that pay to the reliability of a repairable undertaken system. However, it requires a large and well-curated dataset, as well as expertise in deep learning techniques.

VII. Method

Optimization of a repairable system undertaken for analysis using deep learning typically involves the following steps:

- Data collection: Collect data on the input parameters and output metrics of the system. The input parameters could include factors such as the system's design, operating conditions, and maintenance schedule. The output metrics could include measures such as system availability, MTSF, and busy period in show table 2 included.
- Data preprocessing: Clean and preprocess the data, splitting it into training, validation, and test sets. Normalize the input variables to ensure that they are on the same scale.
- Model selection: Choose appropriate deep learning optimization techniques (Adam, SGD, RMS prop) for the sensitivity analysis. Some options contain feed forward neural systems, convolutional neural systems, and regular neural networks. Consider influences such as the size of the dataset, the difficulty of the input-output connection, and the computational capitals existing.
- Model training: Train the selected model on the training data. Use techniques such as stochastic gradient descent and back propagation to minimize the bust time. Monitor the performance of the model on the validation data, and adjust the hyper parameters as needed.
- Model evaluation: Evaluate the trained model on the test data. Calculate metrics such as mean absolute bust time and mean squared error to assess the model's performance of deep learning optimization in show table 1 and table 2.

- Optimization: Use the trained model to perform optimization on the input parameters. Vary the value of one input parameter at a time while holding the others constant. Record the effect on the output metric of interest. Repeat this process for each participation parameter to determine the of the output metric to changes in each parameter.
- Interpretation of results: Analyze the fallouts of the optimization examination to identify which input limits must the utmost impact on the output metric of interest. Use practices such as article importance and incomplete dependence plots to advance insights into the association amid the input limits and output metric [9, 10, 11].
- Several other methodologies were adopted by different scholars in the literature, some include optimization and estimations for the system parameters in reliability allocation and selective maintenance problems [12-15]. Others considered system availability under preventive maintenance [16] While others concentrated on the analysis of multiple hardware-software with Failure Interaction [17].

Overall, performing optimization of Repairable system under discussion using deep learning involves a combination of data collection, preprocessing, model selection, training, evaluation, and analysis.

Table 2: *Performance of model*

Model	-MTSF	Expected Number of Inspections by the repair man	-Busy Period	Availability
Adam	0.923	.9067	0.8012	0.9345
SGD	0.9123	0.9000	0.8123	0.9123
RMS prop	0.9012	0.8912	0.8103	0.9245

It can be a commanding tool for understanding the influences that underwrite to the reliability of the system.

VIII. Results and discussion

The results and discussion of a Optimization of undertaken repairable system parameters using deep learning will depend on the specific system and dataset analyzed. However, here are general insights that could be gained from such an analysis:

- Identification of critical system parameters: The optimization could reveal which input parameters require the greatest effect on the output metric of interest. For example, it could show that system availability is most optimization to the frequency of care or the quality of the components used in the organization.
- Understanding of the non-linear relationship amongst input strictures and output metrics: The deep learning model used in the analysis can capture non-linear relationships amongst input restrictions and output metrics, which could not detect using traditional statistical methods. The optimization can provide insights into the shape and magnitude of these relationships.
- Validation of existing models and assumptions: The results of the optimization applied to validate or challenge existing models and assumptions about the system. For example, the analysis could show that a certain parameter has a much greater impact on system performance than previously thought.
- Prediction of system behavior under different scenarios: The deep learning model applied to predict system performance under different setups, such as vagaries in operating conditions or

maintenance schedules. This can support decision-makers assess the impact of changed strategies and style informed verdicts.

Overall, Optimization of System Parameters of 2: 3 Good Serial System Using Deep Learning Methods can provide valuable insights into the factors that affect system performance, **(MTSF),** Expected Number of Inspections by the repair man, **Busy Period** and **Availability of the System are shown in figure 2, 3, 4 and 5**.

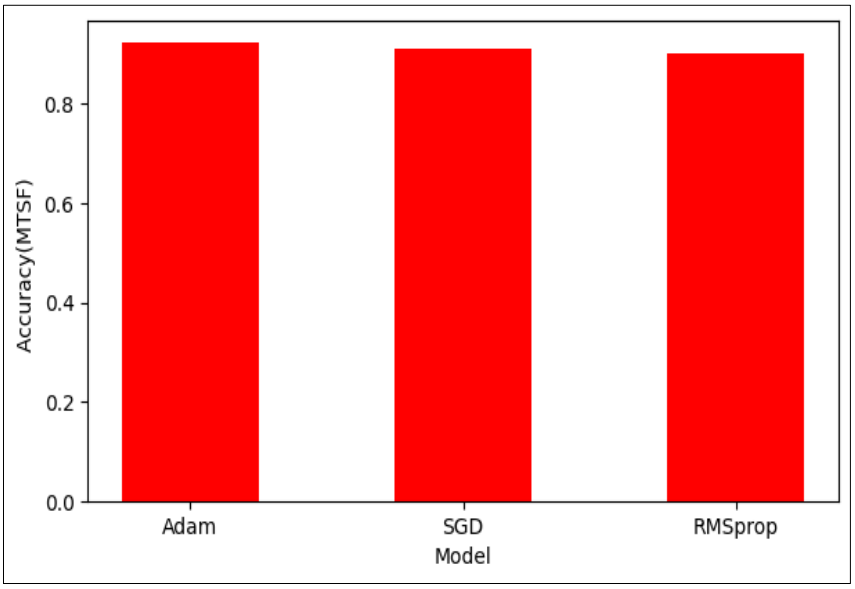


Figure 2: *comparing between models according to MTSF*

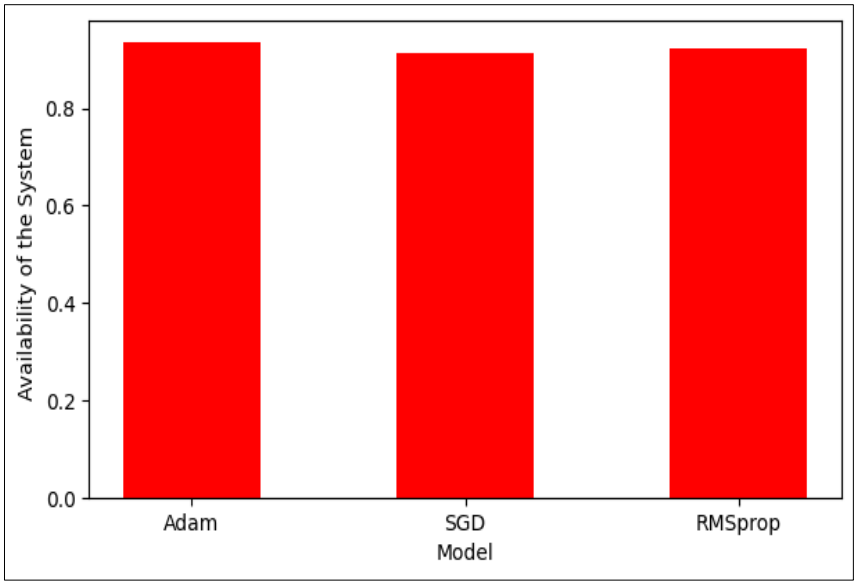


Figure 3: *comparing between models according to Availability*

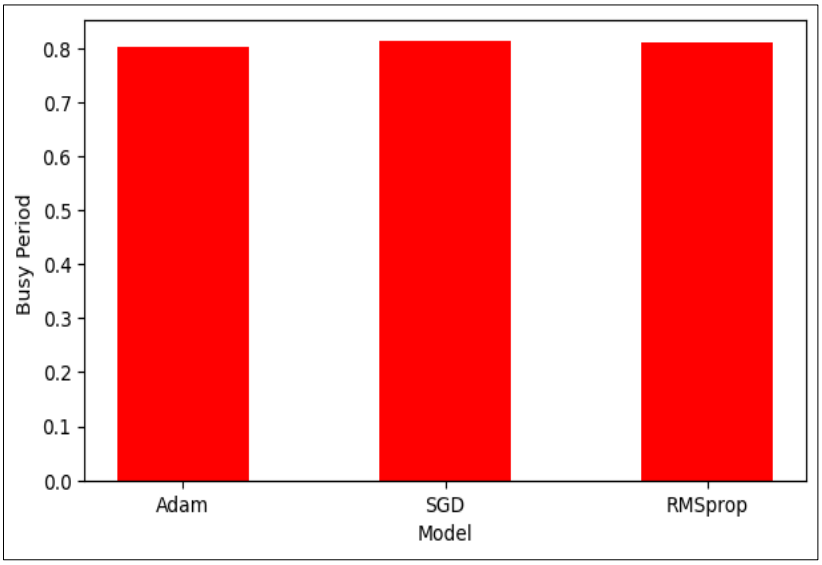


Figure 4: *comparing between models according to busy period*

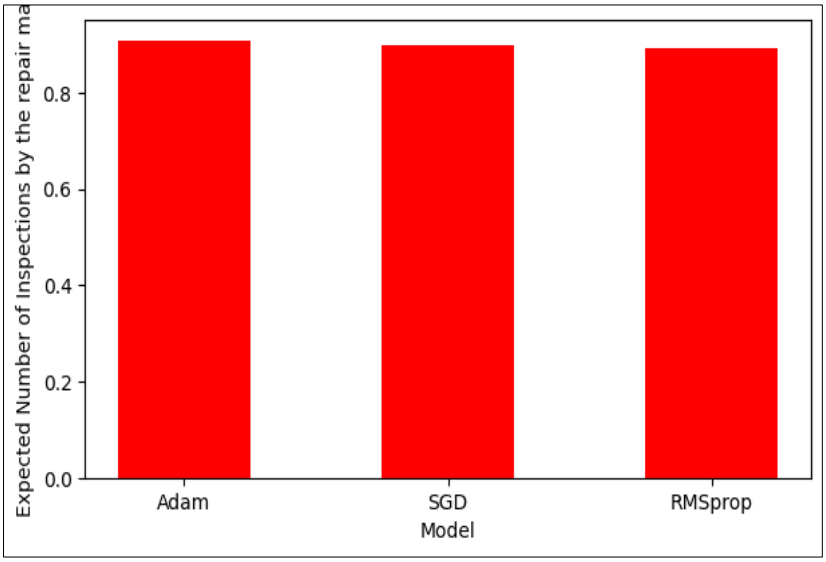


Figure 5: *comparing between models according to Expected Number of Inspections by the repair man*

MTSF between the different model is Adam is best performance among them. And busy time of Adam is better among them of model.

IX. Conclusion

In early research on the repair-replacement problem, the study of the model for a straightforward repairable system was the main focus of the repair-replacement models. However, it appears more logical to predict that the system's subsequent working times after repairs will get shorter and shorter while the system's subsequent repair durations after failure go longer and longer for a simple system that is deteriorating. In this paper, Optimization of System Parameters of 2: 3 Good Serial System Using Deep Learning Methods is carried out, along with external supporting systems for preventive maintenance and a single server that may also fail. The system consists of three similar units, out of which two units are working, at which point the system is operating at full capacity, and the third unit is kept in cold standby, which is switched in with the help of a perfect switch over system. The results of the optimization applied to validate or challenge existing models and assumptions about the system. For example, the analysis could show that a certain parameter has a much greater impact

on system performance than previously thought. It can help optimize maintenance strategies, improve system design, and reduce downtime and maintenance costs.

References

[1] Fries, C., Wiendahl, H.,H. and Assadi, A.A. (2020). Design concept for the intralogistics material supply in matrix productions. Procedia CIRP 2020; 91: 33–8.

[2] Hsieh, Y.,C. and Chen, T.C. (2004). Reliability lower bounds for two-dimensional consecutive-k-out-of-n: F systems. Computers & Operations Research 31(8):1259–72

[3] John, Y. M., Sanusi, A., Yusuf, I., and Modibbo, U., M. (2023). Reliability Analysis of Multi-Hardware–Software System with Failure Interaction. *Journal of Computational and Cognitive Engineering*, *2*(1), 38-46.

[4] Kumar, A. (2020). Reliability And Sensitivity Analysis of Linear Consecutive 2-out-of- 4: F System. *European Journal of Molecular & Clinical Medicine*, *7*(07), 2020

[5] Kumar, A. (2022). Sensitivity Analysis Of Urea Fertilizer Plant. Journal of Reliability Theory and Applications Volume 17, RT&A No. 2 (68)

[6] Kumar, A. (2021) Performance Analysis of the Water Treatment Reverse Osmosis Plant. Journal of Reliability Theory and Applications Volume 16, RT&A No. 3 (63)

[7] Kim, H.K., Lee, S. and Yun, K.S. (2011). Capacitive tactile sensor array for touch screen application. Sensors and Actuators A: Physical 165:2–7.

[8] Koutras, M.V., Padadopoulos, G.K. and Papastavridis, S.G. (1994). Reliability of 2-Dimensional Consecutive-k-out-of-n: F Systems. IEEE Transactions on Reliability, 658–61.

[9] Kamal, M., Modibbo, U. M., AlArjani, A., and Ali, I. (2021). Neutrosophic fuzzy goal programming approach in selective maintenance allocation of system reliability. *Complex & intelligent systems*, *7*(2), 1045-1059.

[10] Khan, M. F., Modibbo, U. M., Ahmad, N., and Ali, I. (2022). Nonlinear optimization in bi-level selective e maintenance allocation problem. *Journal of King Saud University-Science*, *34*(4), 101933.

[11] Singla, S., Lal, A.K., and Bhatia,S.S (2011). Comparative study of the subsystems subjected to independent and simultaneous failure. *Eksploatacja I Niezawodnosc-Maintenance and Reliability*, *4*, 63-71.

[12] Raghav, Y. S., Varshney, R., Modibbo, U. M., Ahmadini, A. A. H., and Ali, I. (2022). Estimation and optimization for system availability under preventive maintenance. *IEEE Access*, *10*, 94337-94353.

[13] Singha, A. K., Pathak, N., Sharma, N., Gandhar, A., Urooj, S., Zubair, S., and Nagalaxmi, G. (2022). An Experimental Approach to Diagnose Covid-19 Using Optimized CNN. Intelligent Automation & Soft Computing, 34(2).

[14] Singla, S., and Kumari, S.(2022). Behavior and profit analysis of a thresher plant under steady stateB International Journal of System Assurance Engineering and Management **13**, 166–171 .

[15] Singla, S., Lal, A. K., and Bhatia, S.S. (2021). Reliability analysis of poly tube industry using supplementary variable Technique *Applied Mathematics and Computation* 281, 3981–3992.

[16] Salvia, A. A. and Lasher, W.C. (1990). 2-dimensional consecutive-k-out-of-n: F models. IEEE Transactions on Reliability 39(3):382–5

[17] Singla, S., Modibbo, U. M., Mijinyawa, M., Malik, S., Verma, S., and Khurana, P. (2022). Mathematical Model for Analysing Availability of Threshing Combine Machine Under Reduced Capacity. *Yugoslav Journal of Operations Research*, *32*(4), 425-437.

[18] John, Y. M., Sanusi, A., Yusuf, I. and Modibbo, U. M. (2023). Reliability Analysis of Multi-Hardware–Software System with Failure Interaction. *Journal of Computational and Cognitive Engineering*, *2*(1), 38-46.

[18] John, Y. M., Sanusi, A., Yusuf, I., & Modibbo, U. M. (2023). Reliability Analysis of Multi-Hardware–Software System with Failure Interaction. *Journal of Computational and Cognitive Engineering*, *2*(1), 38-46.

RAP AND AVAILABILITY ANALYSIS OF MANUFACTURING SYSTEM: SMO AND PSO

Sarita Devi[1], Nakul Vashishth[2], Deepika Garg[3,*]

•

1,2,3 GD Goenka University, Gurugram, Haryana
sarita.fame@gmail.com, nakulsir@gmail.com
Corresponding author: *deepika.garg@gdgu.org

Abstract

In today's scenario manufacturing industries are highly complex and prone to failure. That's why redundancy allocation problem (RAP) and time dependent availability analysis plays a major role for the successful life cycle of a manufacturing industry. RAP is a Np-hard problem which is very difficult to solve by traditional methods. Therefore in this paper, RAP for the Manufacturing system is solved by Spider monkey optimization. SMO is recent meta-heuristic technique. Till now it is not used to solve RAP. Further results are compared with the Particle swarm optimization algorithm and comparison validates the better performance of SMO in this problem. As mentioned above, for avoiding the complete breakdown of the manufacturing system time- dependent availability is analyzed in this study. Firstly failure and repair data is collected from the manufacturing system then with the help of this information transition diagram is developed. Further equations are developed from transition diagram by using Markov birth death process then equations are solved with the use of Runge-Kutta method. This methodology is implemented in MATLAB.

Keywords: Availability, Reliability, RAP, Markov modelling, SMO, PSO, Manufacturing System.

1. Introduction

The demand for high-quality products and system reliability is increasing day by day in the current world business competitive market. It is observed that only those products will stand in the market, which is up to the desired market level and consumer's satisfaction level. Along with the increasing demand of quality products/systems in practical engineering the importance of reliability optimization is also increased. With the limitation of constraints (costs/ weight), important measurements can be used to increase system reliability. There are several measurements which can be used for increasing the system reliability but redundancy allocation is the important strategy for improving the reliability of the system with constraints. Redundancy allocation problem is a reliability optimization problem. It involves selection of components with appropriate levels of redundancy or reliability to maximize the system reliability under the defined constraints. The redundancy allocation problem was solved by heuristic algorithm(HA) and constraint optimization algorithm (COGA) [1].

In this study, RAP of the manufacturing system is solved by PSO and SMO techniques. A few years back, PSO technique which was a peculiarity has now become the zest of researchers in the world. PSO is a nature-inspired and population-based stochastic technique. It was first introduced in the year 1995 [2]. PSO was applied for solving RAP due to its robustness and its simplicity [3]. Hybrid PSO with local search algorithm was applied for solving RAP [4]. Comparative analysis is also done with Tabu search and Multi weighted objectives solutions. A hybrid PSO algorithm with

local search was proposed for solving RAP in series-parallel system [5]. The reformulate of a crisp optimization problem from FMOOP has been done and then applied PSO for solving fuzzified MOOP under a number of constraints [6]. A hybrid particle swarm optimization with constraint optimization genetic algorithm was proposed for solving a RAP [7]. A bare bones PSO and sensitivity based clustering was proposed for solving multi-objective RAP [8]. Multi-objective PSO (MOPSO) was applied for solving RAP in an interval environment [9]. The effectiveness of the algorithm is demonstrated by two numerical examples in their study.

An improved PSO was applied for solving RAP [10]. Inertia and acceleration coefficients of the classical PSO are improved by considering normal distribution for the coefficients which improved the results. The formulation of RAP with global reliability (g-reliability) has been done and applied improved PSO algorithm with a specific particles under RAP [11]. The optimization of the reliability of the system was done with the allocation of the redundancy of the manufacturing systems using hybrid genetic and particle swarm algorithm [12]. The redundancy allocation problem was solved by using hybrid genetic simulating annealing algorithm and a comparative study is presented in this research [13]. The reliability of the pharmaceutical plant was optimized by using heuristic algorithm [14].

SMO algorithm is introduced by J.C. Bansal in the year 2014 [15]. SMO is stimulated by social behavior of a special kind of monkeys called as spider monkeys. Spider monkeys have been classified as animals with fission-fusion social structure. These monkeys follow fission-fusion social systems as they initially work in a large group and but as their needs change over time, they split into smaller forage groups, each led by an adult female. Consequently, the suggested approach can be roughly categorized as drawing inspiration from the intelligent foraging behavior of spider monkeys with fission-fusion social structures. The initialize phase, the global leader phase (GLP), the local leader phase (LLP), and the decision phase are the main four phases of the SMO algorithm.

SMO was used as an optimization technique for the community of electromagnetic [16]. SMO was enforced to figure out the optimal PIDA controller parameters to control the induction motor [17]. This was the initial strategy to attain such a goal using SMO. The outcomes were compared to the Dorf and PSO technique, and it was noticed that SMO outperformed both of other methods.

An effort was made to solve confined consistent optimization problem by employing the SMO algorithm for restricted optimization problems [18]. A new search feature was developed in SMO called Power Law-based Local Search (PLLS) [19]. The SMO technique was applied to classify diabetes and developed SMO that may be used to design the SM-Rule-Miner, an effective rule miner for diabetes diagnosis [20]. Comparing SM-Rule-Miner to other meta-heuristic based rule mining techniques, it was found that it had the second-best average classification performance rating and the highest accuracy sensitivity rating.

A survey was conducted on SMO, its applications, and variants, and compared the findings to other algorithms [21]. Recently, a modern Discrete Spider Monkey Optimization (DSMO) technique was applied to solve the problem of travelling salesmen [22]. SMO and deep neural networks were applied for the future forecasting of brands for marketing purposes by using Twitter data [23]. Further investigation showed that SMO was more reliable and computationally efficient than other techniques.

Most of the products are manufactured in manufacturing plants, which consist of several subsystems performing different operations. These subsystems are made up of mechanical parts that may fail due to wear and tear and also due to usage with the passage of time. The failure is a random aspect that is always related to the operative condition of any physical system. Its causes are either deterioration in the components of the system and human errors.

As a result, the primary priority is to maintain system performance measures such as reliability and availability and redundancy in order to achieve high-profit targets and productivity in terms of system failures. Reliability optimization and availability are closely related concepts. Reliability

optimization involves improving the system's ability to operate without failure, while availability is a measure of the system's ability to be operational when it is required.

The behavior analysis of time-dependent availability of the manufacturing system is measured using the Runge-Kutta fourth-order method in MATLAB. The area synthesis procedure's availability was considered for the compost industry [24]. The reliability and quality attributes for two stochastic models of a framework were discussed which have two non-indistinguishable components, arranged in series, each unit having cold standby of same/equal capacity [25]. The time-dependent availability of repairable m-out of –n and cold standby systems were investigated using arbitrary distributions and repair facilities [26]. An instructional study on reasonable strategies for Markov modeling was considered [27]. The analysis of availability formulations of standby frameworks of parallel units was done [28].

The study of a two-unit warm reserve framework was discussed that expects a bivariate exponential thickness for the joint circulation of sub-units failure/repair rates [29]. The number of operational stages of a repairable Markov framework during interim finite time was studied [30]. Finite Markov processes were utilized to model a repairable framework with time-independent transition rates concerning individual conditions in reliability analysis. The complex system with the imperfect switching using various techniques such as the Markov-method and supplementary variable method was considered [31]. Failure/repair rates are considered as constant. Further, the N sub-unit framework was analyzed in which M sub-unit are warm standby and R sub units are used to repair the failed sub-units [32]. A closed equation is developed to discover reliability parameters under certain constraints and conclusion is made on the basis of this study that without a repairman, framework reliability diminishes. The shortest path study in stochastic systems; acquainted another methodology and getting the reliability capacity of time-subordinate frameworks with standby mode was discussed [33]. The application of pod propulsion that the number of vessel types has been increased consistently over the last two decades was highlighted [34]. A model was developed using Markov process [35]. Regression analysis was utilized in that study to gauge the different transition rates of the model. A cattle feed plant consisting of seven subsystems arrange in series was analyzed by using matrix method [36]. The mathematical model has been developed using the Markov birth-death process and made a transition diagram. The system's behavior was analyzed over an implanted Markov chain method [37]. The discrete-time and continuous-time measures were provided for each of the explicit Markov and two semi-Markov models for thermal availability plant using simulation modeling [38]. The performance of the steam-generating system was evaluated and analyzed for the availability of a thermal plant [39]. In their study, the system consists three subunits with a high-pressure heater, boiler drum, and economizer associated in series, parallel, and combination of these. Further, Markov model was used for analyzing the reliability of coal crushing unit of Badarpur thermal power plant [40]. Transition diagram and differential equations are developed and solved by a recursive approach. Markov modeling was used for the fertilizer plant and analyzed the reliability of the system [41]. Their study deals with Markov birth-death procedure, and the failure rates and repair rates of every system. The RPGT technique was used for the behavioral study of two units [42].

A stochastic model was analyzed that was to be considered two-unit redundant framework [43]. In this framework, the operational unit's software and hardware elements on failure are substituted by the cold standby unit, and replacement may be possible on fractional failure also. Markov model was used to analyze the reliability of the phased-mission system (PMS) [44]. Mathematical modeling is done by utilizing the state merging method. The analyze the reliability of the manufacturing system is done by Fault tree analysis [45].

2. Problem Description and Formulation

The following assumptions and notations are taken for solving RAP and to draw a transition diagram depicting the various possible working, reduced, and failed states, together with the transitional failure and repair rates of subsystems to analyze and discuss the behavior of time-dependent availability of the manufacturing system.

2.1. Assumptions

The assumptions are follows:

- Failure and repair rates are considered constant and statistically independent of each other. Not more than one unit can fail at a time.
- The repaired units work as if a new one.
- There is only one subsystem which has units in parallel; hence it can work in a reduced capacity, on the failure of all units of this subsystem, there is the complete failure of the system.
- All other subsystems (except one given above) have subunits in series, so if a single unit of these subsystems fail, then the whole of the system fails completely.
- No unit fails further when the system is reached in failed state.

2.2. Notations

The notations used in this research work are as follows:

Rs(n):	System reliability
g(.):	A function that yields the system reliability, based on unique subsystems, and which depends on the configuration of the subsystems
ni:	number of ith subsystems
h(n_i):	Cost of ith subsystem
n:	number of subsystems
Rs:	Rupees
y*:	(n1,n2, n3,…, n7) is optimal solution
A, B, C, D, E, F, G:	Indicate that the subsystems are working in full capacity.
$\bar{B}$:	Indicate the reduced state of the subsystem B.
λ_i (i=1, 2, 3, …, 7, 8):	Failure rates of the subsystems A, B, C, D, …, G, $\bar{B}$respectively.
μ_i(i=1, 2, 3, …, 7, 8):	Repair rates of the subsystems a, $\bar{B}$, c, d, …, g, brespectively.
a, b, c, d, e, f, g	Indicate the failure state of the subsystems A, B, C, D, E, F, G
$P_i(t)$ (i=0, 1, 2, 3, …, 14)	probability that system is in the ith state.
	Failed state
	Full working state
	Reduced state
S0	ABCDEFG
S2	A$\bar{B}$CDEFG
S1	aBCDEFG
S3	ABcDEFG
S_4	ABCdEFG
S_5	ABCDeFG
S_6	ABCDEfG
S_7	ABCDEFg
S_8	a$\bar{B}$CDEFG
S_9	AbCDEFG
S_{10}	A$\bar{B}$cDEFG

S_{11}	$A\bar{B}CdEFG$
S_{12}	$A\bar{B}CDeFG$
S_{13}	$A\bar{B}CDEfG$
S_{14}	$A\bar{B}CDEFg$

2.3. System Description

The system is described as:

- Data is collected from the manufacturing system.
- Data related to subsystems (machines) and the number of units connected in that subsystem, regarding subsystems, their configuring, and working.
- The system consists of seven subsystems: Overhead Crane, Roller, Blanking Machine, Stacker machine, Press machine, Molding, and Packing. These subsystems are connected in series for the proper functioning of the system at full capacity initially.
- Overhead Crane Subsystem (A): There is only one overhead crane having subunits in series.
- Roller Subsystem (B): The roller consists of one unit having subunits in parallel. It rolls the steel roll connected to the blanking machine.
- Blanking Machine (C): The blanking machine cuts the blanks having subunits in series.
- Stacker Subsystem (D): Blanks are stacked with the help of a stacker having subunits in series.
- Press Subsystem (E): Blanks are pressed for the desired shape by the Press subsystem having subunits in series.
- Molding Machine (F): The molding machine molds the pressed blanks for the desired shape having subunits in series.
- Packing Subsystem (G): Packing machine packs the final product having subunits in series.

Structural Representation of the system

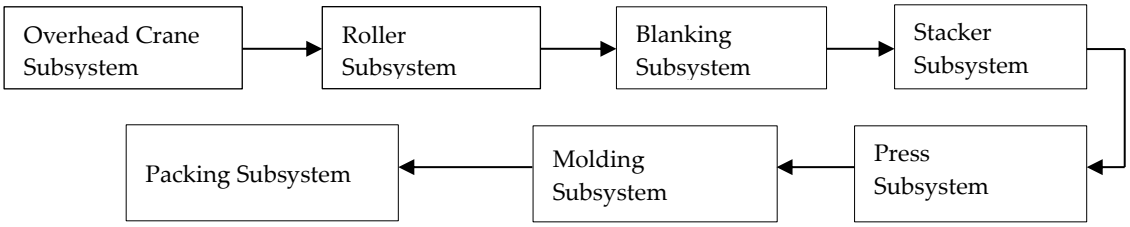


Figure 1: *System Structure*

Problem Statement

In this study, the data related to the subsystem's reliability and cost are presented in Table 1 of the manufacturing system. The given cost constraint is C= Rs 30490000. The objective function for the manufacturing system is to maximize the reliability and subject to cost constraint. The problem is represented by Equation (2), and the cost constraint presented by Equation (3).

The problem is to maximize :

$$R_s(n) = g\,(R_1(n_1), \ldots., R_n(n_n)) = \prod_{i=1}^{7} R_i\,(n) \quad (1)$$

$$R_i(n_i) = \prod_{i=1}^{7}[1 - [Q_i(n_i)]^{n_i}] \quad (2)$$

$$\sum_{i=1}^{7} h(n_i) * n_i \leq 30490000 \quad (3)$$

Table 1: *Reliability and Cost*

Subsystem	n_1	n_2	n_3	n_4	n_5	n_6	n_7
Reliability of subsystem $R_i(n_i)$	0.99	0.9762	0.9188	0.8155	0.8655	0.9287	0.9453
Cost of subsystem(Rs) $h(n_i)$	1280000	960000	2500000	1050000	10500000	950000	250000

Optimization using PSO and SMO Algorithms

Inspired by swarm intelligence in nature, PSO was developed by Kennedy and Eberhart in 1995 [2].

PSO is motivated by the social behavior of bird flocking and applies this behavior to guide the particles for searching the global optimal solutions. Mostly, in PSO, the particle population are randomly spread throughout the search space. The particles are considered to be flying in the search space. PSO searches for optima by updating generations after being initialized with a collection of random particles (solutions).

Each particle is modified by the following the two best values in every iteration. The first one is the best solution (fitness) it has obtained so far. This value is known as pbest. Another "best" value that is obtained by the particle swarm optimizer is the best value obtained so far by any particle in the population. This global best value is known as gbest. When a particle takes part of the population as its topological neighbors, the best value is a local best and is called lbest. The position and velocity updated with the help of two equations:

$$v_i = wv_i + c_1 r_1 (p_{best,i} - y_i) + c_2 r_2 (g_{best} - y_i) \quad (4)$$

$$y_i = y_i + v_i \quad (5)$$

Results of PSO

The results of the redundancy allocation solved by PSO are represented in tabular form. Table 2 shows different redundant units for various subsystems obtained by using PSO.

Table 2: *Results of PSO for the Manufacturing System*

Subsystems	n_1	n_2	n_3	n_4	n_5	n_6	n_7
Redundancy of the subsystems	2	1	2	3	1	3	3

SMO is stimulated by social behavior of a special kind of monkeys called as spider monkeys. Spider monkeys have been classified as animals with fission-fusion social structure. These monkeys follow fission-fusion social systems as they initially work in a large group and but as their needs change over time, they split into smaller forage groups, each led by an adult female. Consequently, the suggested approach can be roughly categorized as drawing inspiration from the intelligent foraging behavior of spider monkeys with fission-fusion social structures. The initialize phase, the global leader phase (GLP), the local leader phase (LLP), and the decision phase are the main four phases of the SMO algorithm.

Results of SMO

The results of RAP of the manufacturing system obtained by SMO are also represented in tabular form. Optimum redundant units of each subsystems redundant of each subsystem are represented in Table 3.

Table 3: *Results of SMO for the Manufacturing System*

Subsystems	n_1	n_2	n_3	n_4	n_5	n_6	n_7
Redundancy of the subsystems	3	2	3	3	1	2	5

Comparative analysis of the results of RAP by PSO and SMO

The comparative analysis of the results obtained by two algorithms PSO and SMO for the Manufacturing system in this research.

The comparative analysis of the results of RAP using PSO and SMO are represented in Table 3. Obtained results demonstrate the increase in the reliability of the manufacturing system.

Table 4: *Comparison of the best optimal solution by PSO and SMO*

Subsystems	n_1	n_2	n_3	n_4	n_5	n_6	n_7	Increase in the reliability
Redundancy								
PSO	2	3	2	5	1	3	5	
SMO	3	2	3	5	2	2	5	

The comparative analysis of the results obtained by these algorithms are described as follows:

The reliability value of the system before applying these two algorithms is 0.5236.
The values of the redundant units of each subsystem before using these algorithms are n*=(1,1,1,1,1,1,1).
The optimal solution of RAP using PSO is Rs(n*) = 0.8332 and the redundancy units n* = (2, 1, 2, 3, 1, 3, 3).
The increment in the value of reliability can be 51%.
The final solution of RAP using SMO is Rs(n*) = 0.8504 and redundant units of each subsystem n*= (3, 2, 3, 3, 1, 2, 4).
The reliability value of the system can be increased by 55% using SMO.
The reliability value of the system is increased by using both the algorithms.
The redundancy units of the subsystems n* are distinct for every subsystem by both the algorithms used in this research.
The larger increment in the reliability value of the system using SMO which is 55%.
The results shows that the batter results obtained by SMO and reliability is also improved by SMO than PSO.

Time dependent availability

The state transition diagram of the manufacturing system is drawn as follows:

2.4. State Transition Diagram

The first order Markov-process is used to develop the state transition diagram. A state transition diagram, together with transition rates, is drawn to describe the various states of the system, is given below in Figure 1.

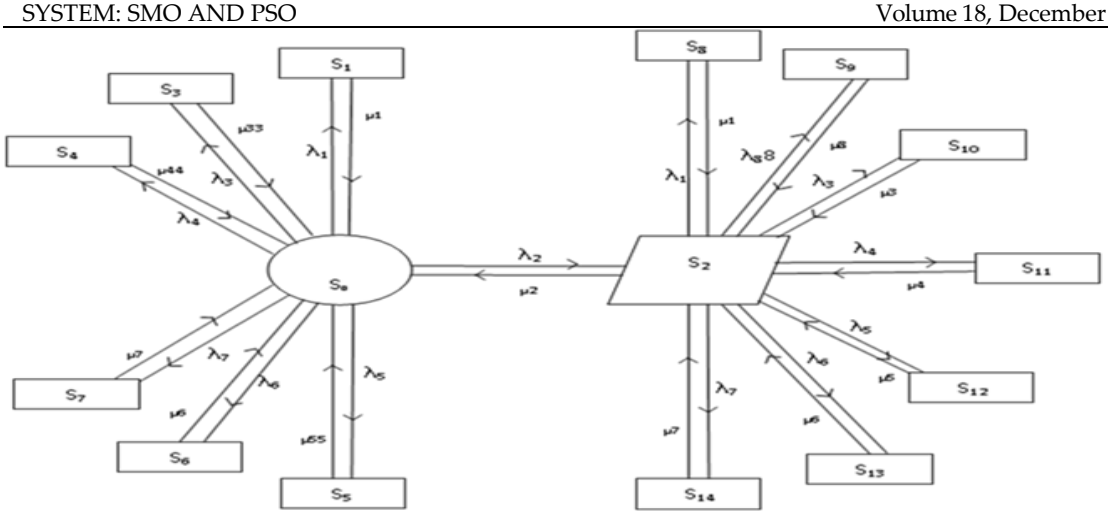


Figure 2: *State Transition Diagram of Manufacturing System*

2.5. Methodology- Runge-Kutta Fourth Order

Differential equations associated with the system are written using the Markov birth-death method, which is further solved using Runge-Kutta fourth-order method by ODE- 45 in MATLAB. Table and graph are drawn to represent the results of the time-dependent availability of the manufacturing system.

Mathematical Modelling:

The differential equations governing the system using the Markov birth-death process are as follows:

$$P_0'(t) + (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7)P_0(t) = \mu_1 P_1(t) + \mu_2 P_2(t) + \mu_3 P_3(t) + \mu_4 P_4(t) + \mu_5 P_5(t) + \mu_6 P_6(t) + +\mu_7 P_7(t) \quad (6)$$

$$\mu_1 P_1(t) = \lambda_1 P_0(t) \quad (7)$$

$$P_2'(t) + (\lambda_1 + \mu_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8)P_2(t) = \lambda_2 P_0(t) + \mu_1 P_8(t) + \mu_8 P_9(t) + \mu_3 P_{10}(t) + \mu_4 P_{11}(t) + \mu_5 P_{12}(t) + \mu_6 P_{13}(t) + \mu_7 P_{14}(t) \quad (8)$$

$$\mu_3 P_3(t) = \lambda_3 P_0(t) \quad (9)$$

$$\mu_4 P_4(t) = \lambda_4 P_0(t) \quad (10)$$

$$\mu_5 P_5(t) = \lambda_5 P_0(t) \quad (11)$$

$$\mu_6 P_6(t) = \lambda_6 P_0(t) \quad (12)$$

$$\mu_7 P_7(t) = \lambda_7 P_0(t) \quad (13)$$

$$\mu_1 P_8(t) = \lambda_1 P_2(t) \quad (14)$$

$$\mu_8 P_9(t) = \lambda_8 P_2(t) \quad (15)$$

$$\mu_3 P_{10}(t) = \lambda_3 P_2(t) \quad (16)$$

$$\mu_4 P_{11}(t) = \lambda_4 P_2(t) \quad (17)$$

$$\mu_5 P_{12}(t) = \lambda_5 P_2(t) \quad (18)$$

$$\mu_6 P_{13}(t) = \lambda_6 P_2(t) \quad (19)$$

$$\mu_7 P_{14}(t) = \lambda_7 P_2(t) \quad (20)$$

3. Results

The results of time-dependent availability obtained using the Runge-Kutta fourth-order method by ODE 45 in MATLAB are represented below in Table 5 and Figure 2.

Table 5: *Time-dependent availability*

Time	Availability
0	1
20	0.827
40	0.8203
60	0.7932
80	0.7757
100	0.7644
120	0.7561
140	0.7515
160	0.7451
180	0.7328
200	0.7215

The availability of the system is also determined after the reliability of the system is improved by the technique SMO.

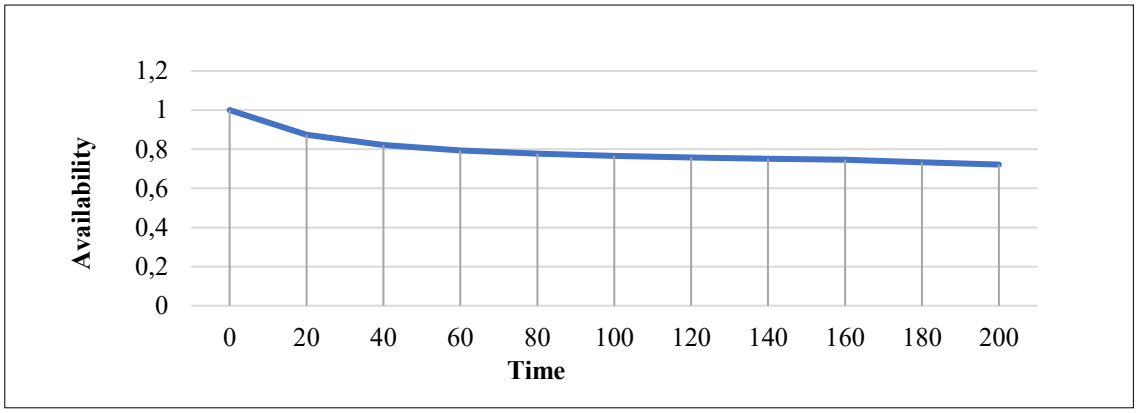


Figure 3: *Time-Dependent Availability*

4. Conclusion

In this study, redundancy is used as a key strategy for reliability optimization to improve system performance, minimize failures, and enhance the overall reliability and availability of a system or process. RAP is a mathematical optimization problem that involves determining the

optimal allocation of redundancy to subsystems in order to improve the system reliability and availability. RAP is solved by SMO and results are compared with PSO.

RAP of the manufacturing system is solved by two optimization techniques PSO and SMO. The reliability of the manufacturing system is improved by both the techniques. But the increment of 55% in reliability of the manufacturing system by SMO which is greater than 51% by PSO.

The availability of the manufacturing system is also analyzed using Markov modelling. The calculations are done using Runge-Kutta method of forth order in MATLAB using the tool ode 45. It is concluded that system availability is a decreasing function of time. As time increases, the availability value of the system decreases.

5. Author Contribution Statement

In this manuscript all authors "contributed equally" to the study.

References

[1] Devi, S., Garg, D. (2017), "Redundancy-allocation in Neel Metal Products Limited", *Indian Journal of Science and Technology*,10(30): 1-5.

[2] Kennedy, J., Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*, IEEE, 4: 1942-1948.

[3] Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization: An overview. *Swarm intelligence*, *1*, 33-57.

[4] Beji, N., Jarboui, B., Eddaly, M., & Chabchoub, H. (2010). A hybrid particle swarm optimization algorithm for the redundancy allocation problem. *Journal of Computational Science*, *1*(3), 159-167.

[5] Wang, Y., Li, L. (2012). Heterogeneous redundancy allocation for series-parallel multi-state systems using hybrid particle swarm optimization and local search. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 42(2): 464-474.

[6] Garg, H., Sharma, S. P. (2013). Multi-objective reliability-redundancy allocation problem using particle swarm optimization. *Computers & Industrial Engineering*, *64*(1), 247-255.

[7] Devi, S., Sahu, A., Garg, D. (2017), "Redundancy optimization problem via comparative analysis of H-PSOCOGA", *IEEE Explore*, 18-23.

[8] Zhang, E., Wu, Y., & Chen, Q. (2014). A practical approach for solving multi-objective reliability redundancy allocation problems using extended bare-bones particle swarm optimization. *Reliability Engineering & System Safety*, *127*, 65-76.

[9] Zhang, E., & Chen, Q. (2016). Multi-objective reliability redundancy allocation in an interval environment using particle swarm optimization. *Reliability Engineering & System Safety*, *145*, 83-92.

[10] Marouani, H. (2021). Optimization for the redundancy allocation problem of reliability using an improved particle swarm optimization algorithm. *Journal of Optimization*, *2021*, 1-9.

[11] Li, S., Chi, X., & Yu, B. (2022). An improved particle swarm optimization algorithm for the reliability–redundancy allocation problem with global reliability. *Reliability Engineering & System Safety*, *225*.

[12] Devi, S., Garg, D. (2020). Hybrid genetic and particle swarm algorithm: redundancy allocation problem. *International Journal of System Assurance and Engineering Management*, 11(6): 313-319.

[13] Garg, D., Devi, S. (2021). RAP via hybrid genetic simulating annealing algorithm. *International Journal of System Assurance and Engeering Management*, Springer.

[14] Dahiya, T., Garg, D., Devi, S., Kumar, R. (2021). Reliability optimization using heuristic algorithm in pharmaceutical plant. *Reliability: Theory & Applications*, 16(3): 195-205.

[15] Bansal, J. C., Sharma, H., Jadon, S. S., & Clerc, M. (2014). Spider monkey optimization algorithm for numerical optimization. *Memetic computing*, *6*, 31-47.

[16] Al-Azza, A. A., Al-Jodah, A. A., & Harackiewicz, F. J. (2016). Spider monkey optimization: A novel technique for antenna optimization. *IEEE Antennas and Wireless Propagation Letters*, *15*, 1016-1019.

[17] Sharma, A., Sharma, H., Bhargava, A., & Sharma, N. (2016). Optimal design of PIDA controller for induction motor using spider monkey optimization algorithm. *International Journal of Metaheuristics*, *5*(3-4), 278-290.

[18] Gupta, K., Deep, K., Bansal, J.C. (2017). Spider monkey optimization algorithm for constrained optimization problems. *Soft Comput. Springer Berlin Heidelberg*, 21: 6933–62.

[19] Sharma, A., Sharma, H., Bhargava, A., & Sharma, N. (2017). Power law-based local search in spider monkey optimisation for lower order system modelling. *International Journal of Systems Science*, *48*(1), 150-160.

[20] Cheruku, R., Edla, D. R., & Kuppili, V. (2017). SM-Rule Miner: Spider monkey based rule miner using novel fitness function for diabetes classification. *Computers in biology and medicine*, *81*: 79-92.

[21] Agrawal, V., Rastogi, R., & Tiwari, D. C. (2018). Spider monkey optimization: a survey. *International Journal of System Assurance Engineering and Management*, *9*: 929-941.

[22] Akhand, M. A. H., Ayon, S. I., Shahriyar, S. A., Siddique, N., & Adeli, H. (2020). Discrete spider monkey optimization for travelling salesman problem. *Applied Soft Computing*, *86*.

[23] Kothamasu, L. A., & Kannan, E. (2022). Sentiment analysis on twitter data based on spider monkey optimization and deep learning for future prediction of the brands. *Concurrency and Computation: Practice and Experience*, *34*(21), e7104.

[24] Singh, J., Pandey, P.C., and Kumar, D. (1990). Designing for reliable operation of urea synthesis in the fertilizer industry. *Microelectronics Reliability*, 30 (6): 1021 – 1024.

[25] Moafib, S. E., Goel, L. R., and Gupta, R. (1991). Comparison of two stochastic models for two-unit series system with cold standbys. *Microelectronic Reliability,* 31 (6): 1105 - 1111.

[26] Gurov, S. V., Utkin, L.V. (1995), **"**A new method to compute reliability of repairable series systems by arbitrary distributions". *Microelectronics Reliability*, 35 (1): 81-85.

[27] Kitchin, J. K. (1988). Practical Markov modelling for reliability analysis. *In Proceedings of the Annual Reliability and Maintainability Symposium,* 290 -296.

[28] Aven, T. (1990). Availability formulae for standby systems of similar units that are preventively maintained. *IEEE Transactions on Reliability,* 39 (5): 603 - 606.

[27] Goel, L. R., and Shrivastava, P. A.(1992). Warm standby redundant system with correlated failures and repairs. *Microelectron Reliability*, 32 (6): 793 -797.

[28] Csenki, A. (1994), "Joint availability of systems modelled by finite semi-Markov processes". *Applied Stochastic Models and Data Analysis*, 10(4): 279-293.

[29] Goel, P.(1997). Availability Analysis and Cost Optimization of Complex System having Imperfect Switch. Ph.D. Thesis, Kurukshetra, India.

[30] Arulmozhi, G.(2002). Reliability of M-out of N Warm Standby System with R Repair Facilities. *Operational Research Society of India*, 39 (2): 77 - 88.

[31] Azaron, A., Katagiri, H., Sakawa, M., and Modarres, M.(2005). Reliability function of a class of time dependent systems with standby redundancy. *European Journal of Operational Research,* 16 (4): 378 – 386.

[32] Aksu, S., and Turan, O.(2006). Reliability and availability of Pod propulsion system. *Journal of Reliability Engineering International*, 22: 41 - 58.

[33] Mishari, S.T., Suliman, S. (2008), "Integrating Six-sigma with other reliability improvement methods in equipment reliability and maintenance applications. Journal of Quality in Maintenance Engineering, 14(1): 59-70.

[34] Garg, D., Kumar, K., and Singh, J. (2009). Availability analysis of a cattle feed plant using matrix method. *International Journals of engineering*, 3 (2): 127 - 133.

[35] Umemura, T. and Dohi, T. (2010). Availability Analysis of an Intrusion Tolerant Distributed Server System with Preventive Maintenance. *IEEE Transactions on Reliability*, 59 (1): 18 - 29.

[36] Gupta, S., and Tewari, P. C.(2011). Simulation modeling in an availability thermal power plant. *Journal of engineering science and Technology review*, 4(2): 110 - 117.

[37] Kumar, R., Sharma, A. K., and Tewari, P. C.(2012). Markov approach to evaluate the availability simulation model for power generation system in athermal power plant. *nternational Journal of Industrial Engineering Computations,* 3 (5): 743 - 750.

[38] Kumar, J., Kadyan, M. S. and Malik, S. C.(2013). Profit Analysis of a 2 out of 2 Redundant System with Single Standby and Degradation of the Units after Repair. *International Journal of System Assurance Engineering and Management,* 424 -434.

[39] Aggarwal, A., Kumar, S., and Singh, V.(2015). Markov modeling and reliability analysis of urea synthesis system of a fertilizer plant. *Journal of industrial engineering international,* 11: 1-14.

[40] Goyal, V., Goel, R., and Goel, P.(2015). Behavioral Analysis of Two Unit System with Preventive Maintenance in Both Units and Degradation in One Unit. *International Journal of Engineering Technology Management and Applied Sciences*, 3: 190 - 197.

[41] Kumar, A., Garg, D., and Goel, P. (2017). Mathematical modelling and profit analysis of an edible oil refinery industry. *Airo International Research Journal*, 12: 1 - 8..

[44] Hua, Y., Li., G., Lie, Q., and Ping, W.(2018). Simplified Markov Model for Reliability Analysis of Phased-Mission System Using States Merging Method. *J. Shanghai Jiao Tong Univ. (Sci.),* 23 (3): 418 - 422.

[45] Devi, S., Garg, D. (2019), "Reliability Analysis of Manufacturing Plant Via Fault Tree Analysis", *Journal of Advances and Scholarly Researches in Allied Education*, 16(5): 256-259.

INTEGRATING TESTING COVERAGE, EFFORT AND CHANGE POINT IN A SOFTWARE RELIABILITY GROWTH MODEL: A COMPREHENSIVE ANALYSIS

Sudeep Kumar
•
Department of Mathematics,
AIAS, Amity University, Uttar Pradesh – 201303 (India).
Corresponding author: sudeepdeepak777@gmail.com

Anu G Aggarwal
•
Department of Operational Research,
University of Delhi, Delhi (India)
E-mail: anuagg17@gmail.com

Abstract

Software reliability growth models (SRGMs) are essential for forecasting and controlling the reliability of software systems. In present article, we propose an enhanced SRGM that incorporates three important factors: testing coverage, testing effort, and change point detection. We introduce a novel testing coverage function that captures the delayed S-shaped behaviour commonly observed in software reliability growth. Weibull distribution is utilized to model the testing effort. Finally, we address the impact of change points in software reliability. To assess how well our suggested model works, we conducted experiments using real-world software failure data provide by Tandem computers. The results demonstrate that our model outperforms existing SRGMs by providing more accurate predictions and a better understanding of the interplay between testing coverage, testing effort, and change points.

Keywords: software reliability growth model, testing coverage, change point detection, reliability modelling, testing effort.

I. Introduction

Software reliability is a critical aspect of software systems, as it directly impacts their quality, user satisfaction, and overall success. To ensure reliable software, it is essential to accurately manage and predict its reliability throughout the development lifecycle. SRGMs have been widely employed for this purpose, aiming to estimate the number of remaining defects and predict the software's reliability over time. Yamada et al. [1] offer two software reliability evaluation models with imperfect debugging. Poonam and Ravneet [2] suggest a method for estimating software's remaining faults using both imperfect and perfect software reliability growth models. Aggarwal et

al. [3] propose a non-homogeneous Poisson process (NHPP) based SRGM by combining imperfect debugging.

Traditional SRGMs often assume simplistic reliability growth patterns, such as the popular J-shaped or S-shaped curves. However, these models fail to capture the complex dynamics of software reliability growth, particularly the delayed S-shaped behaviour commonly observed in practice. In reality, the effectiveness of testing activities varies over time, with initial testing efforts being more efficient in detecting defects, followed by a period of diminishing returns as the software matures. Therefore, there is a need to incorporate a delayed S-shaped testing coverage function into SRGMs to better reflect the actual reliability growth process. With the variability of operational conditions, Chang et al. [4] create a new testing-coverage based SRGM. In the work, Chatterjee and Shukla [5] introduced and integrated a temporal variant fault detection probability into the s-shaped coverage SRGM. Using two distinct testing-time functions, Minamino et al. [6] offered to expand the current univariate SRGMs. Based on testing efforts that consider different testing coverage functions, Bibyan et al. [7] and Kumar et al. [8] developed three SRGMs.

Furthermore, testing effort plays a crucial role in improving reliability of the software. The allocation and distribution of testing effort throughout the development lifecycle significantly impact the detection and removal of defects. The Weibull distribution has proven to be a versatile tool for modeling testing effort, as it can accommodate various shapes of effort curves, such as increasing, decreasing, or constant effort over time. By integrating the Weibull distribution into an SRGM, we can capture the diverse characteristics of testing effort and enhance the accuracy of reliability predictions. Yamada [9] and Kapur et al. [10] suggested a NHPP based SRGM that is adaptable enough to characterise different software failure/reliability curves. For software reliability modeling, both time-dependent fault detection rate (FDR) and testing efforts are taken into account. Kapur et al. [11] and Khatri et al. [12] and Jin and Jin [13] proposed an SRGM that includes a testing-effort function. Pradhan et al. [14] offer SRGMs that incorporate the key notion of testing-effort function.

Additionally, change points pose a significant challenge in software reliability modeling. Change points occur when there are substantial modifications in the software or testing environment, leading to shifts in the software reliability growth pattern. Failure to identify and account for these change points can result in inaccurate reliability estimates and poor decision-making regarding software reliability management. Therefore, it is crucial to incorporate change point detection techniques into SRGMs to effectively handle these shifts and enhance the reliability estimation process. Zhao et al. [15], Inoue et al. [16] Arora et al. [17], Dhaka and Nijhawan [18] and Huang et al. [19] explored an SRGM that was based on the Non-Homogeneous Poisson Process while taking into account the phenomenon of change point.

This research article is to develop an enhanced SRGM by incorporating a delayed S-shaped testing coverage function, utilizing the Weibull distribution for modeling testing effort, and integrating change point detection techniques, the aim to provide a comprehensive approach to software reliability modeling. The proposed model will contribute to more accurate reliability predictions, better understanding of the interplay between testing coverage, testing effort, and change points, and ultimately improve software reliability management practices.

This article is organized into several sections. Section 2 presents the formulation of the proposed model. In this section, the model is described in detail, outlining its key components and mathematical formulation. In Section 3, the results of the experiments are presented, and their implications are discussed. This section highlights the findings and outcomes of the study, examining how well the proposed model performed in predicting software reliability. The implications of these results for software development and management are also explored. Lastly, Section 4 concludes the article and provides directions for future research.

II. Model Description and Development

Maecenas Software reliability estimation plays a critical role in ensuring the dependability of software systems. It is essential to accurately predict the growth of software reliability and effectively manage testing efforts for successful software development and deployment. However, many existing software reliability growth models have limitations that hinder their accuracy and effectiveness. Furthermore, change points, which represent significant shifts in the reliability growth pattern, are often disregarded in existing models. Change points can occur due to various factors, such as changes in development methodologies, software updates, or modifications in the system environment. Failing to account for change points can lead to inaccurate reliability predictions and ineffective management of testing efforts.

To address these limitations, research is focused on developing more advanced software reliability growth models. These models aim to incorporate the delayed S-shaped nature of testing coverage, variations in testing effort over time, and the impact of change points. By considering these factors, these models can provide more accurate and reliable predictions of software reliability growth.

I. Notations

a:	initial number of faults
b:	fault detection rate
ϕ:	constant
m, m(t), m(W):	mean value function (MVF)
W/ W(t):	testing effort function
γ, ζ :	Weibull distribution scale, shape parameter
c/ c(W):	testing coverage function
τ	change point
β	scale parameter of logistic distribution function (constant)
α	fault generation rate
$\overline{W}$:	total testing effort available (constant)
b_1/b_2:	fault detection rate before/after change point

II. Model Assumptions

- During the testing phase, the rate at which faults are detected and removed may vary at any given point.
- The number of faults detected is directly proportional to the number of faults that remain undiscovered in the system.
- The level of testing coverage is influenced by the amount of effort invested in testing.
- Delayed S-shaped testing coverage function is incorporating.
- Fault detection rate for removal is logistically distributed $\left(\frac{b}{1+\beta e^{-bW(t)}}\right)$.

- The testing coverage can be quantified in terms of the rate of fault detection, represented as $\frac{c'(W(t))}{1-c(W(t))}$, where $c(W(t)) = \left[\frac{1-(1+bW)e^{-bW(t)}}{1+\beta e^{-bW(t)}}\right]$ represents the proportion of code covered during testing.
- The extent of code coverage during testing directly impacts the number of faults that can be detected.

III. Model Development

In accordance with the previously mentioned assumptions, we have formulated an NHPP based SRGM. The Mean Value Function (MVF) of our proposed model is expressed as follows:

$$\frac{dm(t)}{dt} = \frac{dm}{dc}\frac{dc}{dW}\frac{dW}{dt} \tag{1}$$

The first component represents the relationship between the level of testing coverage and number of detected faults. This component can be represented using a model that incorporates the delayed S-shaped testing coverage function, capturing the changing effectiveness of testing activities over time. The second component represents the rate at which the amount of testing coverage changes over time. This component can be modeled using the Weibull distribution for testing effort, which allows for varying patterns of testing effort, such as increasing, decreasing, or constant effort over time.

$$\frac{dm}{dc} = \frac{c'(W)}{1-c(W)}(a - m) \tag{2}$$

$$\frac{dc}{dW} = \phi(\text{constant}) \tag{3}$$

with the initial condition

$$m(0) = 0 \tag{4}$$

By combining these two components, the MVF of our proposed SRGM can be accurately represented and used to predict and manage software reliability growth.

$$m(W) = a\left[1 - \left(1 - c(W)\right)^{\phi}\right] \tag{5}$$

The testing effort is represented by Weibull function in this paper can be expressed as:

$$W(t) = \overline{W}\left(1 - e^{-\gamma t^{\zeta}}\right) \tag{6}$$

SRGM 1: Proposed model with perfect debugging

Let us consider the concept of perfect debugging without change point in proposed SRGM, we get the following mean value function:

$$m(t) = a\left[1 - \left(\frac{(1+bw(t)+\beta)e^{-bw(t)}}{1+\beta e^{-bw(t)}}\right)^{\phi}\right] \tag{7}$$

SRGM 2: Proposed model with imperfect debugging

Let us consider the concept of imperfect debugging without change point in proposed SRGM. The fault content of the software undergoes a constant rate of change over time. The rate of

removal of faults is directly proportional to this changing fault content. It can be expressed mathematically as:

$$a(t) = a + \alpha m(t)$$

and the corresponding mean value function is:

$$m(t) = \frac{a}{(1-\alpha)}\left[1 - \left(\frac{(1+bw(t)+\beta)e^{-bw(t)}}{1+\beta e^{-bw(t)}}\right)^{(1-\alpha)\phi}\right] \quad (8)$$

SRGM 3: Proposed model with perfect debugging and change point

Let us consider the concept of perfect debugging without change point in suggested SRGM. With the change-point at time (τ) as the detection point, the detection rate of faults is as follows:

$$b(t) = \begin{cases} b_1; 0 \le t \le \tau \\ b_2; t > \tau \end{cases} \quad (9)$$

while satisfying the condition $W(t = \tau) = W(\tau)$.
where

$$W(t - \tau) = W(t) - W(\tau)$$

The resulting mean value function is:

$$m(t) = \begin{cases} a\left[1 - \left(\frac{(1+b_1 w(t)+\beta)e^{-b_1 w(t)}}{1+\beta e^{-b_1 w(t)}}\right)^{\phi}\right]; \ 0 \le t \le \tau \\ a\left[1 - \left(1 - \left(1 - \left(\frac{(1+b_1 w(\tau)+\beta)e^{-b_1 w(\tau)}}{1+\beta e^{-b_1 w(\tau)}}\right)^{\phi}\right)\right)\left(\left(\frac{1+\beta e^{-b_2 w(\tau)}}{1+\beta e^{-b_2 w(\tau)}}\right)\left(\frac{(1+b_2 w(t)+\beta)}{(1+b_2 w(\tau)+\beta)}\right)e^{-b_2 w(t-\tau)}\right)^{\phi}\right)\right]; \ t > \tau \end{cases} \quad (10)$$

III. Model Validation and Comparison Criteria

To exemplify the estimation process of both the existing and new SRGMs, the analysis used a dataset of software failures provided by tandem computer [20] that was tested over a 20-week period and contained a total of 100 faults that were found. 10,000 units of effort were used in the first release. SRGM parameters were computed using SPSS, a statistical programme. The proposed model's estimations are evaluated using four goodness-of-fit metrics, which are shown in Table 3. A change-point analyser is used to identify the change-point in the dataset. According to the analysis, the transition point happened in the eighth week of testing, and the corresponding efforts are 4.606 For k observations in the data, the actual number of detected faults, and their predicted values are represented by x_j and $m(t_j)$ respectively. The explanation of the criteria and mathematical formulae in Table 1.

On the basis of the software failure data, the estimated parameter values for the testing effort function that follows the Weibull distribution are shown in Table 2. To determine the optimal values for the Weibull function shape and scale parameters, a non-linear regression technique was used in the estimate procedure. The estimated parameter values and fitting of the proposed SRGM's 1, 2 and 3 are shown in Table 3 and Figure 1.

Figure 1 provides an illustration of how the proposed software reliability growth models (SRGMs) are fitted to Tandem Computers' software testing data. The figure visually represents the alignment between the predicted values generated by the SRGMs and the actual observed data

points.

Table 1: Goodness of fit criteria for model

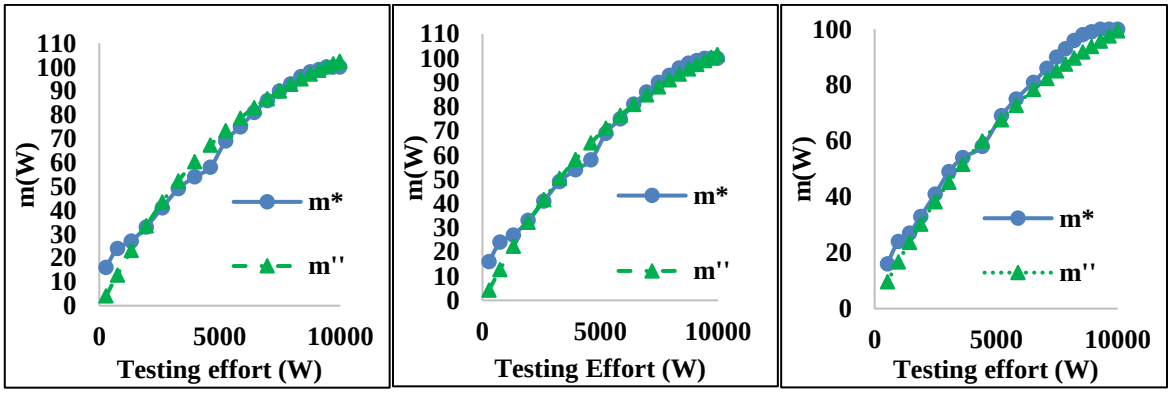
Performance Criteria	R^2	MSE	PRR	PP
Interpretation	larger the value of R^2, indicates model fitting is better to the data	smaller the value of MSE indicates a better fit of the model to the data, indicates model fitting is better to the data	smaller the value of PRR, indicates model fitting is better to the data	Smaller the value of PP, model fitting is better to the data
Expression	$1 - \frac{\text{residual SS}}{\text{actual SS}}$ SS: - sum of squares	$\frac{1}{k}\sum_{j=1}^{k}(m(t_j) - x_j)^2$	$\sum_{j=1}^{k}\left(\frac{m(t_j) - x_j}{m(t_j)}\right)^2$	$\sum_{j=1}^{k}\left(\frac{m(t_j) - x_j}{x_j}\right)^2$

Table 2: Weibull testing effort function parameter estimation

Effort function	Estimated Parameters	Estimated values
Weibull	W′	11740.754
	γ	0.024
	ζ	1.460

Table 3: Estimation results of parameter of proposed SRGM's

Proposed SRGM	a	b/b1	b2	ϕ	α	β
SRGM 1	125	.022	-	.008	-	.179
SRGM 2	101	.040	-	.005	.220	.099
SRGM 3	136	.133	.131	.001	-	.900



(i) (ii) (iii)

Figure 1: Goodness of fit curves for SRGM 1, 2 and 3 respectively

Table 4, on the other hand, presents a comprehensive comparison of the proposed SRGMs in terms of goodness-of-fit curves and metrics, considering various coverage functions. The table provides a quantitative evaluation of the models' performance by comparing different metrics such as mean square error (MSE), predictive power (PP), prediction ratio risk (PRR), and other relevant measures.

By examining the goodness-of-fit curves and metrics in Table 4, we can assess the

effectiveness of each SRGM and identify the most suitable model based on their specific requirements and evaluation criteria. This table serves as a concise summary of the performance evaluation results and facilitates a comparative analysis of the proposed SRGMs across different coverage functions.

Both Figure 1 and Table 4 contribute to the comprehensive assessment of the proposed SRGMs, offering visual and quantitative insights into their fitting accuracy, predictive capabilities, and overall goodness of fit to Tandem Computers' software testing data. These visual and tabular representations aid in understanding and interpreting the results obtained from the comparison of the SRGMs, assisting researchers and practitioners in making informed decisions regarding model selection and further analysis.

Table 4: Goodness of fit criteria for proposed SRGM's

Proposed SRGM	MSE	PP	PRR	R^2
SRGM 1	23.58422	0.848339775	9.489643	0.987
SRGM 2	19.40284	0.813731	8.381536	0.988
SRGM 3	17.79812	0.3138299	0.71271163	0.991

The inclusion of a change point in SRGM 3 contributes to its improved performance. The lower PP, MSE, and PRR values indicate that the predictions made by SRGM 3 align more closely with the actual observed data. This suggests that the model's estimates of software reliability are more accurate and precise when compared to SRGM 1 and SRGM 2. Additionally, the greater coefficient of determination (R^2) in SRGM 3 indicates that a larger proportion of the variability in the software reliability data can be explained by the model. This signifies a stronger relationship between the predicted values and the actual values, further supporting the superior accuracy of SRGM 3. Based on the results, it can be concluded that the proposed model, which incorporates a change point, offers a higher level of precision and accuracy compared to the model without a change point. These findings have implications for software development and management, as they highlight the potential benefits of considering change points in software reliability growth models for more reliable predictions and decision-making.

IV. Conclusion

In this research article, we have suggested an enhanced SRGM that integrates a delayed S-shaped testing coverage function, the Weibull distribution for testing effort, and change point detection techniques. To test the models, the researchers have utilized software failure data from a literature source, and the goodness of fit curves demonstrate the level of similarity between the actual and estimated values of the proposed model. The results of our research demonstrate the effectiveness and accuracy of the suggested SRGM. By integrating the delayed S-shaped testing coverage function, the Weibull distribution for testing effort, and change point detection techniques, our model provides valuable insights for the estimation and management of reliability of the software system. This research article presents a significant advancement in software reliability modeling by incorporating a delayed S-shaped testing coverage function, the Weibull distribution for testing effort, and change point detection techniques. The proposed model offers a comprehensive and accurate framework for software reliability estimation and management, contributing to the development of more reliable software systems and improving overall software quality.

References

[1] Aggarwal, A. G., Gandhi, N., Verma, V. and Tandon, A. (2019). Multi-release software reliability growth assessment: an approach incorporating fault reduction factor and imperfect debugging. *International Journal of Mathematics in Operational Research*, *15*(4), 446. https://doi.org/10.1504/IJMOR.2019.103006

[2] Arora, R., Tandon, A., Aggarwal, A. G. and Verma, V. (2019). Testing Resource based Optimal Release Policy for Software System Incorporating Fault Reduction Factor and Change point. *International Journal of Recent Technology and Engineering*, *8*(4), 1909–1920. https://doi.org/10.35940/ijrte.C5277.118419

[3] Bibyan, R., Anand, S., Aggarwal, A. G. and Kaur, G. (2023). Multi-release software model based on testing coverage incorporating random effect (SDE). *MethodsX*, *10*, 102076. https://doi.org/10.1016/j.mex.2023.102076

[4] Chang, I. H., Pham, H., Lee, S. W. and Song, K. Y. (2014). A testing-coverage software reliability model with the uncertainty of operating environments. *International Journal of Systems Science: Operations & Logistics*, *1*(4), 220–227. https://doi.org/10.1080/23302674.2014.970244

[5] Chatterjee, S. and Shukla, A. (2016). Effect of Test Coverage and Change Point on Software Reliability Growth Based on Time Variable Fault Detection Probability. *Journal of Software*, *11*(1), 110–117. https://doi.org/10.17706/jsw.11.1.110-117

[6] Dhaka, V. and Nijhawan, N. (2022). Effect of change in environment on reliability growth modeling integrating fault reduction factor and change point: a general approach. *Annals of Operations Research*, 1–35. https://doi.org/10.1007/S10479-022-05084-6/FIGURES/23

[7] Huang, Y.-S., Chiu, K.-C and Chen, W.-M. (2022). A software reliability growth model for imperfect debugging. *Journal of Systems and Software*, *188*, 111267. https://doi.org/10.1016/j.jss.2022.111267

[8] Inoue, S., Fukuma, K. and Yamada, S. (2010). Two-Dimensional Change-Point Modeling for Software Reliability Assessment. *International Journal of Reliability, Quality and Safety Engineering*, *17*(06), 531–542. https://doi.org/10.1142/S0218539310003925

[9] Jin, C. and Jin, S.-W. (2016). Parameter optimization of software reliability growth model with S-shaped testing-effort function using improved swarm intelligent optimization. *Applied Soft Computing*, *40*, 283–291. https://doi.org/10.1016/j.asoc.2015.11.041

[10] Kapur, P. K., Goswami, D. N., Bardhan, A. and Singh, O. (2008). Flexible software reliability growth model with testing effort dependent learning process. *Applied Mathematical Modelling*, *32*(7), 1298–1307. https://doi.org/10.1016/j.apm.2007.04.002

[11] Kapur, P. K., Pham, H., Anand, S. and Yadav, K. (2011). A Unified Approach for Developing Software Reliability Growth Models in the Presence of Imperfect Debugging and Error Generation. *IEEE Transactions on Reliability*, *60*(1), 331–340. https://doi.org/10.1109/TR.2010.2103590

[12] Khatri, S. K., Kumar, D., Dwivedi, A. and Mrinal, N. (2012). Software Reliability Growth Model with testing effort using learning function. *2012 CSI Sixth International Conference on Software Engineering (CONSEG)*, 1–5. https://doi.org/10.1109/CONSEG.2012.6349470

[13] Kumar, S., Aggarwal, A. G. . and Gupta, R. (2023). Modeling the Role of Testing Coverage in the Software Reliability Assessment. *International Journal of Mathematical, Engineering and Management Sciences*, *8*(3), 504–513. https://doi.org/10.33889/IJMEMS.2023.8.3.028

[14] Minamino, Y., Inoue, S. and Yamada, S. (2022). Extension of software reliability growth models by several testing-time functions. In *System Assurances* (pp. 155–174). Elsevier. https://doi.org/10.1016/B978-0-323-90240-3.00010-2

[15] Panwar, P. and Kaur, R. (2016). Effect of Imperfect Debugging on Prediction of Remaining Faults in Software. In *Advances in Intelligent Systems and Computing* (pp. 175–185). https://doi.org/10.1007/978-981-10-0448-3_14

[16] Pradhan, V., Dhar, J. and Kumar, A. (2022). Testing-Effort based NHPP Software Reliability Growth Model with Change-point Approach. *Journal of Information Science and Engineering*, *38*(2), 343–355. https://doi.org/10.6688/JISE.202203_38(2).0004

[17] Yamada, S. and Ohtera, H. (1990). Theory and Methodology Software reliability growth models for testing-effort control. *European Journal Of Operational Research*, *46*, 343–349.

[18] Yamada, S., Tokuno, K. and Osaki, S. (1992). Imperfect debugging models with fault introduction rate for software reliability assessment. *International Journal of Systems Science*, *23*(12), 2241–2252. https://doi.org/10.1080/00207729208949452

[19] Zhao, J., Liu, H.-W., Cui, G. and Yang, X.-Z. (2006). Software reliability growth model with change-point and environmental function. *Journal of Systems and Software*, *79*(11), 1578–1587. https://doi.org/10.1016/j.jss.2006.02.030

[20] Wood, A. (1996). Predicting software reliability. *Computer*, *29*(11), 69–77.

https://doi.org/10.1109/2.544240

PREDICTION OF RELIABILITY CHARACTERISTICS OF THRESHER PLANT BASIS ON GENERAL AND COPULA DISTRIBUTION

Urvashi [1], Shikha Bansal [2*]

•

1. Research Scholar, Department of Mathematics, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, 201204, India
urvashigodara8@gmail.com
2*. Assistant Professor, Department of Mathematics, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, 201204, India
* Corresponding author
srbansal2008@gmail.com

Abstract

In the agriculture field industry, farming tools play an important role. Any type of machinery's performance is influenced by factors including dependability, accessibility, and operating conditions. Different types of modern machinery are being used in today's modern world, so the farming system has become very easy. Thresher plants are essential equipment in the agriculture field industry, and these plants have many uses. A transition diagram for the system is used to develop a mathematical model of the thresher plant. Partial differential equations are created associated with the help of a transition diagram and solved using Laplace transforms and the supplementary variables approach to assessing the system's reliability. The copula approach was used to design the experiment, and the same methodology was used to assess the outcomes. The main aim of the present article is to evaluate the reliability factors, Profit, and sensitive analysis of a threshers plant. It is also possible to compute the dependability factor with the aid of general distributions and compare it to that copula distribution.

Keywords: Reliability, MTTF, sensitivity analysis, copula and general distribution, availability, profit.

1. Introduction

Especially for emerging nations, agriculture is a crucial component of the global economy. India is an agricultural field country. Here many people do farming and use different tools of farming machinery. In this field, the thresher farmer tool plays a very important role in farming. Different kinds of threshers have been used in today's machinery and the modern technological world. That's why now farming has become very easy for the farmer. It is necessary for good farming that the farming tools should work well. Basically, there are several uses for a thresher plant; it is used for cutting grass and threshing grain. This plant is mainly used for threshing grain. It is necessary for good farming results that all its equipment or its subsystem should work well. For this plant subsystem to properly, without fail, they have to be more reliable. Reliability theory's primary goal is to evaluate measurement mistakes and offer suggestions on improving the tests so that errors are reduced, maximizing profit, availability, and reliability. System reliability plays a very important role in working with any system or its component. When the system is failed, then repair with a repair facility, after repair system how is available. Analysis of a system's or

component's ability to function under particular conditions is known as reliability. In essence, it is a probability that a system will function correctly or not at a given moment.

Our civilization is becoming more complicated, and with that complexity come more pressing reliability issues that need to be resolved, and reliability issues are one of them. The area of reliability engineering is now receiving significant attention from many researchers and experts. Here is a brief description of these researchers and their research.

Singh, P.c and Pande [1] their model is based on the Chapman-Kolmogorov equation and uses five subsystems of the crystallizing system in the series form to analyze the availability and dependability of the system in sugar factories. Dhillon and B. S. [2] the essay examines the dependability and availability of a two-unit parallel system with warm standby and common-cause failures. Veera Raghavan, Trivedi, and K. S [3] Combinatorial models have been used to simulate the availability and dependability of complex systems without incurring the cost associated with massive Markov models. Perman and other [4] analyses using the power plant that When the transition probabilities are compared, it can be shown that the likelihood of the system refitting converges to its limiting value more rapidly than it does in the Markov model. Marquez and other authors [5] to calculate the reliability and availability of a complicated cogeneration plant employ the Monte Carlo and continuous time Monte Carlo simulation technique; in that paper, a case study for cogeneration plants is also presented. Bansal S [6] compares two independent repairable subsystems based on their availability to assess the complex system using the supplemental variable technique. Shikha Bansal and Sohan tyagi [7] calculated the reliability of a screw plant using the Boolean algebra approach and the orthogonal matrix method. Kumar, Modgil and others [8] use particle swarm optimization and genetic algorithm for availability analysis of ethanol manufacturing system; they analyze the result that PSO gives more accurate availability compared to genetic algorithm. Vinod, Amit, and others [9] introduce a new hybrid bacterial foraging algorithm and compare BFO and PSO optimization algorithms for optimizing the performance and availability of paint manufacturing systems. Yusuf, I., Ismai [10] analyze the reliability characteristic of the multi-computer systems using Laplace transforms and supplementary variable techniques. Saini, M., Raghav [11] using genetic and particle swarm optimization increase the urea decomposition system's dependability and availability. Tyagi and other researchers [12], using the Markov birth-death process, create a mathematical model of a leaf spring production facility and optimize the availability of a plant with regard to time utilizing C programming techniques. Sarwar and other researchers [13] determined sugar-producing plants' maintainability, availability, and dependability. Dionysiou, Bolbo, and other researchers [14] identify the best-fit distribution for a plant's failure and repair rates using various statistical features like skewness, kurtosis, and others. On a cruise ship, the lubricating oil system was examined in order to determine how to increase its dependability, availability, and safety. Bansal, Tyagi, and Verma [15] analyse the deviation of availability for screw plants with the help of the Markov birth-death Process and Matlab tool. Godara and Bansal [16] evaluate the reliability and availability of steam turbine generator plants using the boolean function technique and neural network approach. Tyagi and Bansal [17] optimize the performance of the Wastewater Treatment Process plant using the Runge-Kutta method. Chaudhary and Bansal [18] evaluate the reliability charismatic of the Hydro-Electric Power Station plant using Hydro-Electric Power Station.

There has not been a thorough analysis and sensitivity analysis of the availability of thresher plant subsystems. However, a few articles are accessible on the economics of production and maintenance management in the thresher plant.

Our primary goal is to study the system's availability and reliability utilizing alternative distributions; utilizing them maximises the thresher plant system availability. In this research, we analyze three possible states of the system: good, reduce, and failed. In this article, we also discuss the profit and sensitivity analysis of the system; however, we maximise the plant's profit and how to vary the sensitivity analysis of the thresher plant.

2. Notation

t	Time frame with a time variable
s	Expression of a Laplace transformation
τ_i	Subsystem i failure frequency.
ξ_i	Subsystem repair with copula repair facility.
$\omega_i(x)$	Subsystem repair with general repair facility
$P_i(x,t)$	The system state probability in i^{th} state.
$P_i(x,t)$	Laplace transform of i^{th} state probability.
$S_\omega(x)$	Standard general distribution.
$C_\theta(\xi_(x))$	Standard Copula distribution.
$E_p(t)$	Predicted profit throughout the time period [0,t].
M_1, M_2	Per unit of time, revenue and service costs.

3. Assumption

The following assumption is taken through the model description.

- Initially, the whole system is completely operable or working state.
- Good, reduced, and failed states are the three possible system states.
- System is repairable after complete failed and partial failed state, i.e. repair facility available in the system.
- After repair system works properly in a good state, and there are two types of repair facilities in the system: general repair and copula repair facility.
- When systems partially fail, follow the general standard distribution.
- When systems are completely failed, then follow Copula standard distribution.

4. Description of The System

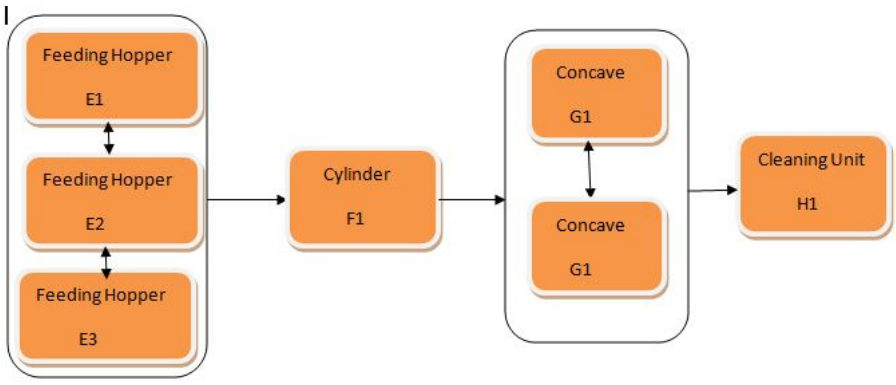


Figure 1: *Block diagram for Thresher Plant*

The machine's availability affects how well it performs in agricultural applications, the equipment's environment, its procedure, and how effectively it is maintained. The block diagram of the thresher plant is shown in Figure 1. There are four subsystems in the plant. In this part, we give a brief description of the thresher plant. It involves removing grain from the plant by rupturing, trampling, and striking. It is the most crucial element of agricultural automation.

1. ***FeedingHopper*** **:** — It is placed on the top of the threshing cylinder. Grain is first entered through the hopper. To deliver uniform (equally) sized grains to the drum, this machine uses a revolving star wheel mechanism between the hopper and the threshing drum.

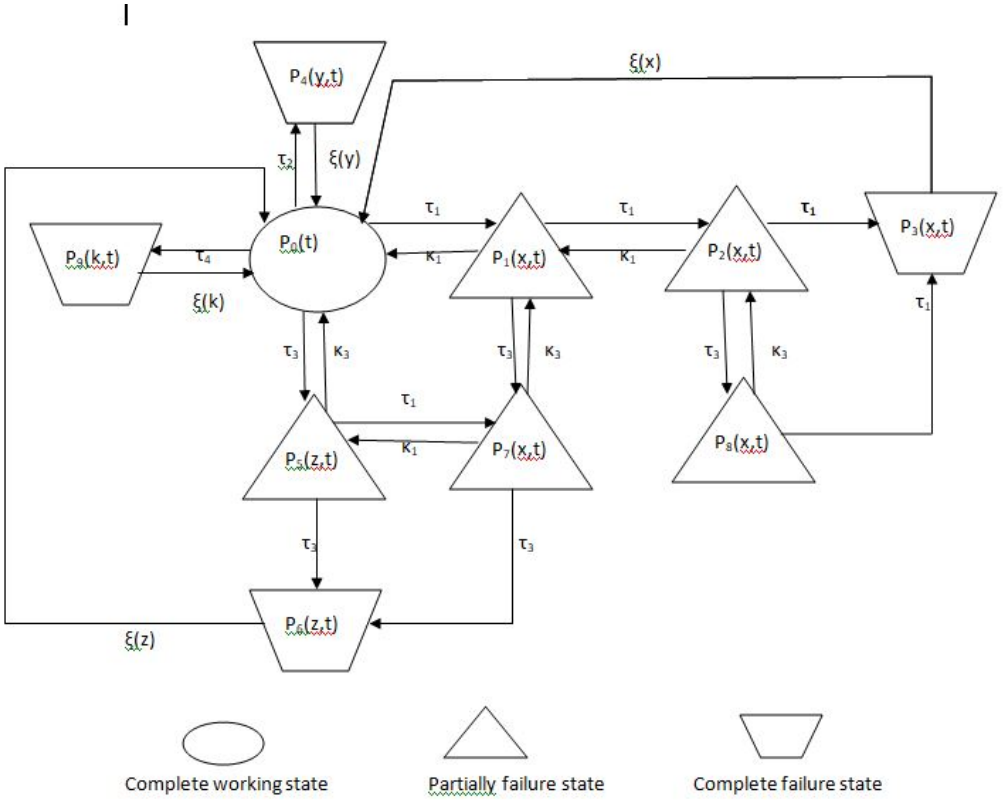


Figure 2: *Transition diagram for Thresher Plant*

2. ***Cylinder*** **:** —There are different types of cylinders used in the thresher plant, e.g., wire loop type, spike tooth type, and rasp bar type cylinder. These primarily used cylinders in the threshing plant. Grain enters in the moving direction of the cylinder.

3. ***Concave*** **:** —It is a curved piece of iron and steel situated next to the threshing cylinder. It removes grain from the straw and separates grain from the harvest. Only grain and a limited quantity of chaff are allowed to pass through it, holding the feed crop inside the threshing chamber. According to the crop size, the distance between the concave and the cylinder can be adjusted.

4. ***CleaningUnit*** **:** — With the use of blowers and aspirators, the grain is separated from the chalk.

5. State Description

- S_0 **:** —In the initial state, Units E_1, F_1, G_1, and H_1 are operational. Units E_2, E_3, and G_2 are currently in standby mode.
- S_1 **:** — In this state , the unit E_1 failed and under repair. And the elapsed repair time is (x,t). While the unit E_2, F_1, G_1 and H_1, are on operation and the units E_3 and G_2 are on standby.
- S_2 **:** — Units E_1 and E_2 failed and were being repaired in this state. The repair time that has already passed is (x,t). While the units E_2,F_1,G_1,H_1, and E_3 are in operation and on standby, respectively.
- S_3 **:** — In both the units E_1 and G_1, there has been a failure. The units E_2 and E_3 are on standby while E_2, F_1, G_2, and H_1 are in operation.
- S_4 **:** —In this instance, the subsystem 1 units E_1 and E_2 are involved. Additionally, G_1 from subsystem three has failed and is being repaired. When performing operations, the unit E_3, from subsystem 1, F_1, from subsystem 2, G_2, from subsystem 3, and H_1 from subsystem 4.

- S_5 : — Due to subsystem 2's failure, state S_5 is a fully failed state.
- S_6 : — Due to subsystem 1's failure, state S_6 is a fully failed state.
- S_7 : — Due to subsystem 3 failure, state S_5 is a fully failed state.
- S_8 : — Due to subsystem 4 failure, state S_5 is a fully failed state.

6. Formulation and Solution of The Mathematical Model

The mathematical model solves the following collection of distinct differential equations. These differential equations are derived with the help of a transition diagram. Hence the mathematical model is as follows.

$$[\frac{\partial}{\partial t}+\tau_1+\tau_2+\tau_3+\tau_4]P_0=\int_0^\infty \kappa_1 P_1(y,t)+\int_0^\infty \xi(x)P_4(x,t)+\int_0^\infty \kappa_3 P_5(z,t) + \int_0^\infty \xi(k)P_9(y,t)+\int_0^\infty \xi(x)P_3(x,t)+\int_0^\infty \xi(z)P_5(z,t) \tag{1}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\tau_1+\tau_3+\kappa_1)P_1(x,t)=0 \tag{2}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\tau_1+\tau_3+\kappa_1)P_2(x,t)=0 \tag{3}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\xi(x))P_3(x,t)=0 \tag{4}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\xi(y))P_4(y,t)=0 \tag{5}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial z}+\tau_1+\tau_3+\kappa_3)P_5(x,t)=0 \tag{6}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial z}+\xi(z))P_6(z,t)=0 \tag{7}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\tau_3+2\kappa_1)P_7(x,t)=0 \tag{8}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\tau_1+\kappa_3)P_8(x,t)=0 \tag{9}$$

$$(\frac{\partial}{\partial t}+\frac{\partial}{\partial k}+\xi(k))P_9(y,t)=0 \tag{10}$$

BoundaryCondition

$$P_1(0,t)=\tau_1 P_0(t) \tag{11}$$

$$P_2(0,t)=\tau_2^2 P_0(t) \tag{12}$$

$$P_3(0,t)=(\tau_1^2+\tau_1^3\tau_3)P_0(t) \tag{13}$$

$$P_4(0,t)=\tau_2 P_0(t) \tag{14}$$

$$P_5(0,t)=\tau_3 P_0(t) \tag{15}$$

$$P_6(0,t)=(\tau_3^2+2\tau_3^2\tau_1)P_0(t) \tag{16}$$

$$P_7(0,t)=2\tau_1\tau_3 P_0(t) \tag{17}$$

$$P_8(0,t)=\tau_1^2\tau_3 P_0(t) \tag{18}$$

$$P_9(0,t)=\tau_4 P_0(t) \tag{19}$$

InitialConditions

$$P_0(t)=\begin{cases}1 & \text{initial}\\ 0 & \text{Otherwise}\end{cases} \tag{20}$$

We use the Laplace technique to solve the (1)-(10) equation. Now taking Laplace transfer of equation (1)-(19) with the help of equation (20), we get the following equation.

$$\begin{aligned}\big[s+\tau_1+\tau_2+\tau_3+\tau_4\big]\bar{P}_0(t) = \int_0^\infty \kappa_1\bar{P}_1(x,s)dx + \int_0^\infty \xi(y)\bar{P}_4(y,s)dy + \\ \int_0^\infty \kappa_3\bar{P}_5(z,s)dz + \int_0^\infty \xi(k)\bar{P}_9(k,s)dk + \int_0^\infty \xi(x)\bar{P}_3(x,s)dx + \int_0^\infty \xi(z)\bar{P}_6(z,s)dz\end{aligned} \tag{21}$$

$$(s+\frac{\partial}{\partial x}+\tau_1+\tau_3+\kappa_1)\bar{P}_1(x,s)=0 \tag{22}$$

$$(s+\frac{\partial}{\partial x}+\tau_1+\tau_3+\kappa_1)\bar{P}_2(x,s)=0 \tag{23}$$

$$(s+\frac{\partial}{\partial x}+\xi(x))\bar{P}_3(x,s)=0 \tag{24}$$

$$(s+\frac{\partial}{\partial y}+\xi(y))\bar{P}_4(y,s)=0 \tag{25}$$

$$(s+\frac{\partial}{\partial z}+\tau_1+\tau_3+\kappa_3)\bar{P}_5(z,s)=0 \tag{26}$$

$$(s+\frac{\partial}{\partial z}+\xi(z))\bar{P}_6(z,s)=0 \tag{27}$$

$$(s+\frac{\partial}{\partial x}+\tau_3+2\kappa_1)\bar{P}_7(x,s)=0 \tag{28}$$

$$(s+\frac{\partial}{\partial x}+\tau_1+\kappa_3)\bar{P}_8(x,s)=0 \tag{29}$$

$$(s+\frac{\partial}{\partial k}+\xi(k))\bar{P}_9(k,s)=0 \tag{30}$$

Now taking Laplace's transfer of boundary conditions

$$\bar{P}_1(0,s)=\tau_1\bar{P}_0(s) \tag{31}$$

$$\bar{P}_2(0,s)=\tau_1^2\bar{P}_0(s) \tag{32}$$

$$\bar{P}_3(0,s)=\big(\tau_1^3+\tau_1^3\tau_3\big)\bar{P}_0(s) \tag{33}$$

$$\bar{P}_4(0,s)=\tau_2\bar{P}_0(s) \tag{34}$$

$$\bar{P}_5(0,s)=\tau_3\bar{P}_0(s) \tag{35}$$

$$\bar{P}_6(0,s)=(\tau_3^2+2\tau_3^2\tau_1)\bar{P}_0(s) \tag{36}$$

$$\bar{P}_7(0,s)=2\tau_1+\tau_3\bar{P}_0(s) \tag{37}$$

$$\bar{P}_8(0,s)=\tau_1^2\tau_3\bar{P}_0(s) \tag{38}$$

$$\bar{P}_9(0,s)=\tau_4\bar{P}_0(s) \tag{39}$$

Through the use of boundary conditions (31) to (39) and solving equations (22) to (30) using Laplace shifting properties. For calculating P_{up} and P_{down} with the help of a transition diagram. As we know that P_{up} is the probability for a good and partially failed state because the repair facility is available in our model, so the repair rate repairs a partially failed state, and it converts to an operable state. Below also used the total sum of probability theorem.

$$\bar{P}_{up}(s)=\bar{P}_0(s)+\bar{P}_1(s)+\bar{P}_2(s)+\bar{P}_5(s)+\bar{P}_7(s)+\bar{P}_9(s) \tag{40}$$

$$\begin{aligned}\bar{P}_{up}(s) = [1+\tau_1\{\frac{1-\bar{S_{\kappa_1}}(s+\tau_1+\tau_3)}{s+\tau_1+\tau_3}\}+\tau_1^2\{\frac{1-\bar{S_{\kappa_1}}(s+\tau_1+\tau_3)}{s+\tau_1+\tau_3}\}+ \\ \tau_3\{\frac{1-\bar{S_{\kappa_3}}(s+\tau_1+\tau_3)}{s+\tau_1+\tau_3}\}+2\tau_3\tau_1\frac{1-S_{2\kappa_1}(s+\tau_3)}{S+\tau_3}\}+\tau_1^2\tau_3\{\frac{1-S_{\kappa_3}(s+\tau_1)}{S+\tau_1}\}]P_0(t)\end{aligned} \tag{41}$$

$$\bar{P}_{down}(s)=1-\bar{P}_{up}(s) \tag{42}$$

7. Reliability Characteristic Analysis

7.1. Availability analysis of the system with general distribution

For calculating the availability of the system, we use general repair distribution. General repair distribution is defined as

$$\bar{P}_{\xi}(s) = \frac{\xi}{s+\xi} \tag{43}$$

Taking failure and repair rate as follows: $\tau_1 = 0.0001, \tau_2 = 0.0002, \tau_1 = 0.0001, \tau_3 = 0.0003, \tau_4 = 0.0004, \xi(x) = \xi(y) = \xi(z) = \xi(k) = \kappa_1(x) = \kappa_2(y) = \kappa_3(z) = \kappa_4(k) = 1$ Using the failure and repair rates mentioned above, use the inverse Laplace transfer of equation (41).

$$\begin{aligned} P_{up}(t) = 0.00012009e^{-1.00020005t} - 0.000000030029e^{-2.0001t} + \\ 0.0004793028e^{-1.00119998t} + 0.9994007089e^{-0.00000006985t} \end{aligned} \tag{44}$$

Table 1: *Availability analysis corresponding to time*

Time(t)In days	Availability
0	1
20	0.999399
40	0.999397
60	0.999396
80	0.999395
100	0.999393
120	0.999392
140	0.999390
160	0.999389
180	0.999388
200	0.999386

Time t=0,20,40,60,80,100,120,140,160,180 is considered to determine the system's availability, and Table 1 is provided.

7.2. Availability analysis of the system with Copula distribution

We use general repair and Copula repair distribution to calculate the system's availability. Copula repair distribution is defined as

$$S_{\alpha}(s) = \frac{exp[x^{\delta} + \{log\xi(x)\}^{\delta}]^{\frac{1}{\delta}}}{s + exp[x^{\delta} + \{log\xi(x)\}^{\delta}]^{\frac{1}{\delta}}} \tag{45}$$

General repair distribution is expressed as,

$$\bar{P}_{\xi}(s) = \frac{\xi}{s+\xi} \tag{46}$$

Taking failure and repair rate as follows: $\tau_1 = 0.0001, \tau_2 = 0.0002, \tau_1 = 0.0001, \tau_3 = 0.0003, \tau_4 = 0.0004, \xi(x) = \xi(y) = \xi(z) = \xi(k) = \kappa_1(x) = \kappa_2(y) = \kappa_3(z) = \kappa_4(k) = 1$ Using the failure and repair rates mentioned above, use the inverse Laplace transfer of equation (41)

$$\begin{aligned} \bar{P}_{up}(t) = 0.99977958e^{-0.0000000699t} - 0.00000030918e^{-1.0008t} \\ -0.00000002999e^{-2.0001t} + 0.0000000330e^{-2.718581739t} + 0.000220730e^{-2.71888215t} \end{aligned} \tag{47}$$

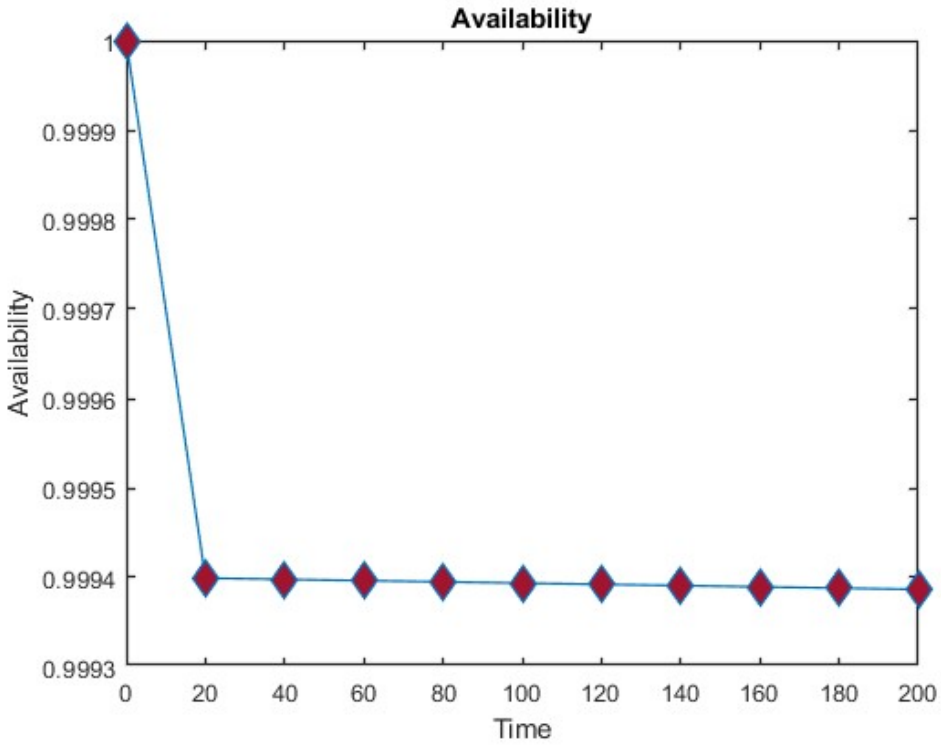


Figure 3: *Availability with general distribution*

Table 2: *Availability analysis with Copula repair corresponding to time*

Time(t)In days	Availability
0	1
20	0.999778
40	0.999776
60	0.999775
80	0.999773
100	0.999772
120	0.999771
140	0.999769
160	0.999768
180	0.999766
200	0.999765

Time t=0,20,40,60,80,100,120,140,160,180 is taken into consideration to determine the system's availability, and Table 2 is provided.

7.3. Reliability Analysis

When doing an analysis or evaluating the system's reliability, we'll assume there isn't a repair facility. In this instance, we assume that there is no repair facility. i.e., all repair rates are zero $\xi(x) = \xi(y) = \xi(z) = \xi(k) = \kappa_1(x) = \kappa_2(y) = \kappa_3(z) = \kappa_4(k) = 0$ and failure rate are $\tau_1 = 0.0001, \tau_2 = 0.0002, \tau_3 = 0.0003, \tau_4 = 0.0004$ Using equation (45) and (46), taking the Laplace transfer of equation (41). One can evaluate the reliability expression in (12) to the computed

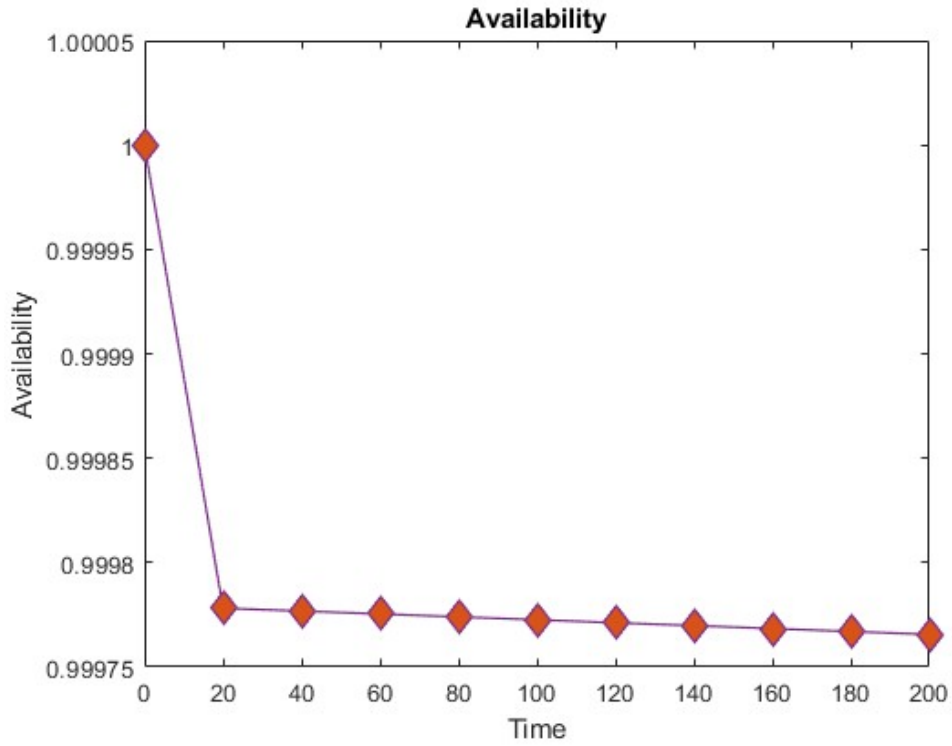


Figure 4: *Availability with Copula distribution*

values that are shown in Table 3.

$$R(t) = 0.000007e^{-0.00001t} + 0.66667e^{-0.00004t} + 0.3333250e^{-0.0001t} \tag{48}$$

Table 3: *Reliability analysis corresponding to Time*

Time(t)In days	Reliability
0	1.000
20	0.988
40	0.976
60	0.964
80	0.953
100	0.942
120	0.931
140	0.920
160	0.909
180	0.898
200	0.888

7.4. MEAN TIME TO FAILURE (MTTF) OF THE SYSTEM

The mean time to failure is the predicted time elapsed between the system's normal functioning and its first failure. Simply put, the average amount of time between system failures is known as the mean time to failure MTTF. In this, we considered that there is no repair facility. There we

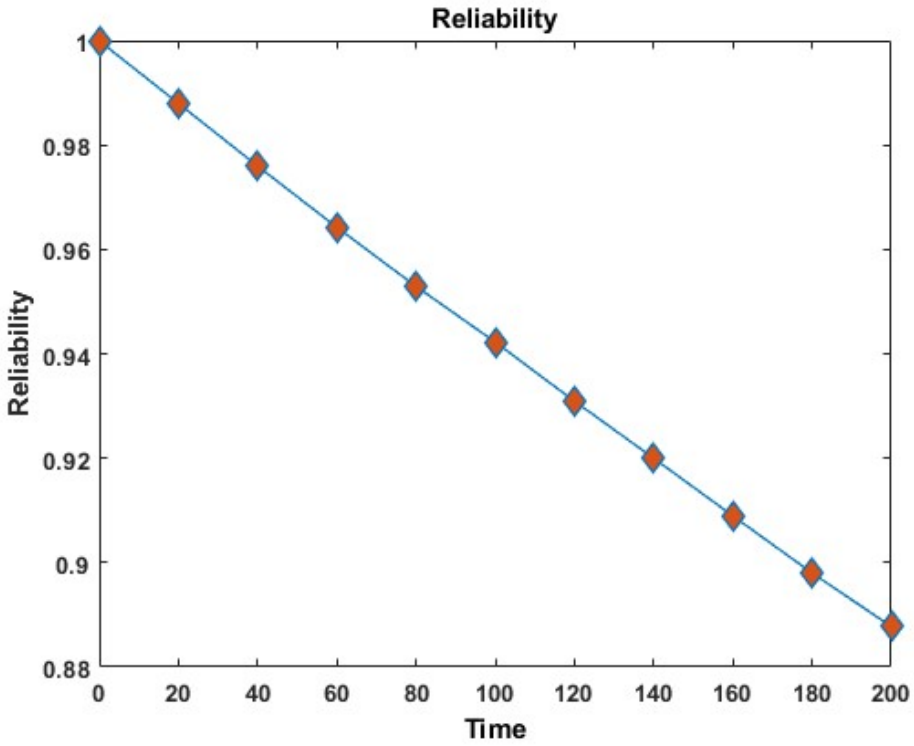


Figure 5: *Reliability Corresponding To Time*

assign all repair rate is zero, and in equation (45) limit tend to be zero; the MTTF of the system is obtained in equation (49).

$$MTTF = \left(\frac{\tau_1 + \tau_1^2 + \tau_3}{\tau_1 + \tau_3} + \frac{2\tau_1\tau_3}{\tau_3} + \frac{\tau_1^2\tau_3}{\tau_1}\right)\frac{1}{\tau_1\tau_2\tau_3\tau_4} \tag{49}$$

Table 4: *System MTTF corresponding to failure rate*

Failure rate	MTTFτ_1	MTTFτ_2	MTTFτ_3	MTTFτ_4
0.0001	2000.2	2222.5	2500.3	2857.5
0.0002	1818.6	2000.2	2222.5	2500.3
0.0003	1667.3	1818.4	2000.2	2222.5
0.0004	1539.3	1666.9	1818.4	2000.2
0.0005	1429.5	1538.6	1666.8	1818.4
0.0006	1334.4	1428.7	1538.6	1666.9
0.0007	1251.2	1333.5	1428.7	1538.6
0.0008	1177.8	1250.1	1333.5	1428.7
0.0009	1112.5	1176.6	1250.1	1333.5

Fix the failure rate $\tau_1 = 0.001, \tau_2 = 0.002, \tau_3 = 0.003, \tau_4 = 0.004$ and fluctuate the failure rate. $\tau_1, \tau_2, \tau_3, \tau_4$ one after the other respectively as $\tau = 0.001, 0.002, 0.003, 0.004,$ $0.005, 0.006, 0.007, 0.008, 0.009$ is taken into consideration to determine the system's MTTF, and Table 4 is provided.

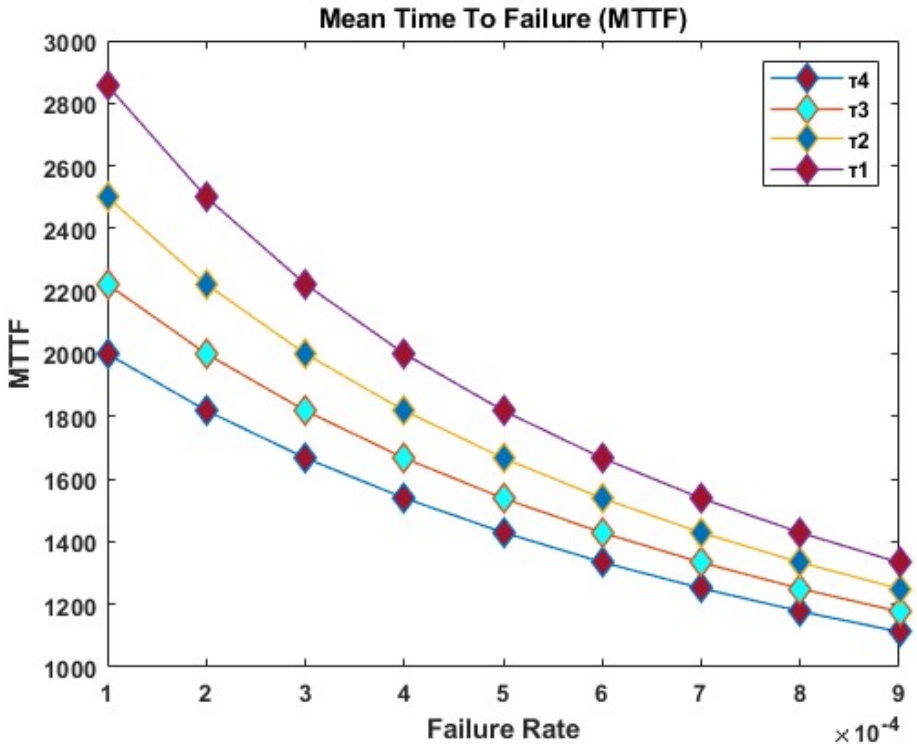


Figure 6: *MTTF Corresponding To Failure rate*

7.5. Analysis of sensitivity for the MTTF

Sensitivity analysis is the partial derivative of MTTF with respect to the failure rate. Now taking the partial derivative of the equation (49) with respect to the failure rate. The sensitivity analysis of the system is given below in Table 5

Table 5: *Sensitivity analysis corresponding to failure rate*

Failure rate	$\frac{\partial(MTTF)}{\tau_1}$	$\frac{\partial(MTTF)}{\tau_2}$	$\frac{\partial(MTTF)}{\tau_3}$	$\frac{\partial(MTTF)}{\tau_4}$
0.0001	$-19978 * 10^2$	$-24694 * 10^2$	$-31257 * 10^2$	$-40821 * 10^2$
0.0002	$-16509 * 10^2$	$-20002 * 10^2$	$-24695 * 10^2$	$-31254 * 10^2$
0.0003	$-13871 * 10^2$	$-16531 * 10^2$	$-20003 * 10^2$	$-24694 * 10^2$
0.0004	$-11819 * 10^2$	$-1389 * 10^2$	$-16531 * 10^2$	$-20002 * 10^2$
0.0005	$-1019 * 10^3$	$-11836 * 10^2$	$-13891 * 10^2$	$-16531 * 10^2$
0.0006	$-88767 * 10^1$	$-10205 * 10^2$	$-11836 * 10^2$	$-1389 * 10^3$
0.0007	$-78017 * 10^1$	$-88899 * 10^1$	$-10205 * 10^2$	$-11836 * 10^2$
0.0008	$-69108 * 10^1$	$-78134 * 10^1$	$-88899 * 10^1$	$-1.0205e + 06$
0.0009	$-61642 * 10^1$	$-69212 * 10^1$	$-78134 * 10^1$	$-88899 * 10^1$

7.6. COST AND PROFIT ANALYSIS FOR THE MODEL

***Case*1 :** Profit Analysis For The Model with General Distribution
when the capacity for the service is available; after that, the equation provides the predicted profit in the interval [0, t].

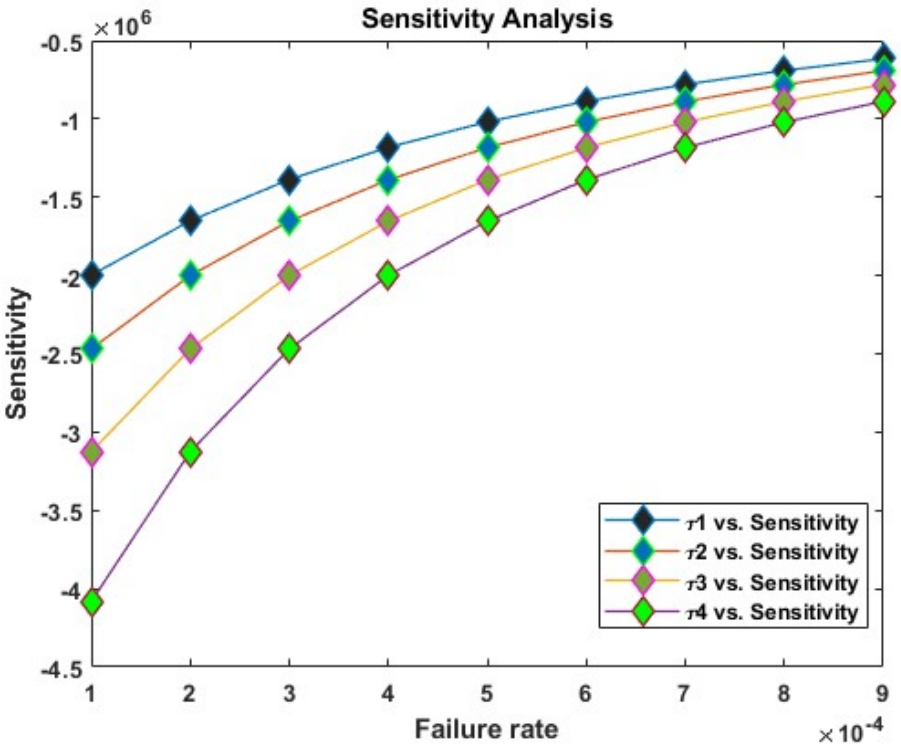


Figure 7: *Sensitivity Analysis Corresponding To Failure rate*

$$E_p(t) = C_1 \int_0^t P_{up}(t)dt - C_2 t. \tag{50}$$

Then using equation (44), equation (50) can be written as

$$E_p(t) = C1(2.5385 * 10^{-25} e^{-0.004t} - 5.3288 * 10^{-20} e^{-2252250286237519t} - 1.8937 * 10^{-16} e^{5277504776533377t} + 1.5014 * 10^{-8} e^{-2.0001t} - 4.7873 * 10^{-4} e^{-1.0012t} + 1.4308 * 10^{7}) - C2t \tag{51}$$

Assuming $C_1 = 1$ and $C_2 = 0.9, 0.8, 0.7, 0.6, 0.5$ and 0.4 respectively and t= 0,20,40,60,80,100,120,140 unit of time than expected profit is given in the Table 6 with the help of equation

Table 6: *Profit estimation with general distribution*

Time	$E_p(t), C_2 = 0.9$	$E_p(t), C_2 = 0.8$	$E_p(t), C_2 = 0.7$	$E_p(t), C_2 = 0.6$	$E_p(t), C_2 = 0.5$
0	0	0	0	0	0
20	1.9886	3.9886	5.9886	7.9886	9.9886
40	3.9766	7.9766	11.9766	15.9766	19.9766
60	5.9645	11.9645	17.9645	23.9645	29.9645
80	7.9524	15.9524	23.9524	31.9524	39.9524
100	9.94032	19.94032	29.94032	39.94032	49.94032
120	11.92818	23.92818	35.92818	47.92818	59.92818
140	13.9160	27.9160	41.9160	55.9160	69.9160
160	15.9038	31.9038	47.9038	63.9038	79.9038
180	17.8916	35.8916	53.8916	71.8916	89.8916

***Case*2 :** Profit Analysis For The Model with Copula Distribution

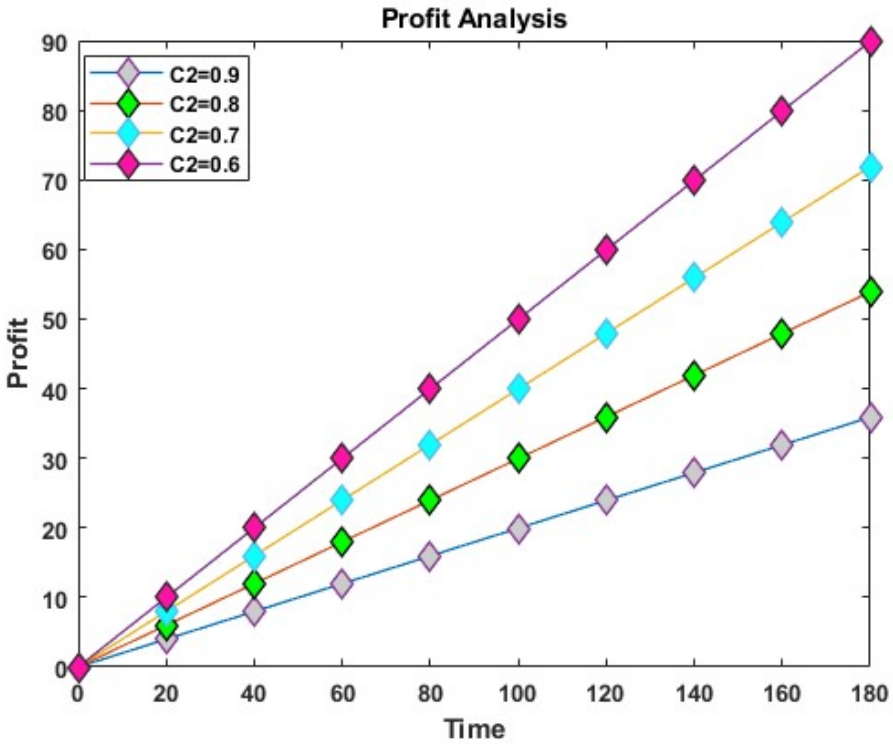


Figure 8: *Profit Corresponding to time with general distribution*

when the capacity for the service is available, thereafter, the equation provides the predicted profit in the interval [0, t].

$$E_p(t) = C_1 \int_0^t P_{up}(t)dt - C_2 t. \tag{52}$$

Then using equation (47), equation (52) can be written as

$$\begin{aligned} E_p(t) = C1(1.4992 * 10^{-18} e^{-2.0001t} - 1.2148 * 10^{-8} e^{-2.7186t} - 3.6164 * 10^{-5} \\ e^{6.1035t} - 1.4306 * 10^{7} e^{-6.9884 * 10^{-8} t} + 3.0894 * 10^{-7} e^{-1.0008t} + 1.4306 * 10^{7}) - C2 * t \end{aligned} \tag{53}$$

Assuming $C_1 = 1$ and $C_2 = 0.9, 0.8, 0.7, 0.6, 0.5$ and 0.4 respectively and t= 0,20,40, 60,80,100,120,140 unit of time than expected profit is given in Table 7 with the help of equation

Table 7: *Profit estimation with Copula distribution*

Time	$E_p(t), C_2 = 0.9$	$E_p(t), C_2 = 0.8$	$E_p(t), C_2 = 0.7$	$E_p(t), C_2 = 0.6$	$E_p(t), C_2 = 0.5$
0	0	0	0	0	0
20	1.9957	3.9957	5.9957	7.9957	9.9957
40	3.9912	7.9912	11.991	15.991	19.991
60	5.9867	11.987	17.987	23.987	29.987
80	7.9822	15.982	23.982	31.982	39.982
100	9.9777	19.978	29.978	39.978	49.978
120	11.973	23.973	35.973	47.973	59.973
140	13.969	27.969	41.969	55.969	69.969
160	15.964	31.964	47.964	63.964	79.964
180	17.959	35.959	53.959	71.959	89.959

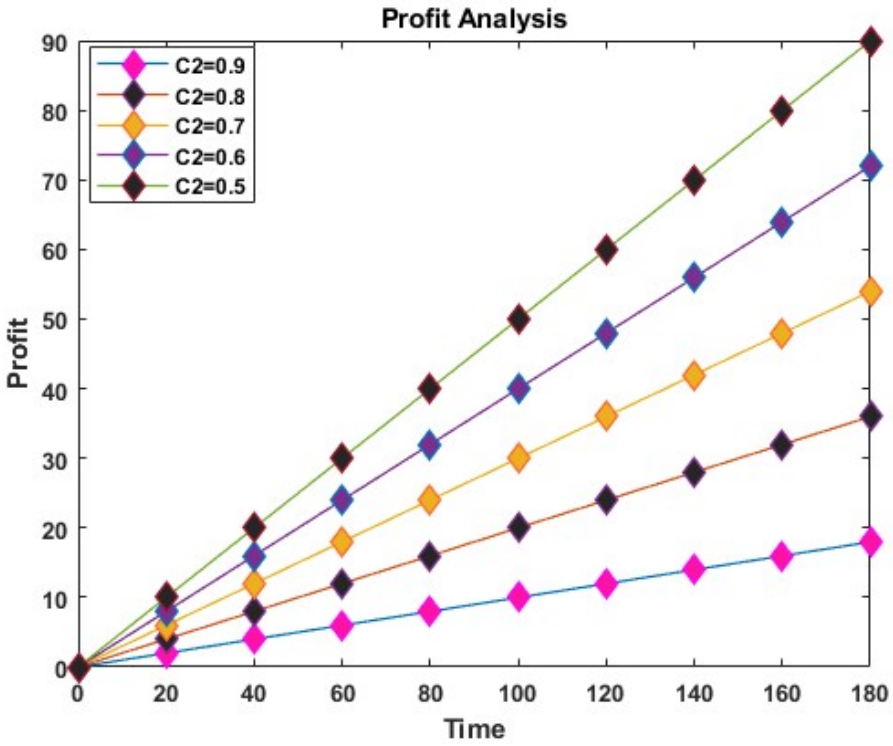


Figure 9: *Profit Corresponding to time with copula distribution*

8. Conclusion

When failures follow a general distribution, information on the system's availability is provided in Figure 3 and Table 1 for each time period. On the basis of it, we draw the conclusion that as time increases, system availability slowly decreases. When failures follow a copula distribution, information on the system's availability is provided in Figure 4 and Table 2 for each time period. On the basis of it, we draw the conclusion that as time increases, system availability slowly decreases. Tables 1 and 2 allow us to conclude that the copula distribution scheme provides more availability than the general distribution. We can predict system availability with respect to time using Tables 1 and 2. For a long time, system availability tends to be zero.

In the absence of a repair facility, information on the system's reliability is provided in Figure 5 and Tables 3 for each time period. On the basis of it, we draw the conclusion that as time increases, system reliability rapidly decreases. Figures 3,4 and 5 show that the system's performance is improved greatly when repair facilities are available.

Table 4 and Figure 6 provide information about the system's mean time to failure(MTTF); for the different failure rate parameters, MTTF gives different values. we draw the conclusion that as the failure rate increases, system MTTF rapidly decreases.

Sensitivity analysis for the MTTF corresponding to failure rate is shown in Table 5 and Figure 6. Tables 6 and 7 and Figures 8 and 9 show the system's profit and cost analysis corresponding to time; here, revenue cost C1=1 is fixed, and service cost(c2) is variable between 0.5 and 0.9. That leads us to the conclusion that the system's profit increases quickly whenever time increases while service costs remain low. The yield is larger when the service cost (c2) is 0.5, and the profit is lowest when the service cost is high (c2=0.9). Conclusion: To maximize profit, service costs must be kept to a minimum.

References

[1] Kumar, D., Singh, J., Pandey, P. C. (1992). Availability of the crystallization system in the sugar industry under common-cause failure. IEEE Transactions on Reliability, 41(1), 85-91.
[2] Dhillon, B. S. (1993). Reliability and availability analysis of a system with warm standby and common cause failures. Microelectronics Reliability, 33(9), 1343-1349.
[3] Veeraraghavan, M., & Trivedi, K. S. (1993, October). An approach for combinatorial performance and availability analysis. In Proceedings of 1993 IEEE 12th Symposium on Reliable Distributed Systems (pp. 24-33). IEEE.
[4] Perman, M., Senegacnik, A., Tuma, M. (1997). Semi-Markov models with an application to power-plant reliability analysis. IEEE Transactions on Reliability, 46(4), 526-532.
[5] Marquez, A. C., Heguedas, A. S., Iung, B. (2005). Monte Carlo-based assessment of system availability. A case study for cogeneration plants. Reliability Engineering System Safety, 88(3), 273-289.
[6] Bansal, Shikha. "Availability Analysis of a Repairable Redundant System Under Preemptive-Resume Repair Discipline." International Journal of Mathematics And its Applications 6.1-D (2018): 665-671.
[7] Bansal, Shikha, and Sohan Tyagi. "Reliability analysis of screw manufacturing plant using orthogonal matrix method." Pertanika Journal of Science Technology 26.4 (2018): 1789-1800.
[8] Kumar, A., Kumar, V., Modgil, V. (2018). Performance optimization for ethanol manufacturing system of distillery plant using particle swarm optimisation algorithm. International Journal of Intelligent Enterprise, 5(4), 345-364.
[9] Kumar, A., Kumar, V., Modgil, V. (2019). Performance modelling and optimization for a complex repairable system of paint manufacturing unit using a hybrid BFO-PSO algorithm. International Journal of Quality & Reliability Management.
[10] Yusuf, I., Ismail, A. L., Singh, V. V., Ali, U. A., Sufi, N. A. (2020). Performance analysis of multi-computer system consisting of three subsystems in series configuration using copula repair policy. SN Computer Science, 1(5), 1-11.
[11] Saini, M., Raghav, Y. S., Kumar, A., Chandnani, D. (2021). Availability and performance optimization of urea decomposition system using genetic algorithm and particle swarm optimization. Life Cycle Reliability and Safety Engineering, 10(3), 285-293.
[12] Tyagi, S. L., Bansal, S., Agarwal, P., & Yadav, A. S. (2021). Mathematical Modeling and Availability Analysis of Leaf Spring Manufacturing Plant. Pertanika Journal of Science & Technology, 29(2).
[13] Saini, M., Kumar, A., & Sinwar, D. (2022). Parameter estimation, reliability and maintainability analysis of sugar manufacturing plant. International Journal of System Assurance Engineering and Management, 13(1), 231-249.
[14] Dionysiou, K., Bolbot, V., & Theotokatos, G. (2022). A functional model-based approach for ship systems safety and reliability analysis: Application to a cruise ship lubricating oil system. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, 236(1), 228-244.
[15] Bansal, S., Tyagi, S. L., & Verma, V. K. (2022). Performance Modeling and Availability Analysis of Screw Manufacturing Plant. Materials Today: Proceedings, 57, 1985-1988.
[16] U. Godara and S. Bansal, "Performance of Reliability Factors in Steam Turbine Generator Power Plant Using Boolean Function Technique and Neural Network Approach," 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, 2023, pp. 507-512.
[17] S. L. Tyagi and S. Bansal, "Optimization Model for Wastewater Treatment Process," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 165-168.
[18] P. Chaudhary and S. Bansal, "Assessment of the Reliability Performance of Hydro-Electric Power Station," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 148-153.

APPLICATION OF MACHINE LEARNING ALGORITHMS IN THE PROBLEMS OF IMPROVING MODE RELIABILITY OF MODERN POWER SYSTEMS

Viktor Kurbatsky[1], Huseyngulu Guliyev[2], Nikita Tomin[1],
Famil Ibrahimov[2], Nijat Huseynov[3]

•

[1]Melentyev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences,
Irkutsk, Russia,
664033, Lermontov str., 120
e-mail: kurbatsky@isem.irk.ru, tomin@isem.irk.ru

[2]Azerbaijan Technical University, Baku, Azerbaijan
AZ1073, 25 Hussein Javid pros.
e-mail: huseyngulu@mail.ru, amfanet@mail.ru

[3]Sumgait State University, Sumgait, Azerbaijan,
AZ1073, 43 quarter, SDU.
e-mail: nicathuseynov739@gmail.com

Abstract

In order to increase the regime reliability of energy systems, the experience of applying machine learning algorithms and models for various issues of operative-dispatching and counter-accident management was reviewed. It is indicated that an effective solution to this problem is the use of machine learning algorithms and models that are able to learn to predict and control the operating modes of the power system, taking into account many changing influencing factors. The experience of using machine learning technology in the tasks of operational dispatch and emergency control of EPS is presented, which clearly shows the prospects of such studies for subsequent practical implementation in the work of various automated control systems for electric power networks of power systems. Until recently, models based on neural network structures have remained the most popular among machine approaches in predictive problems. The advantages of using this structure are shown, first of all, by the fact that the neural network structure makes it possible to obtain models with good approximating abilities. A comparative analysis of the effectiveness of various models in predicting electricity consumption is given. The issues of voltage and reactive power regulation in the electrical network of power systems using an artificial neural network are considered and the effectiveness of this approach is shown. A model and algorithm for estimating voltage stability in power system nodes under various influencing factors is proposed, as well as results are presented that confirm the reliability of the results obtained.

Keywords: power system, electric power system, artificial intelligence, machine learning, operational control of power system modes, mode prediction, artificial neural network, voltage and reactive power regulation

1. Introduction

A characteristic feature of the trend in the development of modern energy systems is the

increase in the share of non-traditional energy sources - these are alternative energy sources, such as wind turbines, solar PV installations, geothermal, converters, fuel cells and other types of renewable sources. The generation of power from non-traditional sources such as wind and solar power plants is stochastically variable, and for this reason the system operator cannot control their generation.

The development of artificial intelligence (AI) methods made it possible to significantly speed up and automate the solution of this and a whole range of tasks for managing the modes of electric power systems (EPS) in the context of integrating renewable energy sources (RES) [1]. In recent decades, one of the actively developed areas within the Smart Grid technology platform is the use and implementation of machine learning (ML) technology, which includes methods for constructing algorithms that can learn. The use of various types of learning ML models: with a teacher (Supervisedlearning), without a teacher (Unsupervisedlearning), with reinforcement (Reinforcementlearning), deep learning (Deeplearning), etc., made it possible to create original adaptive, trainable software modules regulation and control of both individual devices and subsystems of the EPS, and its mode as a whole. Their main advantages are speed, high adaptability, the ability to approximate nonlinear functions and the presence of a certain kind of machine intelligence, which allows developing the most autonomous systems capable of independent decision-making based on experience and original generalization properties.

The article reflects the experience of using the authors of algorithms and models of ML for various tasks of operational dispatch and emergency control of EPS, concomitant with an increase in the regime reliability of power systems.

II. Review of the application of ml in problems of operational control of power system modes

A necessary condition for the effective management of modern EPS is the availability of a short-term forecasting tool. The relevance of this task is predetermined by the fact that when analyzing and forming control actions only according to past data, a delayed reaction to the behavior of the EPS is obtained, which causes overestimated requirements for static and dynamic stability reserves, and also reduces the efficiency of control systems.

Traditionally, the task of short-term forecasting of parameters in the EPS was carried out mainly using statistical methods based, for example, on the autoregression of the integrated moving average (ARMA), the Kalman filter [2]. An analysis of modern scientific works indicates that for the problem of short-term forecasting of EPS parameters, ML methods, primarily artificial neural networks (ANN) [5-7], support vector machine (SVM) [5], decision trees (DT) are widely used [6].

Until recently, ANN-based models remained the most popular among machine approaches in predictive problems. The success of ANN application is due, first of all, to the fact that the neural network structure makes it possible to obtain models with good approximating abilities. Various ways of combining neurons with each other and organizing their interaction led to the creation of different types of ANNs, among which the most widely used for short-term forecasting of EES parameters was found: a multilayer perceptron (MP), an ANN based on radial basis functions, an Elman network and a generalized regression network.

Along with ANNs, predictive models based on SW demonstrate high efficiency [7]. Methodologically, SVM, like ANN, is based on the well-known Cover theorem [8] on increasing the probability of linear separability of images when transforming a nonlinear problem of classifying images into a space of higher dimension. In [9], a comparative analysis of the effectiveness of various models in predicting electricity consumption in one of the energy Hong

Kong. The measurements were carried out daily throughout the year. After processing the data, three models for predicting energy consumption were built: linear regression, ANN, DT. As can be seen from Table 1, the best results are given by the DT and ANN models with a deviation from the true value of 5-6%.

Table 1: *Electricity consumption forecast results in one of the power districts of Hong Kong based on various models of ML*

Forecasting model	Average relative error, %
Linear regression	7.6-8.2
Artificial neural network	5.5-6.7
Decision Trees	4.8-5.6

DT|-based models are a promising technology for predicting complex non-stationary implementations that cannot always be processed based on neural network models. The structure of DT is "leaves" and "branches". The edges ("branches") of the decision tree contain the attributes on which the objective function depends, the "leaves" contain the values of the objective function, and the remaining nodes contain the attributes by which the cases are distinguished. The DT model makes it possible to obtain stable solutions comparable to SVM and ANN without using the large computing power required by previous models.

At the same time, the most popular modification of the AR algorithm used for forecasting problems is the "random forest" (from the English RandomForest) models, which allow you to build many trees on different subsets of the training sample and, due to the law of large numbers, get the best results by choosing the average of all prediction trees. In [10], a large-scale study of DT committee (ensemble) models is presented - random forest and gradient boosting for intra-day wind speed forecasts for wind power stations (WPPs) of Sotavento (Spain) and wind generation throughout Spain. The performance of DT-based models was compared with MOW and MT as the most popular MT models for wind generation forecasting. However, the results obtained did not reveal a clear "winner" (Table 2).

At the forecast level for a single WPP, the models based on the DR committee outperform the MOU model. At the same time, forecasts for the entire EPS of Spain showed that the results of the DT are close to the accuracy indicators of the SRM. And finally, the lowest forecast accuracy for these cases was recorded when using MP models.

Table 2: *Forecast errors, MAE for WPP Sotavento and EPS of Spain*

Models	WPP "Sotavento"		Wind power for Spain's EPS	
	Training	Test	Training	Test
Support Vector Machine	5.62	7.80	1.01	3.13
Random forest	2.68	7.68	1.29	3.68
Gradient boosting	3.69	7.84	1.06	3.41
Extreme gradient boosting	2.43	7.72	0.90	3.22
Multilayer perceptron	5.61	7.78	1.22	3.69

The development and application of ML methods made it possible to significantly speed up and automate the solution of a whole range of tasks of regime and emergency control of EPS. At the same time, most of the solutions in this direction are associated with the transformation of the classical optimization problem into a regression/classification problem, which can significantly reduce the calculation time while maintaining acceptable accuracy. The most successful developments have been obtained in the field of application of various DT algorithms for monitoring and controlling the operational reliability of EPS: online DT models for voltage/reactive

power regulation [11], random forest models for monitoring the dynamic operational reliability of EPS (separate developments of the Hydro-Québec energy companies), Canada [12] and Energinet.dk, Denmark [13]).

Significant progress in the control of EPS modes was obtained on the basis of a group of ML methods with reinforcement, such as Monte Carlo methods, dynamic programming, learning based on time differences (SARSA, Q-learning), etc. [14]. These methods involve learning what to do, how to map situations into actions in order to maximize some reward (reward) signal that takes numerical values. The model being trained (the agent) is not told what action to take, as is the case in most ML approaches. Instead, they must try different actions and find out for themselves which ones will bring him the greatest reward.

As a result, this approach allows the agent to choose an EPS control strategy not randomly, but to take into account the experience of previous interaction with the system based on the assessment of the utility function Q. Agents trained offline based on the Q-algorithm successfully control individual components of the power system within the ODE and / or PAHs such as dynamic brake, thyristor-controlled series capacitor, synchronous generators, individual or aggregated loads, etc. for optimal control strategies. Applications of the reinforcement learning method have shown good results in a whole range of ODE and PAC tasks: wind turbine generation control, load control, power system restoration, micro-EPS modes control, voltage regulation, cascade accident prevention, power consumption forecasting, etc. [15,16].

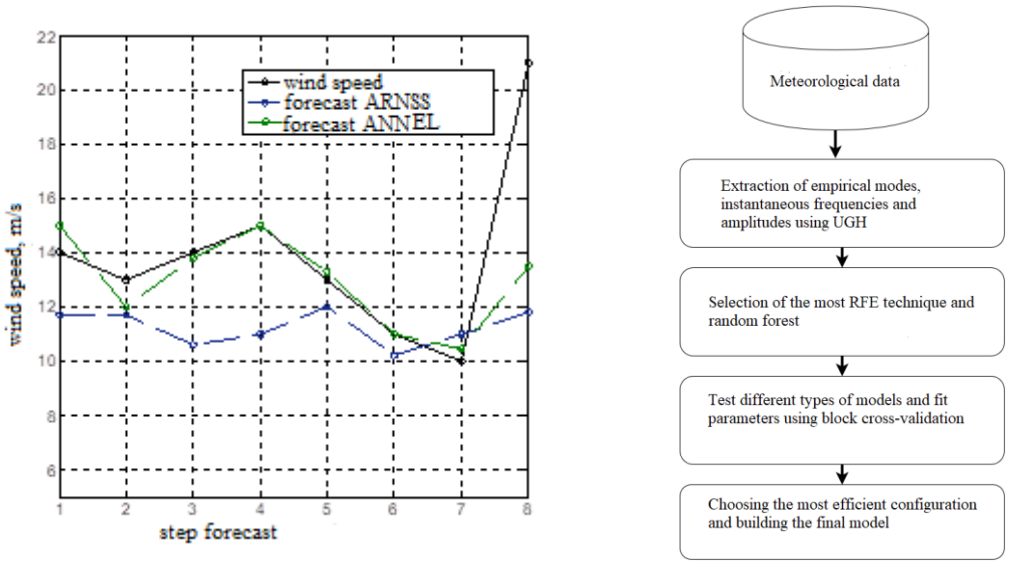
Despite the undoubted advantages of the noted ML models, the question of the high efficiency of these methods is still open and requires further research. One of the most promising areas is the use of hybrid approaches. In [17], ANN and SW models are considered in combination with the Hilbert-Huang transform (HCT). In [18], the Box-Jenkins methodology was supplemented with ANN. Neural network models together with fuzzy logic methods were proposed in [19]. Models combining ANN and burst theory are presented in [20]. The success of hybrid approaches is explained by dividing the task of constructing a predictive model into two basic stages: the stage of data preprocessing in order to identify features that are most significant for the prediction, and the stage of identifying a dynamic predictive model itself.

III. Short-term forecasting of power system parameters using ml models

The scientific groups of ISEM SB RAS (Irkutsk, Russia), "Azerenergy" and AzNIiPIIE (Baku, Azerbaijan) conducted a study of autoregression models and neural networks in predicting the power generation of offshore wind turbines in various sectors of the Caspian Sea [21]. To build and train models, we used data from measurements of wind speeds and a number of other additional parameters (Fig. 1a). To predict the power of wind turbines in intraday cycles "for 1 hour ahead"', the prehistory of data for one week was used. (averaging 5 minutes). According to the test results, the Artificial Neural Network for Extreme Learning (ANN EL) gave a significantly higher prediction accuracy (MAE = 2.538, RMSE = 2.686) than the autoregressive ARNSS model (MAE = 6.649, RMSE = 7.178).

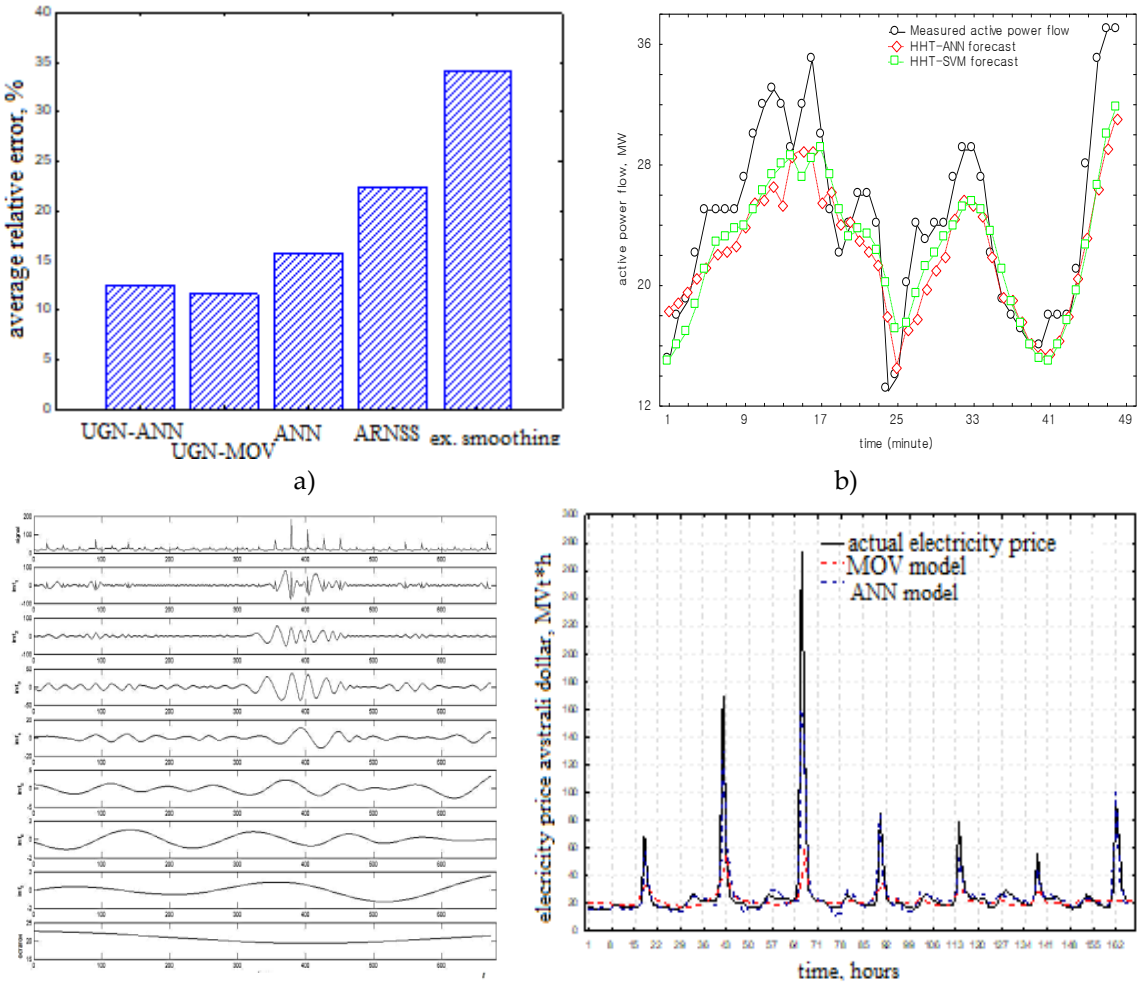
The staff of ISEM SB RAS, together with the research team of the Irish National University in Cork, developed a hybrid approach (Fig. 1b) to short-term forecasting of EPS parameters (load, generation, power flows, electricity prices) based on an effective apparatus for analyzing non-stationary time series UGH and ML algorithms [17,22].

UGH can be divided into two parts: empirical mode decomposition (SEM) and the application of the Hilbert transform to them. This makes it possible to obtain a set of frequencies and amplitudes localized in time. There are various SEM-based hybrid approaches that include modifications of ML algorithms for dependency recovery. Various types of ANN (MP, ANN EL), MOV and others are used as such algorithms.



а) б)

Figure 1: *Results of wind speed forecast based on the ARPSS and ANN EL models (a) and the general diagram of the hybrid approach for creating forecast models (b)*



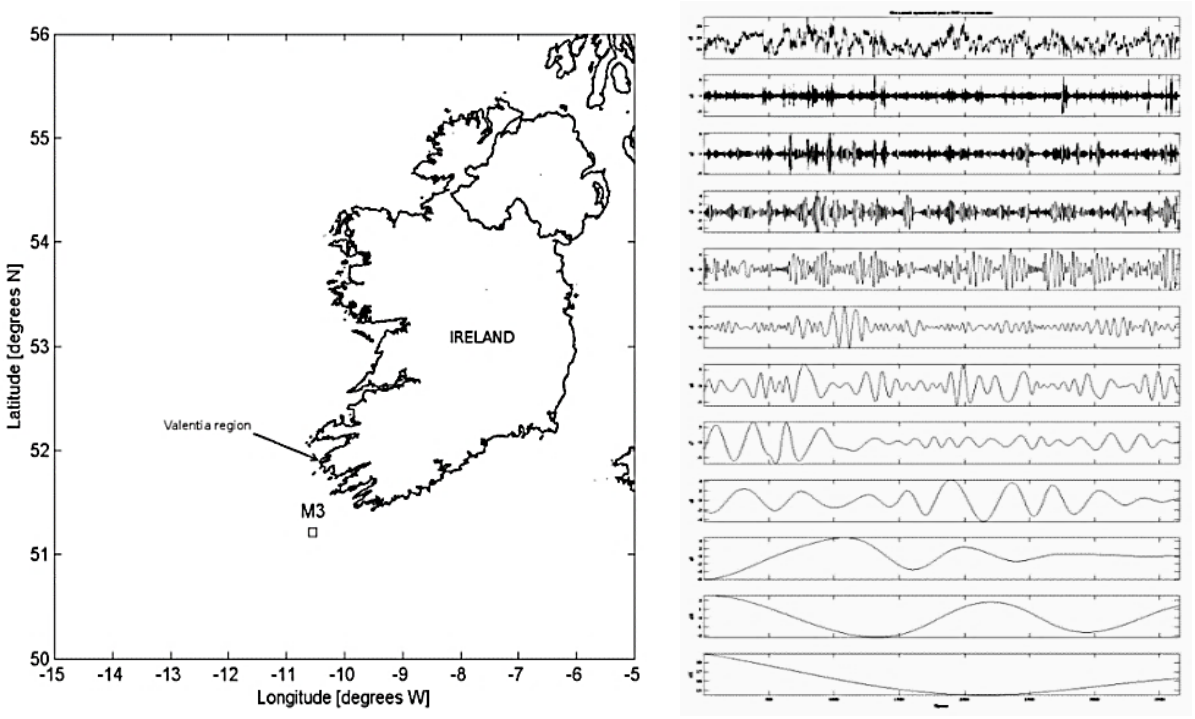
a) b)

c) d)

Figure 2: *Application of the hybrid approach: a, b - the results of forecasting the flow of active power "for 1 hour ahead"; c - the results of the SEM of electricity prices for the South Wales area; d - forecast schedule for electricity prices in the Australian market "1 hour ahead"*

The effectiveness of this hybrid approach has been demonstrated on real data in various test tasks: in ultra-short-term forecasting of active power flows in the section of the traction substation "Gidrostroitel-Korshunikha", Russia, Irkutsk region (Fig. 2, a, b) [23] and in forecasting electricity prices “1 hour ahead” according to the Australian National Energy Exchange (Fig. 2, c, d) [17].

Additionally, the hybrid approach was tested on real data when predicting wind speed “for 24 hours ahead” for the tasks of controlling the modes of wind power plants in the region of Valentia, Ireland (Fig. 3) [22]. In addition to retrospective data on wind speed and direction, such characteristics as atmospheric pressure, wave height and period, maximum gust speed, relative humidity, and water temperature were used. All parameters were measured hourly from 12/15/2020 to 05/15/2022.



a) b)

Figure 3: *Location of the M3 weather station in the Valentia region (a) and a fragment of the decomposition of the implementation of the wind speed into empirical modes (b)*

Models were built to make forecasts 4, 6 and 24 hours ahead. The results of model testing were compared for various methods, such as random forest, SVM, DT based on gradient boosting, ANN EL. As can be seen from fig. 6, the use of the hybrid approach makes it possible to improve the accuracy of the wind generation forecast compared to the use of simple neural network forecasting.

Studies have shown that for short-term and ultra-short-term forecasts of wind generation, the most suitable of the considered methods are ANN EL and gradient boosting tree. Forecasting for a longer period requires the use of more complex and robust methods such as gradient boosting tree and random forest.

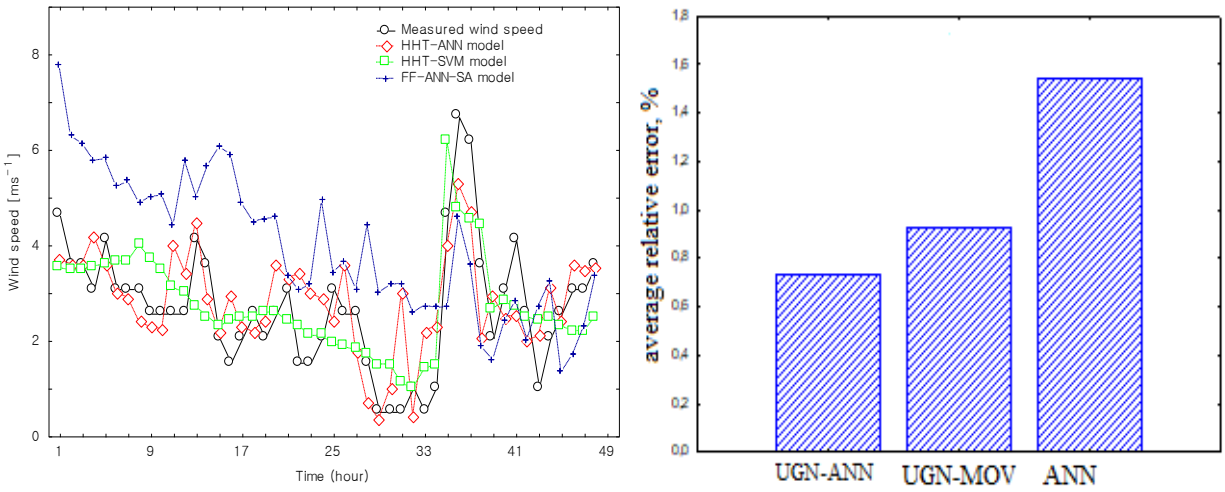


Figure 4: *Forecasting wind speed "for 24 hours ahead"(a), Valentia, Ireland for hybrid models UGH-ANN, UGH-MOV and MP (b)*

Table 3: *Comparison of errors of various ML models*

Model	Period, hour	RMSE	MAE
Naive forecast	4	1.834	1.392
ANN EL		1.359	1.076
gradient boosting		1.416	1.112
MoU		1.848	1.463
random forest		1.429	1.127
Naive forecast	6	2.208	1.708
ANN EL		1.669	1.322
gradient boosting		1.676	1.312
MoU		2.597	2.206
random forest		1.677	1.344
Naive forecast	24	2.757	2.081
ANN EL		2.259	1.844
gradient boosting		2.157	1.708
MoU		3.480	2.620
random forest		2.235	1.835

Detailed information about the quality of models obtained by different ML algorithms for different prediction periods is presented in Table 3. As a result, for ultra-short-term (up to 4 hours ahead) and short-term forecasts for wind energy (up to 6 hours ahead), the most suitable of the considered methods are ANNs of extreme learning and gradient boosting tree. Forecasting for a longer period (up to 24 hours ahead) requires the use of more complex and robust methods such as gradient boosting tree and random forest.

IV. Voltage and reactive power regulation in Eps with integrated res using ml models

On fig. Figure 5 shows the schedule of daily power generation on one of the typical days during the month [24-26]. As can be seen from this figure, the power generation by wind farms differs significantly by day - the difference is for calm and windy days (0 - 5) % and (70 - 90) %, respectively. Differences can also be significant in power generation in the morning and afternoon. The variability of production can be so significant that the difference from the average monthly

day or the average annual day is not consistent with stochasticity. The output varies throughout the day from zero to maximum. In this case, the modeling of the development characteristic in the form of a normal distribution seems to be unsuccessful, and therefore the consideration of the development process as stochastic is incorrect. In similar situations, controlling voltage and reactive power and ensuring reliability, as well as voltage stability of the EPS by traditional means, turns out to be difficult and inefficient. Therefore, the problem of applying ML to solve this problem is considered.

In [27-29], a method for estimating the EPS stability limit by the variability of the voltage profile described by the ANN (Fig. 6), developed by the scientific groups of "Azerenergy" and AzSRDSIE (Baku, Azerbaijan), is given.

In this case, the place and volume of the additionally included reactive power injection is determined by analyzing the sensitivity of the voltage stability limit value relative to the voltage variability in each network node:

$$\frac{\partial \Pi_U}{\partial U} = \frac{\partial \psi}{\partial E}\left(E^{\circ}\right)\cdot\left(\sum_{i=1}^{H} W_2(i)\frac{\partial \varphi_i}{\partial r_i}\left(r_i^{\circ}\right)\cdot\sum_{j=1}^{h}\left(W_1(i,j)\cdot T(j,u)\right)\right) \quad (1)$$

Where $H-$ number of hidden neurons in the ANN model;$h-$ number of nodes in the electrical network diagram; $W_1(i,j)-$ weight coefficient of connection of the j-th neuron of the input layer with the i-th neuron of the hidden layer; $W_2(i)$ – weight coefficient of the connection of the output neuron with the i-th neuron of the hidden layer; r_i, φ_i – input and output of i-th hidden neuron, respectively; E, ψ – input and output of the output neuron, respectively; r_i^0- initial output value of the i-neuron of the hidden layer; E^0 – the initial value of the output of the output neuron; $u-$ number of unsupervised nodes (bus PQ);$T(j,u)$ – transformation matrix element.

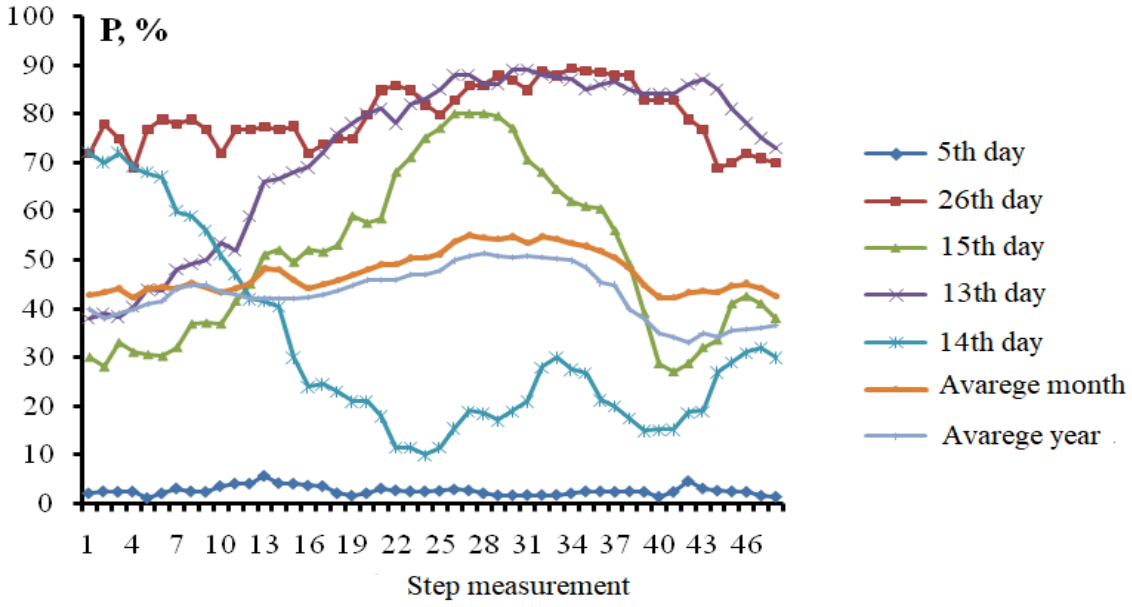


Figure 5: *Uncertainty of daily power generation wind farms on one of the typical days of the month*

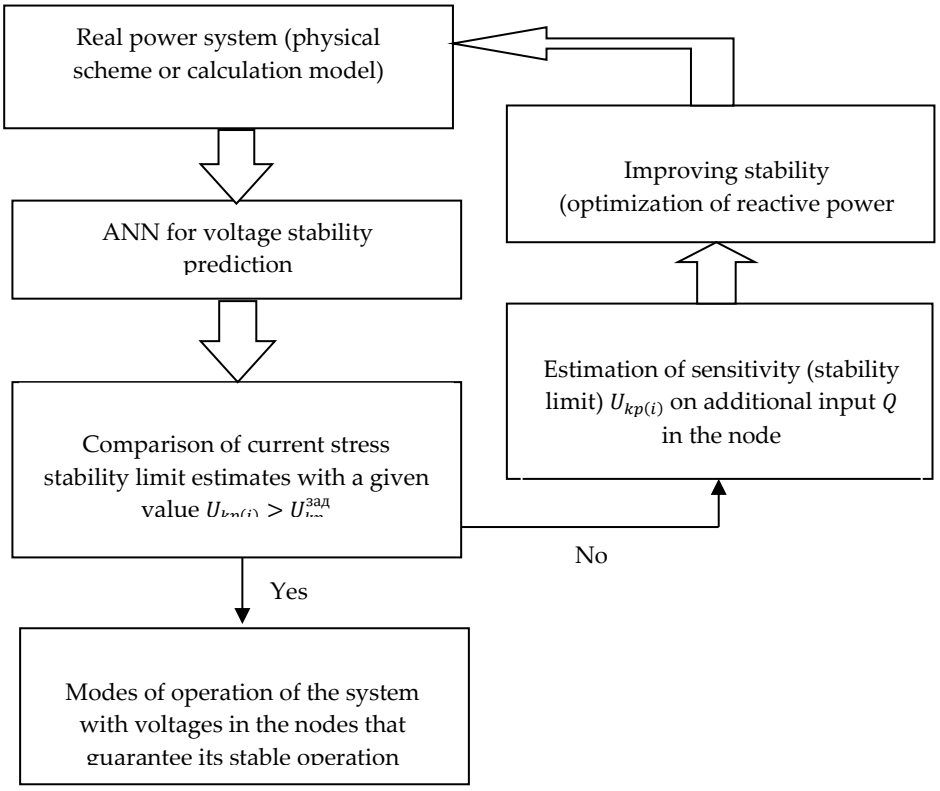


Figure 6: *Schematic diagram of monitoring the current value of the voltage stability limit using the ANN model*

Using the reduced Jacobi matrix, one can determine the sensitivity of the voltage stability limit with respect to the reactive power injection on the busbars:

$$\Delta \Pi_{U_k} = \sum_{i=1}^{N_u} \frac{\partial \Pi}{\partial U_i} \Delta U_i = \sum_{i=1}^{N_u} \frac{\partial \Pi}{\partial U_i} j_R^*(i,k) Q_{1nj,k} \quad (2)$$

$$S_k^{\Pi_u} = \frac{\Delta \Pi_k}{Q_{1nj,k}} = \sum_{i=1}^{N_u} \frac{\partial \Pi}{\partial U_i} j_R^*(i,k) \quad (3)$$

where, $N_u -$ total number of non-steered or PQ tires;

$Q_{1nj,k} -$ injection of reactive power introduced into node "k";

$j_R^*(i,k) -$ element of the reduced Jacobi matrix.

In order to increase Π_U to the desired value $\Pi_U^{\lim}$, it is necessary to introduce reactive power Q_{1nj} in the node with high sensitivity, established by the estimate from equation (3).

It should be noted that the process of improving the Π_U value by introducing additional injection reactive power must be performed for each node. In other words, equation (2) determines the final change in the value of Π_U, which is achieved by summing the reactive power injections at different nodes. At each mode point, the desired value of $\Pi_U^{\lim}$ is determined as a fraction of the current load value in p.u. or in %:

$$\Pi_U^{\lim} = \beta P_0 \quad (4)$$

where, $P_0 -$ current mode load power;

$\Pi_U^{\lim}$ – voltage stability limit in the current mode;

$\beta -$ reserv factor

In order to improve voltage stability, it is necessary to ensure effective control of the reactive power source in the nodes identified as sensitive from the analysis of the regimes of the ML stability model. The studies were carried out on a test 14-node IEEE scheme, as well as on a real scheme of the energy system of Azerbaijan.

For a 14-node scheme, training samples were obtained by calculating the flow distribution at different loads, which varied in steps of 5% from P_0 to the limiting $P_{\lim}$ in terms of load capacity. As for the samples corresponding to certain load increments, each of them is characterized by a profile and limiting stress values in terms of stability.

Table 4 shows the number of training samples, validation, and experimental use to predict the stress stability limit.

Table 4: *Parameters of the ML model for estimating the stress stability limit*

Number of training samples	Number of Validation Samples	Number of testing samples	Number of hidden neurons	Studying time sec.
1600	400	2000	15	29, 17

As can be seen from Table 5, by reactive power compensation in the network (nodes 3,6,9), it is possible to increase the system load from 290.04 MW to 613.83 MW. In this case, the total power of the compensating devices will be 248.91 MVAr.

The scientific group of ISEM SB RAS has developed a hybrid system for voltage/reactive power control in normal modes and prevention of voltage avalanche in EPS using online algorithms of ML and multi-agent system (MAS) [11]. The main idea of the development is to train and tune intelligent algorithms to recognize in a timely manner the characteristic indicators of EPS stability (coefficients of the Jacobian matrix of the steady state, L-index) in order to implement joint measures to prevent serious accidents.

Table 5: *Values of the maximum load of the system when the limit voltage is reached under the conditions of reactive power compensation*

Load power in the initial normal mode, MW	Before compensation MW		After compensation MW		Compensation nodes Π_U most sensitive to change	Optional reactive power injection, MVAr
	According to the INS model	traditional calculations oft ETAP method	According to the INS model	traditional calculations oft ETAP method		
310,8	290,0	277,2	613,8	491	3 6 9	92,04 51,37 105,5

Together with the MAS, which implements decentralized voltage regulation by changing the settings of AVR generators, to regulate reactive power sources in load nodes, a model was proposed based on online DT - an online random forest algorithm, which was trained to determine additional reactive power injections for various EPS operating modes. The basis for training the model is the results of solving a system of

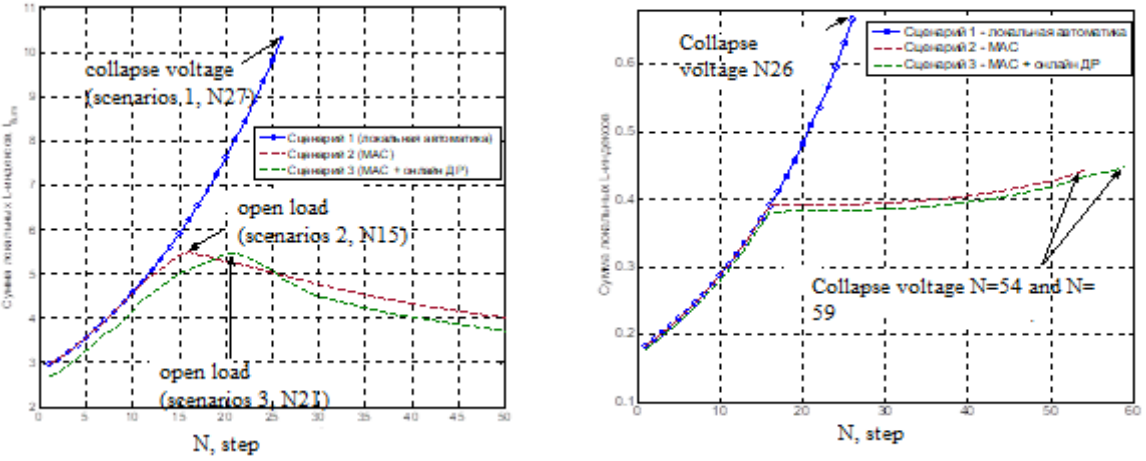
equations with respect to partial derivatives of the function of the sum of local L-indices, by reactive power injections

$$\frac{\partial L_{sum}}{\partial \Delta Q} = \begin{bmatrix} \frac{\partial L_{sum}}{\partial \Delta Q_1} \\ \cdots\cdots\cdots \\ \frac{\partial L_{sum}}{\partial \Delta Q_m} \end{bmatrix} = 0 \quad (5)$$

where $L_{sum} = L_1 + L_2 + \cdots + L_m$ – sum of local indices, $L = \max_{j \in \alpha_L}(L_j) = \left(\frac{S_j^+}{Y_{jj}^+ \cdot U_j^2} \right)$ – global L-index proposed in [25] as an indicator of impending voltage collapse.

The proposed online random forest algorithm, PDSRF, has the ability to independently and adaptively rebuild in real time in the event of serious changes in incoming information without reducing the accuracy of identifying EPS modes. The developed intelligent system was tested both on standard IEEE circuits (IEEE 6, IEEE 118) and on a real circuit of the Bodaibo power district of the Irkutsk power system, which has problems with voltage stability.

Figure 7 clearly shows that the additional use of intelligent models based on MAC and ML, along with traditional voltage regulation by local automatics, makes it possible to maintain system stability in heavy modes and prevent voltage collapse. As a result, such a hybrid control system based on intelligent tools provides an automatic solution to the problems of secondary regulation and emergency control, thereby eliminating the human factor and ensuring the continuity of the processes of operational and automatic emergency control.



а) б)

Figure 7: *Changes L_{sum} under different control scenarios: a) IEEE 118 test pattern; b) real scheme of the Bodaibo energy district, Irkutsk region, Russia*

Conclusion

The ever-increasing saturation of modern PS with new stochastic components (high-power wind turbines, solar panels, demand management systems, etc.), the introduction of new technologies (distributed generation, flexible power transmission equipment, energy storage systems, micro-EPS, hybrid AC/DC networks, digital substations, etc.), as well as the formation of market principles of regulation, significantly complicate the tasks of forecasting and controlling the modes of modern EPS. An effective solution to this problem is the use of algorithms and ML models that are able to learn to predict and control the operating modes of the EPS, taking into account many changing factors. The results of the application of the ML technology in the tasks of operational dispatch and emergency control of EPS clearly show the prospects of such studies for subsequent practical implementation in the work of various automated control systems for electrical networks.

References

[1] Negnevitsky, M. Artificial intelligence: a guide to intelligent systems. 3rdedn, Addison Wesley, Harlow, England, 201 I. 504 p.

[2] Gamm A.3., Glazunova A.M., Grishin Yu.A., Kurbatsky V.G., Sidorov D.N., Spiryaev V.A., Tomin N.V. Methods of forecasting the parameters of the regime of electric power systems for the purposes of monitoring and control, Electricity, 2011, No. 5, pp. 17-26.

[3] Hippert H.S. Neural Networks for Short-Term Load Forecasting: A Review and Evaluation / H. SteinherzHippert, C. Eduardo Pedreira, R. Castro Souza, IEEE Trans. OnPowerSystems, 2001, Vol. 16, No. 1.

[4] Shumilova G. P., Gotman N. E., Startseva T. B., Forecasting of electrical loads during operational management of electric power systems based on neural network structures. Ekaterinburg: УРОРАН, 2008, p.89.

[5] Mohandes M.A., Halawani T.O., Rehman S. and Hussain A. Support Vector Machines for wind speed prediction, Renewable Energy, Vol. 29, No. 6, May 2004, pp. 939-947.

[6] Yiqian Liu Machine Learning for Wind Power Prediction, 2016

[7] Yusifbayli N.A., Nasibov V., Guliyev H.B., Aghaliyev N. Assessment of penetration level of variable renewable generations on the flexibility of power system, Electroenergetics, Electrotechnics, Electromechanics + Control (EEEC), Scientific – industrial journal, Vol.11, No.1, 2021, pp.31-44.

[8] Mammadov C.F., Aliyev I.R. et.al. Algorithmic support for the management of the computer-aided design of flexible manufacture system and its equipment, Cybernetyc and system analyze. Springer, 2021. Vol. 57, N 6. P. 118–127.

[9] Geoffrey K..F. Tso Kelvin K.W. Yau. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks, Energy, 2007, Vol. 32, No. 9, pp. 1761-1768.

[10] Torres-Barrana A., Alonso A., Dorronsoro J.R. Regression tree ensembles for wind energy and solar radiation prediction, Neurocomputing, 2015, https://doi.org/10.1016/j.neucom.2017.05.104

[11] Voropai, N.I., Tomin, N.V., Sidorov, D.N. et al. A Suite of Intelligent Tools for Early Detection and Prevention of Blackouts in Power Interconnections, Autom Remote Control (2018) 79, pp. 1741-1755.

[12] Kaci, A et al.Synchrophasor Data Baselining and Mining for Online Monitoring of Dynamic Security Limits, IEEE Trans, on Power Systems, 2014. Vol. 29, No. 6, pp. 2681-2695

[13] Liu, C et al. Dynamic security assessment of western Danish power system based on ensemble decision trees, in Proc. of the 12th 1ET International Conference on Developments in Power System Protection. Copenhagen, 2014

[14] Саттон Р.С., Барто Э.Г. Обучение с подкреплением, пер. с англ. 2-е изд. (эл.). М. : БИНОМ. 2014, 402 с.

[15] Glavic, M.,Fonteneau R., Ernst D. Reinforcement Learning for Electric Power System Decision and Control: Past Considerations and Perspectives, IFAC-PapersOnLine, 2017, Vol. 50, No. 1, pp. 6918-6927

[16] Mocanu E., Nguyen P.H.,Gibescu M. Deep Learning for Power System Data Analysis, In Big Data Application in Power Systems, ed. by Reza Arghandeh and Yuxun Zhou, Elsevier, 2018, pp. 125-158

[17] Kurbatsky V. G., Sidorov D. N., Spiryaev V.A., Tomin N.V. Forecasting nonstationary time series based on Hilbert-Huang transform and machine learning, Autom. Remote Control, 75:5, 2014, pp. 922-934

[18] Areekul Ph. A Hybrid ARIMA and Neural Network Model for Short-Term Price Forecasting in Deregulated Market / Ph. Areekul, T. Senjyu, H. Toyama, A. Yona, IEEE Trans. Power Syst., 2010, Vol. 25, No. 1, pp. 524-530.

[19] Khotanzad A.Z. A neuro-fuzzy approach to short-term load forecasting in a price sensitive environment / A. Zhou Khotanzad, H. Elragal, IEEE Trans. Power Syst., 2002, Vol. 17, No. 4, pp. 1273-1282.

[20] Sinha N. Wavelet-GA-ANN Based Hybrid Model for Accurate Prediction of Short-Term Load Forecast / N. Sinha, L.L. Lai, P. Kumar Ghosh, Y. Ma, Proc. the IEEE Inter. Conf. onISAP., 2007, TokiMesse, Niigata, pp. 1-8.

[21] Rakhmanov H.P., Kurbatsky V.G., Guliyev H.B., Tomin N.V. Short-term forecasting of power generation of wind turbines to ensure the reliability of electrical networks, Proceedings of the 87th International Scientific Seminar "Methodological Issues of Researching the Reliability of Large Energy Systems", Minsk, Belarus, 7-1 September 1, 2015

[22] N. Tomin, D. Sidorov, V. Kurbatsky, V. Spiryaev, A. Zhukov, P. Leahy. A Hybrid Wind Speed Forecasting Strategy based on Hilbert-Huang Transform and Machine Learning Algorithms, in Proc. of 2014 International Conference on Power System Technology (POWERCON 2014), Chengdu, China, 20-22 Oct. 2014

[23] Kurbatsky V.G. On a neural network approach to forecasting non-stationary time series based on the Hilbert-Huang transform / V. G. Kurbatsky, D. N. Sidorov, V. A. Spiryaev, N. V. Tomin, Avtomat, i Telemekh., 2011, No. 7, pp. 58-68

[24] Mammadov C.F., Huseynov A.H., Hasanova Y.M., M.R. Salimova M.R. Algorithmic and mathematical support for topology research of the corporative networks, Advanced Mathematical Models & Applications Vol.7, No.1, 2022, pp. 85-92.

[25] Rakhmanov, N.R. Rakhmanov N.R., Kurbatsky V.G., Guliyev H.B., Tomin N.V. Application of algorithms and models of machine learning in the problems of modern energy systems control. News of Azerbaijan Higher Technical Schools, Baku: 2019, ASSU, Vol. 21, No. 2, pp. 61-72.

[26] Stennikov V.A. Kurbatskiy V.G., Rakhmanov N.R., Guliyev H.B. и др. Regional aspects of wind energy. Russian Academy of Sciences, Siberian Department of Energy Systems Institute Melentieva, Novosibirsk, 2020, p.296.

[27]Rakhmanov N.R, Guliev G.B. Application of the neural network model to estimate the current value of the limit of resistance of the electrical system by voltage, Elektrichestvo, 2015, No. 4, pp. 43-52

[28]Guliyev H.B., Ilyasov O.V. Justification of the efficiency of the phase rotary transformer in repair and post-emergency modes (on the example of Azereneji), Rudenko international conference “Methodological problems in reliability study of large energy systems” (RSES 2021), Conference date: 13–17 September 2021, Location: Volzhsky, Russia, ISBN: 978-0-7354-4307-5, Volume number: 2552, Published: Jan 5, 2023

[29] Guliyev H.B. Method for determining the critical parameters of the sustainability of an energy system with integrated renewable energy sources, Rudenko International Conference “Methodological Problems in Reliability Study of Large Energy Systems” (RSES 2022), E3S Web of Conf., Volume 384, 2023, pp.1-5.

OPTIMIZATION OF PRIORITY SERVICE WITH EFFICIENT COORDINATION OF ADMISSION CONTROL, EMERGENCY VACATION OF AN UNRELIABLE SERVER

[1]G. Ayyappan, [2]S. Nithya

•

Department of Mathematics
Puducherry Technological University
Puducherry-605014, India
ayyappan@ptuniv.edu.in, nithyamouttouraman@gmail.com

Abstract

In this paper, we establish a single server retrial queueing system with two types of customers, admission control, balking, emergency vacation, differentiate breakdown, and restoration. There are two distinct factors which must be considered when classifying priority and ordinary customers. The non-preemptive priority discipline proposed by this concept. Ordinary and priority customers arrive in accordance with Poisson processes. For both priority and ordinary customers, the server continuously offers a single service that is distributed arbitrarily. In this study, we compute the Laplace transforms of the time-dependent probabilities of system states using a probability generating function and the supplementary variable technique. The sensitivity analysis of system descriptions is assisted by study of numerical findings.

Keywords: Batch Arrivals; Priority Queues; Admission Control; Emergency Vacation; Balking; Differentiate Breakdown; Restoration.
AMS Subject Classification (2010): 60K25, 68M30, 90B22.

1. Introduction

When a customer arrives and discovers the server is busy, they are asked to leave the service area and join a trial queue known as orbit, which separates queues with repeated efforts. Once a certain period has gone, the orbiting the customer may resubmit their service request. Any customer in the orbit can repeatedly request services without affecting the other customers in the orbit. In computer and communication systems, such queues have a unique function. Numerous researchers explored two various customer types in retrial queues Wang ([26]; Dimitriou [12]; Wu and Lian [27]; Rajadurai et al.[23]).

Several aspects of daily life involve priority queues, especially when specific groups of individuals are given special attention, e.g. telecommunications field. Priority systems are a inestimable scheduling tool that allows messages of several types to receive a wide range of service. Due to this, the priority queue has earned a lot of attention in the literature Ayyappan and Thilagavathy [5], Kim et al. [19], Bhagat and Madhu Jain [8], Rismawati et al. [24], Pandey and Pal [22].

The service channel will temporarily fail if the regularly busy server breaks down, which could happen at any time. In other words, the server is temporarily unavailable. Fiems et al. [13] focused at the single-server queueing system with two different kinds of server disruption. Jain and Agarwal [16] considered an unreliable server $M^{[X]}/M/1$ queueing system with multiple

types of server breakdowns. Zhang and Zhu [29] investigated retrial queueing model with vacations and two types of breakdowns. Jayakumar and Senthilnathan [17] described the server breakdown without interruption in a batch arrival queueing system with multiple vacations and closedown. Yi-chih Hsieh and Andersland [28] explored steady-state queue length distribution and mean queue length of Markov queueing systems subject to random breakdowns. Ayyappan and Sankeetha [7] discussed single server that provides both regular and optional service with vacation, breakdown and repair.

In 1957, Haight conducted the first study on the phenomenon of balking, in which customers decide not to wait in queue if the server is not present. Customers may become frustrated in several types of scenarios, including call centres, computer systems, websites, and telephone switchboards. Artalejo and Lopez-Herrero [3] introduced an M/G/1 retrial queue with balking. A $M^{[X]}/G/1$ queue with variable vacations and balking was researched by Ke [18]. An $M/G/1$ retrial queue with non-persistent customers was mentioned by Gao and Wang [14] where the server was susceptible to failure because of the negative arrivals.

The following admission control policy was examined in this study. An arriving batch of low priority customers may be allowed to join the orbit with probability a or may not be allowed with probability $(1-a)$. For instance, the company may not be able to select one candidate from all the applicants during an interview. Some selection criteria could be used, such as a screening process, a group discussion, etc. A single server batch arrival queueing system with two service phases, an admission control system, and Bernoulli vacation was examined by Choudhury [9]. Artalejo et al. [4] generalised this queue to discrete-time. A $M/E_k/1$ queuing system's control approach was created by Madhu and Indhu [20]. A single server batch arrival retrial queueing system with two service phases and a Bernoulli admission algorithm was obtained by Choudhury and Deka [10].

Recent research regarding various vacation policies adopted by service providers has produced a considerable impact on queueing systems. This is a result of its widespread application in many kinds of real-life situations, particularly in flexible production systems, communication systems, and computer systems. Shekar et al. [25] introduced a single server queueing system's emergency vacation. This vacation policy states that the working server may take a vacation in an emergency without finishing the service to customers who are waiting in service interruption. A priority retrial queueing system that involves working vacations and vacation interruptions, emergency vacations, negative arrivals, and delayed repairs was studied by Ayyappan and Thamizhselvi [6]. With two different vacation possibilities, Anna Bagyam and Udhaya Chandrika [2] investigated a single server retrial queueing system. [1] has highlighted the transient behaviour of a multiple vacations queue with frustrated clients. [11] investigated a $M^{[X]}/G/1$ queueing system with a single vacation policy. Ayyappan and Meena developed the phase type queueing model with degrading service, breakdown and Vacation.

In this study, we investigate a single server retrial priority queueing system with admission control, balking, emergency vacation, differentiate breakdown and restoration. Incoming ordinary customers have the option of entering the orbit or exiting the system if the server is down. If the server is busy suddenly they go for vacation and the interrupted customer wait in the queue and get fresh service after return from vacation. The regular busy period server may breakdown at any instance. Hard and soft failure are the two kind of system failures. Hard failure is defined as an equipment failure that requires a repairman with specialized knowledge to be physically present, which is a time-consuming process. While soft failure is defined as failure brought on by circumstances rather than a physical problem and usually resolved by restarting the system. After breakdown that the system will take some time for its refunctioning. This recovering period is called the restoration.

The rest of the paper is organized as follows: Mathematical model is described in Section 2 and queue size distribution is analyzed in Section 3. An explicit expression for governing equation is

enlisted in Section 4. Steady state analysis is discussed in Section 5. Stability condition discussed in Section 6. Particular cases are obtained in Section 7. The effect of system performance measures is illustrated in Section 8. Numerical and graphical results are derived and conclusion is obtained in Section 9 and 10.

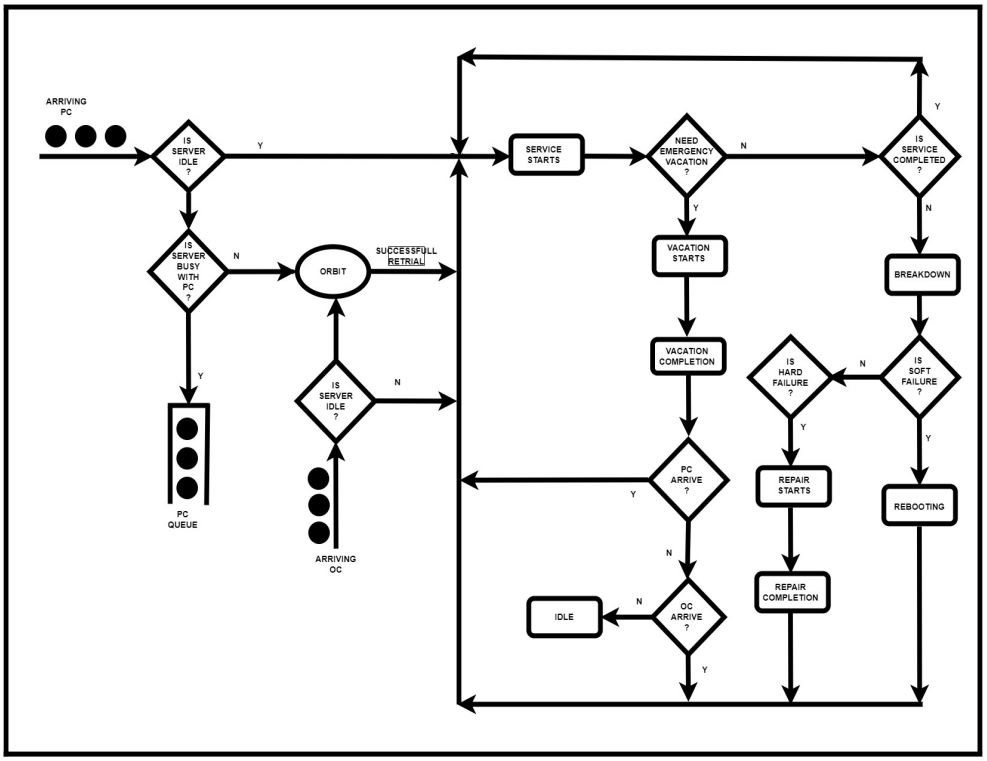


Figure 1: *Schematic representation*

2. Description of the Model

- **Arrival Process :** Two different types of units arrive in batches with independent Poisson compound process. Let $\lambda_p, \lambda_o > 0$ represent the corresponding arrival rates for priority and ordinary customers. Assume that the first order probabilities for priority and ordinary customers $\lambda_p c_i dt$ $(i = 1, 2, 3, ...)$ and $\lambda_o c_j dt$ $(j = 1, 2, 3, ...)$ respectively. The system has i and j batch size customers enters within a short period of time $(t, t + dt)$. Here, $0 \leq c_i \leq 1$, $\sum_{i=1}^{\infty} c_i = 1$, $0 \leq c_j \leq 1$, $\sum_{j=1}^{\infty} c_j = 1$.
- **Retrial Service Process :** Ordinary customers are known as retrial customers. These customers will go back to the orbit and will request repeatedly for their service after some time if the server is busy or unavailable. Retrial customers service time has rate $\beta(\nu)$ that follows the general distribution.
- **Regular Service Process :** Ordinary and priority customers ordinate in batches with distinct queues. Service rate follows general distribution and server renders single service for priority customers and ordinary customers with service rate $\mu_i(\nu)$, $i = 1, 2$ respectively. When the priority queue is empty, the service for ordinary customers begins.
- **Admission Control:** The server will follow the admission control policy for ordinary customers if it is overloaded in priority or ordinary customers. The server may grant ordinary customers admission with probability a or restricted entry with probability $(1 - a)$.
- **Differentiate Breakdown and restoration :** The rates of hard and soft failure are exponentially distributed with rate α_1 and α_2 respectively. For soft failure, the restoration time follows exponential distribution with rate η_1 and for hard failure restoration time follows

general distribution with rate $\eta_2(\nu)$.

- **Balking:** Incoming ordinary customers have the option of entering the orbit with probability b or exiting the system with probability $1-b$, if the server is down.
- **Emergency Vacation:** During the service term, the server has the opportunity to take an unexpected vacation at the exponentially distributed rate θ. The emergency vacation time for interruptions of priority and ordinary customers follows general distributions with rate $\gamma(\nu)$.

3. Analysis of queue size distribution

This section deals with the derivation of governing equations. On account of non-Markovian queueing system, probability generating function and supplementary variable have been used to solve this model.
Let
$N_1(t)$ = Number of priority customers in the queue at time t,
$N_2(t)$ = Number of ordinary customers in the orbit at time t,
$Y(t)$ = State of the server at time t.
Here $M^0(t)$, $B_i^0(t)$ for $i = 1, 2.$, $(R^{(2)})^0(t)$ and $E^0(t)$ indicates elapsed retrial time, elapsed service time for priority and ordinary customers, elapsed restoration period, elapsed emergency vacation at time t.

To obtain a bivariate Markov process $\{N_1(t), N_2(t), Y(t), t > 0\}$, $Y(t)$ denotes the server state. Here $Y(t)$ = (0,1,2,3,4,5), which mean as follows: 0, the server is idle; 1, server is in retrial state; 2, busy with priority customers; 3, busy with ordinary customers; 4, restoration; 5, on Emergency vacation.

Let us assume that, $M^0(0) = 0$, $M^0(\infty) = 1$, $B_i^0(0) = 0$, $B_i^0(\infty) = 1$, $(R^{(2)})^0(0) = 0$, $(R^{(2)})^0(\infty) = 1$ and $E^0(0) = 0$, $E^0(\infty) = 1$ be continuous at $\nu = 0$ for $i = 1, 2$.

If the elapsed time is ν, let function $\beta(\nu)$, $\mu_1(\nu)$, $\mu_2(\nu)$, $\eta_2(\nu)$ and $\gamma(\nu)$ represent the conditional probability of completion rates for the retrial period, high priority and low priority customer's service period, restoration period, and emergency vacation period.

$$\beta(\nu) = \frac{dM(\nu)}{1 - M(\nu)}; \quad \mu_i(\nu) = \frac{dB_i(\nu)}{1 - B_i(\nu)}, i = 1, 2 \quad \eta_2(\nu) = \frac{dR^{(2)}(\nu)}{1 - R^{(2)}(\nu)}$$

$\gamma(\nu) = \dfrac{dE(\nu)}{1 - E(\nu)}$; are the hazard rate functions of $M(.)$, $Bi(.)$, $R^{(2)}(.)$ and $E(.)$ respectively.

The probability $I_{0,n}(\nu, t) = PrN_1(t) = 0, N_2(t) = 0, Y(t) = 0$ and probability densities are as follows:

$$\begin{aligned}
I_{0,n}(\nu,t)d\nu &= \Pr\{N_1(t) = 0, N_2(t) = n, Y(t) = 1; \nu \le I^0(t) \le \nu + d\nu\}, n \ge 1 \\
P^{(1)}_{m,n}(\nu,t)d\nu &= \Pr\{N_1(t) = m, N_2(t) = n, Y(t) = 2; \nu \le B_1^0(t) \le \nu + d\nu\}, \\
P^{(2)}_{m,n}(\nu,t)d\nu &= \Pr\{N_1(t) = m, N_2(t) = n, Y(t) = 3; \nu \le B_2^0(t) \le \nu + d\nu\}, \\
R^{(2)}_{m,n}(\nu,t)d\nu &= \Pr\{N_1(t) = m, N_2(t) = n, Y(t) = 4; \nu \le R^0(t) \le \nu + d\nu\}, \\
E_{m,n}(\nu,t)d\nu &= \Pr\{N_1(t) = m, N_2(t) = n, Y(t) = 5; \nu \le E^0(t) \le \nu + d\nu\}
\end{aligned}$$

for $\nu \ge 0$, $t \ge 0$, $m \ge 0$ and $n \ge 0$.

4. Equation Governing the System

$$\begin{aligned}\frac{\partial}{\partial t}P^{(1)}_{m,n}(\nu,t)+\frac{\partial}{\partial \nu}P^{(1)}_{m,n}(\nu,t) =& -(\lambda_p+a\lambda_o+\alpha_1+\alpha_2+\mu_1(\nu)+\theta)P^{(1)}_{m,n}(\nu,t)\\ &+\lambda_p(1-\delta_{0m})\sum_{i=1}^{m}c_iP^{(1)}_{m-i,n}(\nu,t)\\ &+a\lambda_o(1-\delta_{0n})\sum_{j=1}^{n}c_jP^{(1)}_{m,n-j}(\nu,t)\quad \text{for } m,n\geq 1.\end{aligned} \tag{1}$$

$$\begin{aligned}\frac{\partial}{\partial t}P^{(2)}_{m,n}(\nu,t)+\frac{\partial}{\partial \nu}P^{(2)}_{m,n}(\nu,t) =& -(\lambda_p+a\lambda_o+\alpha_1+\alpha_2+\mu_1(\nu)+\theta)P^{(2)}_{m,n}(\nu,t)\\ &+\lambda_p(1-\delta_{0m})\sum_{i=1}^{m}c_iP^{(2)}_{m-i,n}(\nu,t)\\ &+a\lambda_o(1-\delta_{0n})\sum_{j=1}^{n}c_jP^{(2)}_{m,n-j}(\nu,t)\quad \text{for } m,n\geq 1.\end{aligned} \tag{2}$$

$$\frac{\partial}{\partial t}I_{0,n}(\nu,t)+\frac{\partial}{\partial \nu}I_{0,n}(\nu,t) = -(\lambda_p+\lambda_o+\beta(\nu))I_{0,n}(\nu,t)\quad \text{for } n\geq 1. \tag{3}$$

$$\begin{aligned}\frac{\partial}{\partial t}E_{m,n}(\nu,t)+\frac{\partial}{\partial \nu}E_{m,n}(\nu,t) =& -(\lambda_p+b\lambda_o+\gamma(\nu))E_{m,n}(\nu,t)\\ &+\lambda_p(1-\delta_{0m})\sum_{i=1}^{m}c_iE_{m-i,n}(\nu,t)\\ &+b\lambda_o(1-\delta_{0n})\sum_{j=1}^{n}c_jE_{m,n-i}(\nu,t)\quad \text{for } m,n\geq 1.\end{aligned} \tag{4}$$

$$\begin{aligned}\frac{d}{dt}R^{(1)}_{m,n}(\nu,t)+\frac{d}{d\nu}R^{(1)}_{m,n}(\nu,t) =& -(\lambda_p+b\lambda_o+\eta_1)R^{(1)}_{m,n}(t)+\lambda_p(1-\delta_{0m})\sum_{i=1}^{m}c_iR^{(1)}_{m-i,n}(t)\\ &+\alpha_1\int_0^\infty(P^{(1)}_{m,n}(\nu,t)+P^{(2)}_{m,n}(\nu,t))d\nu\\ &+b\lambda_o(1-\delta_{0n})\sum_{j=1}^{n}c_jR^{(1)}_{m,n-i}(t)\quad \text{for } m,n\geq 1.\end{aligned} \tag{5}$$

$$\begin{aligned}\frac{\partial}{\partial t}R^{(2)}_{m,n}(\nu,t)+\frac{\partial}{\partial \nu}R^{(2)}_{m,n}(\nu,t) =& -(\lambda_p+b\lambda_o+\eta_2(\nu))R^{(2)}_{m,n}(\nu,t)\\ &+\lambda_p(1-\delta_{0m})\sum_{i=1}^{m}c_iR^{(2)}_{m-i,n}(\nu,t)\\ &+b\lambda_o(1-\delta_{0n})\sum_{j=1}^{n}c_jR^{(2)}_{m,n-i}(\nu,t)\quad \text{for } m,n\geq 1.\end{aligned} \tag{6}$$

$$\begin{aligned}\frac{d}{dt}I_{0,0}(t) =& -(\lambda_p+\lambda_o)I_{0,0}(t)+\int_0^\infty P^{(1)}_{0,0}(\nu,t)\mu_1(\nu)d\nu+R^{(1)}_{0,0}(t)\eta_1\\ &+\int_0^\infty P^{(2)}_{0,0}(\nu,t)\mu_2(\nu)d\nu+\int_0^\infty R^{(2)}_{0,0}(\nu,t)\eta_2(\nu)d\nu\\ &+\int_0^\infty E_{0,0}(\nu,t)\gamma(\nu)d\nu.\end{aligned} \tag{7}$$

Define, the boundary conditions at $\nu = 0$

$$\begin{aligned} I_{0,n}(0,t) &= \int_0^\infty E_{0,n}(\nu,t)\gamma(\nu)d\nu + \int_0^\infty P_{0,n}^{(1)}(\nu,t)\mu_1(\nu)d\nu + R_{0,n}^{(1)}(t)\eta_1 \\ &+ \int_0^\infty P_{0,n}^{(2)}(\nu,t)\mu_2(\nu)d\nu + \int_0^\infty R_{0,n}^{(2)}(\nu,t)\eta_2(\nu)d\nu, \quad \text{for } n \geq 0. \end{aligned} \tag{8}$$

$$\begin{aligned} P_{m,n}^{(1)}(0,t) &= \int_0^\infty P_{m+1,n}^{(1)}(\nu,t)\mu_1(\nu)d\nu + R_{m+1,n}^{(1)}(t)\eta_1 + \int_0^\infty P_{m+1,n}^{(2)}(\nu,t)\mu_2(\nu)d\nu \\ &+ \int_0^\infty E_{m+1,n}(\nu,t)\gamma(\nu)d\nu + \int_0^\infty R_{m+1,n}^{(2)}(\nu,t)\eta(\nu)d\nu + \lambda_p c_{m+1} I_{0,n}(t), \end{aligned} \tag{9}$$

$$P_{0,0}^{(2)}(0,t) = \lambda_o c_1 I_{0,0}(t) + \int_0^\infty I_{0,1}(\nu,t)\beta(\nu)d\nu \tag{10}$$

$$\begin{aligned} P_{0,n}^{(2)}(0,t) &= \lambda_o c_{n+1} I_{0,0}(t) + \int_0^\infty I_{0,n+1}(\nu,t)\beta(\nu)d\nu + \sum_{i=1}^{n} \lambda_o C_i(\nu,t) \\ &+ \int_0^\infty I_{0,n+1-i}(\nu,t)d\nu \quad \text{for } n \geq 1. \end{aligned} \tag{11}$$

$$\begin{aligned} P_{0,n}^{(2)}(0,t) &= \lambda_o c_{n+1} I_{0,0}(t) + \int_0^\infty I_{0,n+1}(\nu,t)\beta(\nu)d\nu + \sum_{i=1}^{n} \lambda_o C_i(\nu,t) \\ &+ \int_0^\infty I_{0,n+1-i}(\nu,t)d\nu \quad \text{for } n \geq 1. \end{aligned} \tag{12}$$

$$R_{m,n}^{(2)}(0,t) = \alpha_2 \int_0^\infty P_{m,n}^{(1)}(\nu,t)\mu_1(\nu)d\nu + \alpha_2 \int_0^\infty P_{m,n}^{(2)}(\nu,t)d\nu \quad \text{for } m,n \geq 0. \tag{13}$$

$$\begin{aligned} &P_{m,n}^{(1)}(0) = P_{m,n}^{(2)}(0) = E_{m,n}(0) = R_{m,n}^{(1)}(0) = R_{m,n}^{(2)}(0) = 0, \quad \text{for } m,n \geq 0 \text{ and } I_{0,0} = 1, \\ &I_{0,n}(0) = 0, \quad \text{for } n \geq 1 \text{ are the initial conditions.} \end{aligned} \tag{14}$$

Now, we define the Probability Generating Function (PGF),

$$\begin{aligned} &I(\nu,t,z_o) = \sum_{n=1}^{\infty} z_o^n I_{0,n}(\nu,t); \qquad A(\nu,t,z_p,z_o) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} z_o^m z_p^n A_{m,n}(\nu,t); \\ &A(\nu,t,z_p) = \sum_{m=0}^{\infty} z_p^m A_m(\nu,t); \qquad A(\nu,t,z_o) = \sum_{n=0}^{\infty} z_o^n A_n(\nu,t); \end{aligned} \tag{15}$$

here $A = P^{(1)},\ P^{(2)},\ E,\ R^{(1)},\ R^{(2)}$.
By applying Laplace transforms to equations (1) to (13) and by using (14) and (15) , we obtain the following equations:

$$\overline{I}_0(\nu,s,z_o) = \overline{I}_0(0,s,z_o)e^{-(s+\lambda_p+\lambda_o)\nu - \int_0^\nu \beta(t)dt}, \tag{16}$$

$$\overline{P}^{(1)}(\nu,s,z_p,z_o) = \overline{P^{(1)}}(0,s,z_p,z_o)e^{-\phi_1(s,z)\nu - \int_0^\nu \mu_1(t)dt}, \tag{17}$$

$$\overline{P}^{(2)}(\nu,s,z_p,z_o) = \overline{P^{(2)}}(0,s,z_p,z_o)e^{-\phi_1(s,z)\nu - \int_0^\nu \mu_2(t)dt}, \tag{18}$$

$$\overline{E}(\nu,s,z_p,z_o) = \overline{E}(0,s,z_p,z_o)e^{-\phi_2(s,z)\nu - \int_0^\nu \gamma(t)dt}, \tag{19}$$

$$\overline{R}^{(2)}(\nu,s,z_p,z_o) = \overline{R}^{(2)}(0,s,z_p,z_o)e^{-\phi_2(s,z)\nu - \int_0^\nu \eta_2(t)dt}. \tag{20}$$

where,

$$\phi_1(s,z) = s + \lambda_p(1 - C(z_p)) + a\lambda_o(1 - C(z_o)) + \alpha_1 + \alpha_2 + \theta, \tag{21}$$

$$\phi_2(s,z) = s + \lambda_p(1 - C(z_p)) + b\lambda_o(1 - C(z_o)), \tag{22}$$

$$\phi_3(s,z) = s + \lambda_p(1 - C(z_p)) + b\lambda_o(1 - C(z_o)) + \eta_1, \tag{23}$$

$$\overline{P}^{(2)}(0,s,z_o)=\frac{\left\{\begin{array}{l}\lambda_o C(z_o)\overline{I}_{0,0}(s)[1-\lambda_p C(g(z_o))\\ \Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]]-[\overline{I}_{0,0}(s)(s+\lambda_p+\lambda_o)-1]\\ [\overline{M}(s+\lambda_p+\lambda_o)+C(z_o)\lambda_o\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]]\end{array}\right\}}{\left\{\begin{array}{l}z_2[1-\lambda_p C(g(z_o))\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]]\\ -[\overline{M}(s+\lambda_p+\lambda_o)+C(z_o)\lambda_o\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]]\\ [\overline{B}_2(\sigma_1(z,s))+\theta z_o\overline{E}(\sigma_2(z,s))\Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]+\\ \Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]\Big[\frac{\alpha_1}{\sigma_3(z,s)}+\alpha_2\overline{R}^{(2)}(\sigma_2(z,s))\Big]]\end{array}\right\}}, \tag{24}$$

$$\overline{I}(0,s,z_o)=\frac{\left\{\begin{array}{l}\lambda_o C(z_o)\overline{I}_{0,0}(s)[\overline{B}_2(\sigma_1(z,s)+\theta z_o\overline{E}(\sigma_2(z,s))\\ \Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]+\Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]\\ \Big[\frac{\alpha_1}{\sigma_3(z,s)}+\alpha_2\overline{R}^{(2)}(\sigma_2(z,s))\Big]]\\ -[\overline{I}_{0,0}(s)(s+\lambda_p+\lambda_o)-1]z_o\end{array}\right\}}{\left\{\begin{array}{l}z_2[1-\lambda_p C(g(z_o))\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]]\\ -[\overline{M}(s+\lambda_p+\lambda_o)+C(z_o)\lambda_o\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]]\\ [\overline{B}_2(\sigma_1(z,s))+\theta z_o\overline{E}(\sigma_2(z,s))\Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]+\\ \Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]\Big[\frac{\alpha_1}{\sigma_3(z,s)}+\alpha_2\overline{R}^{(2)}(\sigma_2(z,s))\Big]]\end{array}\right\}}, \tag{25}$$

$$\overline{P}^{(1)}(0,s,z_p,z_o)=\frac{\left\{\begin{array}{l}\lambda_p[C(z_p)-C(g(z_o))]\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]\overline{I}_0(0,s,z_o)\\ +[\overline{B_2}(\phi_1(z,s))-\overline{B_2}(\sigma_1(z,s))+\theta z_o\overline{E}(\phi_2(z,s))\\ \Big[\frac{1-\overline{B}_2(\phi_1(z,s))}{\phi_1(z,s)}\Big]-\theta z_o\overline{E}(\sigma_2(z,s))\Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]\\ +\Big[\frac{\alpha_1}{\phi_3(z,s)}+\alpha_2\overline{R}^{(2)}(\phi_2(z,s))\Big]\Big[\frac{1-\overline{B_2}(\phi_1(z,s))}{\phi_1(z,s)}\Big]\\ -\Big[\frac{1-\overline{B}_2(\sigma_1(z,s))}{\sigma_1(z,s)}\Big]\Big[\frac{\alpha_1}{\sigma_3(z,s)}+\alpha_2\overline{R}^{(2)}(\sigma_2(z,s))\Big]]\end{array}\right\}}{\left\{\begin{array}{l}z_p-[\overline{B_1}(\phi_1(z,s))+\theta z_o\overline{E}(\phi_2(z,s))\Big[\frac{1-\overline{B}_1(\phi_1(z,s))}{\phi_1(z,s)}\Big]\\ +\Big[\frac{\alpha_1}{\phi_3(z,s)}+\alpha_2\overline{R}^{(2)}(\phi_2(z,s))\Big]\Big[\frac{1-\overline{B_1}(\phi_1(z,s))}{\phi_1(z,s)}\Big]]\end{array}\right\}}, \tag{26}$$

$$\begin{aligned}\overline{E}(0,s,z_p,z_o)=\theta z_p\overline{P}^{(1)}(0,s,z_p,z_o)\Big[\frac{1-\overline{B_1}(\phi_1(z,s))}{\phi_1(z,s)}\Big]\\ +\theta z_o\overline{P}^{(2)}(0,s,z_o)\Big[\frac{1-\overline{B_2}(\phi_1(z,s))}{\phi_1(z,s)}\Big],\end{aligned} \tag{27}$$

$$\overline{R}^{(2)}(0,s,z_p,z_o) = \alpha_2\overline{P}^{(1)}(0,s,z_1,z_2)\Big[\frac{1-\overline{B_1}(\phi_1(z,s))}{\phi_1(z,s)}\Big] + \alpha_2\overline{P}^{(2)}(0,s,z_2)\Big[\frac{1-\overline{B_2}(\phi_1(z,s))}{\phi_1(z,s)}\Big], \tag{28}$$

$$\sigma_1(s,z) = s+\lambda_p(1-C(g(z_o)))+a\lambda_o(1-C(z_o))+\alpha_1+\alpha_2+\theta,$$
$$\sigma_2(s,z) = s+\lambda_p(1-C(g(z_o)))+b\lambda_o(1-C(z_o)),$$
$$\sigma_3(s,z) = s+\lambda_p(1-C(g(z_o)))+b\lambda_o(1-C(z_o))+\eta_1.$$

Theorem.1 When the system is in regular service, breakdown, emergency vacation and repair by using the Laplace transforms the probability generating function of the number of customers in the respective queue is given by.

$$\overline{I}_0(s,z_o) = \overline{I}_0(0,s,z_o)\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big], \tag{29}$$

$$\overline{P}^{(1)}(s,z_p,z_o) = \overline{P}^{(1)}(0,s,z_p,z_o)\Big[\frac{1-\overline{B}_1(\phi_1(s,z))}{\phi_1(s,z)}\Big], \tag{30}$$

$$\overline{P}^{(2)}(s,z_p,z_o) = \overline{P}^{(2)}(0,s,z_o)\Big[\frac{1-\overline{B}_2(\phi_1(s,z))}{\phi_1(s,z)}\Big], \tag{31}$$

$$\overline{E}(s,z_p,z_o) = \overline{E}(0,s,z_p,z_o)\Big[\frac{1-\overline{E}(\phi_2(s,z))}{\phi_2(s,z)}\Big], \tag{32}$$

$$\overline{R}^{(2)}(s,z_p,z_o) = \overline{R}^{(2)}(0,s,z_p,z_o)\Big[\frac{1-\overline{R}^{(2)}(\phi_2(s,z))}{\phi_2(s,z)}\Big]. \tag{33}$$

Proof: Integrating the preceding equations (29) to (33) with respect to ν and applying the solution of renewal theory we obtain the following

$$\int_0^\infty \big[1-H(\nu)\big]e^{-s\nu}d\nu = \frac{1-\overline{h(s)}}{s}. \tag{34}$$

Here, the LST of the distribution function of a random variable H(ν) is denoted as $\overline{h(s)}$. The overall results of the probability generating functions for the following states, $\overline{I}_0(s,z_o)$, $\overline{P}^{(1)}(s,z_p,z_o)$, $\overline{P}^{(2)}(s,z_p,z_o)$, $\overline{E}(s,z_p,z_o)$, and $\overline{R}^{(2)}(s,z_p,z_o)$ are obtained by using equation (29) to (33).

5. Steady State Analysis

According to Tauberian property,

$$\lim_{s\to 0} s\overline{f}(s) = \lim_{t\to\infty} f(t).$$

Despite of the state of the system, the probability generating function of the queue size is as follows:

$$W_q(z_1,z_2) = \frac{Nr(z_1,z_2)}{Dr(z_1,z_2)}, \tag{35}$$

where

$$Nr(z_p,z_o) = N_3(z)D_1(z)\phi_1(z)\phi_2(z)\phi_3(z)\Big[\frac{1-\overline{M}(s+\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big] + N_1(z)D_2(z)(1-\overline{B_1}(\phi_1(z,s)))F_1(z) + N_2(z)D_1(z)(1-\overline{B_2}(\phi_1(z,s)))F_2(z)$$

$$Dr(z_p,z_o) = D_1(z)D_2(z)\phi_1(z)\phi_2(z)\phi_3(z),$$

$$N_1(z) = \lambda_p[C(z_p)-C(g(z_o))]\Big[\frac{1-\overline{M}(\lambda_p+\lambda_o)}{\lambda_p+\lambda_o}\Big]\overline{I}_0(0,z_o) + [\overline{B_2}(\phi_1(z)) - \overline{B_2}(\sigma_1(z)) + \theta z_o\overline{E}(\phi_2(z)) + \Big[\frac{\alpha_1}{\phi_3(z)}\Big[\frac{1-\overline{B}_2(\phi_1(z))}{\phi_1(z)}\Big] - \theta z_o\overline{E}(\sigma_2(z))\Big[\frac{1-\overline{B}_2(\sigma_1(z))}{\sigma_1(z)}\Big] + \alpha_2\overline{R}^{(2)}(\phi_2(z))\Big]\Big[\frac{1-\overline{B_2}(\phi_1(z))}{\phi_1(z)}\Big] - \Big[\frac{1-\overline{B}_2(\sigma_1(z))}{\sigma_1(z)}\Big]\Big[\frac{\alpha_1\eta_1}{\sigma_3(z)} + \alpha_2\overline{R}^{(2)}(\sigma_2(z))\Big]],$$

$$N_2(z) = -[\lambda_p(1-C(g(z_o))) + b\lambda_o(1-C(z_o))][1-\lambda_pC(g(z_o))\Big[\frac{1-\overline{M}(\lambda_p+\lambda_o)}{\lambda_p+\lambda_o}\Big]] [\overline{M}(\lambda_p+\lambda_o) + C(z_o)\lambda_o\Big[\frac{1-\overline{M}(\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big],$$

$$N_3(z) = -[\lambda_p(1-C(g(z_o))) + b\lambda_o(1-C(z_o))][\overline{B}_2(\sigma_1(z)) + \theta z_o\overline{E}(\sigma_2(z)) \Big[\frac{1-\overline{B}_2(\sigma_1(z))}{\sigma_1(z)}\Big] + \Big[\frac{1-\overline{B}_2(\sigma_1(z))}{\sigma_1(z)}\Big]\Big[\frac{\alpha_1\eta_1}{\sigma_3(z)} + \alpha_2\overline{R}^{(2)}(\sigma_2(z))\Big]],$$

$$D_1(z) = (z_p - [\overline{B_1}(\phi_1(z)) + \theta z_o\overline{E}(\phi_2(z))\Big[\frac{1-\overline{B}_1(\phi_1(z))}{\phi_1(z)}\Big] + \Big[\frac{\alpha_1\eta_1}{\phi_3(z)} + \alpha_2\overline{R}^{(2)}(\phi_2(z))\Big]\Big[\frac{1-\overline{B_1}(\phi_1(z))}{\phi_1(z)}\Big]],$$

$$D_2(z) = z_o[1-\lambda_pC(g(z_o))\Big[\frac{1-\overline{M}(\lambda_p+\lambda_o)}{s+\lambda_p+\lambda_o}\Big]] - [\overline{M}(\lambda_p+\lambda_o) + C(z_o)\lambda_o\Big[\frac{1-\overline{M}(\lambda_p+\lambda_o)}{\lambda_p+\lambda_o}\Big]][\overline{B}_2(\sigma_1(z)) + \theta z_o\overline{E}(\sigma_2(z)) \Big[\frac{1-\overline{B}_2(\sigma_1(z))}{\sigma_1(z)}\Big] + \Big[\frac{1-\overline{B}_2(\sigma_1(z))}{\sigma_1(z)}\Big]\Big[\frac{\alpha_1\eta_1}{\sigma_3(z)} + \alpha_2\overline{R}^{(2)}(\sigma_2(z))\Big]].$$

6. Stability Condition

We apply the normalising condition to determine $I_{0,0}$.

$$I_{0,0} + I_0(1) + P^{(1)}(1,1) + P^{(2)}(1,1) + E(1,1) + R^{(1)}(1,1) + R^{(2)}(1,1) = 1 \tag{36}$$

$$I_{0,0} = \frac{D_2'\phi_1\phi_2'\phi_3 - N_2'F'(1-\overline{B}_2(\phi_1(z)))}{D_2'\phi_1\phi_2'\phi_3} \tag{37}$$

and the utilization factor is given by

$$\rho = \frac{N_2'F'(1-\overline{B}_2(\phi_1(z)))}{D_2'\phi_1\phi_2'\phi_3} \tag{38}$$

The stability condition for the model under which steady state exists is $\rho < 1$

7. Performance Measures

The expected queue size for priority customer is as follows:

$$L_{q_1} = \frac{d}{dz_p} W_q(z_p, 1)|_{z_p=1} \tag{39}$$

The expected orbit size for ordinary customer is as follows:

$$L_{q_2} = \frac{d}{dz_o} W_q(1, z_o)|_{z_o=1} \tag{40}$$

then

$$L_{q_1} = \frac{Dr''(1)Nr^{(''')}(1) - Dr^{(''')}(1)Nr''(1)}{3(Dr''(1))^2}, \tag{41}$$

$$L_{q_2} = \frac{Dr'''(1)Nr^{(iv)}(1) - Dr^{(iv)}(1)Nr'''(1)}{4(Dr'''(1))^2}. \tag{42}$$

The expected waiting time for priority queue is as follows:

$$W_{q_1} = \frac{L_{q_1}}{\lambda_p} \tag{43}$$

The expected waiting time for orbit is as follows:

$$W_{q_2} = \frac{L_{q_2}}{\lambda_o}. \tag{44}$$

8. Particular Cases

Case 1:
Without a priority queue, all arriving customers are accepted into the system; customers do not balk to orbit, take vacations and no failures, the above model becomes

$$I_0(z) = \frac{I_{0,0}[C(z_o)\overline{B}(\phi(z)) - z_o][1 - \overline{M}(\lambda_o)]}{\overline{B}(\phi(z))[C(z_o) + \overline{M}(\lambda_o)(1 - C(z_o))] - z_o},$$

$$P^{(2)}(z) = \frac{I_{0,0}(1 - \overline{B}(\phi(z)))\overline{M}(\lambda_o)}{\overline{B}(\phi(z))[(1 - C(z_o))\overline{M}(\lambda_o) - z_o]}.$$

This result is associated with Gomez-Corral [15].

Case 2:
In the absence of priority queue, when there is no breakdown, no retrial, no balking, no repair and no vacation then the above model becomes

$$P^{(2)}(z) = \frac{I_{0,0}(1 - \overline{B}_2(\phi(z)))}{\overline{B}_2(\phi(z)) - z_o}$$

This result is associated with Medhi [21].

9. Numerical Results

The numerical and graphical analyses of this model are covered in this section. We assumed that the distribution of service time, breakdown, repair, and vacation time are all exponential.

Table 1: *Effect of priority arrival rate (λ_p)*

λ_p	$I_{0,0}$	ρ	L_{q_1}	W_{q_1}	L_{q_2}	W_{q_2}
0.5	0.9971	0.0029	1.3928	2.7856	0.2013	0.1006
0.6	0.9963	0.0037	1.8271	3.0452	0.2497	0.1248
0.7	0.9955	0.0045	2.4404	3.4863	0.3016	0.1508
0.8	0.9945	0.0055	3.3321	4.1651	0.3574	0.1787
0.9	0.9934	0.0066	4.6809	5.2010	0.4174	0.2087
1.0	0.9922	0.0078	6.8391	6.8391	0.4819	0.2410

Table 1 demonstrates that the probability of the server being idle reduces as the arrival rate (λ_p) of priority customers for the priority queue rises. However, average queue lengths, busy period and customers average waiting times all rises: we assume the values as $\lambda_o = 2$, $\alpha_1 = 0.3$, $\alpha_2 = 0.5$ $\mu = 5$, $\eta_1 = 3$,$\eta_2 = 5$,$\theta = 0.5$, $\beta = 15$,$\gamma = 10$, $a = 0.5$,$b = 0.6$ and $\lambda_p = 0.5$ to 1.0.

Table 2: *Effect of service rate (μ)*

μ	$I_{0,0}$	ρ	L_{q_1}	W_{q_1}	L_{q_2}	W_{q_2}
2	0.6611	0.3389	3.7025	1.9418	7.0960	4.3513
3	0.9627	0.0373	2.2242	1.8339	6.3020	3.1121
4	0.9906	0.0094	1.7731	1.5863	5.7105	2.3861
5	0.9966	0.0034	0.3434	1.3856	5.2686	1.1717
6	0.9985	0.0015	0.1792	1.2542	4.9155	0.0896

Table 2 indicates that when the service rate (μ) increases, the probability of the server being busy reduces. However, average queue lengths, idle time, and customers' average waiting times all reduces: we assume the values as $\lambda_1 = 0.5$, $\lambda_2 = 2$, $\alpha_1 = 0.3$, $\alpha_2 = 0.5$, $\eta_1 = 1$, $\eta_2 = 6$ $\beta = 10$, $\theta = 0.5$,$\gamma = 10$, $\beta = 15$,$a = 0.5$,$b = 0.6$and $\mu = 2$ to 6.

Table 3: *Effect of retrial rate (β)*

β	$I_{0,0}$	ρ	L_{q_1}	W_{q_1}	L_{q_2}	W_{q_2}
10	0.9725	0.0275	2.6038	5.2076	0.7172	0.3586
11	0.9724	0.0276	2.3409	4.6817	0.6998	0.3499
12	0.9722	0.0278	2.1251	4.2503	0.6856	0.3428
13	0.9721	0.0279	1.9449	3.8899	0.6737	0.3369
14	0.9720	0.0280	1.7922	3.5843	0.6636	0.3318
15	0.9719	0.0281	1.6610	3.3220	0.6549	0.3275

Table 3 indicates the probability of the server being busy period rises as the retrial rate rises. However, average queue lengths, idle time, and customers' average waiting times all reduces: we assume the values as $\lambda_p = 0.5$,$\lambda_o = 2$, $\alpha_1 = 0.3$,$\alpha_2 = 0.6$, $\eta_1 = 3$,$\mu = 3$,$\theta = 0.5$,$\eta_2 = 5$,$a = 0.5$,$b = 0.6$ and $\beta = 10to15$.

We assumed to be follow the Erlang-2 distribution for service time, breakdown, repair, and vacation time in graphical representations. The two-dimensional graphs are shown in Figure 2 - 4. Figure 2 exhibits that the expected length of the queue (L_{q_1}, L_{q_2}) rises, the expected length of the queue extends together with the priority arrival rate (λ_p). The behaviour of the queue sizes (L_{q_1}, L_{q_2}), which depends on the service rate (μ), is shown in Figure 3, the length of the queue as

the service rate decreases. Figure 4 shows the expected queue length (L_{q_1}, L_{q_2}), which depends on the retrial rate (β), the length of the queue as the service rate decreases.

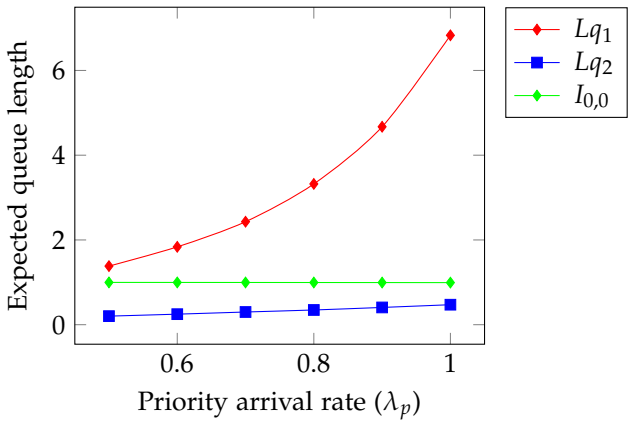


Figure 2: *Expected queue length vs priority arrival rate λ_p*

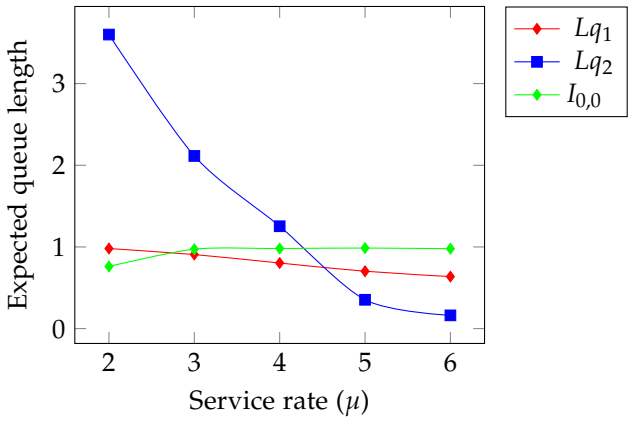


Figure 3: *Expected queue length vs service rate μ*

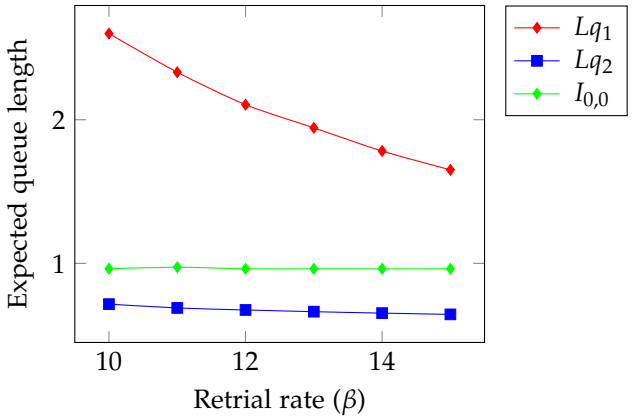


Figure 4: *Expected queue length vs Retrial rate β*

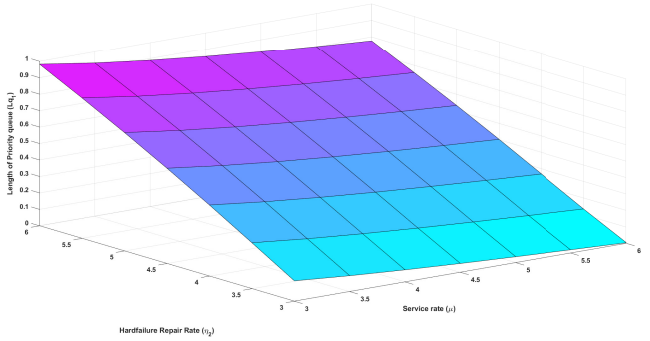


Figure 5: *Lq_1 Vs μ and η_2*

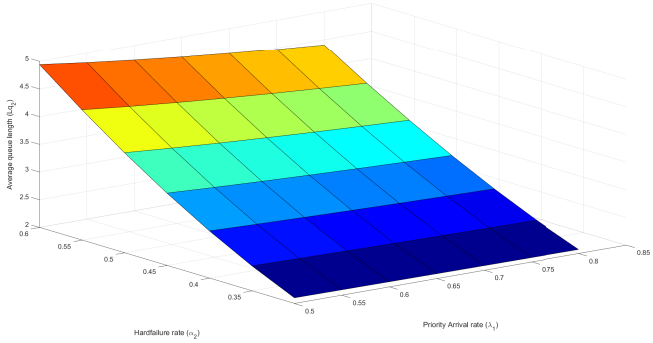


Figure 6: *Lq_2 Vs λ_p and α_2*

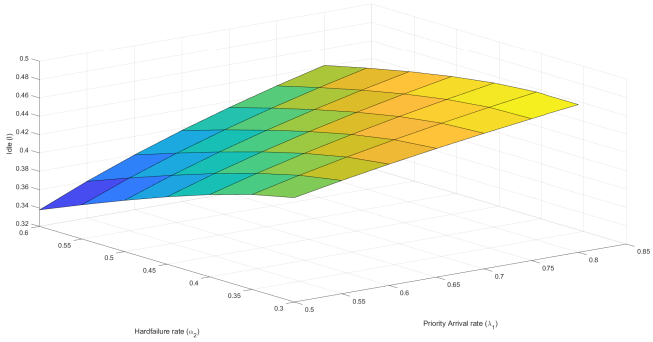


Figure 7: *Idle Vs λ_p and α_2*

Graphs in three dimensions can be found in Figures 5 - 7. Figure 5 in the reference indicate that the service rate (μ) and hard failure repair rate (α_2) increase, the expected queue size (Lq_2) decease. Figure 6 in the reference indicate that as the priority arrival rate (λ_1) and breakdown(hard failure) rate (α_2) increase, the expected queue size (Lq_2) rises. Figure 6 in the reference indicate that as the priority arrival rate (λ_p) and breakdown (hard failure) rate (α_2) increase, the idle is $(I_0, 0)$ rises.

10. Conclusion

In this inquiry, we investigated a single server retrial queueing system with admission control, balking, non-preemptive priority service, and emergency vacation where the server is susceptible

to various breakdown and restoration periods. The analytical findings that are supported by numerical examples can be applied to design outputs in a variety of real-world scenarios. The supplementary variable technique is used to determine the PGFs for the number of customers in the system when it is free, busy and restoration period.The average queue length of the orbit and system contains explicit expressions. The mean busy period and other significant system performance measures are obtained. Finally, it is demonstrated that this retrial queueing method works well with the conditional decomposition law. Our technique is more adaptable in dealing with real-time systems of numerous sectors in many real-life queueing scenarios. This work can be expanded in various directions by considering the concept of:

- Multi server batch arrival priority queueing model with production inventory system.
- Batch arrival bulk service double orbit retrial queueing system with priority service.

References

[1] Ammar, S.I. (2017). Transient solution of an $M/M/1$ vacation queue with a waiting server and impatient customers, *Journal of the Egyptian Mathematical Society*, 25(3):337-342.

[2] Anna Bagyam, J.E. and Udhaya Chandrika, K. (2010). Single server retrial queueing system with two different vacation policies, *International Journal of Contemporary Mathematical Sciences*, 5(32):1591-1598.

[3] Artalejo, J.R. and Lopez-Herrero, M.J. (2014). On The Single Server Retrial Queue With Balking, *INFOR: Information Systems and Operation Research*, 38(1):33-50.

[4] Artalejo, J. R., Atencia, I. and Moreno, P. (2005). A discrete time $Geo[X]/G/1$ retrial queue with control of admission, *Applied Mathematical Modelling*, 29(11):1100-1120.

[5] Ayyappan, G. and Thilagavathy, K. (2021). Analysis of $MAP(1), MAP(2)/PH/1$ Non-preemptive Priority Queueing Model Under Classical Retrial Policy with Breakdown, Repair, Discouragement, Single Vacation, Standby Server, Negative Arrival and Impatient Customers, *International Journal of Applied and Computational Mathematics*, 184(7):1-24.

[6] Ayyappan, G. and Thamizhselvi, T. (2018). Transient analysis of $M^{[X_1]}, M^{[X_2]}/G_1, G_2/1$ retrial queueing system with priority services, working vacations and vacation interruption, emergency vacation, negative arrival and delayed repair, *International Journal of Applied and Computational Mathematics*, 77(4):1-35.

[7] Ayyappan, G. and Sankeetha, S. (2022). $MAP/PH/1$ Queue with vacation, customer introduced intrruption, optional service, breakdown and repair completion, *Reliability: Theory and Applications*, 17(2):367-386.

[8] Bhagat, A. and Madhu Jain. (2020). Retrial queue with multiple repairs, multiple services and non preemptive priority, *OPSEARCH* , 57:787-814.

[9] Choudhury, G. (2007). A two phase batch arrival retrial queueing system with Bernoulli vacation schedule, *Applied Mathematics and Computation*, 188(2):1455-1466.

[10] Choudhury, G. and Deka, K. (2009). An $M^{[X]}/G/1$ unreliable retrial queue with two phases of service and Bernoulli admission mechanism, *App. Math. and Computation*, 215(3):936-949.

[11] Choudhury, G. (2002). A batch arrival queue with a vacation time under single vacation policy, *Computers & Operation Research*, 29(4):1941-1955.

[12] Dimitriou, I. (2013). A preemptive resume priority retrial queue with state dependent arrivals, unreliable server and negative customers, *TOP*, 21(3):542-571.

[13] Fiems, D., Maertens, T. and Bruneel, H. (2008). Queueing systems with different types of server interruptions, *Europian Journal of Operation Research* , 188:838-845.

[14] Gao, S. and Wang, J. (2014). Performance and reliability analysis of an $M/G/1$-G retrial queue with orbital search and non-persistent customers, *Europian Journal of Operation Research*, 236(2):561-572.

[15] Gomez-Corral, A. (1999). Stochastic analysis of a single server retrial queue with general retrial times, *Naval Research Logistics*, 46:61-581.

[16] Jain, M. and Agrawal, P.(2008). Optimal policy for bulk queue with multiple types of server breakdown, *International Journal of Operation Research*, 4(1):35-54.

[17] Jeyakumar, S. and Senthilnathan, B. (2012). A study on the behaviour of the server breakdown without interruption in a $M^{[X]}/G^{(a,b)}/1$ queueing system with multiple vacations and closedown time, *Applied Mathematics and Computation*, 219(5):2618-2633.

[18] Ke, J.C. (2007). Operating characteristic analysis on the $M^{[X]}/G/1$ system with a variant vacation policy and balking, *Applied Mathematical Modelling*, 31:1321-1337.

[19] Kim, B., Kim, J. and Ole Bueker.(2021). Non-preemptive priority $M/M/m$ queue with server's vacations, *Computers & Industrial Engineering*, 160:107390.

[20] Madhu, J. and Indu, D. (2001). Control policy for $M/E_k/1$ queueing system', *Journal of Statistics and Management Systems*, 4(1):73-82.

[21] Medhi, J. (2003). Stochastic models in queueing theory, *Academic Press, Boston, Second edition*.

[22] Pandey, D.C. and Pal, A.K. (2014). Delay Analysis of a Discrete-Time Non-Preemptive Priority Queue with Priority Jumps, *Applications and Applied Mathematics*, 9(1):1-12.

[23] Rajadurai, P., Chandrasekaran, V.M. and Saravanarajan, M.C. (2015). Analysis of an$M^{[X]}/G/1$ unreliable retrial G-queue with orbital search and feedback under Bernoulli vacation schedule, *OPSEARCH*, 53(1):197-223.

[24] Rismawati, D., Mangku, W. and Sumarno, H.(2023). $M/M/1$ Non-preemptive Priority Queuing System with Multiple Vacations and Vacation Interruptions, *IOPSEARCH* , 7(3):751-764.

[25] Shekhar, C., Varshney, S. and Amit kumar. (2020). Optimal control of a service system with emergency vacation using bat algorithm, *Journal of Computational and Applied Mathematics*, 364:1-14.

[26] Wang, J. (2008). On the single server retrial queue with priority subscribers and server breakdowns, *Journal of Systems Science and Complexity*, 21(2):304-315.

[27] Wu, J. and Lian, Z. (2013). A single-server retrial G-queue with priority and unreliable server under Bernoulli vacation schedule, *Computational and Industrial Engineering*, 64(1):84-93.

[28] Yi-Chih Hsieh and Andersland, S.(1995). Repairable single server systems with multiple breakdown modes, *Microelectronics Reliability*, 35(2):309-318.

[29] Zhang, F. and Zhu, Z.(2013). A Discrete-Time $Geo/G/1$ Retrial Queue with Vacations and Two Types of Breakdowns, *Journal of Applied Mathematics*, 2013:1-11.

PERFORMANCE ANALYSIS OF BULK ARRIVAL GENERAL SERVICE QUEUE WITH FEEDBACK, IMPATIENT CUSTOMERS AND SECOND OPTIONAL SERVICE

P. Vijaya Laxmi[1,*], Hasan A. Qrewi[2], E. Girija Bhavani [3]

[1,2] Department of Applied Mathematics, Andhra University,
Visakhapatnam, India.
[3] MVGR College of Engineering, Vizianagaram, India
* Corresponding Author. vijayalaxmiau@gmail.com

Abstract

This paper analyzes the steady state behavior of batch arrival non-Markovian service queue with feedback, balking, reneging, and second optional service (SOS). The steady-state probabilities are computed using the probability generating function. After completing the first essential service (FES), if a customer is unsatisfied with it, he may choose to rejoin the system (feedback), opt for the SOS, or depart from the system with specific probabilities. Once a customer arrives, he decides immediately to join the queue or refuses to join (balking). Furthermore, after joining the queue if a customer does not get service within a specific time, may become impatient, and decide to leave the line without getting any service (reneging). Reneging time follows exponential distribution while service time (FES and SOS) follow general distribution. Also, the cost model was presented to determine the optimal service rates to minimize the expected cost. Finally, various performance measures and numerical illustrations are provided.

Keywords: Batch arrival; Steady State; Non Markovian; Feedback; Balking; Reneging; First essential service, Second optional service; Queue

I. Introduction

In queueing theory, items may arrive in batches. Known as batch arrival queueing models. A perfect example of such models is a digital communication system as [1] studied batch arrival queue systems with breakdown and repairs in which the services are performed in two different stages. At the end of each second phase of service, the server takes a compulsory vacation. The service times of the two stages follow general distributions. The expected number of units in the system has been obtained using the probability generating function. In [2] the probability generating functions have been used to study the transient and the steady state behavior of a batch arrival system and batch service with SOS. The service time distribution of both FES and SOS are exponential. [3] analyzed the steady state of $M^X/G/1$ queue with a retrial and two stages of heterogeneous services with admission, feedback, and general retrial time. The arrivals join with dependent admission due to the server state. The supplementary variable approach has been used to derive the stationary equations, the generating functions of the number of customers in the system and the orbit, and the mean queue size in the system and the orbit. Prominent research papers on the batch arrival queues can be found in [4], [5], [6], [7], [8], [9] and the references therein.

Many authors have studied customer behavior in the queueing system whereby some customers, upon arrival, decide to join the queue or refuse to join the queue. This situation is referred to as balking. The other situation is reneging where a customer upon joining the queue and

waiting a specific period of time without getting service, may get impatient and may leave the queue. These two terminologies of balking and reneging are referred to as impatience behavior. [10] analyzed a single server queue model with impatience where the customers lose patience if the wait is more than the threshold they fixed. Later in [11] a study on batch arrival queue system with vacation and breakdown is done. The server provides two stages of service one by one in succession, and the customer may renege during breakdown or vacation period. Recently [12] studied batch arrival queueing system with balking, three types of heterogeneous service, and vacation. The impatient customers are assumed to balk during the period when the server is activated on the system or when the server is on vacation. Many related studies on balking are found in [13], [14], [15], [16], [17], [18], etc.

Several researchers have studied queueing systems with feedback, such as [19] investigated a batch arrival system with two-phase heterogeneous service, breakdown, and compulsory server vacation. After a customer completes two stages of services and if feels unsatisfied with the service, then he may join the tail of the queue as a feedback customer for receiving another service with a certain probability otherwise he leave the system. Later, [20] studied an M/G/1 with feedback and vacation. They consider the service times as independent and identically distributed with different rates when the customer is served with feedback or without feedback. Recently, in [21] the authors have investigated an $M/M^b/1$ with SOS and feedback. The customers are served in batches with batch size of maximum capacity b. After customers complete FES, if they are unsatisfied, they will rejoin the queue and retake the service; otherwise, they opt for SOS or leave the system. Other studies on feedback are found in [22], [23], [24], [25], [26], etc.

In queueing literature, we found studies on batch arrival non-Markovian queue systems, which include some assumptions such as feedback, balking, and reneging. The queue systems with balking, reneging, and feedback have many applications in our lives. For example, inventory and production, call centers, computer networks, etc.Therefore, adding SOS to the model which includes feedback, balking, and reneging will make the model more adaptable, and motivates us to explore its behavior under a steady state environment. We use the probability generating function to obtain the steady-state probabilities. Some important performance measures are obtained. Also, some interesting special cases were discussed. The cost analysis is derived by using the method of Quasi-Newton method. Finally, some numerical results are presented in the form of tables and graphs to show the effect of parameters on the performance measures.

This paper is structured as follows: description of the model and governing equations are presented in Section 2. In Section 3, we study the steady-state solution. Some performance measures are obtained in Section 4. In Section 5, we discuss some particular cases. Cost analysis and numerical illustrations are presented in Section 6. Finally, Section 7 concludes our paper.

II. Model Description and Mathematical Formulation

In this paper, we study an $M^X/G/1$ queue with SOS, balking, reneging and feedback. A brief description of the model is presented in the following lines:

- Customers arrive in bathes of the random size, say X , say X, in a compound Poisson process with probability $P(X = j) = c_j$, so that $\lambda c_j dt$ is the probability of first order that j $(j = 1, 2, ...)$ customers (units) arrives at the system during a short interval of time $(t,\ t + dt]$. Further, $\sum_{j=1}^{\infty} c_j = 1$, $0 \leq c_j \leq 1$ for all j, where $\lambda > 0$ is the mean arrival rate of batches.
- The first-come, first-served (FCFS) discipline of service is followed.
- The service time for FES and SOS are assumed to follow general arbitrary distribution with distribution functions $F(x)$ and $H(x)$ and the density functions are $f(x)$ and $h(x)$,

respectively. Let $\mu(x)dx, \beta(x)dx$ be the conditional probabilities of the completion of FES and SOS, respectively during the interval $(x, x+dx]$ with elapsed service time x, so that

$$\mu(x) = \frac{f(x)}{1-F(x)} \quad \text{and} \quad f(s) = \mu(s)e^{-\int_0^s \mu(x)dx},$$
$$\beta(x) = \frac{h(x)}{1-H(x)} \quad \text{and} \quad h(v) = \beta(v)e^{-\int_0^v \beta(x)dx}.$$

- When a customer arrives, he/she joins the line with probability b or refuses to join the line (balking) with probability $1-b$.
- We assume that customers may leave the system after joining the queue without getting any service (renege) during FES and SOS and the reneging times is assume to follow exponential distribution with parameter α.
- After completion of FES, a customer may join the SOS with probability r_0 or depart from the system with probability r_1 or rejoin the system (feedback) if not satisfied with FES with probability r_2 where $r_0 + r_1 + r_2 = 1$.
- All various stochastic processes included in the system are mutually independent .

Formulation of Mathematical Model

The state of the system at time t is defined by the Markov process as

$$\{(L_q(t), M(t), \varepsilon_i(t)); i = 1, 2,\ t \geq 0\},$$

where $L_q(t)$ is the queue length at time t, $M(t)$ be the state of the server at time t which is given by

$$M(t) = \begin{cases} 0, & \text{the server is idle and the queue is empty at time t,} \\ 1, & \text{the server is operating FES at time } t, \\ 2, & \text{the server is operating SOS at time } t. \end{cases}$$

and $\varepsilon_i(t)$ is the elapsed service time of a batch in service ($i = 1$ for FES and $i = 2$ for SOS) at time t. The state space of the Markov process is given as follows:

$$\Omega = \{\{0,0\} U \{n, i, \varepsilon_1\} U \{n, i, \varepsilon_2\}; n \geq 0, i = 1, 2.\}$$

The probabilities involved in this model are defined as

- $Q(t)$ is the probability that the system is empty and the server is in idle.
- $P_{n,i}(x,t)$ is the probability of n $(n \geq 0)$ units in the queue, with one unit in the service, elapses service time is x and the server is providing FES for $i = 1$ and SOS for $i = 2$.

According to the description that is given in the previous section, the differential-difference

equations are formulated as follows:

$$\frac{d}{dt}Q(t) + \lambda Q(t) = r_1 \int_0^\infty P_{0,1}(x,t)\mu(x)dx + \int_0^\infty P_{0,2}(x,t)\beta(x)dx, \tag{1}$$

$$\frac{\partial}{\partial x}P_{0,1}(x,t) + \frac{\partial}{\partial t}P_{0,1}(x,t) = -(\lambda b + \mu(x))P_{0,1}(x,t) + \alpha P_{1,1}(x,t), \tag{2}$$

$$\frac{\partial}{\partial x}P_{n,1}(x,t) + \frac{\partial}{\partial t}P_{n,1}(x,t) = -(\lambda b + \mu(x) + \alpha)P_{n,1}(x,t) + \lambda b \sum_{i=1}^{n} c_i P_{n-i,1}(x,t) + \alpha P_{n+1,1}(x,t), \quad n \geq 1, \tag{3}$$

$$\frac{\partial}{\partial x}P_{0,2}(x,t) + \frac{\partial}{\partial t}P_{0,2}(x,t) = -(\lambda b + \beta(x))P_{0,2}(x,t) + \alpha P_{1,2}(x,t), \tag{4}$$

$$\frac{\partial}{\partial x}P_{n,2}(x,t) + \frac{\partial}{\partial t}P_{n,2}(x,t) = -(\lambda b + \beta(x) + \alpha)P_{n,2}(x,t) + \lambda b \sum_{i=1}^{n} c_i P_{n-i,2}(x,t) + \alpha P_{n+1,2}(x,t), \ n \geq 1. \tag{5}$$

Equations (1)-(5) must be solved at $x = 0$ with the following boundary conditions

$$P_{n,1}(0,t) = \lambda c_{n+1} Q(t) + r_1 \int_0^\infty P_{n+1,1}(x,t)\mu(x)dx + r_2 \int_0^\infty P_{n,1}(x,t)\mu(x)dx + \int_0^\infty P_{n+1,2}(x,t)\beta(x)dx, \ n \geq 0, \tag{6}$$

$$P_{n,2}(0,t) = r_0 \int_0^\infty P_{n,1}(x,t)\mu(x)dx, \ n \geq 0. \tag{7}$$

. At steady state, i.e, as $t \to \infty$, the above probabilities are denoted by Q, $P_{n,i}(x)$ and their derivatives with respect to time t vanish.

III. Steady State Solution of the Model

Considering the model in steady state, the state equations (1) - (7) are given as follows:

$$\lambda Q = r_1 \int_0^\infty P_{0,1}(x)\mu(x)dx + \int_0^\infty P_{0,2}(x)\beta(x)dx, \tag{8}$$

$$\frac{\partial}{\partial x}P_{0,1}(x) + (\lambda b + \mu(x))P_{0,1}(x) = \alpha P_{1,1}(x), \tag{9}$$

$$\frac{\partial}{\partial x}P_{n,1}(x) + (\lambda b + \mu(x) + \alpha)P_{n,1}(x) = \lambda b \sum_{i=1}^{n} c_i P_{n-i,1}(x) + \alpha P_{n+1,1}(x), \quad n \geq 1, \tag{10}$$

$$\frac{\partial}{\partial x}P_{0,2}(x) + (\lambda b + \beta(x))P_{0,2}(x) = \alpha P_{1,2}(x), \tag{11}$$

$$\frac{\partial}{\partial x}P_{n,2}(x) + (\lambda b + \beta(x) + \alpha)P_{n,2}(x) = \lambda b \sum_{i=1}^{n} c_i P_{n-i,2}(x) + \alpha P_{n+1,2}(x) \ n \geq 1. \tag{12}$$

The boundary conditions are given by

$$P_{n,1}(0) = \lambda c_{n+1} Q + r_1 \int_0^\infty P_{n+1,1}(x)\mu(x)dx + r_2 \int_0^\infty P_{n,1}(x)\mu(x)dx + \int_0^\infty P_{n+1,2}(x)\beta(x)dx, \ n \geq 0, \tag{13}$$

$$P_{n,2}(0) = r_0 \int_0^\infty P_{n,1}(x)\mu(x)dx, \ n \geq 0. \tag{14}$$

Generating Functions of the Queue Length

The main purpose of this subsection is to solve the equations (8) - (14) using bi-variate probability generating functions (PGFs). The PGFs are defined as follows:

$$P_i(x,z) = \sum_{n=0}^{\infty} P_{n,i}(x)z^n, \quad |z| \le 1, \ x > 0, \ i = 1,2. \tag{15}$$

$$P_i(0,z) = \sum_{n=0}^{\infty} P_{n,i}(0)z^n, \quad |z| \le 1, \ i = 1,2. \tag{16}$$

$$C(z) = \sum_{j=1}^{\infty} c_j z^j, \quad |z| \le 1. \tag{17}$$

lemma 1. *For $x > 0$ we have*

$$(I) \ \frac{\partial}{\partial x}P_1(x,z) + \big(\lambda b(1-C(z)) + \mu(x) + \alpha - \frac{\alpha}{z}\big)P_1(x,z) = 0, \tag{18}$$

$$(II) \ \frac{\partial}{\partial x}P_2(x,z) + \big(\lambda b(1-C(z)) + \beta(x) + \alpha - \frac{\alpha}{z}\big)P_2(x,z) = 0. \tag{19}$$

Proof. (I) Multiplying equations (9) and (10) by appropriate power z^n, summing them from $n = 0$ to $n = \infty$, and using the definition of PGFs, we get the result.
(II) Similarly, from equations (11) and (12), we get the desired result. □

lemma 2. *For $x > 0$, we have*

$$(I) \ P_1(x,z) = P_1(0,z)e^{-[\eta(z)]x - \int_0^x \mu(t)dt}, \tag{20}$$

$$(II) \ P_2(x,z) = P_2(0,z)e^{-[\eta(z)]x - \int_0^x \beta(t)dt}, \tag{21}$$

where $\eta(z) = \lambda b(1-C(z)) + \alpha - \frac{\alpha}{z}$.

Proof. Integrating equations (18) and (19) in the interval $[0,x]$, we get the desired result. □

lemma 3. *For $x > 0$, we have*

$$(I) \ \int_0^{\infty} P_1(x,z)\mu(x)dx = P_1(0,z)F^*(\eta(z)). \tag{22}$$

$$(II) \ \int_0^{\infty} P_2(x,z)\beta(x)dx = P_2(0,z)H^*(\eta(z)). \tag{23}$$

where $F^*[\eta(z)]$, $H^*[\eta(z)]$ are the Laplace-Steiltjes transform (LST) of the service times $F(x)$ and $H(x)$, respectively.

$$F^*[\eta(z)] = \int_0^{\infty} e^{-(\eta(z))x}dF(x),$$
$$H^*[\eta(z)] = \int_0^{\infty} e^{-(\eta(z))x}dH(x).$$

Proof.
Multiplying equations (20) and (21) by $\mu(x)$ and $\beta(x)$, respectively and integrating with respect to x, we get the result. □

lemma 4. *The PGFs $P_i(z)$, $i = 1,2$ are given by*

$$(I) \ P_1(z) = \frac{\lambda(C(z)-1)\big[1-F^*(\eta)\big]Q}{\big[z - r_1F^*(\eta) - r_2zF^*(\eta) - r_0F^*(\eta)H^*(\eta)\big]\eta(z)}, \tag{24}$$

$$(II) \ P_2(z) = \frac{r_0\lambda(C(z)-1)F(\eta(z))\big[1-H^*(\eta(z))\big]Q}{\big[z - r_1F^*(\eta(z)) - r_2zF^*(\eta(z)) - r_0F^*(\eta(z))H^*(\eta(z))\big]\eta(z)}. \tag{25}$$

where $P_i(z) = \int_0^{\infty} P_i(x,z)dx, \ \ i = 1,2.$

Integrating equations (20) and (21) by parts, we get

$$P_1(z) = P_1(0,z)\left(\frac{1-F^*(\eta(z))}{\eta(z)}\right), \tag{26}$$

$$P_2(z) = P_2(0,z)\left(\frac{1-H^*(\eta(z))}{\eta(z)}\right). \tag{27}$$

Now, we have to find $P_1(0,z), P_2(0,z)$.
Multiplying equation (13) by appropriate powers of z^n, summing them from $n=0$ to ∞, and using the definition of PGFs, we get

$$zP_1(0,z) = \lambda C(z)Q + r_1\int_0^\infty P_1(x,z)\mu(x)dx + zr_2\int_0^\infty P_1(x,z)\mu(x)dx + \int_0^\infty P_2(x,z)\beta(x)dx - \left[r_1\int_0^\infty P_{0,1}(x)\mu(x)dx + \int_0^\infty P_{0,2}(x)\beta(x)dx\right] \tag{28}$$

Substituting equation (8) into equation (28), we get

$$zP_1(0,z) = \lambda C(z)Q + r_1\int_0^\infty P_1(x,z)\mu(x)dx + r_2z\int_0^\infty P_1(x,z)\mu(x)dx + \int_0^\infty P_2(x,z)\beta(x)dx - \lambda Q. \tag{29}$$

Substituting equations (22) and (23) in equation (29), we get

$$zP_1(0,z) = \lambda(C(z)-1)Q + r_1F^*(\eta(z))P_1(0,z) + r_2zF^*(\eta(z))P_1(0,z) + P_2(0,z)H^*(\eta(z)), \tag{30}$$

Similarly, multiplying equation (14) by appropriate powers of z^n, summing them from $n=0$ to ∞, and using the definition of PGFs, we get

$$P_2(0,z) = r_0\int_0^\infty P_1(x,z)\mu(x)dx. \tag{31}$$

Substituting equation (22) in equation (31), we obtain

$$P_2(0,z) = r_0F^*(\eta(z))P_1(0,z). \tag{32}$$

Substituting equation (32) in equation (30), we get

$$zP_1(0,z) = \lambda(C(z)-1)Q + r_1F^*(\eta(z))P_1(0,z) + r_2zF^*(\eta(z))P_1(0,z) + r_0F^*(\eta(z))H^*(\eta(z))P_1(0,z). \tag{33}$$

After algebraic calculations, we get

$$P_1(0,z) = \frac{\lambda(C(z)-1)Q}{z - r_1F(\eta(z)) - r_2zF(\eta(z)) - r_0F(\eta(z))H(\eta(z))}. \tag{34}$$

Substituting equation (34) in equation (32), we get

$$P_2(0,z) = \frac{r_0\lambda(C(z)-1)F(\eta(z))Q}{z - r_1F(\eta(z)) - r_2zF(\eta(z)) - r_0F(\eta(z))H(\eta(z))}. \tag{35}$$

After substituting equations (34) and (35) in equations (26) and (27) respectively, and some algebraic calculations, the equations (24) and (25) are obtained.

lemma 5. *The PGF of the queue size is given by*

$$P_q(z) = \frac{\left[\lambda(C(z)-1)Q\right]\left[1-F^*(\eta(z)) + r_0F^*(\eta(z)) - r_0F^*(\eta(z))H^*(\eta(z))\right]}{\left[\lambda b(1-C(z)) + \alpha - \frac{\alpha}{z}\right]\left[z - r_1F^*(\eta(z)) - r_2zF^*(\eta(z)) - r_0F^*(\eta(z))H^*(\eta(z))\right]} \tag{36}$$

Proof. Let us suppose the PGF of the queue size irrespective of the state of the system be given by

$$P_q(z) = P_1(z) + P_2(z) \tag{37}$$

Substituting equations (24) and (25) in equation (37), we get the result. □

lemma 6. *Based on the previous results, we have*

$$Q = \frac{(-\lambda bE(X)+\alpha)\big[1-r_2+(-\lambda bE(X)+\alpha)[E(S)+r_0E(V)]\big]}{-[-\lambda E(X)(1-b)-\alpha](-\lambda bE(X)+\alpha)[E(S)+r_0E(V)]+(-\lambda bE(X)+\alpha)\big[1-r_2\big]}, \tag{38}$$

Proof.
To obtain Q, we have to use the normalizing condition

$$P_q(1) + Q = 1. \tag{39}$$

Now, clearly $z = 1$ brings P_q in equation (39) to indeterminate $(\frac{0}{0})$ form. Therefore using L'Hospital's rule, we obtain

$$P_q(1) = \lim_{z\to 1} P_q(z) = \frac{\lambda C'(1)(-\lambda bC'(1)+\alpha)\big[F^{*\prime}(0)+r_0H^{*\prime}(0)\big]Q}{(-\lambda bC'(1)+\alpha)\big[1-r_2+(-\lambda bC'(1)+\alpha)F^{*\prime}(0)+r_0[(-\lambda bC'(1)+\alpha)H^{*\prime}(0)\big]}. \tag{40}$$

Substituting $C(1) = 1$, $C'(1) = E(X)$, $F^*(0) = 1$, $F^{*\prime}(0) = -E(S)$, $H^*(0) = 1$, $H^{*\prime}(0) = -E(V)$ in (39), we get

$$P_q(1) = \frac{-\lambda E(X)(-\lambda bE(x)+\alpha)\big[E(S)+r_0E(V)\big]Q}{(-\lambda bE(X)+\alpha)\big[1-r_2-(-\lambda bE(X)+\alpha)[E(S)+r_0E(V)]\big]}. \tag{41}$$

where $E(S)$ and $E(V)$ are the mean service times for FES and SOS, respectively. $E(X)$ is the mean batch size of the arriving units.
Substituting the equation (41) in (39), the equation (38) is derived. □

IV. Performance Measures

In this section, using the PGF of the queue size distribution that we obtained in previous section, we get the mean queue size and the waiting time of a customer in the queue. Let L_q be the mean queue size which is define as following

$$L_q = \lim_{z\to 1} \frac{d}{dz} P_q(z), \tag{42}$$

where $P_q(z)$ denote the PGF of the queue size. Taking the limit of derivative of $P_q(z)$ at $z = 1$ brings equation (41) to indeterminate $(\frac{0}{0})$ form. Then using L'Hospital's rule and carrying out the derivatives at $z = 1$, we obtain

$$L_q = \frac{M''(1)N'''(1) - N''(1)M'''(1)}{3(M''(1))^2}. \tag{43}$$

Let us derive the second the third derivatives at z=1 with some algebra calculations, we get

$$
\begin{aligned}
N(z) &= \big[\lambda(C(z)-1)Q\big]\big[1-F^*(\eta(z))+r_0F^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))\big],\\
N'(1) &= 0\\
N''(1) &= -2\lambda E(X)(-\lambda bE(X)+\alpha)[E(S)+r_0E(V)]Q,\\
N'''(1) &= -3\lambda E(X(X-1))(-\lambda bE(X)+\alpha)\big[E(S)+r_0E(V)\big]Q\\
&\quad -3\lambda E(X)\Big[(-\lambda bE(X(X-1))+2\alpha)E(S)+2(-\lambda bE(X)+\alpha)^2E(S^2)\\
&\quad +2r_0(-\lambda bE(X)+\alpha)^2(E(S))(E(V))\\
&\quad +r_0\big[-(\lambda bE(X(X-1))+2\alpha)E(V)+2(-\lambda bE(X)+\alpha)^2E(V^2)\big]\Big]Q,\\
M(z) &= \big[\lambda b(1-C(z))+\alpha-\frac{\alpha}{z}\big]\big[z-r_1F^*(\eta(z))-r_2zF^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))\big],\\
M'(1) &= 0,\\
M''(1) &= 2[-\lambda bE(X)+\alpha][1-r_2-(-\lambda bE(X)+\alpha)(E(S)+r_0E(V))],\\
M'''(1) &= -3[\lambda bE(X(X-1))+2\alpha][1-r_2-(-\lambda bE(X)+\alpha)[E(S)+r_0E(V)]]\\
&\quad -3\Big(-\lambda bE(X)+\alpha\Big)\Big[-(\lambda bE(X(X-1))+2\alpha)E(S)+2(-\lambda bE(X)+\alpha)^2E(S^2)]\\
&\quad +2r_2(-\lambda bE(X)+\alpha)(E(S))+2r_0(-\lambda bE(X)+\alpha)^2(E(S)E(V))\\
&\quad +r_0\big[-(\lambda bE(X(X-1))+2\alpha)E(V)+2(-\lambda bE(X)+\alpha)^2E(V^2)\big]\Big].
\end{aligned}
$$

where $K''(1)=E(X(X-1))$ is the second factorial moment of the batch size of the arriving units, $E(S^2)$ and $E(V^2)$ are the second moment of the service time for FES and SOS, respectively. Now substituting N'', N''', M'', M''' in (43) we obtain L_q in closed form
Let W_q is the mean of waiting time of a customer in the queue. Using Little's formula we have

$$W_q=\frac{L_q}{\lambda bE(X)}. \tag{44}$$

V. Particular Cases

In this Section, we derive some particular cases from the main results obtained in this paper.
Case 1:

(1) We assume that the service time (FES and SOS) are following exponential distribution. Here, we take

$$
\begin{aligned}
E(S)&=\frac{1}{\mu} \quad , \quad E(S^2)=\frac{2}{(\mu)^2}\\
E(V)&=\frac{1}{\beta} \quad , \quad E(V^2)=\frac{2}{(\beta)^2}
\end{aligned}
$$

(2) We assume that the service time (FES and SOS) are following hyper-exponential distribution. Here, we take

$$
\begin{aligned}
E(S)&=\frac{p}{\mu_1}+\frac{1-p}{\mu_2}, \quad E(S^2)=2\Big(\frac{p}{(\mu_1)^2}+\frac{1-p}{(\mu_2)^2}\Big)\\
E(V)&=\frac{p}{\beta_1}+\frac{1-p}{\beta_2}, \quad E(V^2)=2\Big(\frac{p}{(\beta_1)^2}+\frac{1-p}{(\beta_2)^2}\Big)
\end{aligned}
$$

(3) We assume that the service time (FES and SOS) are following Erlang-k distribution. Here, we take

$$E(S) = \frac{1}{\mu} \quad , \quad E(S^2) = \frac{k+1}{k(\mu)^2}$$
$$E(V) = \frac{1}{\beta} \quad , \quad E(V^2) = \frac{k+1}{k(\beta)^2}$$

Case 2: we assume the costumer may not renege during FES or SOS i.e ($\alpha = 0$), the model reduces to $M^X/G/1$ queueing system with balking, feedback and SOS.
Using this assumption in the main result of the paper, we get

$$P_q(z) = \frac{(-Q)\big[1 - F^*(\eta(z)) + r_0F^*(\eta(z)) - r_0F^*(\eta(z))H^*(\eta(z))\big]}{b\big[z - r_1F^*(\eta(z)) - r_2zF^*(\eta(z)) - r_0F^*(\eta(z))H^*(\eta(z))\big]}.$$
$$Q = \frac{b\big[1 - r_2 - \lambda bE(X)E(S) - r_0\lambda bE(X)E(V)\big]}{(1-b)(\lambda bE(X))\big[E(S) + r_0E(V)\big] + b(1-r_2)},$$
$$L_q = \lim_{z\to 1}\frac{d}{dz}P_q(z) = \frac{M'(1)N''(1) - N'(1)M''(1)}{2(M'(1))^2},$$

where N', N'', M', M'' are given in the flowing equations:

$$N'(1) = (Q)(-\lambda bE(X) + \alpha)\big[E(S) + r_0E(V)\big],$$
$$\begin{aligned} N''(1) = (Q)\Big[&(-\lambda bE(X(X-1)) + 2\alpha)E(S) + 2(-\lambda bE(X) + \alpha)^2E(S^2) \\ &+ 2r_0(-\lambda bE(X) + \alpha)^2(E(S))(E(V)) \\ &+ r_0\big[-(\lambda bE(X(X-1)) + 2\alpha)E(V) + 2(-\lambda bE(X) + \alpha)^2E(V^2)\big]\Big], \end{aligned}$$
$$M'(1) = b[1 - r_2 - (-\lambda bE(X) + \alpha)E(S) - r_0(-\lambda bE(X) + \alpha)E(V)],$$
$$\begin{aligned} M''(1) = -b\Big[&2r_2(-\lambda bE(X) + \alpha)E(S)(-\lambda bE(X(X-1)) + 2\alpha)E(S) + 2(-\lambda bE(X) + \alpha)^2E(S^2) \\ &+ 2r_0(-\lambda bE(X) + \alpha)^2(E(S))(E(V)) \\ &+ r_0\big[-(\lambda bE(X(X-1)) + 2\alpha)E(V) + 2(-\lambda bE(X) + \alpha)^2E(V^2)\big]\Big]. \end{aligned}$$

Case 3: Consider $r_0 = 0$ (no SOS), $b = 1$ (no balking), $\alpha = 1$ (no reneging) a feedback model in $M^X/G/1$ queue is obtained.

$$Q = \frac{1 - r_2 - \lambda E(X)E(S)}{1 - r_2},$$
$$L_q = \lim_{z\to 1}\frac{d}{dz}P_q(z) = \frac{M'(1)N''(1) - N'(1)M''(1)}{2(M''(1))^2},$$

where N', N'', M', M'' is given in the flowing equations:

$$N'(1) = -[(-\lambda E(X) + \alpha)E(S)]Q$$
$$N'''(1) = -[(\lambda E(X(X-1)) + 2\alpha)E(S) + 2(-\lambda E(X) + \alpha)^2E(S^2)]Q$$
$$M'(1) = [1 - r_2 - (-\lambda E(X) + \alpha)E(S)],$$
$$M''(1) = \big[-((\lambda E(X(X-1)) + 2\alpha)E(S) + 2(-\lambda E(X) + \alpha)^2E(S^2)) - 2r_2(-\lambda E(X) + \alpha)E(S)\big].$$

We note that this result agrees as special case with the result of $M^X/G/1$ queue with feedback and optional server vacations (see [4])

VI. Numerical Results and Discussion

In this section, Some numerical illustrations with discussion based on Q, L_q and W_q are provided with the purpose to illustrate the effect of the parameters (λ, μ, β, b, r_0, r_1, r_2) on Q, L_q and W_q.

In Table 1, we show the impact of the probability of feedback (r_2) and the probability of join SOS (r_0) on the L_q. For the fixed probability of the departure (r_1), as r_2 increases and (r_0) decreases, the situation leads to an increase in L_q. This indicating that more customers feel unsatisfied and decide to rejoin the queue. We take; the service time (FES and SOS) follow Exponential distribution and $\lambda = 2,\ \mu = 5,\ \beta = 4,\ \alpha = 1, r_0 = 0.1,\ r_2 = 0.5,\ b = 0.10,\ E(X) = 1,\ E(X(X-1)) = 0$.
Also, we show in (Table 2) the impact of the mean arrival rate of batches λ and mean of reneging α on the (L_q). We observe that L_q decreases as mean reneging α increases .Thus more customers leave the the queue. For the fixed mean reneging (α), as λ increases L_q increases. We take; the service times (FES and SOS) to follow exponential distribution and $r_0 = 0.6,\ r_2 = 0.2,\ \mu = 4,\ \beta = 3, b = 0.20, E(X) = 1, E(X(X-1)) = 0$.
We show in (Table 3) the effect of batch arrival rate λ on Q and L_q when the service times (FES and SOS) are following general distribution (exponential , Erlang-κ , hyper -exponential). We observe that server's idle time Q decreases and the L_q increases as batch arrival rate λ increases. Here, when the service times (FES and SOS) to follow exponential distribution we take; $\mu = 5,\ \beta = 3,\ r_0 = 0.5,\ r_2 = 0.3,\ \alpha = 1,\ b = 0.25,\ E(X) = 1,\ E(X(X-1)) = 0$ and when they follow Erlang-κ we take $\kappa = 5,\ \mu = 5,\ \beta = 3,\ r_0 = 0.5,\ r_2 = 0.3,\ \alpha = 1,\ b = 0.25,\ E(X) = 1,\ E(X(X-1)) = 0$, and when they follow hyper-exponential $p = 0.5,\ \mu_1 = 5,\ \mu_2 = 4,\ \beta_1 = 3,\ \beta_2 = 2,,\ r_0 = 0.5,\ r_2 = 0.3,\ \alpha = 1,\ b = 0.10,\ E(X) = 1,\ E(X(X-1)) = 0$.
In Figure 1, we show the effect of batch arrival rate λ on L_q in different joining probability b. We observe that L_q increases as λ or b increases. We take; the service times (FES and SOS) to follow exponential distribution and $r_0 = 0.5,\ r_2 = 0.3,\ \mu = 5,\ \beta = 4,\ \alpha = 1,\ E(X) = 1,\ E(X(X-1) = 0$.
Also in figures 2, and 3, we show the effect of the service rate (FES and SOS) on L_q in different joining probability b. We observe that L_q decreases when the FES rate and SOS rate increase as we expected. Further, we notice that as b increases, the L_q increases i.e. additional customers joining the queue.
We take; the service times (FES and SOS) to follow exponential distribution and $\beta = 4, \alpha = 1,\ r_0 = 0.5,\ r_2 = 0.3,\ b = 0.10,\ E(X) = 1,\ E(X(X-1) = 0$, in Figure 2 and $\mu = 5, \alpha = 1,\ r_0 = 0.5,\ r_2 = 0.3,\ b = 0.10,\ E(X) = 1,\ E(X(X-1) = 0$, in Figure 3

Table 1: *The impact of r_0 and r_2 on Q, L_q and W_q.*

r_2	r_0	Q	ρ	Lq	Wq
0.1	0.5	0.640884	0.359116	0.0296485	0.148243
0.2	0.4	0.634146	0.365854	0.0304878	0.152439
0.3	0.3	0.625850	0.374150	0.0312408	0.156204
0.4	0.2	0.615385	0.384615	0.0317308	0.158654
0.5	0.1	0.601770	0.398230	0.0315591	0.157795

Table 2: *Impact of λ and α on Q, L_q and W_q.*

λ	α	Q	ρ	L_q	W_q
$\lambda = 1.0$	$\alpha = 1$	0.720497	0.279503	0.0373921	0.186960
	$\alpha = 2$	0.781553	0.218447	0.0321112	0.160556
	$\alpha = 3$	0.820717	0.179283	0.0276886	0.138443
$\lambda = 1.5$	$\alpha = 1$	0.543147	0.456853	0.0977276	0.244319
	$\alpha = 2$	0.628099	0.371901	0.0759388	0.189847
	$\alpha = 3$	0.686411	0.313589	0.0623758	0.155940
$\lambda = 2.0$	$\alpha = 1$	0.420601	0.579399	0.1789000	0.298166
	$\alpha = 2$	0.514388	0.485612	0.1304270	0.217379
	$\alpha = 3$	0.582043	0.417957	0.1035010	0.172502

Table 3: *The impact of batch arrival rate λ on Q and L_q in General distribution service time and repair time.*

	exponential		*Erlang* $-\kappa$		*hyper* $-$ *exponential*	
λ	Q	Lq	Q	Lq	Q	Lq
1.0	0.726708	0.0384350	0.726708	0.0537750	0.702857	0.0447747
1.5	0.628169	0.0701825	0.628169	0.0884330	0.598972	0.0830000
2.0	0.546392	0.1097060	0.546392	0.128442	0.514019	0.1311680
2.5	0.477435	0.1571710	0.477435	0.174245	0.443255	0.1897080
3.0	0.418502	0.2131230	0.418502	0.226523	0.383399	0.2596380

Figure 1: *The effect of batch arrival rate (λ) on (L_q) in different joining probability b*

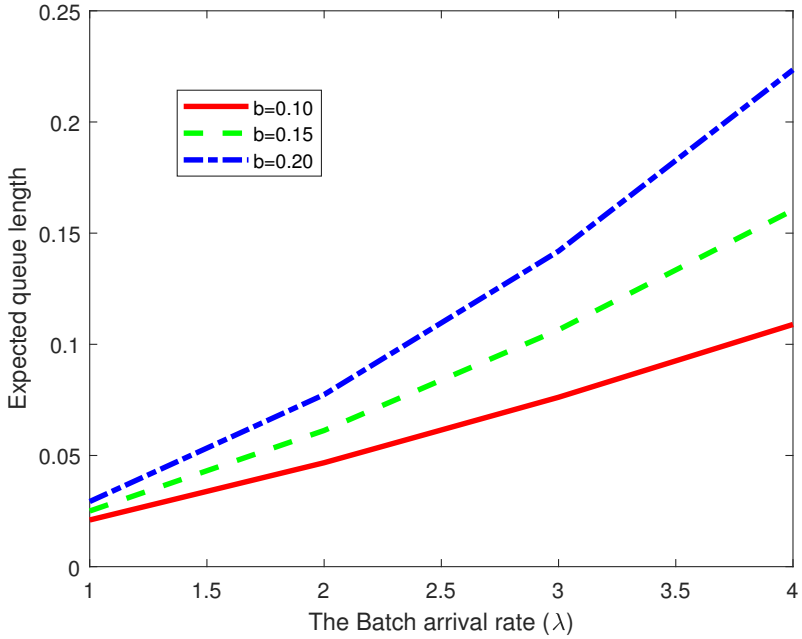


Figure 2: *The effect of the FES rate (μ) on (L_q) in different joining probability b*

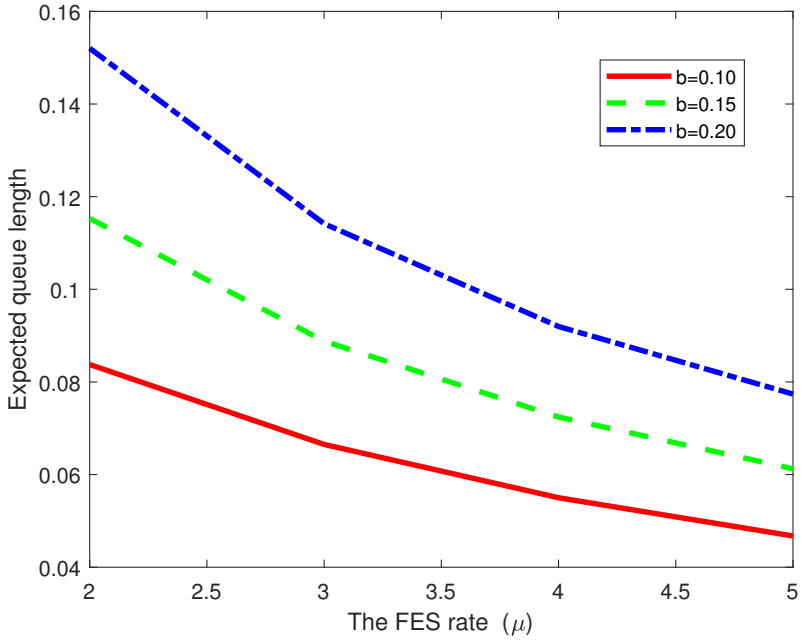
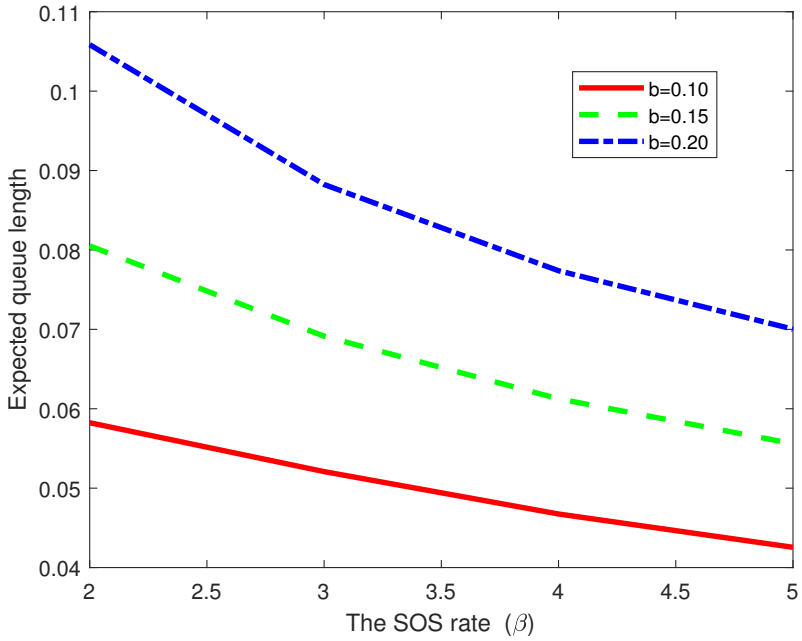


Figure 3: *The effect of SOS rate (β) on (L_q) in different joining probability b*



The Cost Model

To achieve the optimal service rate in FES and SOS with a minimum expected cost function, we have developed the expected cost function per unit time as :

$$f(\mu, \beta) = CL + C_1\mu + C_2\beta + C_r\alpha, \tag{45}$$

where :

- C = cost per unite time per customer present in the queue.
- C_1 = cost per unite time during FES.
- C_2 = cost per unite time during SOS.
- C_r = cost per unite time when the customer renege.

The cost minimization problem $f(\mu, \beta)$ can be presented mathematically as

$$f(\mu^*, \beta^*) = \underset{s.t\mu,\beta>0}{\text{Minimize}} f(\mu, \beta). \tag{46}$$

We use the Quasi- Newton method to search for (μ, β) until the minimum of $f(\mu, \beta)$ is obtained. For details of Quasi- Newton method, one may refer Lewis and Overton [27].

Table 4: *Impact of r_0 and r_2 on the expected cost*

r_0	r_2	μ^*	β^*	$f(\mu^*,\ \beta^*)$
	r_2 = 0.20	1.41917	0.917929	51.2880
r_0= 0.2	r_2 = 0.40	1.74319	1.05108	60.1129
	r_2 = 0.60	2.34484	1.29993	76.0659
	r_2 = 0.20	1.44822	1.07991	54.1650
r_0= 0.2	r_2 = 0.40	1.78370	1.24700	63.5951
	r_2 = 0.60	2.40812	1.56344	80.7229
	r_2 = 0.20	1.47416	1.21697	56.5886
r_0= 0.3	r_2 = 0.40	1.81941	1.41499	66.5526
	r_2 = 0.60	2.46299	1.79318	84.7184

From Table 4, we notice that for fixed r_0, $(\mu^*,\ \beta^*)$ and $f(\mu^*,\ \beta^*)$ increase with the increase of r_2. This is because many customers have not satisfied with the service and repeat the service, leading to high-cost implications.
Similarly, for fixed r_2, as r_0 increases, we observe that both $(\mu^*,\ \beta^*)$ and $f(\mu^*,\ \beta^*)$ increase . This is due to the fact that as r_0 increases, customers tend to enter SOS service, thereby increasing the service rate, which in turn results in an increase of cost. We take the service times (FES and SOS) to follow exponential distribution and $\lambda = 2,\ \mu = 2,\ \beta = 1,\ \alpha = 0.1,\ b = 0.2,\ E(X) = 1,\ E(X(X-1)) = 0$.

Table 5: *Impact of α and b on the expected cost*

α	b	μ^*	β^*	$f(\mu^*,\ \beta^*)$
	b = 0.20	1.47416	1.21697	56.5886
α= 0.10	b = 0.25	1.76159	1.45299	64.9474
	b = 0.30	2.03931	1.67967	72.8668
	b = 0.20	1.39146	1.14110	56.955
α= 0.15	b = 0.25	1.67757	1.37735	65.2815
	b = 0.30	1.95431	1.60419	73.1769
	b = 0.20	1.30875	1.06509	57.3181
α= 0.20	b = 0.25	1.59351	1.30166	65.6134
	b = 0.30	1.86928	1.52869	73.4854

Table 5 shows the impact of reneging rate α on the minimum expected cost function $f(\mu^*, \beta^*)$for different values of joining probability b. In this table, we observe that the optimal service rates (μ^*, β^*) and expected cost $f(\mu^*, \beta^*)$ increase as both α and b increase.
Particularly, For fixed b as α increases, customers departure from the queue which leads to decrease the service rates μ^* , β^* and increase cost, so that to balance the system profitability.We take; the service times (FES and SOS) to follow exponential distribution and ($\lambda = 2,\ \mu = 2,\ \beta = 1,\ r_0 = 0.4,\ r_2 = 02,\ E(X) = 1,\ E(X(X-1)) = 0.$)

VII. Conclusion

In this paper, we analyzed the steady state behavior of a single server batch arrival non -Markovian batch service queue with a second optional service, balking, reneging and feedback using the supplementary variable technique to get the probability generating function of the number of customers in the system. The mean of the queue size and waiting time of a customer in the queue were obtained. Some interesting special cases were discussed. We assumed general distribution for the service time. The cost model was presented to determine the optimal service rates to minimize the expected cost. Finally, the numerical results through graphical illustrations and tables were presented.

References

[1] Rajan, B. S., Ganesan, V. and Rita, S. (2020). Batch arrival poisson queue with breakdown and repairs. *International Journal of Mathematics in Operational Research*. 17(3), 424–435. https://doi.org/10.1504/IJMOR.2020.110033.

[2] P. Vijaya Laxmi, G. Andwilile Abrahamu and E. Girija Bhavani. (2021). Performance of a Single Server Batch Queueing Model with Second Optional Service under Transient and Steady State Domain.*Reliability: Theory & Applications*. 16 (4 (65)), 226-238. https://doi:10.24412/1932-2321-2021-465-226-238.

[3] Abdollahi, S., Salehi Rad and M. R. (2021). Analysis of a batch arrival retrial queue with Two-Phase Services, Feedback and admission. *Bulletin of the Iranian Mathematical Society*. 48(3), 791–804.

[4] Madan, K.C. and Al-Rawwash, M. (2005). On the $M^X/G/1$ queue with feedback and optional server vacations based on a single vacation policy.*Applied Mathematics and Computation*. 160(3), 909–919. https://doi.org/10.1016/j.amc.2003.11.037

[5] Maraghi, F. A., Madan, K.C. and Darby-Dowman, K. (2010). Batch Arrival Vacation Queue with Second Optional Service and Random System Breakdowns. *Journal of Statistical Theory and Practice*, 4(1), 137–153. https://doi.org/10.1080/15598608.2010.10411977.

[6] Bouchentouf, A.A. and Guendouzi. (2019). A. Cost optimization analysis for an$M^X/M/c$ vacation queueing system with waiting servers and impatient customers. *SeMA Journal*. 76, 309–341. https://doi.org/10.1007/s40324-018-0180-2.

[7] Khalaf, R., Madan, K.C. and Lucas, C. (2012). On an $M^X/G/1$ Queue with, Random Breakdowns, server vacations, delay Times and standby server. *International Journal of Operational Research* . 15(1), 30–46. https://doi.org/10.1504/IJOR.2012.048290.

[8] P. Vignesh, S. Srinivasan, and S. Maragatha Sundari (2019). Analysis of a non-Markovian single server batch arrival queueing system of compulsory three stages of services with fourth optional stage service. service interruptions and deterministic server vacations. *International Journal of Operational Research*, 4(1), 28-53. https://doi.org/10.1504/IJOR.2019.096937.

[9] Vignesh, P., Srinivasan, S., Sundari, S.M. and Eswar, S.K. (2023). An investigation on $M^X/G/1$ queuing model of interrupted services in the manufacturing of edible cutlery process. *International Journal of Mathematical Modelling and Numerical Optimisation*, 13(2), 173-201.

[10] Choudhury, A. (2008). newblockImpatience in Single Server Queueing Model. *American Journal of Mathematical and Management Sciences*. 28(1-2), 177–211. https://doi.org/10.1080/01966324.2008.10737723.

[11] Baruah, M., Madan, K.C. and Eldabi, T. (2013). A Two Stage Batch Arrival Queue with Reneging during Vacation and Breakdown Periods.*American Journal of Operations Research*.. 3(6), 570–580.

[12] Enogwe, Samuel, Onyeagu, Sidney and Obiora-Ilouno, Happiness. (2021). On single server batch arrival queueing system with balking, three types of heterogeneous service and Bernoulli schedule server vacation. *Mathematical Theory and Modeling*. 11(5), 40–68.

[13] Vijaya Laxmi, P., Goswami, V. and Jyothsna, K. (2013). Optimization of Balking and Reneging Queue with Vacation Interruption under N-Policy. *Journal of Optimization*, 2013, 1-9.

[14] Baruah, M., Madan, K.C. and Eldabi, T. (2013). A batch arrival queue with second optional service and reneging during vacation periods. *Investigacion Operacional*. 34(3), 244–258.

[15] Sivagnanasundararam M. and Santhanagopalan S., (2014). A Non-Markovian Multistage Batch Arrival Queue with Breakdown and Reneging.*Mathematical Problems in Engineering*. 2014. https://doi.org/10.1155/2014/519579.

[16] Singh, C.J., Jain, M. and Kumar, B. (2014). Analysis of $M^X/G/1$ queueing model with balking and vacation. *International Journal of Operational Research*. 19(2). 154–170. https://doi.org/10.1504/IJOR.2014.058952

[17] P., V. L., and George, A. A. (2020). Transient Analysis of Batch Service Queue with Second Optional Service and Reneging. *International Research Journal on Advanced Science Hub*, 2(10), 29-38. doi: 10.47392/irjash.2020.185.

[18] Cherfaoui, M., Bouchentouf, A.A. and Boualem, M. (2023). Modelling and simulation of Bernoulli feedback queue with general customers' impatience under variant vacation policy. *International Journal of Operational Research*, 46(4), 451-480.

[19] Saravanarajan, M.C. and Chandrasekaran, V.M. (2014). Analysis of $M^X/G/1$ feedback queue with two-phase service, compulsory server vacation and random breakdowns.*OPSEARCH*. 51, 235–256. https://doi.org/10.1007/s12597-013-0141-6.

[20] Shanmugasundaram, S. and Sivaram, G. (2020). M/G/1 feedback queue when server is off and on vacation.*International Journal of Applied Engineering Research.* 15(10), 1025—1028. DOI:10.37622/IJAER/15.10.2020.1025-1028

[21] P. Vijaya Laxmi, Qrewi, H. A. and George, A. A. (2022). Analysis of Markovian batch service queue with feedback and second optional service. *Reliability: Theory & Applications*. 17(2), 507–518. https://doi:10.24412/1932-2321-2022-268-507-518.

[22] Baruah, M., Madan, K.C. and Eldabi, T. (2012). Balking and Re-service in a Vacation Queue with Batch Arrival and Two Types of Heterogeneous Service.*Journal of Mathematics Research*. 4(4), 114–124.

[23] Arivudainambi, D. and Godhandaraman, P. (2012). A batch arrival retrial queue with two phases of service,feedback and K optional vacations. *Applied Mathematical Sciences (Ruse)*. 6(22), 1071-1087.

[24] Govindan, A. and Shyamala, s. (2016). Transient solution of an $M^X/G/1$ queueing model with feedback, random breakdowns, Bernoulli schedule server vacation and random setup time. *International Journal of Operational Research* . 25(2), 196–211. https://doi.org/10.1504/IJOR.2016.073956.

[25] Jain, M. and Singh, M. (2020). Transient Analysis of a Markov Queueing Model with Feedback, Discouragement and Disaster. *Int. J. Appl. Comput. Math*. 6(31), https://doi.org/10.1007/s40819-020-0777-x.

[26] Bouchentouf, A.A. and Guendouzi. (2021). A. Single Server Batch Arrival Bernoulli Feedback Queueing System with Waiting Server, K-Variant Vacations and Impatient Customers. *Operations Research Forum*, 2(1), https://doi.org/10.1007/s43069-021-00057-0.

[27] Lewis,A.S. and Overton, M.L. (2013). Nonsmooth optimazation via quasi-Newton methods. *Mathematical programming*.141, 135–163.

RELIABILITY MODELLING OF A PARALLEL-COLD STANDBY SYSTEM WITH REPAIR PRIORITY

Puran Rathi[1], Anuradha[2] & S.C. Malik[3]

rathipuran777@gmail.com[1], anudixit083@gmail.com[2] & sc_malik@rediffmail.com[3]
Department of Statistics, Maharshi Dayanand University Rohtak, India

Abstract

This paper deals with the reliability modelling of a parallel cold standby system of four units. The units operate in two phases; phase-I and phase-II. In phase-I, two identical units (called main units) work in parallel and the other two identical units (called duplicate units) have been taken as spare in cold standby. The units of phase-I and of the phase-II are not identical. The priority to repair the units of phase-I has been given over the repair of the units of the phase-II. However, no priority is given for operation of the units of both phases. There is a single repair facility which tackles all types of faults whenever occurred in the system. After repair each unit works as new and the switches devices are considered as perfect. The repair time of the units follows arbitrary probability distribution while the failure time of the units is assumed as constant. The behaviour of mean sojourn time (MST), transition probabilities, mean time to system failures (MTSF), availability, expected number of repairs for both phase-I and phase-II units, expected number of visits of the server, busy period of the server and finally the profit function are obtained in steady state by making use of well-known semi-Markov process (SMP) and Regenerative Point Technique (RPT) for arbitrary values of the parameters in steady state. Novelty and Application: A four-unit system is configured in two phases namely; phase-I and phase-II under some novel assumptions with a practical visualization in metallic bush manufacturing industries.

Keywords: Parallel-Cold Standby System, Phase wise Non-identical Units, Reliability Measures, Priority and Profit Analysis

1. Introduction

The advancement in technology played a great role in economic development and thus in the overall growth of a country. This plays a fundamental role in wealth creation, improvement of the quality of life, and real economic growth and transformation in any society. Technology has a great impact on everyone's life, including industries, which are dependent on machines for their chores. The great challenge for researchers and engineers is to produce highly reliable products at minimum cost. Thus, a basic need in the fast-growing industries is to select highly reliable systems subject to the cost. Many attempts from the researchers, engineers and industrialists have been made to improve the performance and designing of existing machines. Moreover, it is challenging for researchers and engineers to produce high quality products at minimum cost. Thus, reliability and profit analysis play a key role in defining quality of systems. Various techniques for improving performance and reliability of maintainable systems operating under different environmental conditions have been suggested by the researchers from time to time.

Barak and Malik [1] performed the cost benefit analysis of computer system with priority to

preventive maintenance using the concepts of maximum operation and repair times. Pundir et al. [2], [3] performed a Bayesian analysis by using some prior information for two non-identical cold standby system and stochastic analysis of two non-identical unit parallel system with priority in repair. Kadyan et al. [4,5] discussed a non-identical repairable system of three units with priority for operation and priority to main unit for operation and repair with the simultaneous working of cold-standby units. Using the concept of periodic switching approach, reliability modelling of two-unit cold standby system is performed by Behboudi et al. [6]. Fryilmaz and Finkelsteil [7] discussed the reliability of two-unit system with the revisit of standby system. Stochastic analysis of a computer system with redundant and priority to hardware and repair subject to failure of service facility have been discussed by Yadav and Malik [8]. Anuradha and Malik [10] have obgtained the reliability measures of a 2-out-3 systems under the conditioner service facility. A cold standby system subject to refreshment was studied by kumar at el. [11]. The models discussed by different researchers focus either on the identical units in parallel or one non- identical unit in spare. But there can be situations where one non-identical unit is not capable enough to work at place of failed unit and two or more units are required to work simultaneously in order to meet the system expectations. Therefore, the reliability analysis of a system model of four unit operating in two phases with simultaneous working of parallel and cold standby units has been analysed to add significant insight into reliability literature.

2. System description and Notations

I. Notations

Table 1: *Symbol Description*

MST	Mean sojourn time
MTSF	Mean time to system failure
O	System is operative
Dc	Cold-standby unit
•	Regenerative point
M/D	Phase-I unit (main units)/ Phase-II unit (duplicate units)
MFur/MFwr	Phase-I unit is failed and under repair/waiting for repair
MFUR/ MFWR	Phase-I unit is failed and under repair/waiting for repair continuously from the previous state
DFur/DFwr	Phase-II unit is failed and under repair//waiting for repair
DFUR/ DFWR	Phase-II unit is failed and under repair/waiting for repair continuously from the previous state
λ/λ_1	Failure rate of phase-I unit/ phase-II unit
g(t)/G(t)	pdf/cdf of the repair rate of the phase-I unit
f(t)/F(t)	pdf/cdf of the repair rate of the phase-II unit
$\overline{G(t)}/\overline{F(t)}$	cdf of repair rate of Phase I /phase II units that repair will not be completed in (0,t]
*/**	Symbol for Laplace transformation/Laplace-Stieltjes
$q_{ij}(t)/Q_{ij}(t)$	pdf/cdf of passage time from regenerative state 'i' to a regenerative state 'j' or to a failed state 'j' without visiting any other regenerative state in (0,t]
$q_{ij.k,r}(t)/Q_{ij.k,r}(t)$	pdf/cdf of direct transition time from regenerative state 'i' to a regenerative state 'j' or to a failed state 'j' visiting state k, r once in (0,t]
q_{ij.{k(r,s)}^n}/ $Q_{ij.k(r,s)^n}$	pdf/cdf of direct transition time from regenerative state 'i' to a regenerative state 'j' or to a failed state 'j' visiting state k once and n-times states r and s
Ⓢ/©	Symbol for Stieltjes convolution / Laplace Convolution.
LIT/ LT/LST	Laplace Inverse Transform/ Laplace Transform/Laplace Stieltjes Transform
$A_i(t)$	Probability that the system is in up-state at instant time 't'
$V_i^S(t)$	Expected number of visits of server
$R_i^M(t)/R_i^D(t)$	Expected number of repairs of phase-I /phase-II Units
$B_i^R(t)$	Busy period of server due to repair
$W_i(t)$	Probability that the server is busy in the state Si up to time 't' without making any transition to any other regenerative state or returning to the same state via one or more non-regenerative states

II. The state transition diagram of the system model

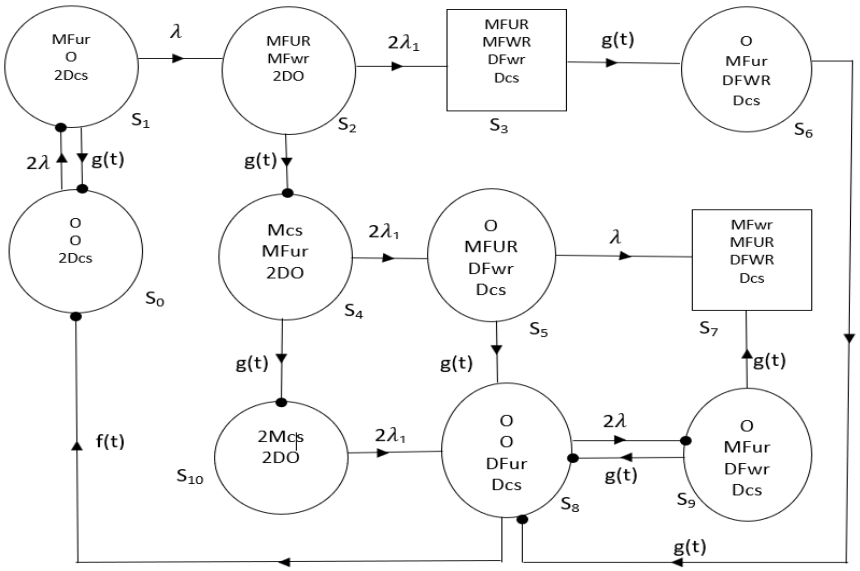


Figure 1: *State Transition Diagram*

3. State Transition Probabilities and Mean Sojourn Time

Table 2: *Transition State Description*

S_0	O O 2Dcs	The initial state in which the phase-I units are in operation and phase-II units are in cold standby
S_1	MFur O 2Dcs	The second state in which one of the phase-I units is in operation and other unit is failed under repair and phase-II units are in cold-standby
S_2	MFUR MFwr 2DO	The third state in which one of the phase-I units is continuously failed under repair and other is failed waiting for repair and the phase-II units are in operation mode
S_3	MFUR MFWR DFwr Dcs	The fourth state in which system is completely failed. One phase-I units are failed under repair and waiting for repair continuously from the previously state and phase-II unit is failed waiting for repair, other is in cold standby mode.
S_4	Mcs MFur 2DO	The fifth state in which one of the phase-I unit is in spare and other unit failed under repair and the phase-II units are in operation
S_5	O MFUR DFwr Dcs	The sixth state in which one phase-I unit is in operation, other is failed under repair continuously and in phase-II one unit is failed waiting for repair, other unit is cold standby mode
S_6	O MFur DFWR Dcs	The seventh state in which system is in operation. One of phase-I unit is operative, other is failed under repair and one of the phase-II unit is failed waiting for repair continuously, other unit is in cold standby mode
S_7	MFwr MFUR DFWR Dcs	The eighth state in which system is completely failed. One unit of phase-I is failed under repair continuously, other unit is failed waiting for repair and one of the phase-II unit is continuously failed under repair, other is in cold standby mode
S_8	O O DFur Dcs	The ninth state in which phase-I units are in operation and in phase-II, one unit is failed under repair and other is in cold-standby mode

S_9	O MFur DFwr Dcs	The tenth state in which one phase-I unit is in operation, other is failed under repair and one of phase-II unit is failed waiting for repair, other is in cold standby.
S_{10}	2Mcs 2DO	The eleventh state in which phase-I units are in spare and phase-II units are in operation

4. Reliability Measures

I. Transition Probabilities

The expressions for transition probabilities from state i to j as follow:

$p_{ij}=Q_{ij}(\infty)=\int_0^\infty q_{ij}(t)dt$ $p_{ij}=\lim_{n\to\infty} Q_{ij}(t)dt$

$dQ_{01}(t)=2\lambda e^{-2\lambda t}dt$ $dQ_{10}(t)=e^{-\lambda t}g(t)\,dt$ $dQ_{12}(t)=\lambda e^{-\lambda t}\overline{G(t)}dt$

$dQ_{23}(t)=2\lambda_1 e^{-2\lambda_1 t}\overline{G(t)}dt$ $dQ_{24}(t)=e^{-2\lambda_1 t}g(t)dt$ $dQ_{36}(t)=g(t)dt$

$dQ_{45}(t)=2\lambda_1 e^{-2\lambda_1 t}\overline{G(t)}dt$ $dQ_{4'10}(t)=e^{-2\lambda_1 t}g(t)dt$ $dQ_{58}(t)=e^{-\lambda t}g(t)\,dt$

$dQ_{57}(t)=\lambda e^{-\lambda t}\overline{G(t)}dt$ $dQ_{68}(t)=\lambda e^{-\lambda t}\overline{G(t)}dt$ $dQ_{67}(t)=e^{-\lambda t}g(t)\,dt$

$dQ_{76}(t)=g(t)dt$ $dQ_{80}(t)=e^{-2\lambda t}f(t)\,dt$ $dQ_{89}(t)=2\lambda e^{-2\lambda t}\overline{F(t)}dt$

$dQ_{97}(t)=\lambda e^{-\lambda t}\overline{G(t)}dt$ $dQ_{98}(t)=e^{-\lambda t}g(t)\,dt$ $dQ_{10,8}(t)=2\lambda_1 e^{-2\lambda_1 t}dt$

$dQ_{14.2}(t)=dQ_{12}(t)\textcircled{s}dQ_{24}$ $dQ_{18.2,3,6}(t)=dQ_{12}(t)\textcircled{s}dQ_{23}\textcircled{s}dQ_{36}\textcircled{s}dQ_{68}$

$dQ_{18.2,3(6,7)^n}(t)=\frac{dQ_{12}(t)\textcircled{s}dQ_{23}\textcircled{s}dQ_{36}\textcircled{s}dQ_{67}\textcircled{s}dQ_{76}\textcircled{s}dQ_{68}}{1-dQ_{67}(t)dQ_{76}(t)}$

$dQ_{48.5}(t)=dQ_{45}(t)\textcircled{s}dQ_{58}$ $dQ_{48.5,7,6}(t)=dQ_{45}(t)\textcircled{s}dQ_{57}\textcircled{s}dQ_{76}\textcircled{s}dQ_{68}$

$dQ_{48.5(6,7)^n}(t)=\frac{dQ_{45}(t)\textcircled{s}dQ_{57}\textcircled{s}dQ_{76}\textcircled{s}dQ_{67}\textcircled{s}dQ_{76}\textcircled{s}dQ_{68}}{1-dQ_{67}(t)dQ_{76}(t)}$

$dQ_{98.7,6}(t)=dQ_{97}(t)\textcircled{s}dQ_{76}\textcircled{s}dQ_{68}$

$dQ_{98.(6,7)^n}(t)=\frac{dQ_{97}(t)\textcircled{s}dQ_{76}\textcircled{s}dQ_{67}\textcircled{s}dQ_{76}\textcircled{s}dQ_{68}}{1-dQ_{67}(t)dQ_{76}(t)}$

$p_{ij}=\lim_{t\to\infty} Q_{ij}(t)dt=\lim_{s\to 0} Q_{ij}^{**}(s)$

$p_{01}=p_{36}=p_{76}=p_{76}=1$

$p_{10}=p_{58}=p_{68}=p_{98}=g^*(\lambda)$ $p_{12}=p_{57}=p_{67}=p_{97}=1-g^*(\lambda)$

$p_{24}=p_{4,10}=g^*(2\lambda_1)$ $p_{23}=p_{45}=1-g^*(2\lambda_1)$

$p_{80}=f^*(2\lambda)$ $p_{89}=1-f^*(2\lambda)$

$p_{36}=p_{76}=g^*(0)=1$ $p_{14.2}=p_{12}p_{24}$

$p_{18.2,3,6}=p_{12}p_{23}p_{36}p_{68}$

$P_{18.2,3(6,7)^n}=\frac{p_{12}p_{23}p_{36}p_{67}p_{76}p_{68}}{1-p_{67}p_{76}}=p_{12}p_{23}p_{67}$ $p_{48.5}=p_{45}p_{58}$

$p_{48.5,7,6}=p_{45}p_{57}p_{68}$

$p_{48.5(7,6)^n}=\frac{p_{45}p_{57}p_{76}p_{67}p_{76}p_{68}}{1-p_{67}p_{76}}=p_{45}p_{57}p_{67}$

$p_{98.7,6}=p_{97}p_{68}$

$p_{98.(7,6)^n}=\frac{p_{97}p_{76}p_{67}p_{76}p_{68}}{1-p_{67}p_{76}}=p_{97}p_{67}$

Also, it is verified that

$p_{01}=p_{10}+p_{12}=p_{23}+p_{24}=p_{36}=p_{45}+p_{4,10}=p_{57}+p_{58}=p_{67}+p_{68}=p_{76}=p_{80}+p_{89}$

$=p_{97}+p_{98}=p_{10,8}=p_{10}+p_{14.2}+p_{18.2,3,6}+p_{18.2,3(6,7)^n}=p_{4,10}+p_{48.5}+p_{48.5,7,6}+p_{48.5,7,6}p_{71.10}=p_{98}+p_{98.76}+p_{98.(7,6)^n}=1$

Mean Sojourn Times

$\mu_0=m_{01}=\frac{1}{2\lambda}$ $\mu_1=m_{10}+m_{12}=\frac{1}{\lambda}\left[1-g^*(\lambda)\right]$

$\mu_2=m_{23}+m_{24}=\frac{1}{2\lambda_1}[1-g^*(2\lambda_1)]$ $\mu_3=m_{36}=-g^{*'}(0)$

$\mu_4=m_{45}+m_{4,10}=\frac{1}{2\lambda_1}[1-g^*(2\lambda_1)]$ $\mu_5=m_{57}+m_{58}=\frac{1}{\lambda}\left[1-g^*(\lambda)\right]$

$\mu_6=m_{67}+m_{68}=\frac{1}{\lambda}\left[1-g^*(\lambda)\right]$ $\mu_7=m_{76}=-g^{*'}(0)$

$\mu_8=m_{80}+m_{89}=\frac{1}{2\lambda}\left[1-f^*(\lambda)\right]$ $\mu_9=m_{97}+m_{98}=\frac{1}{\lambda}\left[1-g^*(\lambda)\right]$

$\mu_{10}=\frac{1}{2\lambda_1}$

$\mu_1'=\mu_1+\mu_2p_{12}+p_{12}p_{23}\left(\mu_3+\mu_6\right)+\frac{p_{12}p_{23}p_{67}(\mu_6+\mu_7)}{p_{68}}$

$\mu_4'=\mu_4+\mu_5p_{45}+\frac{p_{45}p_{57}(\mu_6+\mu_7)}{p_{68}}$

$\mu_9'=\mu_9+p_{97}(\mu_6+\mu_7)+\frac{p_{97}p_{67}(\mu_6+\mu_7)}{p_{68}}$

II. Mean Time to System Failure and Reliability

Let $\phi_i(t)$ be the cdf of the first passage time from regenerative state 'i' to a failed state, regarding failed state as absorbing state, we have

$\phi_0(t)=Q_{01}(t)\textcircled{s}\phi_1(t)$ (1)

$\phi_1(t)=Q_{10}(t)\textcircled{s}\phi_0(t)+Q_{14.2}(t)\textcircled{s}\phi_4(t)+Q_{13.2}(t)$ (2)

$\phi_4(t)=Q_{4,10}(t)\textcircled{s}\phi_{10}(t)+Q_{48.5}(t)\textcircled{s}\phi_8(t)+Q_{47.5}(t)$ (3)

$\phi_8(t)=Q_{80}(t)\textcircled{s}\phi_0(t)+Q_{97}(t)$ (4)

$\phi_9(t)=Q_{98}(t)\textcircled{s}\phi_8(t)+Q_{71.10}(t)\textcircled{s}\phi_1(t)+Q_{79.10}(t)$ (5)

$\phi_{10}(t)=Q_{10,8}(t)\textcircled{s}\phi_8(t)$ (6)

By taking LST of the above expressions the reliability of the system model can be obtained as:

$$\phi_0^{**}(s)=\frac{\Delta_1}{\Delta}=\frac{\left(1-Q_{89}^{**}(s)Q_{98}^{**}(s)\right)\left[Q_{13.2}^{**}(s)Q_{01}^{**}(s)-Q_{01}^{**}(s)Q_{14.2}^{**}(s)Q_{47.5}^{**}\right]+(Q_{01}^{**}(s)Q_{14.2}^{**}(s)Q_{89}^{**}Q_{97}^{**})[Q_{48.5}^{**}(s)+Q_{4,10}^{**}(s)Q_{10,8}^{**}(s)]}{\left(1-Q_{89}^{**}(s)Q_{98}^{**}(s)\right)\left(1-Q_{01}^{**}(s)Q_{10}^{**}(s)\right)-(Q_{14.2}^{**}(s)Q_{01}^{**}(s))[(Q_{48.5}^{**}(s)Q_{80}^{**}(s)+Q_{4,10}^{**}(s)Q_{4,8}^{**}(s)Q_{80}^{**}(s)]}$$

$$\text{MTSF}=\lim_{s\to0}\frac{1-\phi_0^{**}(s)}{s}=\frac{N_1}{D_1}$$

$N_1=\left(\mu_0+\mu_1'+\mu_{10}p_{14.2}p_{4,10}\right)\left(1-p_{89}p_{98}\right)+\left(\mu_4'p_{14.2}\right)\left(p_{80}+p_{97}p_{89}\right)+\left(\mu_8+\mu_9p_{89}-\mu_0p_{80}\right)\left(p_{4,10}+p_{48.5}\right)p_{14.2}$

$D_1=\left(1-p_{10}\right)\left(1-p_{89}p_{98}\right)-p_{14.2}p_{80}\left(p_{48.5}+p_{4,10}\right)$; *where,*

$\mu_1'=\mu_1+p_{12}\mu_2$

$\mu_4'=\mu_4+p_{45}\mu_5$

The reliability of the system is determined as:

$R^*(s)=\frac{1-\phi_0^{**}(s)}{s}$ and $R(t)=L^{-1}\left(R^*(s)\right)$

III. Long Run Availability of the System

Let $A_i(t)$be the probability that the system is in up-state at instant 't' given that the system entered regenerative state 'i' at time instant t=0. We have

$A_0(t)=M_0(t)+q_{01}(t)\copyright A_1(t)$ (7)

$A_1(t)=M_1(t)+q_{10}(t)\copyright A_0(t)+q_{14.2}(t)\copyright A_4(t)+[q_{18.2,3,6}(t)+q_{18.2,3(6,7)^n}(t)]\copyright A_8(t)$ (8)

$A_4(t)=M_4(t)+q_{4,10}(t)\copyright A_{10}(t)+[q_{48.5}(t)+q_{48.5,7,6}(t)+q_{48.5(7,6)^n}]\copyright A_8(t)$ (9)

$A_8(t)=M_8(t)+q_{80}(t)\copyright A_0(t)+q_{89}(t)\copyright A_9(t)$ (10)

$A_9(t)=M_9(t)+q_{98}\copyright A_8(t)+[q_{98.7,6}+q_{98.(7,6)^n}]\copyright A_8(t)$ (11)

$A_{10}(t)=M_{10}(t)+q_{10,8}\copyright A_8(t)$ (12)

Where $M_i(t)$ is the probability that the system is up initially in state i is up at time t without visiting to any other regenerative state.

$M_0(t)=e^{-2\lambda t}dt$ $M_1(t)=e^{-\lambda t}\overline{G(t)}dt$ $M_4(t)=e^{-2\lambda_1 t}\overline{G(t)}dt$

$M_8(t)=e^{-2\lambda t}dt$ $M_9(t)=e^{-\lambda t}\overline{G(t)}dt$ $M_{10}(t)=e^{-2\lambda_1 t}dt$

Taking LT of above equations and solving for $A_0^*(s)$, the steady state availability is given by:

$A_0(\infty)=\lim_{s\to 0} s\,A_0^*(s)=\frac{N_2}{D_2}$; where

$N_2=(\mu_0+\mu_1)p_{80}+p_{80}p_{14.2}\mu_4+\mu_8p_{12}+\mu_9p_{12}p_{89}+\mu_{10}p_{14.2}p_{4,10}p_{80}$

$D_2=(\mu_0+\mu'_1)p_{80}+\mu'_4p_{14.2}p_{80}+\mu_8p_{12}+\mu'_9p_{12}p_{89}+\mu_{10}p_{14.2}p_{4,10}p_{80}$

$\mu'_1=\mu_1+\mu_2p_{12}+p_{12}p_{23}(\mu_3+\mu_6)+\frac{p_{12}p_{23}p_{67}(\mu_6+\mu_7)}{p_{68}}$

$\mu'_4=\mu_4+\mu_5p_{45}+\frac{p_{45}p_{57}(\mu_6+\mu_7)}{p_{68}}$

$\mu'_9=\mu_9+p_{97}(\mu_6+\mu_7)+\frac{p_{97}p_{67}(\mu_6+\mu_7)}{p_{68}}$

IV. Expected Number of Repairs of Phase-I Units

Let $R_i^M(t)$be the expected number of repairs of phase-I units given to the server in (0,t] such that the system entered regenerative state 'i' at t=0. We have

$R_0^M(t)=Q_{01}(t)\,\textcircled{S}\,R_1^M(t)$ (13)

$R_1^M(t)=Q_{10}(t)\,\textcircled{S}\,[1+R_0^M(t)]+Q_{14.2}(t)\,\textcircled{S}\,[1+R_4^M(t)]+[Q_{18.2,3,6}(t)+Q_{18.2,3(6,7)^n}(t)]\,\textcircled{S}\,[1+R_8^M(t)]$ (14)

$R_4^M(t)=Q_{4,10}(t)\,\textcircled{S}\,[1+R_{10}^M(t)]+[Q_{48.5}(t)+Q_{48.5,7,6}(t)+Q_{48.5(7,6)^n}(t)]\,\textcircled{S}\,[1+R_8^M(t)]$ (15)

$R_8^M(t)=Q_{80}(t)\,\textcircled{S}\,R_0^M(t)+Q_{89}(t)\,\textcircled{S}\,R_9^M(t)$ (16)

$R_9^M(t)=[Q_{98}(t)+Q_{98.7,6}(t)+Q_{98.(7,6)^n}(t)]\,\textcircled{S}\,[1+R_8^M(t)]$ (17)

$R_{10}^M(t)=Q_{10,8}\,\textcircled{S}\,R_8^M(t)$ (18)

Taking LST of above equations and solving for$R_0^{M^{**}}(s)$, by using Cramer's Rule we get the expected number of repairs of phase-I unit as:

$R_0^M(\infty)=\lim_{s\to 0} s\,R_0^{M^{**}}(t)=\frac{\Delta_1}{\Delta'_1}=\frac{N_3}{D_2}$; where

$N_3=P_{80}+p_{12}p_{24}+P_{12}P_{23}P_{89}$ *and*

$D_2=(\mu_0+\mu'_1)p_{80}+\mu'_4p_{14.2}p_{80}+\mu_8p_{12}+\mu'_9p_{12}p_{89}+\mu_{10}p_{14.2}p_{4,10}p_{80}$

V. Expected Numbers of Repairs of Phase-II Units

Let $R_i^D(t)$be the expected number of repairs of phase-II unit given to the server in (0,t] such that the system entered regenerative state 'i' at t=0. We have

$R_0^D(t)=Q_{01}(t)\,\textcircled{S}\,R_1^D(t)$ (19)

$R_1^D(t)=Q_{10}(t)\,\textcircled{S}\,R_0^D(t)+Q_{14.2}(t)\,\textcircled{S}\,R_4^D(t)+[Q_{18.2,3,6}(t)+Q_{18.2,3(6,7)^n}(t)]\,\textcircled{S}\,R_8^D(t)$ (20)

$R_4^D(t)=Q_{4,10}(t)\,\textcircled{S}\,R_{10}^D(t)+[Q_{48.5}(t)+Q_{48.5,7,6}(t)+Q_{48.5(7,6)^n}(t)]\,\textcircled{S}\,R_8^D(t)$ (21)

$R_8^D(t)=Q_{80}(t)\,\textcircled{S}\,[1+R_0^D(t)]+Q_{89}(t)\,\textcircled{S}\,R_9^D(t)$ (22)

$R_9^D(t)=[Q_{98}(t)+Q_{98.7,6}(t)+Q_{98.(7,6)^n}(t)]\,\textcircled{S}\,R_8^D(t)$ (23)

$R_{10}^D(t)=Q_{10,8}\,\textcircled{S}\,R_8^D(t)$ (24)

Taking LST of above equations and solving for $R_0^{D^{**}}(s)$, by using this we get expected number of repairs of phase-II units as:

$R_0^D(\infty)=\lim_{s\to 0} s\,R_0^{D^{**}}(s)=\frac{N_4}{D_2}$

Where $N_4=p_{12}p_{80}$ and

$D_2=(\mu_0+\mu'_1)p_{80}+\mu'_4p_{14.2}p_{80}+\mu_8p_{12}+\mu'_9p_{12}p_{89}+\mu_{10}p_{14.2}p_{4,10}p_{80}$

VI. Expected Number of Visits by the Server

LetV_0^S be the expected number of repairs of duplicate unit by the repairman in (0,t] such that the

system entered regenerative state 'i' at t=0.We have

$V_0^S(t)=Q_{01}(t)\,Ⓢ[1+V_1^S(t)]$ (25)

$V_1^S(t)=Q_{10}(t)\,ⓈV_0^S(t)+Q_{14.2}(t)\,ⓈV_4^S(t)+[Q_{18.2,3,6}(t)+Q_{18.2,3(6,7)^n}(t)]\,ⓈV_8^S(t)$ (26)

$V_4^S(t)=Q_{4,10}(t)\,ⓈV_{10}^S(t)+[Q_{48.5}(t)+Q_{48.5,7,6}(t)+Q_{48.5(7,6)^n}(t)]\,ⓈV_8^S(t)$ (27)

$V_8^S(t)=Q_{80}(t)\,Ⓢ[1+V_0^S(t)]+Q_{89}(t)\,ⓈV_9^S(t)$ (28)

$V_9^S(t)=[Q_{98}(t)+Q_{98.7,6}(t)+Q_{98.(7,6)^n}(t)]\,ⓈV_8^S(t)$ (29)

$V_{10}^S(t)=Q_{10,8}\,Ⓢ[1+V_8^S(t)]$ (30)

Taking LST of above equations and solving for $V_0^{S**}(s)$, by using this we get expected number of visits of the server:

$V_0^S(\infty)=\lim_{s\to 0} s\,V_0^{S**}(s)=\dfrac{N_5}{D_2}$

Where $N_5=p_{80}[1+P_{12}P_{24}P_{4,10}]$

And $D_2=\left(\mu_0+\mu'_1\right)p_{80}+\mu'_4p_{14.2}p_{80}+\mu_8p_{12}+\mu'_9p_{12}p_{89}+\mu_{10}p_{14.2}p_{4,10}p_{80}$

VII. Busy Period Analysis for the Server due to Repair

Let $B_i^R(t)$ be the probability that a server is busy at the time point given that the system entered in the regenerative state 'i' at t=0. We have

$B_0^R(t)=q_{01}(t)©B_1^R(t)$ (31)

$B_1^R(t)=W_1^R(t)+q_{10}(t)©B_0^R(t)+q_{14.2}(t)©B_4^R(t)+[q_{18.2,3,6}(t)+q_{18.2,3(6,7)^n}(t)]©B_8^R(t)$ (32)

$B_4^R(t)=W_4^R(t)+q_{4,10}(t)©B_{10}^R(t)+[q_{48.5}(t)+q_{48.5,7,6}(t)+q_{48.5(7,6)^n}]©B_8^R(t)$ (33)

$B_8^R(t)=W_8^R(t)+q_{80}(t)©B_0^R(t)(t)+q_{89}(t)©B_9^R(t)$ (34)

$B_9^R(t)=W_9^R(t)+q_{98}©A_8(t)+[q_{98.7,6}+q_{98.(7,6)^n}]©B_8^R(t)$ (35)

$B_{10}^R(t)(t)=q_{10,8}©B_8^R(t)$; *where* (36)

$W_1^R(t)=e^{-\lambda t}\overline{G(t)}dt$ $W_4^R(t)=e^{-2\lambda_1 t}\overline{G(t)}dt$

$W_8^R(t)=e^{-2\lambda t}dt$ $W_9^R(t)=e^{-\lambda t}\overline{G(t)}dt$

Taking LT of the above expressions and solving for $B_0^{R^*}(s)$ we have

$B_0^R(\infty)=\lim_{s\to 0} s\,B_0^{R^*}(s)=\dfrac{N_6}{D_2}$

$N_6=W_2^{R^*}(0)p_{80}+W_4^{R^*}(0)p_{12}p_{24}p_{80}+W_8^{R^*}(0)P_{12}+p_{12}p_{24}p_{4,10}\left(1+p_{80}\right)+W_9^{R^*}(0)p_{12}p_{89}(p_{23}+p_{24}p_{45})$

$D_2=\left(\mu_0+\mu'_1\right)p_{80}+\mu'_4p_{14.2}p_{80}+\mu_8p_{12}+\mu'_9p_{12}p_{89}+\mu_{10}p_{14.2}p_{4,10}p_{80}$

5. Profit Analysis

The profit (P) incurred to the system model in steady state can be obtained as:

$P=K_0A_0-K_1B_0^R-K_2R_0^D-K_3R_0^M-K_4V_0^S$ (37)

Where,

K_0 = Revenue per unit up time of the system

K_1=Cost per unit time for which server is busy to repair

K_3 = Cost per unit time repairs of phase-I unit

K_2 = Cost per unit time repairs of phase-II unit

K_4= Cost per unit time visit of the server

For graphical presentation of profit, these constants need some values and for this purpose K_0, K_1, K_2, K_3 and K_4 have been taken to be 15000, 3000, 800, 2000 and 1000 respectively.

6. Results and Graphical Representation of Reliability Measures

The reliability characteristics of the system model have been obtained by assuming arbitrary distributions for repairs rates of the units. The results for these measures have also been obtained for the particular situations where the repair rates of the units follow negative exponential distribution. The behavior of MTSF, Availability and Profit function have been shown numerically and graphically respectively in the tables 1, 2, 3 and in the figures 2, 3 and 4. Here, we take the repair time distribution as negative exponential: $g(t)=\xi e^{-\xi t}$ and $f(t)=\Psi e^{-\Psi t}$.

Table 3: *MTSF Vs Failure Rate of Phase-I Unit*

λ	λ1=0.001,Ψ=2,ξ=1.5	λ1=0.002,Ψ=2.0,ξ=1.5	λ1=0.001,Ψ=2,ξ=2.5	λ1=0.001,Ψ=2.5,ξ=1.5
0.1	76879.65	37210.85	136435.40	91728.39
0.11	64441.61	31295.15	112598.00	77338.17
0.12	54852.02	26717.79	94594.12	66138.10
0.13	47300.84	23100.57	80660.17	57248.89
0.14	41247.20	20190.9	69651.85	50075.23
0.15	36318.62	17814.55	60800.75	44202.02
0.16	32251.73	15848.03	53575.41	39332.56
0.17	28856.03	14201.76	47598.75	35250.15
0.18	25990.91	12809.45	42597.23	31793.53
0.19	23550.81	11621.14	38368.47	28840.73

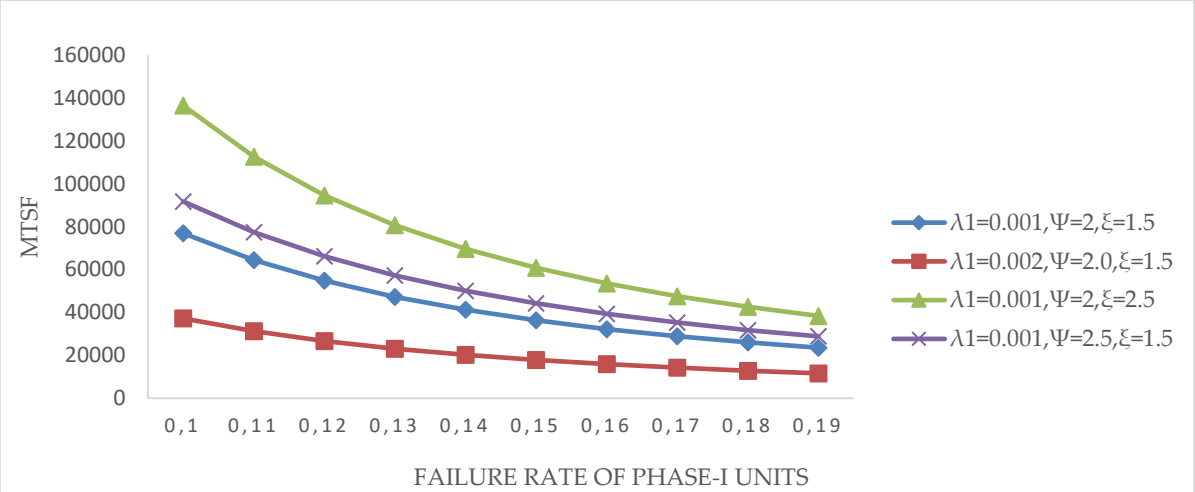


Figure 2: *MTSF Vs Failure Rate of Phase-I Unit*

Table 4: *Availability Vs Failure Rate of Phase-I Unit*

λ	λ1=0.001,Ψ=2,ξ=1.5	λ1=0.002,Ψ=2.0,ξ=1.5	λ1=0.001,Ψ=2,ξ=2.5		λ1=0.001,Ψ=2.5,ξ=1.5
0.1	0.99885337	0.99800616	0.999369592		0.998856064
0.11	0.99882166	0.997912792	0.999344996		0.998825033
0.12	0.99879522	0.997833532	0.999324552		0.998799332
0.13	0.99877269	0.997765415	0.999307327		0.998777608
0.14	0.99875309	0.997706120	0.999292617		0.998758883
0.15	0.99873570	0.997653834	0.999279881		0.998742441
0.16	0.99872001	0.997607137	0.999268708		0.998727751
0.17	0.99870561	0.997564913	0.999258778		0.998714414
0.18	0.99869219	0.997526279	0.999249842		0.998702124
0.19	0.99867952	0.997490532	0.999241707		0.998690646

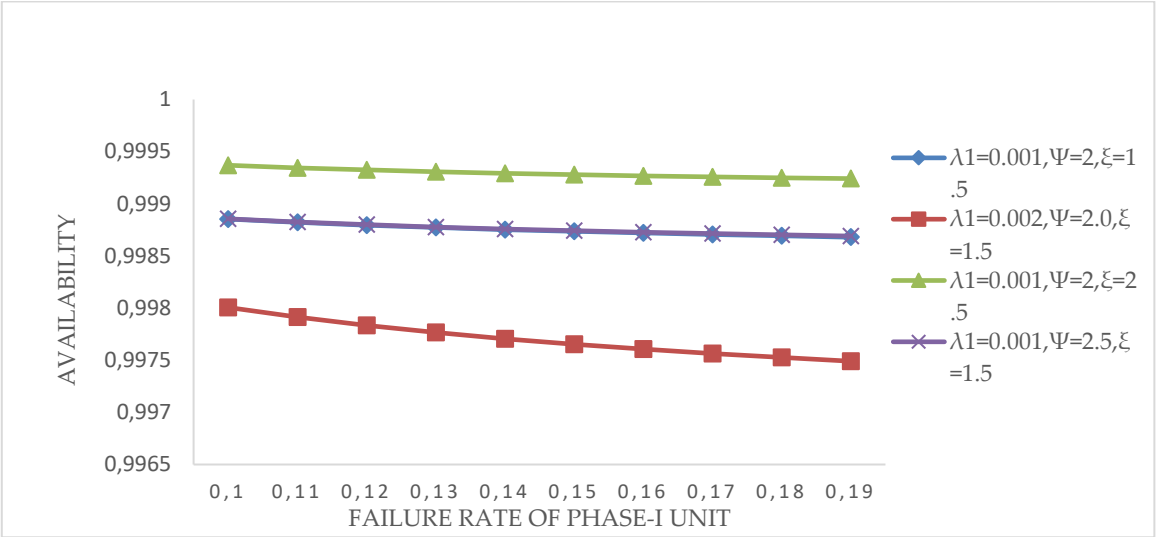


Figure 3: *Availability Vs Failure Rate of Phase-I Unit*

Table 5: *Profit Vs Failure Rate of Phase-I Unit*

λ	λ1=0.001,Ψ=2,ξ=1.5	λ1=0.002,Ψ=2.0,ξ=1.5	λ1=0.001,Ψ=2,ξ=2.5	λ1=0.001,Ψ=2.5,ξ=1.5
0.1	14955.29	14922.42	14961.05	14958.79
0.11	14953.85	14918.44	14959.35	14957.51
0.12	14952.63	14915.00	14957.91	14956.42
0.13	14951.55	14911.99	14956.66	14955.48
0.14	14950.60	14909.32	14955.56	14954.64
0.15	14949.73	14906.92	14954.59	14953.88
0.16	14948.94	14904.76	14953.72	14953.20
0.17	14948.2	14902.77	14952.93	14952.56
0.18	14947.50	14900.93	14952.2	14951.97
0.19	14946.85	14899.22	14951.52	14951.41

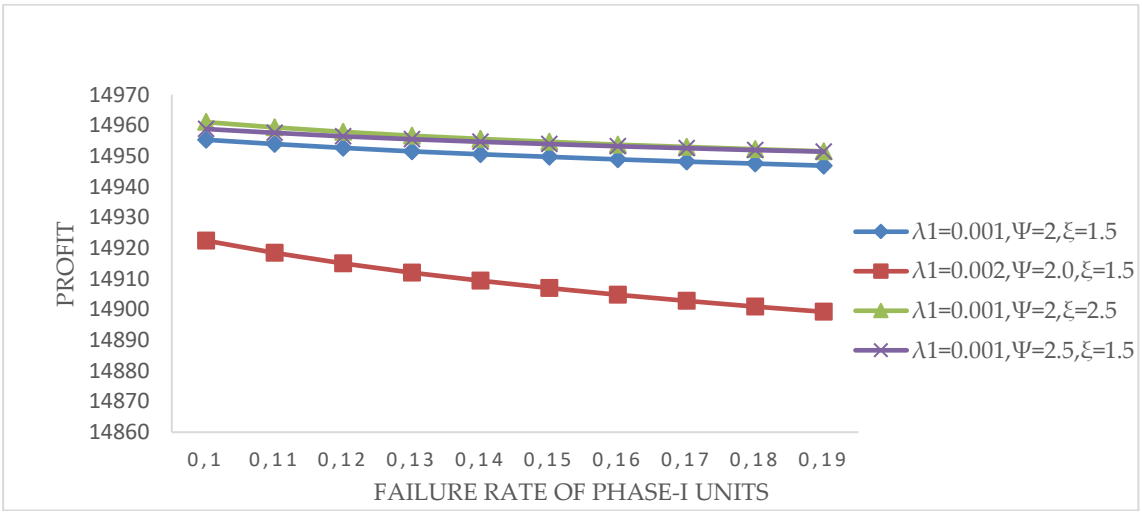


Figure 4: *Profit Vs Failure Rate of Phase-I Unit*

7. Application

The present study has the application in the system of turning and boring operation for making gun metal bushes to be required in production line. To make India as a developed country many industries have been established in last two decades. In industries there is huge requirement of machinery, one of them is CNC (for automatic programmed operation) and Lathe machine (for manual operation) which are used to produce a metallic bush. To meet up the heavy requirement of metallic bush two CNC machine (Phase-I units) and Lathe machine (Phase-II units) are installed. The single CNC machine can fulfil the requirement of production line in case of failure another CNC machine. In case of electric/mechanical failure of CNC machine (Phase-I units), additional arrangement of two lathe machines (Phase-II units) are installed to achieve the same. The system can be shown in the following figure 5:

Fig. 5: Manufacturing of Metallic Components

8. Conclusion

A Parallel-Cold Standby system of four units has been analyzed stochastically with the idea of priority for repair to the phase-I units. There are four units in system comprising two units as phase-I units which work initially in parallel mode and the other two units (called phase-II units) which remain as spare in cold standby mode. The phase-II units can be installed to work simultaneously at the failure of the phase-I units. The important reliability characteristics have been obtained and analyzed for arbitrary values of the parameters. Graphical and tabulated presentations have been studied by taking exponential distributions for the repair time i.e.,$g(t)=\xi e^{-\xi t}$ and $f(t)=\Psi e^{-\Psi t}$. The results are shown graphically in the figures 2,3, and 4 respectively. The following conclusions can be made from the graphical study:

1. From Fig. 2 it is quite evident that MTSF has downward trend with the increase of failure rate of the phase-I unit and phase-II unit while it increases with the increase in repair rate of phase-I. There is little change in the (almost negligible) in the values of availability with the increase of repair rate of Phase-II units. We conclude that a repairable parallel-cold standby system of four units can be made to use in a batter way in terms of reliability by increasing repair rate (from 1.5 to 2.5) of phase-I units.
2. Fig.3 depicts that availability of the system keeps on decreasing with increase of failure rates (0.1 to 0.19) of phase -I units. However, the system availability is more when repair of phase -I units is kept as priority as compare to phase-II units. Availability is increased slightly from 0.998 to 0.999 in case of increased repair rate of phase-I unit. Hence, keeping the repair priority policy for phase-I units is beneficial in case of availability.

3. Fig.4 of profit analysis represents the same trend as that of availability. Repair of phase-I units gives more profit as compare to phase-II units. Therefore, spending money on increasing repair rate of phase-II units will not be fruitful and thus, one should avoid the use of low-quality units in standby.

References

[1] Barak, A.K. and Malik S.C. (2013). Cost-Benefit Analysis of a Computer System with Priority to Preventive Maintenance over Hardware Repair Subject to Maximum Operation and Repair Times, *Indian Journal of Science and Technology*, Vol. 6(3), pp. 1-9.

[2] Pundir, Pramendra Singh; Patawa, Rohit; & Gupta Puneet Kumar. (2018). Stochastic outlook of two non-identical unit parallel system with priority in repair, *Cogent Mathematics & Statistics.* Vol. 5(1), pp. 1-18

[3] Pundir, Pramendra Singh; Patawa, Rohit; & Gupta Puneet Kumar (2020). Analysis of Two Non-Identical Unit Cold Standby System in Presence of Prior Information, *American Journal of Mathematical and Management Science.* Vol. 40(4), pp. 320-335.

[4] Kadyan, S. & Malik, S.C. & Gitanjali (2020) Stochastic analysis of a three-unit non-identical repairable system with priority for operation and repair. *International Journal of Reliability and Safety.* Vol. 14(2/3), pp.182 – 198.

[5] Kadyan, S. & Malik, S. C. & Gitanjali (2021) Stochastic Analysis of a Three-Unit Non-Identical Repairable System with Simultaneous Working of Cold Standby Units. *Journal of Reliability and Statistical Studies.* Vol. 13(2), pp. 385-400.

[6] Behboudi Z., G.R. Mohtashami Borzadaran, M. Asadi, (2021). Reliability modeling of two-unit cold standby systems: A periodic switching approach, *Applied Mathematical Modelling,* Vol. 92, pp. 176-195.

[7] Eryilmaz S, Finkelstein M. (2022). Reliability of the two-unit priority standby system revisited. Proceedings of the Institution of Mechanical Engineers, *Part O: Journal of Risk and Reliability.* Vol. 236(6): pp. 1096-1103.

[8] S. C. Malik, R. K. Yadav: Stochastic Analysis of a Computer System with Unit Wise Cold Standby Redundancy and Priority to Hardware Repair Subject to Failure of Service Facility. *International Journal of Reliability, Quality and Safety Engineering.* Vol. 28(2), 2150013

[9] Anuradha and S.C. Malik (2023). Reliability Measures of 2-Out-3G System with Priority and Faulure of Service Facility during Repair, *Reliability: Theory & Applications,* March 1(72): 214-223.

[10] Kumar, A., Garg, R. & Barak, M.S. (2023). Reliability measures of a cold standby system subject to refreshment. *Int J Syst Assur Eng Manag* Vol. 14, pp. 147–155.

Analysis of $M, MAP/PH_1, PH_2/1$ non-preemptive priority Queueing model with Delayed working vacations, immediate feedback,impatient customer, differentiate breakdown and phase type repair

G. Ayyappan, N. Arulmozhi

•

Department of Mathematics,
Puducherry Technological University,
Puducherry, India.
Chennai 600005, India.
ayyappan@ptuniv.edu.in, arulmozhi.n@pec.edu,

Abstract

The arrival of high priority customers is governed by the Poisson process while that of low priority customers is governed by the Markovian Arrival Process, and the service times are determined by a distinct Phase-type distribution. When the service is finished and the system is empty, the server stays idle for a random period (delay time). If a customer arrives within the delayed period, the server resumes normal service to the customer immediately. Otherwise, at the end of the delayed period, the server will take a working vacation and will instantly provide slow service to customers (high priority customers only). The Matrix analytic method is used to investigate the system. We also discussed the steady-state vector and busy period for our concept. The estimated and visually displayed performance measures of the system

Keywords: Non-preemptive Priority, Working vacation policy, Phase-type repair, Immediate feedback, Differentiate breakdown, Delay time.

AMS Subject Classification (2010): 60K25, 68M30, 90B22.

1. Introduction

For the past two decades, the priority queue hypothesis has been used in communication strategies. Because priority does not come under FCFO, it distinguishes it from a normal queue. It is a special type of queue in which each customer is dealt with priority and served according to its priority. There are two different types of priority service available in a queueing system: preemptive and non-preemptive. Priority customers that arrive early will wait until the service is finished while regular customers are serviced. This belongs under the non-preemptive priority rule. In the event of a preemptive rule, high-priority consumers would frequently interrupt low-priority service.

Ayyappan et al. [21] looked at $M/M/1$ for retrials, with negative arrival while using non-preemptive priority service. Bhagat and Jain [5] described a multi-server, non-preemptive priority service that is susceptible to failure and maintenance. According to Jeganathan et al. [9], the inventory system and non-preemptive priority service for retrials have been discussed. Additionally, discretionary priority service is utilized, taking into account both disciplines. Ayyappan and Somasundaram [3] analyzed discretionary priority service for retrials used $M^{X1}, M^{X2}/G1, G2/1$.

Krishnamoorthy and Divya [13] examined queueing models with MAP and PH distributions, as well as working vacations under N-policy.

In many real-world queueing situations, the server can be seen working during its rest period if necessary. Working vacation means that the server offers service at a lower rate throughout the vacation period rather than entirely shutting down. In the past few decades have seen, queueing systems with server working vacation, owing to similarities between telecommunication system, manufacturing system, and computer system. Yang et al. [22] applied the spectral expansion method to deal with a single server queueing model with delayed working vacations and working breakdown: The author showed the steady-state probability vector, LST of sojourn time, and expected sojourn time. After service completion, the server is idle when there are no customers in the system for a certain amount of time (changeover time) (Pikkala et al. [19], Krishna Reddy and Anitha [11]). The server begins offering service if customers access the system during changeover time; if not, the server goes on vacation at the end of the changeover time.

After obtaining service from the server, customers may be satisfied or unsatisfied. Customers who are satisfied with the system will leave, while those who are not satisfied will get feedback right away. A single server model with starting failures, standby server, single vacation, delayed repair, breakdown, immediate feedback, and impatient customers was extensively analyzed by Ayyappan and Thilagavathy [1], who found the expected results for both the system size and orbit size. In their 2008 study, Badamchi Zadeh and Shahkar [4] examined queuing systems that included two phases of heterogeneous service, optional second service, and feedback for each service. In contrast to the current study, when services are parallel, they had sequential services during their studies. Afterward, performance measures for the Poisson arrival queuing system and probability-generating functions are obtained under the assumption of exponential service times. Ayyappan and Thilagavathy [2] explored closedown, breakdown and multiple vacation used $MAP/PH/1$.

When the system is inactive or when a customer is being served, random failures can happen. The terms "hard failure" and "soft failure" refer to two different kinds of system failure. Hard failure's typically takes a long period and needs the repairman's actual presence. On the other side, soft failure's takes less time because the system may be recovered with a simple reboot. Markovian queueing models with two different forms of server breakdown have already been studied by Jain and Jain (2010) [7], Kalyanaraman (2019) [10], Krishna Kumar(2008)[12], Li (2013) [16], and many others. Using the matrix geometric technique, stability conditions for a single server infinite capacity Markovian queue were obtained. According to Janani [8], the final value theorem of the Laplace transform is used to convert the transient state probabilities of the model into steady-state probabilities.

When customers abandon the line because they have waited too long for service, they are considered impatient customers. Kumar [14] investigated a non-Markovian queue with an unreliable server that first provides an essential service and then one of the m optional services. He has described the balking techniques as well as cost analysis for the objective of model optimization. A single server queueing system with associated reneging, feedback, and balking was investigated by Rakesh Kumar and Soodan [20]. We explored the time-dependent behavior of the model using the Runge-Kutta method. Additionally, they discovered the average waiting time and system size. In modeling, the arrival using a Markovian Arrival Process, a particular type of Versatile Markovian Point Process was proposed by Neuts [18]. Lucantoni et al. [17], with considerable VMPP as BMAP notational simplifications since it started in 1990. Due to its ability to simulate a broad spectrum of real-world events, MAP is an effective point process in stochastic modeling. Chakravarthy [6], describes two parameter matrices of m dimensions, let's say D_0 and D_1. Transitions in the MAP are determined by the generator matrix $D = D_0 + D_1$.

2. Model Formulation

Within this part, our focus is on a system for queueing with a single server, utilizing non-preemptive priority. Customers categorized as high priority (HP) arrive through a Poisson process with rate denoted by λ_2, while low priority (LP) customers arrive via a Markovian Arrival Process represented by (D_0, D_1) of order m. The matrix D_0 means no arrival LP customer, while the matrix D_1 depicts LP customer arrival. HP customers have a limited capacity of K size, while LP customers have unlimited capacity. The fundamental arrival rate, denoted as λ_1, is equivalent to $\pi_1 D_1 e$, where π_1 represents the stationary probability vector. A customer categorized as HP is assumed to have a service time that follows a phase-type distribution with the notation (γ, U) of order n, while an LP customer's service time is assumed to follow a phase-type distribution with the notation (γ', U') of order n'.

Upon completion of the service, if no customer is in the system, then the server will remain inactive for a random duration. That time is referred to as the delayed period. The delayed period follows an exponential distribution with parameter ω. when a customer arrives during the delayed period, prompt resumption of regular service is initiated by the server. However, if the delayed period ends and any customer does not arrive, the server will proceed on a working vacation. The vacation period is generated by an exponentially distributed parameter η. HP customers who arrive during this period will be served at a lower service rate and it is followed by phase-type distribution with representation $(\gamma, \theta U)$, where $0 < \theta < 1$. As such, the mean service rate in normal mode is $\mu_1 = [\gamma(-U)^{-1}e]^{-1}$, and the vacation mode of service rate is $\theta\mu_1$.

After completion of service for HP customer during working vacation, if there exists no HP customer awaiting service, then the server will doemant in vacation mode, irrespective of the presence of LP customers in the system. After the expiration of the vacation clock during a WV, the server shall revert to its normal working mode. At the end of vacation period, LP customers shall be considered for service during no HP customer present in the system. The expected service rate of an LP customer is denoted by $\mu_2 = [\gamma'(-U')^{-1}e]^{-1}$.

The server is affected by soft failure(short time) during idle period and hard failure (long time) during normal busy period (both HP and LP customers). The rates of soft and hard failure are exponentially distributed with parameters ψ_1 and ψ_2. When a soft and hard failure happens, the server repair process starts immediately. The customer who is receiving service at that point must join the front of the waiting queue. If there are any customers in line when the repair is finished, the server will start servicing them. Or else, the server remains idle and repair times follows a phase-type distribution (α, T) of order l for soft failure, where $T^0 + Te = 0$ and (α', T') of order l' for hard failure, where $T'^0 + T'e = 0$. The repair rate is indicated as $\tau_1 = [\alpha(-T)^{-1}e]^{-1}$ and $\tau_2 = [\alpha'(-T')^{-1}e]^{-1}$ respectively.

The arriving LP Customers may balk the system with probability b during working vacation or join the system with probability $(1 - b)$. After receiving normal service (both HP and Lp customers), the satisfied customer leave the system with probability p_1 and if the customer is not satisfied with probability q_1 then they will get feedback immediately.

3. The QBD process infinitesimal generation matrix

Notations

We will need the following notations:

- $\otimes$ -Kronecker product of two matrices of various dimensions resulting in a block matrix.
- $\oplus$ - Kronecker sum of two matrices of various dimensions resulting in a block matrix.
- I_m stand for identity matrix of m x m order.
- e - a column vector of the suitable order. Each of its entries is one.

- e_0=$e_{3m+2Knm+Kl'm+(K+1)lm}$.
- e_1=$e_{Kmn+(K+1)n'm+(K+1)l'm+(K+1)lm+Kmn+m}$.
- $N_1(t)$: the total number of LP customers in the system at epoch t.
- $N_2(t)$: the total number of HP customers in the system at epoch t.
- $J(t)$ represents the server's status at epoch t.
  As a result, the server is in one of the following states at any given time t:

$$J(t) = \begin{cases} 0, & \text{idle during normal mode,} \\ 1, & \text{if the server is offering service to HP customers during normal mode,} \\ 2, & \text{if the server is offering service to LP customers during normal mode,} \\ 3, & \text{hard failure (during normal busy mode),} \\ 4, & \text{delay time,} \\ 5, & \text{soft failure (during idle),} \\ 6, & \text{busy(HP) in working vacation mode,} \\ 7, & \text{idle in working vacation mode.} \end{cases}$$

- $R(t)$ stands for the repair process considered by phases.
- $K(t)$ stands for phases of the service.
- $A(t)$- The Markovian arrival process is considered in phases.
- Let Y=$\{Y(t) : t \geq 0\}$, where $Y(t) = \{N_1(t), N_2(t), J(t), R(t), K(t), A(t)\}$ is a $CTMC$ with state space

$$\Phi = \phi(0) \bigcup_{i=1}^{\infty} \phi(i). \tag{1}$$

where

$$\begin{aligned} \phi(0) =& \{(0,0,0,a) :\ 1 \leq a \leq m\} \cup \{(0,r,1,k_1,a) :\ 1 \leq r \leq K,\ 1 \leq k_1 \leq n,\ 1 \leq a \leq m\} \\ & \cup \{(0,r,3,k_4,a) :\ 1 \leq r \leq K,\ 1 \leq k_4 \leq l',\ 1 \leq a \leq m\} \cup \{(0,0,4,a) :\ 1 \leq a \leq m\} \\ & \cup \{(0,0,5,k_3,a) :\ 0 \leq r \leq K,\ 1 \leq k_3 \leq l,\ 1 \leq a \leq m\} \\ & \cup \{(0,1,6,k_1,a) :\ 1 \leq r \leq K,\ 1 \leq k_1 \leq n,\ 1 \leq a \leq m\} \cup \{(0,0,7,a) :\ 1 \leq a \leq m\}, \end{aligned}$$

and for $i \geq 1$,

$$\begin{aligned} \phi(i) =& \{(i,r,1,k_1,a) :\ 1 \leq r \leq K,\ 1 \leq k_1 \leq n,\ 1 \leq a \leq m\} \\ & \cup \{(i,r,2,k_2,a) :\ 0 \leq r \leq K,\ 1 \leq k_2 \leq n',\ 1 \leq a \leq m\} \\ & \cup \{(i,r,3,k_4,a) :\ 0 \leq r \leq K,\ 1 \leq k_4 \leq l',\ 1 \leq a \leq m\} \\ & \cup \{(i,r,5,k_3,a) :\ 0 \leq r \leq K,\ 1 \leq k_3 \leq l,\ 1 \leq a \leq m\} \\ & \cup \{(i,r,6,k_1,a) :\ 1 \leq r \leq K,\ 1 \leq k_1 \leq n,\ 1 \leq a \leq m\} \cup \{(i,0,7,a) :\ 1 \leq a \leq m\}. \end{aligned}$$

3.1. The Infinitesimal Generator Matrix

The quasi-birth–death process has the generator matrix Q given by

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & \dots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & 0 & 0 & \dots \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots \end{bmatrix} \tag{2}$$

$$B_{00} = \begin{bmatrix} B_{00}^{11} & B_{00}^{12} & 0 & 0 & B_{00}^{15} & 0 & 0 \\ 0 & B_{00}^{22} & B_{00}^{23} & B_{00}^{24} & 0 & 0 & 0 \\ 0 & B_{00}^{32} & B_{00}^{33} & 0 & 0 & 0 & 0 \\ 0 & B_{00}^{42} & 0 & B_{00}^{44} & 0 & 0 & B_{00}^{47} \\ B_{00}^{51} & B_{00}^{52} & 0 & 0 & B_{00}^{55} & 0 & 0 \\ 0 & B_{00}^{62} & 0 & 0 & 0 & B_{00}^{66} & B_{00}^{67} \\ B_{00}^{71} & 0 & 0 & 0 & 0 & B_{00}^{76} & B_{00}^{77} \end{bmatrix},$$

where

$B_{00}^{11} = D_0 - (\lambda_2 + \psi_1)I_m$, $B_{00}^{12} = e_1' \otimes \alpha \otimes \lambda_2 I_m$, $B_{00}^{15} = e_1'(K+1) \otimes \alpha \otimes \psi_1 I_m$,

$$B_{00}^{22} = \begin{bmatrix} L_1 & L_2 & 0 & \dots & 0 & 0 \\ L_3 & L_1 & L_2 & \dots & 0 & 0 \\ 0 & L_3 & L_1 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_1 & L_2 \\ 0 & 0 & 0 & \dots & L_3 & L_1 + L_2 \end{bmatrix}, B_{00}^{23} = I_K \otimes e_n \otimes \alpha' \otimes \psi_2 I_m, \quad B_{00}^{24} = e_1 K \otimes qU^0 \otimes I_m,$$

where $L_1 = (U + pU^0\gamma) \oplus D_0 - (\lambda_2 + \psi_2)I_{nm}$, $L_2 = \lambda_2 I_{nm}$, $L_3 = qU^0\gamma \otimes I_m$.

$$B_{00}^{32} = I_K \otimes T'^0\gamma \otimes I_m, B_{00}^{33} = \begin{bmatrix} L_4 & L_5 & 0 & \dots & 0 & 0 \\ 0 & L_4 & L_5 & \dots & 0 & 0 \\ 0 & 0 & L_4 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_4 & L_5 \\ 0 & 0 & 0 & \dots & 0 & L_4 + L_5 \end{bmatrix},$$

where $L_4 = T' \oplus D_0 - \lambda_2 I_{l'm}$, $L_5 = \lambda_2 I_{l'm}$.

$B_{00}^{42} = e_1' K \otimes \alpha \otimes \lambda_2 I_m$, $B_{00}^{44} = D_0 - (\lambda_2 + \omega)I_m$, $B_{00}^{47} = \omega I_m$, $B_{00}^{51} = e_1(K+1) \otimes T^0 \otimes I_m$,

$$B_{00}^{52} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ T^0\gamma \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & T^0\gamma \otimes I_m & 0 & \dots & 0 & 0 \\ 0 & 0 & T^0\gamma \otimes I_m & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & T^0\gamma \otimes I_m & 0 \\ 0 & 0 & 0 & \dots & 0 & T^0\gamma \otimes I_m \end{bmatrix},$$

$$B_{00}^{55} = \begin{bmatrix} L_6 & L_7 & 0 & \dots & 0 & 0 \\ 0 & L_6 & L_7 & \dots & 0 & 0 \\ 0 & 0 & L_6 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_6 & L_7 \\ 0 & 0 & 0 & \dots & 0 & L_6 + L_7 \end{bmatrix}, B_{00}^{66} = \begin{bmatrix} L_8 & L_9 & 0 & \dots & 0 & 0 \\ L_{10} & L_8 & L_9 & \dots & 0 & 0 \\ 0 & L_{10} & L_8 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_8 & L_9 \\ 0 & 0 & 0 & \dots & L_{10} & L_8 + L_9 \end{bmatrix},$$

where $L_6 = T \oplus D_0 - \lambda_2 I_{lm}$, $L_7 = \lambda_2 I_{lm}$, $L_8 = \theta U \oplus (D_0 + bD_1) - (\eta + \lambda_2)I_{nm}$, $L_9 = \lambda_2 I_{nm}$, $L_{10} = \theta U' \otimes I_m$.

$B_{00}^{62} = I_K \otimes e_n \otimes \eta\gamma \otimes I_m$, $B_{00}^{67} = e_1 K \otimes \theta U^0 \otimes I_m$, $B_{00}^{71} = \eta I_m$, $B_{00}^{76} = e_1' K \otimes \gamma \otimes \lambda_2 I_m$, $B_{00}^{77} = (D_0 + bD_1) - (\eta + \lambda_2)I_m$.

$$B_{01} = \begin{bmatrix} 0 & B_{01}^{12} & 0 & 0 & 0 & 0 \\ B_{01}^{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & B_{01}^{33} & 0 & 0 & 0 \\ 0 & B_{01}^{42} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{01}^{54} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{01}^{65} & 0 \\ 0 & 0 & 0 & 0 & 00 & B_{01}^{76} \end{bmatrix},$$

where
$B_{01}^{12} = e_1'(K+1) \otimes \gamma' \otimes D_1, \quad B_{01}^{21} = I_K \otimes I_n \otimes D_1,$

$$B_{01}^{33} = \begin{bmatrix} 0 & I_{l'} \otimes D_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & I_{l'} \otimes D_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & I_{l'} \otimes D_1 & \dots & 0 & 0 \\ & & & & \ddots & \ddots & \\ 0 & 0 & 0 & 0 & \dots & I_{l'} \otimes D_1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & I_{l'} \otimes D_1 \end{bmatrix}, \; B_{01}^{42} = e_1'(K+1) \otimes \gamma' \otimes D_1,$$

$B_{01}^{54} = I_{K+1} \otimes I_l \otimes D_1, \quad B_{01}^{65} = I_K \otimes I_n \otimes (1-b)D_1, \quad B_{01}^{76} = (1-b)D_1.$

$$B_{10} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & B_{10}^{22} & 0 & B_{10}^{24} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where

$$B_{10}^{22} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ qU'^0\gamma \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & qU'^0\gamma \otimes I_m & 0 & \dots & 0 & 0 \\ 0 & 0 & qU'^0\gamma \otimes I_m & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & qU'^0\gamma \otimes I_m & 0 \\ 0 & 0 & 0 & \dots & 0 & qU'^0\gamma \otimes I_m \end{bmatrix},$$

$B_{10}^{24} = e_1(K+1) \otimes qU'^0 \otimes I_m.$

$$B_{11} = \begin{bmatrix} B_{11}^{11} & B_{11}^{12} & B_{11}^{13} & 0 & 0 & 0 \\ 0 & B_{11}^{22} & B_{11}^{23} & 0 & 0 & 0 \\ B_{11}^{31} & B_{11}^{32} & B_{11}^{33} & 0 & 0 & 0 \\ B_{11}^{41} & B_{11}^{42} & 0 & B_{11}^{44} & 0 & 0 \\ B_{11}^{51} & 0 & 0 & 0 & B_{11}^{55} & B_{11}^{56} \\ 0 & B_{11}^{62} & 0 & 0 & B_{11}^{65} & B_{11}^{66} \end{bmatrix},$$

where

$$B_{11}^{11} = \begin{bmatrix} L_1 & L_2 & 0 & \dots & 0 & 0 \\ L_3 & L_1 & L_2 & \dots & 0 & 0 \\ 0 & L_3 & L_1 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_1 & L_2 \\ 0 & 0 & 0 & \dots & L_3 & L_1 + L_2 \end{bmatrix}, \quad B_{11}^{12} = \begin{bmatrix} qU^0\gamma' \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

$$B_{11}^{13} = \begin{bmatrix} 0 & e_n \otimes \psi_2\alpha' \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & e_n \otimes \psi_2\alpha' \otimes I_m & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & e_n \otimes \psi_2\alpha' \otimes I_m & \dots & 0 & 0 \\ & & & & \ddots & \ddots & \\ 0 & 0 & 0 & 0 & \dots & e_n \otimes \psi_2\alpha' \otimes I_m & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & e_n \otimes \psi_2\alpha' \otimes I_m \end{bmatrix},$$

$$B_{11}^{22} = \begin{bmatrix} L_{11} & L_{12} & 0 & \dots & 0 & 0 \\ 0 & L_{11} & L_{12} & \dots & 0 & 0 \\ 0 & 0 & L_{11} & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & L_{11} & L_{12} \\ 0 & 0 & 0 & \dots & 0 & L_{11} + L_{12} \end{bmatrix}, \quad B_{11}^{23} = I_{K+1} \otimes e_n' \otimes \psi_2\alpha' \otimes I_m,$$

where $L_{11} = (U' + pU'^0\gamma_1) \oplus D_0 - (\lambda_2 + \psi_2)I_{n'm}$, $\quad L_{12} = \lambda_2 I_{n'm}$.

$$B_{11}^{31} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ T'^0\gamma \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & T'^0\gamma \otimes I_m & 0 & \dots & 0 & 0 \\ 0 & 0 & T'^0\gamma \otimes I_m & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & T'^0\gamma \otimes I_m & 0 \\ 0 & 0 & 0 & \dots & 0 & T'^0\gamma \otimes I_m \end{bmatrix},$$

$$B_{11}^{32} = \begin{bmatrix} T'^0\gamma' \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \quad B_{11}^{33} = \begin{bmatrix} L_4 & L_5 & 0 & \dots & 0 & 0 \\ 0 & L_4 & L_5 & \dots & 0 & 0 \\ 0 & 0 & L_4 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_4 & L_5 \\ 0 & 0 & 0 & \dots & 0 & L_4 + L_5 \end{bmatrix},$$

$$B_{11}^{41} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ T^0\gamma \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & T^0\gamma \otimes I_m & 0 & \dots & 0 & 0 \\ 0 & 0 & T^0\gamma \otimes I_m & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & T^0\gamma \otimes I_m & 0 \\ 0 & 0 & 0 & \dots & 0 & T^0\gamma \otimes I_m \end{bmatrix},$$

$$B_{11}^{42} = \begin{bmatrix} T^0\gamma' \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, B_{11}^{44} = \begin{bmatrix} L_6 & L_7 & 0 & \dots & 0 & 0 \\ 0 & L_6 & L_7 & \dots & 0 & 0 \\ 0 & 0 & L_6 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_6 & L_7 \\ 0 & 0 & 0 & \dots & 0 & L_6 + L_7 \end{bmatrix},$$

$$B_{11}^{51} = I_K \otimes e_n \otimes \eta\gamma \otimes I_m,$$

$$B_{11}^{55} = \begin{bmatrix} L_8 & L_9 & 0 & \dots & 0 & 0 \\ L_{10} & L_8 & L_9 & \dots & 0 & 0 \\ 0 & L_{10} & L_8 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \dots & L_8 & L_9 \\ 0 & 0 & 0 & \dots & L_{10} & L_8 + L_9 \end{bmatrix}, \quad B_{11}^{56} = e_1 K \otimes \theta U^0 \otimes I_m,$$

$B_{11}^{62} = e_1'(K+1) \otimes \eta\gamma' \otimes I_m, \quad B_{11}^{65} = e_1'(K) \otimes \alpha \otimes \lambda_2 I_m, B_{11}^{66} = (D_0 + bD_1) - (\eta + \lambda_2) I_m.$

$$B_{12} = \begin{bmatrix} B_{12}^{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & B_{12}^{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & B_{12}^{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{12}^{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{12}^{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{12}^{66} \end{bmatrix},$$

where
$B_{12}^{11} = I_K \otimes I_n \otimes D_1, \quad B_{12}^{22} = I_{K+1} \otimes I_{n'} \otimes D_1, \quad B_{12}^{33} = I_{K+1} \otimes I_{l'} \otimes D_1,$
$B_{12}^{44} = I_{K+1} \otimes I_l \otimes D_1, \quad B_{12}^{55} = I_K \otimes I_n \otimes (1-b) D_1, \quad B_{12}^{66} = (1-b) D_1,$

$$B_{21} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ B_{21}^{21} & 0 & B_{21}^{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where

$$B_{21}^{21} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ qU'^0\gamma \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & qU'^0\gamma \otimes I_m & 0 & \dots & 0 & 0 \\ 0 & 0 & qU'^0\gamma \otimes I_m & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & qU'^0\gamma \otimes I_m & 0 \\ 0 & 0 & 0 & \dots & 0 & qU'^0\gamma \otimes I_m \end{bmatrix},$$

$$B_{21}^{23} = \begin{bmatrix} qU'^0\gamma' \otimes I_m & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

4. Analysis of Stability Condition

We examined our model under the assumption that the system is stable.

4.1. Condition for Stability

Let $A = A_0 + A_1 + A_2$ be the square matrix of order $Kmn + (K+1)n'm + (K+1)l'm + (K+1)lm + Kmn + m$ and it is an infinitesimal generator matrix is an irreducible. Let χ indicate the steady-state probability vector of A satisfying $\chi A = 0$ and $\chi e = 1$. The vector χ is partitioned by

$\chi = (\chi_0, \chi_1, \chi_2, \chi_3, \chi_4, \chi_5)$=$(\chi_{00}, \chi_{01}, \ldots, \chi_{0K-1}, \chi_{0K}, \chi_{11}, \chi_{12}, \ldots, \chi_{1K-1}, \chi_{1K}, \chi_{20}, \chi_{21}, \ldots, \chi_{2K-1}, \chi_{2K}, \chi_{30}, \chi_{31}, \ldots, \chi_{3K-1}, \chi_{3K}, \chi_{40}, \chi_{41}, \ldots, \chi_{4K-1}, \chi_{4K}, \chi_{50}, \chi_{51}, \ldots, \chi_{5K-1}, \chi_{5K})$, where χ_0 is of dimension Kmn, χ_1 is of dimension $(K+1)n'm$, χ_2 is of dimension $(K+1)l'm$, χ_3 is of dimension $(K+1)lm$, χ_4 is of dimension Kmn and χ_5 is of dimension m. The probability vector χ is calculated by solving the following equations:

$$\chi_{00}[(U + pU^0\gamma) \otimes I_m - (\lambda_2 + \eta)I_{nm}] + \chi_{01}(qU^0\gamma \otimes I_m) + \chi_{11}(qU'^0\gamma \otimes I_m) + \chi_{21}(T'^0\gamma \otimes I_m) + \chi_{31}(T^0\gamma \otimes I_m) + \chi_{40}(e_n \otimes \eta\gamma \otimes I_m) = 0.$$

$$\chi_{0j-1}(\lambda_2 I_{nm}) + \chi_{0j}[(U + pU^0\gamma) \otimes I_m - (\lambda_2 + \eta)I_{nm}] + \chi_{0j+1}(qU^0\gamma \otimes I_m) + \chi_{1j+1}(qU'^0\gamma \otimes I_m) + \chi_{2j+1}(T'^0\gamma \otimes I_m) + \chi_{3j+1}(T^0\gamma \otimes I_m) + \chi_{4j}(e_n \otimes \eta\gamma \otimes I_m) = 0, \quad for\ 1 \le j \le K-1.$$

$$\chi_{0K-1}(\lambda_2 I_{nm}) + \chi_{0K}[(U + pU^0\gamma) \otimes I_m - \eta I_{nm}] + \chi_{4K}(e_n \otimes \eta\gamma \otimes I_m) = 0.$$

$$\chi_{00}(qU^0\gamma' \otimes I_m) + \chi_{10}[(U' + pU'^0\gamma') \otimes I_m - (\lambda_2 + \psi_2)I_{n'm}] + \chi_{20}(T'^0\gamma' \otimes I_m) + \chi_{30}(T^0\gamma' \otimes I_m) + \chi_{50}(\eta\gamma' \otimes I_m) = 0.$$

$$\chi_{1j-1}(\lambda_2 I_{n'm}) + \chi_{1j}[(U' + pU'^0\gamma') \otimes I_m - (\lambda_2 + \psi_2)I_{n'm}] + \chi_{5j}(\eta\gamma' \otimes I_m) = 0,\ for\ 1 \le j \le K-1.$$

$$\chi_{1L-1}(\lambda_2 I_{n'm}) + \chi_{1L}[(U' + pU'^0\gamma') \otimes I_m] = 0.$$

$$\chi_{10}[((e'_n \otimes \psi_2\alpha') + qU'^0\gamma') \otimes I_m] + \chi_{20}(T' \otimes I_m - \lambda_2 I_{l'm}) = 0.$$

$$\chi_{0j-1}[e_n \otimes \psi_2\alpha' \otimes I_m] + \chi_{1j}[e'_n \otimes \psi_2\alpha' \otimes I_m] + \chi_{2j-1}(\lambda_2 I_{l'm}) + \chi_{2j}(T' \otimes I_m - \lambda_2 I_{l'm}) = 0, \quad for\ 1 \le j \le K-1.$$

$$\chi_{0K}[e_n \otimes \psi_2\alpha' \otimes I_m] + \chi_{1K}[e'_n \otimes \psi_2\alpha' \otimes I_m] + \chi_{2K-1}(\lambda_2 I_{l'm}) + \chi_{2K}(T' \otimes I_m) = 0.$$

$$\chi_{30}(T \otimes I_m - \lambda_2 I_{lm}) = 0,$$

$$\chi_{3j-1}(\lambda_2 I_{lm}) + \chi_{3j}(T \otimes I_m - \lambda_2 I_{lm}) = 0, \quad for\ 1 \le j \le K-1.$$

$$\chi_{3K-1}(\lambda_2 I_{lm}) + \chi_{3K}(T \otimes I_m) = 0.$$

$$\chi_{40}[\theta U \otimes I_m - (\eta + \lambda_2)I_{nm}] + \chi_{41}(\theta U' \otimes I_m) + \chi_{50}(\alpha \otimes \lambda_2 I_m) = 0.$$

$$\chi_{4j-1}(\lambda_2 I_{nm}) + \chi_{4j}[\theta U \otimes I_m - (\eta + \lambda_2)I_{nm}] + \chi_{4j+1}(\theta U' \otimes I_m) + \chi_{5j}(\alpha \otimes \lambda_2 I_m) = 0, \quad for\ 1 \le j \le K-1.$$

$$\chi_{4K-1}(\lambda_2 I_{nm}) + \chi_{4K}[\theta U \otimes I_m - \eta I_{nm}] + \chi_{5K}(\alpha \otimes \lambda_2 I_m) = 0.$$

$$\chi_{4K-1}(e_1 K \otimes \theta U^0 \otimes I_m) - \chi_{5K}(\eta + \lambda_2)I_m = 0.$$

Subject to normalizing condition

$$\sum_{r=1}^{K} \chi_{0r} e_{nm} + \sum_{r=0}^{K} \chi_{1r} e_{n'm} + \sum_{r=0}^{K} \chi_{2r} e_{l'm} + \sum_{r=0}^{K} \chi_{3r} e_{lm} + \sum_{r=1}^{K} \chi_{4r} e_n m + \chi_{50} e_m = 1.$$

The stability condition $\chi A_0 e < \chi A_2 e$ is obtained after some algebraic simplification, which turns out to be

$$\sum_{r=1}^{K} \chi_{0r}(e_n \otimes D_1 e_m) + \sum_{r=0}^{K} \chi_{1r}(e_{n'} \otimes D_1 e_m) + \sum_{r=0}^{K} \chi_{2r}(e_{l'} \otimes D_1 e_m) + \sum_{r=0}^{K} \chi_{3r}(e_l \otimes D_1 e_m) + \sum_{r=1}^{K} \chi_{4r}(e_n \otimes (1-b)D_1 e_m) + \chi_{50}(1-b)D_1 e_m < \sum_{r=0}^{K} \chi_{1r}(qU'^0 \otimes e_m).$$

4.2. The Stationary Probability Vector

Let y be the stationary probability vector of the infinitesimal generator Q of the process {Y(t): $t \ge 0$}. The subdivision of y by level as, $y = (y_0, y_1, y_2, ...)$, where y_0 is of dimension $(3m + 2Knm + Kl'm + (K+1)lm)$ for $i = 0$ and $y_1, y_2, ...$ are of dimension $Kmn + (K+1)n'm + (K+1)l'm + (K+$

$1)lm + Kmn + m$ for $i \geq 1$. As y is a stationary probability vector satisfies the relation $yQ = 0$ and $ye = 1$. Furthermore, while the stability criterion is satisfied, the equation gives the various levels.

$$y_j = y_1 R^{j-1},\ j \geq 2 \tag{3}$$

where R is the smallest non-negative solution of the quadratic equation

$$R^2 A_2 + R A_1 + A_0 = 0$$

and satisfies the relation $RA_2 e = A_0 e$ and the vector y_0, y_1 are obtained with the help of succeeding equations:

$$y_0 B_{00} + y_1 B_{10} = 0, \tag{4}$$

$$y_0 B_{01} + y_1 [A_1 + R A_2] = 0, \tag{5}$$

subject to normalizing condition

$$y_0 e_0 + y_1 [I - R]^{-1} e_1 = 1. \tag{6}$$

As a result, we can compute matrix R using Logarithmic reduction algorithm in Latouche and Ramaswami[15] and the vector y by using the special structure of something like the coefficient matrices.

5. Busy Period Analysis

• In a single-server queueing demonstration, the word busy period is characterized as the length of time between the entry of a customer into the void system and the first time from that point that the system size reaches zero. As, the first passage epoch to level zero, starting from level one. It is the first return time of level zero, taken after by a least one visit to a few other levels, which is the analog of the busy cycle.

• We have to present an outline of the fundamental period to analyze the busy period. when the QBD process is taken into thought the first passage time from level i to $i-1$, where $i \geq 2$.

• It is worth pointing out that for each level i, $i \geq 2$, there are $(3m + 3nm + lm)$ states. The state (i, j) of level i signifies the j^{th} state of level i when the states are sorted alphabetically.

• Let $\mathsf{G}_{jj'}(u, y)$ represent the conditional probability that the QBD process, starting at time $t = 0$ in the state (i, j) and keep track of the time until the first visit to the level $(i-1)$ but not later than time y. We can modify after exactly u transitions to the left and enter the state (i, j'), $t = 0$.

Let the joint transform matrix

$$\overline{\mathsf{G}}_{jj'}(z, s) = \sum_{u=1}^{\infty} z^u \int_0^{\infty} e^{-sy} d\mathsf{G}_{jj'}(u, y)\ \ ; |z| \leq 1,\ Re(s) \geq 0, \tag{7}$$

and put the matrix $\overline{\mathsf{G}}(z, s) = \overline{\mathsf{G}}_{jj'}(z, s)$. Specifically, computed the matrix $\overline{G}(z, s)$ satisfy the equation,

$$\overline{\mathsf{G}}(z, s) = z(SI - A_1)^{-1} A_2 + (SI - A_1)^{-1} A_0 \overline{\mathsf{G}}^2(z, s). \tag{8}$$

The matrix $\mathsf{G} = \mathsf{G}_{jj'} = \overline{\mathsf{G}}(1, 0)$ is concerned with negating the boundary states during the first passage times. knowing the rate matrix R allows us to use the below result to find the matrix G

$$\mathsf{G} = -(A_1 + R A_2)^{-1} A_2. \tag{9}$$

The matrix G can be found with the assistance of the Logarithmic reduction algorithm [15]. We find the matrix with the succeeding equation

$$\overline{\mathsf{G}}^{(1,0)}(z,s) = z(sI - A_1)^{-1}B_{10} + (sI - A_1)^{-1}A_0\overline{\mathsf{G}}(z,s)\overline{\mathsf{G}}^{(1,0)}(z,s) \tag{10}$$

$$\overline{\mathsf{G}}^{(0,0)}(z,s) = (sI - B_{00})^{-1}B_{01}\overline{\mathsf{G}}^{(1,0)}(z,s). \tag{11}$$

Thus, the moments that obey are calculated using the matrices G, $\overline{\mathsf{G}}^{(0,0)}(1,0)$ and $\overline{\mathsf{G}}^{(1,0)}(1,0)$ are stochastic at $z = 1$ and $s = 0$. We can find the moments as follows:

$$\vec{\mathcal{F}}_1 = -\frac{\partial}{\partial s}\overline{\mathsf{G}}(z,s)e = -[A_1 + A_0(I + \mathsf{G})]^{-1}e, \tag{12}$$

$$\vec{\mathcal{F}}_2 = \frac{\partial}{\partial z}\overline{\mathsf{G}}(z,s)e = -[A_1 + A_0(I + \mathsf{G})]^{-1}A_2e, \tag{13}$$

$$\vec{\mathcal{F}}_1^{(1,0)} = -\frac{\partial}{\partial s}\overline{\mathsf{G}}^{(1,0)}(z,s)e = -[A_1 + A_0\mathsf{G}]^{-1}(A_0\vec{\mathcal{F}}_1 + e), \tag{14}$$

$$\vec{\mathcal{F}}_2^{(1,0)} = \frac{\partial}{\partial z}\overline{\mathsf{G}}^{(1,0)}(z,s)e = -[A_1 + A_0\mathsf{G}]^{-1}(A_0\vec{\mathcal{F}}_2 + B_{10}e), \tag{15}$$

$$\vec{\mathcal{F}}_1^{(0,0)} = -\frac{\partial}{\partial s}\overline{\mathsf{G}}^{0,0}(z,s)e = -B_{00}^{-1}[B_{01}\vec{\mathcal{F}}_1^{(1,0)} + e], \tag{16}$$

$$\vec{\mathcal{F}}_2^{(0,0)} = \frac{\partial}{\partial z}\overline{\mathsf{G}}^{(0,0)}(z,s)e = -B_{00}^{-1}[B_{01}\vec{\mathcal{F}}_2^{(1,0)}]. \tag{17}$$

6. System Performance Measures

- Expected number of LP customers in the system
  $E_{LP} = \sum_{i=1}^{\infty} iy_ie$.
- Probability that the server is idle
  $P_{idle} = \sum_{a=1}^{m} y_{000a}$.
- Probability that the server busy with HP customers
  $P_{Hbusy} = \sum_{i=0}^{\infty}\sum_{r=1}^{K}\sum_{k_1=1}^{n}\sum_{a=1}^{m} y_{ir1k_1a}$
- Probability that the server is on hard failure
  $P_{HF} = \sum_{r=1}^{K}\sum_{k_4=1}^{l'}\sum_{a=1}^{m} y_{0r3k_4a} + \sum_{i=1}^{\infty}\sum_{r=0}^{K}\sum_{k_4=1}^{l'}\sum_{a=1}^{m} y_{ir3k_4a}$
- Probability that the server is Delay time to go for vacation
  $P_{DT} = \sum_{a=1}^{m} y_{004a}$
- Probability that the server is busy during working vacation
  $P_{BWV} = \sum_{i=0}^{\infty}\sum_{r=0}^{K}\sum_{k_1=1}^{n}\sum_{a=1}^{m} y_{ir6k_1a}$

7. Numerical Implementation

To compute numerical outcomes, we have employed distinct MAP representations for the arrival process in a manner that ensures their mean values are 1, as recommended by Chakravarthy [6].

Erlang of order 2 (ERL-A):

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}.$$

Exponential (EXP-A):

$$D_0 = [-1], \quad D_1 = [1].$$

Hyper exponential (HYP-A):

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}.$$

MAP-Negative Correlation (MAP-NC-A):

$$D_0 = \begin{bmatrix} -1.00243 & 1.00243 & 0 \\ 0 & -1.00243 & 0 \\ 0 & 0 & -225.797 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.01002 & 0 & 0.99241 \\ 223.539 & 0 & 2.258 \end{bmatrix}.$$

MAP-Positive Correlation (MAP-PC-A):

$$D_0 = \begin{bmatrix} -1.00243 & 1.00243 & 0 \\ 0 & -1.00243 & 0 \\ 0 & 0 & -225.797 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.99241 & 0 & 0.01002 \\ 2.258 & 0 & 223.539 \end{bmatrix}.$$

Let us consider PH-distributions for the service and repair process as follows:

Erlang of order 2 (ERL-S):

$$\gamma = \gamma' = [1, 0], \quad U = U' = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}.$$

Erlang of order 2 (ERL-R):

$$\alpha = \alpha' = [1, 0], \quad T = T' = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}.$$

Exponential (EXP-S):

$$\gamma = \gamma' = [1], \quad U = U' = [-1].$$

Exponential (EXP-R):

$$\alpha = \alpha' = [1], \quad T = T' = [-1].$$

Hyper exponential (HYP-S):

$$\gamma = \gamma' = [0.8, 0.2], \quad U = U' = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}.$$

Hyper exponential (HYP-R):

$$\alpha = \alpha' = [0.8, 0.2], \quad T = T' = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}.$$

7.1. Illustration 1

We have examined the consequence of the hard failure rate ψ_2 against the Expected number of LP customers in the system(E_{LP}) in the following tables 1 - 3. Fix $\mu_1 = 20$, $\mu_2 = 15$, $K = 5$, $\lambda_1 = 1$, $\lambda_2 = 1.5$, $\eta = 8$, $\omega = 0.5$, $\psi_1 = 0.5$, $\tau_1 = 2$, $\tau_2 = 6$, $\theta = 0.6$, $b = 0.7$, $p_1 = 0.3$, $q_1 = 0.7$ such that the system is stable.

- As the hard failure rate (ψ_2) increases, the variety of arrangements of arrival and service times than the corresponding E_{LP} also increases.
- Observe the arrival times, E_{LP} increases highly in $MAP - PC - A$ and increases much slower in $ERL - A$ than all other arrival times.

7.2. Illustration 2

We investigated the impact of the vacation rate (η) against the probability of the server being idle (P_{idle}) in the following tables 4 - 6. Fix $\mu_1 = 20$, $\mu_2 = 15$, $K = 5$, $\lambda_1 = 1$, $\lambda_2 = 1.5$, $\omega = 0.5$, $\psi_1 = 0.5$, $\psi_2 = 1$, $\tau_1 = 2$, $\tau_2 = 6$, $\theta = 0.6$, $b = 0.7$, $p_1 = 0.3$, $q_1 = 0.7$ such that the system is stable.

- As the vacation rate (η) increases, the variety of arrangements of arrival and service times than the corresponding P_{idle} also increases.
- While comparing to $EXP - S$ and $HYP - S$, P_{idle} increases more rapidly for $ERL - S$. Similarly, P_{idle} increases slowly for $HYP - S$.

7.3. Illustration 3

We analyze the effect of the repair rate (τ_2) on the probability of the server being busy for HP customer (P_{Hbusy}) in the following tables 7 - 9. Fix $\mu_1 = 16$, $\mu_2 = 15$, $K = 5$, $\lambda_1 = 1$, $\lambda_2 = 1.5$, $\eta = 8$, $\omega = 0.5$, $\psi_1 = 0.5$, $\psi_2 = 1$, $\tau_1 = 2$, $\theta = 0.6$, $b = 0.7$, $p = 0.3$, $q = 0.7$ such that the system is stable.

- While maximizing the repair rate (τ_2), P_{Hbusy} minimizes for various possible arrangements of arrival and service times.
- When correlating the distinct arrival times, P_{Hbusy} decreases more quickly in the case of $MAP - PC - A$ whereas slowly in $ERL - A$. Similarly, considering the service times, P_{Hbusy} decreases gradually in $ERL - S$ and highly in $HYP - S$.

7.4. Illustration 4

To determine the existence of the service rate of HP customer (μ_2) versus the expected system size for LP customer (E_{LP}) in Figures 1 - 5. Fix $\mu_2 = 15$, $K = 5$, $\lambda_1 = 1$, $\lambda_2 = 1.5$, $\eta = 8$, $\omega = 0.5$, $\psi_1 = 0.5$, $\psi_2 = 1$, $\tau_1 = 2$, $\tau_2 = 6$, $\theta = 0.6$, $b = 0.7$, $p_1 = 0.3$, $q_1 = 0.7$ such that the system remains stable.

A quick observation from Figures 1 - 5, E_{LP} decreases while increasing the service rate of HP customers for all combinations of arrival and service time groupings. Due to the availability of the HP service rate in the system, the customers will get service successfully which leads to E_{LP} decreases. However, E_{LP} decreases slowly for $ERL - A$ with the combination of $ERL - S$ whereas slowly in $HYP - S$. Likewise, E_{LP} decreases highly for $HYP - A$ in $HYP - S$ whereas slowly in $ERL - S$.

7.5. Illustration 5

To see the features of both the HP service rate (μ_1) and repair rate of hard failure (τ_2) on the expected number of LP customers in the system (E_{LP}) in the Figures 6 - 10. Fix $\mu_2 = 15$, $K = 5$, $\lambda_1 = 1$, $\lambda_2 = 1.5$, $\eta = 8$, $\omega = 0.5$, $\psi_1 = 0.5$, $\psi_2 = 1$, $\tau_1 = 2$, $\theta = 0.6$, $b = 0.7$, $p = 0.3$, $q = 0.7$ such that stability condition is satisfied.

Observation in Figures 6 - 10, we increase the values of both the HP service rate and repair rate of hard failure, then E_{LP} decreases with various arrival groupings. Due to the HP customer increase in the service rate, E_{LP} decreases likewise increase the repair rate of hard failure decrease in the E_{LP}. Let's look at the arrival times, E_{LP} decreases slowly for $ERL - A$ and decreases fastly for $MAP - PC - A$.

Table 1: *Hard Failure rate (ψ_2) vs E_{LP} - ERL-S*

ψ_2	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
1	0.163936	0.206320	0.293281	0.330092	20.674732
1.2	0.172609	0.217533	0.312405	0.342647	21.257457
1.4	0.181624	0.229181	0.332498	0.355637	21.857954
1.6	0.190997	0.241285	0.353617	0.369083	22.477042
1.8	0.200747	0.253867	0.375822	0.383009	23.115590
2	0.210893	0.266951	0.399180	0.397438	23.774519
2.2	0.221454	0.280563	0.423761	0.412394	24.454807
2.4	0.232454	0.294730	0.449638	0.427907	25.157494
2.6	0.243915	0.309480	0.476893	0.444003	25.883685

Table 2: *Hard Failure rate (ψ_2) vs E_{LP} - EXP-S*

ψ_2	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
1	0.174277	0.217929	0.311487	0.340494	20.533617
1.2	0.182994	0.229005	0.330339	0.352724	21.059251
1.4	0.191976	0.240406	0.349917	0.365273	21.596139
1.6	0.201232	0.252143	0.370253	0.378153	22.144696
1.8	0.210771	0.264230	0.391379	0.391377	22.705353
2	0.220605	0.276679	0.413329	0.404958	23.278556
2.2	0.230744	0.289502	0.436138	0.418908	23.864775
2.4	0.241200	0.302715	0.459844	0.433243	24.464498
2.6	0.251986	0.316331	0.484485	0.447976	25.078233

Table 3: *Hard Failure rate (ψ_2) vs E_{LP} - HYP-S*

ψ_2	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
1	0.233555	0.282009	0.402324	0.416602	19.623165
1.2	0.240760	0.290404	0.415758	0.425489	19.866887
1.4	0.247923	0.298753	0.429141	0.434323	20.109536
1.6	0.255059	0.307070	0.442498	0.443121	20.351499
1.8	0.262179	0.315369	0.455850	0.451895	20.593109
2	0.269292	0.323660	0.469217	0.460658	20.834659
2.2	0.276407	0.331955	0.482615	0.469419	21.076406
2.4	0.283532	0.340261	0.496061	0.478189	21.318580
2.6	0.290674	0.348587	0.509566	0.486975	21.561381

Table 4: *Vacation rate (η) vs P_{idle} - ERL-S*

η	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
5	0.082706	0.088338	0.098335	0.092979	0.094606
6	0.084704	0.090362	0.100285	0.094971	0.096400
7	0.086239	0.091918	0.101790	0.096500	0.097789
8	0.087467	0.093162	0.102999	0.097723	0.098910
9	0.088478	0.094188	0.103999	0.098732	0.099840
10	0.089331	0.095053	0.104845	0.099583	0.100629
11	0.090063	0.095797	0.105573	0.100315	0.101311
12	0.090700	0.096444	0.106210	0.100952	0.101907
13	0.091261	0.097015	0.106772	0.101514	0.102435

Table 5: *Vacation rate (η) vs P_{idle} - EXP-S*

η	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
5	0.084405	0.090174	0.100258	0.095251	0.096557
6	0.086462	0.092258	0.102267	0.097311	0.098407
7	0.088037	0.093853	0.103811	0.098887	0.099835
8	0.089292	0.095124	0.105045	0.100142	0.100980
9	0.090321	0.096167	0.106061	0.101172	0.101925
10	0.091185	0.097042	0.106916	0.102036	0.102722
11	0.091922	0.097789	0.107648	0.102775	0.103407
12	0.092561	0.098437	0.108283	0.103415	0.104002
13	0.093121	0.099006	0.108842	0.103977	0.104527

Table 6: *Vacation rate (η) vs P_{idle} - HYP-S*

η	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
5	0.090101	0.096426	0.106953	0.102570	0.103410
6	0.092254	0.098602	0.109047	0.104734	0.105339
7	0.093876	0.100237	0.110625	0.106360	0.106796
8	0.095147	0.101518	0.111863	0.107631	0.107942
9	0.096173	0.102552	0.112863	0.108656	0.108870
10	0.097022	0.103406	0.113691	0.109503	0.109639
11	0.097737	0.104125	0.114389	0.110217	0.110290
12	0.098349	0.104741	0.114987	0.110827	0.110848
13	0.098880	0.105275	0.115507	0.111356	0.111333

Table 7: *Repair rate (τ_2) vs P_{Hbusy} - ERL-S*

τ_2	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
5	0.105099	0.105226	0.105427	0.105573	0.105653
6	0.105061	0.105172	0.105341	0.105479	0.105548
7	0.105034	0.105135	0.105284	0.105412	0.105472
8	0.105014	0.105107	0.105243	0.105362	0.105416
9	0.104998	0.105086	0.105213	0.105323	0.105373
10	0.104986	0.105070	0.105190	0.105292	0.105339
11	0.104976	0.105056	0.105171	0.105267	0.105311
12	0.104968	0.105045	0.105156	0.105246	0.105288
13	0.104961	0.105036	0.105143	0.105228	0.105268

Table 8: *Repair rate (τ_2) vs P_{Hbusy} - EXP-S*

τ_2	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
5	0.103623	0.103745	0.103941	0.104090	0.104179
6	0.103590	0.103698	0.103864	0.104004	0.104081
7	0.103565	0.103663	0.103811	0.103941	0.104009
8	0.103545	0.103637	0.103772	0.103892	0.103954
9	0.103530	0.103616	0.103742	0.103854	0.103911
10	0.103518	0.103600	0.103719	0.103824	0.103876
11	0.103509	0.103587	0.103701	0.103798	0.103848
12	0.103500	0.103576	0.103685	0.103778	0.103824
13	0.103493	0.103567	0.103673	0.103760	0.103805

Table 9: *Repair rate (τ_2) vs P_{Hbusy} - HYP-S*

τ_2	$ERL-A$	$EXP-A$	$HYP-A$	$NC-A$	$PC-A$
5	0.095098	0.095149	0.095235	0.095317	0.095365
6	0.095158	0.095216	0.095310	0.095423	0.095481
7	0.095182	0.095241	0.095332	0.095455	0.095514
8	0.095191	0.095247	0.095335	0.095458	0.095515
9	0.095192	0.095247	0.095329	0.095449	0.095503
10	0.095190	0.095243	0.095321	0.095436	0.095487
11	0.095187	0.095238	0.095313	0.095422	0.095470
12	0.095183	0.095232	0.095304	0.095408	0.095453
13	0.095180	0.095227	0.095296	0.095394	0.095438

(a) ERL-A

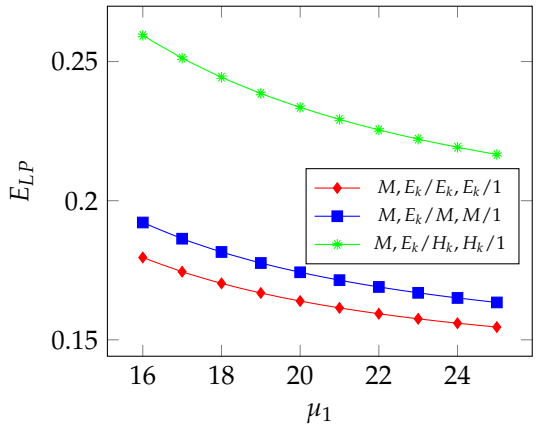


Figure 1: *High priority service rate vs.* E_{LP}

(b) EXP-A

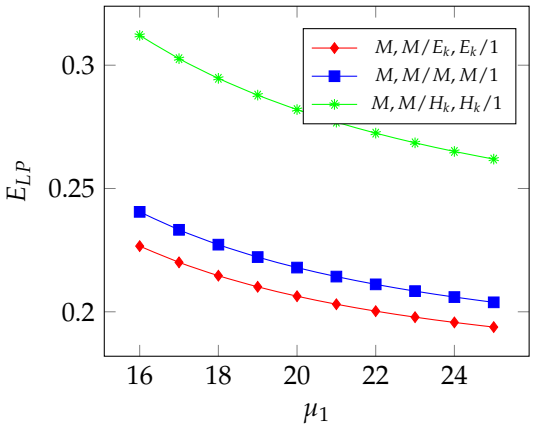


Figure 2: *High priority service rate vs.* E_{LP}

(c) HYP-A

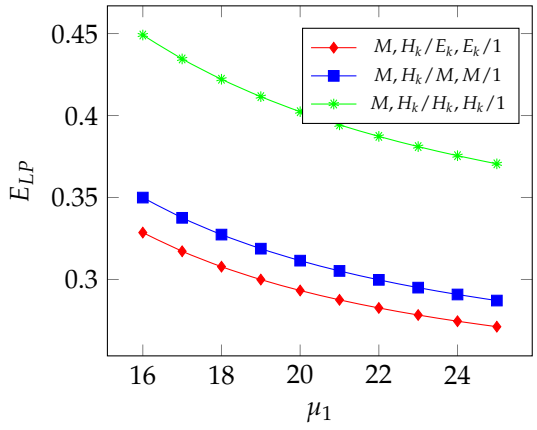


Figure 3: *High priority service rate vs.* E_{LP}

(d) MAP-NC-A

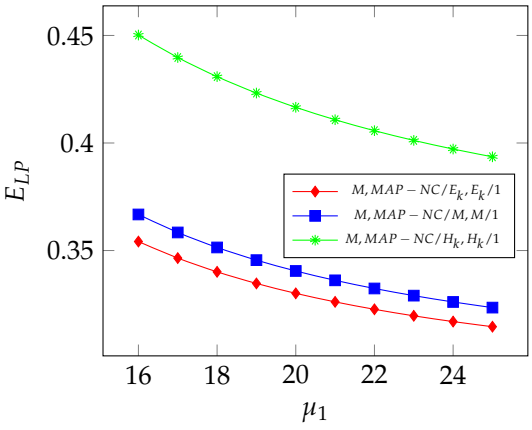


Figure 4: *High priority service rate vs.* E_{LP}

(e) MAP-PC-A

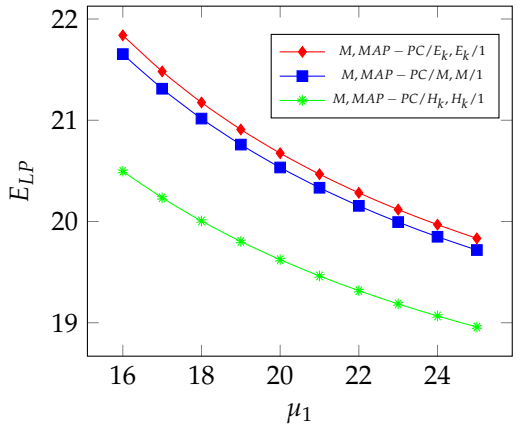


Figure 5: *High priority service rate vs.* E_{LP}

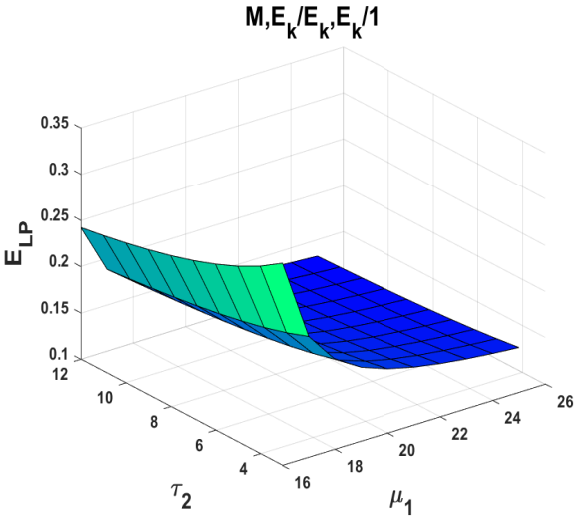


Figure 6: *HP service* (μ_1) *and Repair(HF)* (τ_2) *rates vs.* E_{LP} *- ERL-S*

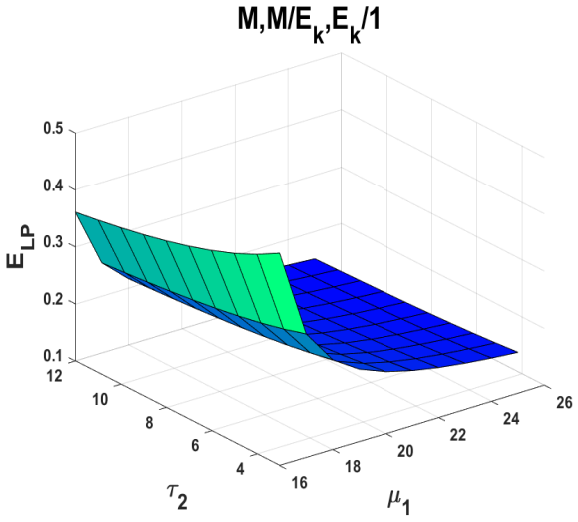


Figure 7: *HP service (μ_1) and Repair(HF) (τ_2) rates vs. E_{LP} - ERL-S*

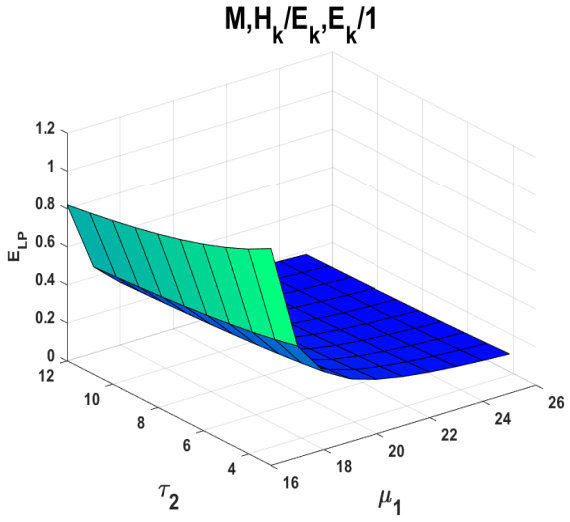


Figure 8: *HP service (μ_1) and Repair(HF) (τ_2) rates vs. E_{LP} - ERL-S*

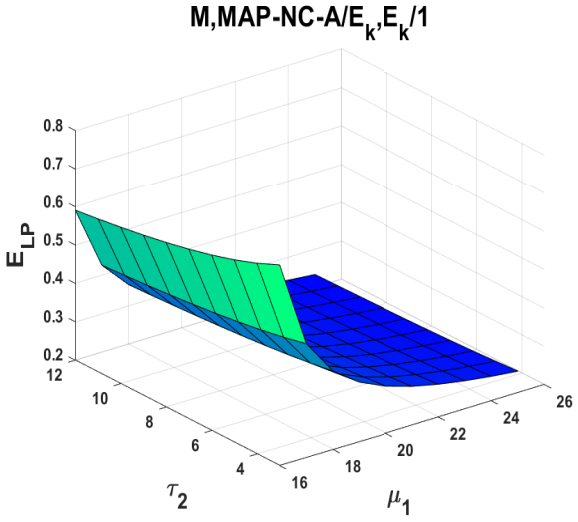


Figure 9: *HP service (μ_1) and Repair(HF) (τ_2) rates vs. E_{LP} - ERL-S*

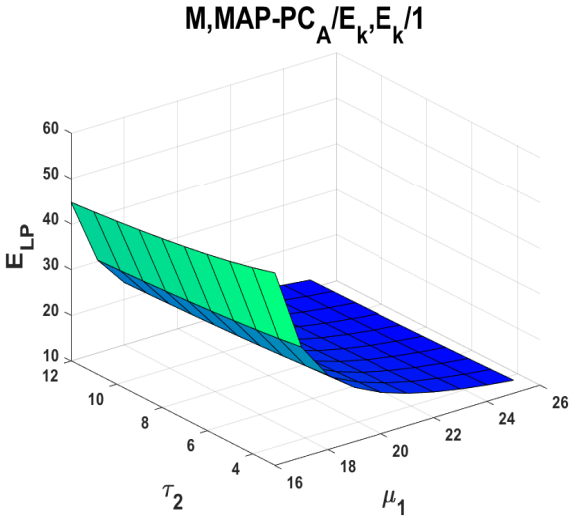


Figure 10: *HP service (μ_1) and Repair(HF) (τ_2) rates vs. E_{LP} - ERL-S*

8. Conclusion

This paper contributes by employing the Matrix analytic method to compute the stationary distribution of the number of customers in the $M, MAP/PH_1, PH_2/1$ queueing system with delayed working vacations under non-preemptive priority. We discussed some system performance measures using steady-state probabilities and also calculated busy period analysis. We used numerical examples to show how different system parameters affect performance measures.

References

[1] Ayyappan, G. and Thilagavathy, K. (2021). Analysis of MAP/PH/1 queueing model with immediate feedback, starting failures, single vacation, standby server, delayed repair, breakdown and impatient customers, *International Journal of Mathematics in Operational Research*, 19(3):269-301.

[2] Ayyappan, G. and Thilagavathy, K. (2021). Analysis of MAP/PH/1 Queueing model with Setup, Closedown, Multiple Vacations, Standby Server, Breakdown, Repair and Reneging, *Reliability: Theory and Applications*, 15(2):104–143.

[3] Ayyappan, G. and Somasundaram, B. (2019). Analysis of two-stage $M^{X_1}, M^{X_2}/G_1, G_2/1$ retrial G-queue with discretionary priority services, working breakdown, Bernoulli vacation, preferred and impatient units, *Applications and Applied Mathematics: An International Journal*, 14(2):640-671.

[4] Badamchi Zadeh, A. and Shahkar, G. H. (2008). A Two Phases Queue System with Bernoulli Feedback and Bernoulli Schedule Server Vacation, *International Journal of Information and Management Sciences*, 19:329–338.

[5] Bhagat, A. and Jain, M. (2020). Retrial queue with multiple repairs, multiple services, and non-preemptive priority, *OPSEARCH*, 57(3):1-28.

[6] Chakravarthy, S.R. (2011). Markovian arrival processes, *Wiley Encyclopedia of Operation Research and Management Science*.

[7] Jain, M. and Jain, A (2010). Working vacation queueing model with multiple types of server breakdowns, *Applied Mathematical Modelling*, 34:1-13.

[8] Janani, B. (2022).Transient Analysis of Differentiated Breakdown Model, *Applied Mathematics and Computation*, 417: p.126779.

[9] Jeganathan, K., Anbazhagan, N. and Kathiresan, J. (2013). A Retrial Inventory System with Non-Preemptive Priority Service, *International Journal of Information and Management Sciences*, 24(1):57-77.

[10] Kalyanaraman, R. and Sundaramoorthy, A. (2019). A Markovian working vacation queue with server state-dependent arrival rate and with partial breakdown, *International Journal of Recent Technology and Engineering*, 7:664-668.

[11] Krishna Reddy, G. V., and Anitha, R. (1998). Markovian bulk service queue with delayed vacations, *Computers and Operations Research*, 25(12):1159–1166.

[12] Krishna Kumar, B., Rukmani, R., Thanikachalam, A., and Kanakasabapathi, V. (2008). Performance analysis of retrial queue with server subject to two types of breakdowns and repairs, *Operations Research*, 18(2):521-559.

[13] Krishnamoorthy, A. and Divya, V.(2020). (M, MAP)/(PH, PH)/1 Queue with Non-preemptive Priority and Working Vacation Under N-Policy, *Journal of the Indian Society for Probability and Statistics*, 21(1): 69-122

[14] Kumar, B. (2018). An unreliable bulk queueing model with optional services, Bernoulli vacation schedule and balking, *International Journal of Mathematics in Operational Research*, 12(3):293-316.

[15] Latouche, G., and Ramaswami, V. (1993). A Logarithmic Reduction Algorithm for Quasi-Birth-Death Processes, *Journal of Applied Probability*, 30:650-674.

[16] Li, L., Wang, J., and Zhang, F. (2013). Equilibrium customer strategies in Markovian queues with partial breakdowns, *Computers and Industrial Engineering*, 66:751–757.

[17] Lucantoni, D.M., Meier-Hellstern, K.S., and Neuts, M.F. (1990). A single server queue with server vacations and a class of non-renewal arrival processes, *Advances in Applied Probability*, 22:676-705.

[18] Neuts, M. F. (1979). A Versatile Markovian Point Process, *Journal of Applied Probability*, 16:764-779.

[19] Pikkala, V. L. and Vepada, S. (2014). Analysis of general input state-dependent working vacation queue with changeover time, *ISRN Computational Mathematics*, 8, Article ID 248704. https://doi.org/10.1155/2014/248704.

[20] Rakesh Kumar and Soodan, B.S. (2019). Transient Numerical Analysis of a Queueing Model with Correlated Reneging, Balking, and Feedbac,. *Reliability: Theory and Applications*, 14(4):46-54.

[21] Subramanian, A.M.G., Ayyappan, G. and Sekar, G. (2009). $M/M/1$ Retrial queueing system with negative arrival under non-preemptive priority service, *Journal of Fundamental Sciences*, 5(2):129-145.

[22] Yang, D.-Y., Chen, Y.-H. and Wu, C.-H. (2021). Sojourn times in a Markovian queue with working breakdowns and delayed working vacations. *Computers and Industrial Engineering* 156(107239):1-13.

Analysis of $MAP/PH_1, PH_2, PH_3/1$ Queueing-Inventory System with Two Commodities

S. Meena, N. Arulmozhi*, G. Ayyappan, K. Jeganathan

•

Department of Mathematics,
Puducherry Technological University,
Puducherry, India.
Ramanujan Institute for Advanced Study in Mathematics,
University of Madras,
Chennai 600005, India.
meenasundar2296@gmail.com, arulmozhisathya@gmail.com,
ayyappan@ptuniv.edu.in, kjeganathan@unom.ac.in

Abstract

In this work, a single server implements a two-commodity inventory queueing system. We assume that both commodities have a finite capacity. Customers arrive by a Markovian Arrival Process, there is a need for a single item, and either or both types of commodities are required, and this requirement is modeled using certain probabilities. The lead times are exponentially distributed, and the service times have a PH distribution. We use matrix analytical techniques to investigate the queueing inventory system and adopt an (s, S)-type replenishment policy that is dependent on the type of commodity. In the steady state, the joint and individual probability distribution of the E_{system}, inventory level, and server status is obtained. A few significant performance measures are attained. Our mathematical concept is then illustrated with a few numerical examples.

Keywords: Queueing-inventory; Markovian Arrival Process; Phase-type distribution; (s, S)-type policy; Two-Commodity.

AMS Subject Classification (2010): 60K25, 68M30, 90B22.

1. Introduction

Many researchers have been interested in the study of queueing inventory systems, and proposals involving two commodities have been made. Sigman and Levi [18] presented the M/G/1 queueing-inventory model with exponentially distributed lead time under light traffic in 1992. Several models with various ordering criteria have been developed to operate such systems. Balintfy [5] and Silver, E.A., [19] both contributed to the development of the joint ordering policy. A two-commodity inventory system with zero lead time and an equal demand process was examined, according to Krishnamoorthy et al. [12] and Anbazhagan and Arivarignan [2].

Neuts [15] developed, studied, and instructed MAP in 1984. Chakravarthy [8] derived the Markovian arrival process by depicting matrix (D_0, D_1) as the guideline for the MAP at the dimension m, where D_0 governs for no arrival, where D_1 governs for arrival. The generator of the matrix Q defined by $D = D_0 + D_1$ is an irreducible stochastic matrix. A single-server inventory system using Markovian Arrival Process (MAP)-based arrivals were studied by Paul Manuel et al. [16].

Yadavalli et al. [22] considered a two-commodity stochastic inventory system with joint and individual ordering policies, Poisson arrivals and lost sales. Anbazhagan et al. [3] for their

consideration of a two-commodity continuous review inventory system with substitutable items and Markovian demands. When the sum of the two commodities' on-hand inventory levels reaches a certain level s, reordering for supply is initiated. A two-commodity inventory problem was studied by Krishnamoorthy and Varghese [13] with no lead time and Markovian shifts in demand for the first, second, and both commodities.

Binitha Benny et al. [6] considered a total cost inventory system with a single server and the buffer capacity will be limited. Customers arrive through a Poisson process, and the probabilities used to determine the demand for each commodity or both commodities depend on which commodity is being purchased. Sivakumar et al. [20] investigated a total cost continuous review inventory system with a demand renewal and ordering policy, a policy combination known as ordering individual commodities and ordering both commodities jointly.

A two-commodity model with a compliment and regular working vacations is examined by Lakshmanan et al. in their study [14]. Each customer orders service at a convenient moment, and both commodities are independent of their ordering procedures. Each customer is given a finite retry orbit when the requested item is out of stock or the server is overloaded. Schwarz et al. [17] looked into a brand-new type of stochastic network that shows a product from steady-state distribution. There, integrated models for networks of service stations and inventories were constructed using stochastic networks. They assume that even though a server with associated inventory stops accepting new customers when the stock is out, lost sales are still recorded in the system.

According to Yadavalli et al. [23], the three types of demand for the two goods are comparable. They looked at a system with a phase-type distributed lead time and perishable items. A Markovian arrival process governs the occurrence of all three different kinds of demands. Each commodity's lifetime has an exponential distribution with unique properties. A continuous-time Markov chain that identified the system was used to give a stability analysis and identify individual ordering strategies. Amirthakodi [1] thought of an inventory system with one server service facility and a limited number of trial feedback customers. An inventory system with a single server, two commodities, queue-dependent services for a finite queue, and an optional retrial facility was examined by Jeganathan et al. [10].

Federgruen et al. [9] investigated a continuous review multi-item inventory system with demands generated by independent compound Poisson processes using the (S, c, s) ordering strategy. One consequence of implementing this approach is the requirement to find three optimal variables for each item. Kalpakam and Arivarignan [11] proposed a policy with fewer variables for making decisions and for an (s, S) policy generated by [11] that is appropriate for related but non-substitutable items, a single reorder level s is determined. The total cost is determined by the average inventory, a customer in queue, and reorder rates, according to Berman [7], who provided a deterministic approximation for their inventory system with a service facility.

The demand for each commodity occurs in independent Poisson processes with a variety of parameters in two-commodity retrial inventory systems with varied ordering strategies has been studied by Sivakumar [21] and Jeganathan and Anbazhagan [4]. The constant retrial policy was taken into consideration in both experiments. In other words, a signal is sent out when there are i demands in the orbit according to an exponential distribution that is independent of the orbit's number.

1.1. Motivation for the proposed model

The main motivating factor for our model is the Textile scenario. Buyers usually go to a Textile shop to purchase one or more (like churidar, sarees, shirts, kurtas, and so on) items or goods. Let's say there are n various items and people are shopping for the product i with probability, p_i, $1 \leq i \leq n$. Customers shop for objects $i_1, ..., i_k$, for $2 \leq k \leq n$ with probability $p_{i_1,...,i_k}$, where $i_1, ..., i_k$ is an element of the set of integers $1, 2, ..., n$. Customers will be served only those products

that are in stock of the ones requested if all of the requested different goods are not in stock. If a Buyer is unable to obtain any product, they will be disappointed. A customer has a $2^n - 1$ different possibility to shop for the products and we will concentrate on the case where $n = 2$.

1.2. Research Gap

Benny et al. [6] worked with two-commodity in the single server queueing inventory system and arrival follows the Poisson process and service follows an exponential distribution. This article examines two-commodity in the inventory with arrival following MAP and service times following Phase-type distributions. The authors handle (s, S) policy, and both individual and joint orders are obtained. In this article, we develop (s, S) policy, both individual and joint orders, and numerical implementation of 2D using Matlab software.

1.3. Viewpoint for This Work

The manuscript for this work is synchronized as follows: A brief explanation of our model is provided in Section 2. Our model's notations and matrix generation are described in Section 3. Section 4 contains our model's steady-state probability. Section 5 provides performance measures. Numerical illustrated in Section 6. The conclusion is given in Section 7.

2. Model Description

Consider a single server queueing model subject to a two-commodity. Customers arrive according to a MAP and each commodity has a single item demand. The MAP is specified by two m x m matrices (D_0, D_1), $D = D_0 + D_1$, which is an irreducible infinitesimal generator. The matrix D_0 means no arrival similarly, the matrix D_1 means arrival.

There is a need for a single unit, and either or both types of commodities are required, and this requirement is modeled using certain probabilities. The lead times are exponentially distributed, and the service times have a PH distribution. Customers may want both commodities or only one, depending on some predetermined probability. Only when services are being offered are the customers' needs disclosed. If the requested item is not available, the customer permanently exits the system. When only one of the requested items is available and both are demanded, the customer is given the one that is in stock. In the case where both commodity inventory levels are 0, customers are not allowed to join the system. However, customers join the system even when the server is operating and no more inventory is available. For the customer's needed item to be provided at the time the item is taken for service, it is planned that the items will be replenished during the current service. When a customer cannot get the commodity they need at the time of service, the customer is also lost.

When taken for service, the customer requests item I_i with probability c_i, for $i = 1, 2$ or both I_1 and I_2 with probability c_3 such that $c_1 + c_2 + c_3 = 1$. After a random period of service, the requested item is delivered to the customer. The service times for processing orders for I_1, I_2 or both $I_1 and I_2$ are PH- distribution with represented by (α_v, T_v), $1 \leq v \leq 3$. Whose matrix is order n_v with $T_v^0 + T_v e = 0$ implies that $T_v^0 = -T_v e$. Here λ is the arrival rate, which is signified as $\lambda = \pi_1 D_1 e$, where π_1 is the steady-state probability vector. The mean service rate is denoted by $\mu_v = [\alpha_v(-T_v)^{-1} e_{n_v}]^{-1}$.

For both commodities, the system has a maximum capacity of S_i items. We utilize a (s_i, S_i) replenishment strategy for the commodity I_i, where $i = 1, 2$. That is, an order is placed for just that item to raise the inventory level of commodity I_i back to S_i, i = 1, 2 at the time of replenishment, anytime it drops to s_i. For parameters, β_i, for $i = 1, 2$, the lead time has an exponential distribution.

3. The QBD process's infinitesimal generation matrix

The following notations and assumptions are used to explain our model of producing QBD processes in this section.

Notations

We will define the following notations:

- $\otimes$ -Kronecker product of two matrices of various dimensions resulting in a block matrix.
- $\oplus$ - Kronecker sum of two matrices of various dimensions resulting in a block matrix.
- I_m stand for identity matrix of m rows and m columns.
- e - A column vector of the suitable order. Each of its entries is one.
- $N(t)$ represents the total number of customers in the queue.
- $V(t)$ represents the server's status at epoch t.

$$V(t) = \begin{cases} 0, & \text{if server is idle} \\ 1, & \text{if server is busy with } I_1 \\ 2, & \text{if server is busy with } I_2 \\ 3, & \text{if server is busy with } I_1 \text{ and } I_2 \end{cases}$$

- $L_i(t)$ stands for the excess inventory level of commodity I_i, $i = 1, 2$.
- $S(t)$ stands for phases of the service.
- $M(t)$- The Markovian arrival process is considered in phases.
- Let Y=$\{Y(t) : t \geq 0\}$, where $Y(t) = \{N(t), V(t), I_1(t), I_1(t), S(t), M(t)\}$ is a $CTMC$ with state space

$$\Phi = \phi(0) \bigcup_{i=1}^{\infty} \phi(i). \tag{1}$$

where
$\phi(0) = \{(0, 0, a_1, a_2, k) : \ 0 \leq a_1 \leq S_1, \ 0 \leq a_2 \leq S_2, \ 1 \leq k \leq m\}$
$\cup\{(0, v, a_1, a_2, j_v, k) : \ 1 \leq v \leq 3, \ 0 \leq a_1 \leq S_1, \ 0 \leq a_2 \leq S_2, \ 1 \leq j_v \leq n_v, \ 1 \leq k \leq m\}$
and for $p \geq 1$,

$\phi(p) = \{(p, v, a_1, a_2, j_v, k) : \ 1 \leq v \leq 3, \ 0 \leq a_1 \leq S_1, \ 0 \leq a_2 \leq S_2, \ 1 \leq j_v \leq n_v, \ 1 \leq k \leq m\}$.

3.1. The Infinitesimal Generator Matrix

The infinitesimal generator matrix of the Markov chain is given by:

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & 0 & \dots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0 & 0 & \dots \\ B_{20} & A_2 & A_1 & A_0 & 0 & 0 & 0 & \dots \\ B_{30} & A_3 & A_2 & A_1 & A_0 & 0 & 0 & \dots \\ B_{40} & A_4 & A_3 & A_2 & A_1 & A_0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \dots \end{bmatrix} \tag{2}$$

The following describes Markov chain transitions and the corresponding rates:
The matrix B_{00} governs,

- $(0, v, a_1, a_2, j_v, k) \to (0, 0, a_1, a_2, k)$ with rate ${T_v}^0 \otimes I_m$ for $1 \leq v \leq 3$, $0 \leq a_1 \leq S_1$, $0 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,

- $(0,0,a_1,a_2,k) \rightarrow (0,0,S_1,a_2,k)$ with rate $\beta_1 I_m$ for $0 \leq v \leq 3$, $0 \leq a_1 \leq s_1$, $0 \leq a_2 \leq S_2$, $1 \leq k \leq m$,
- $(0,v,a_1,a_2,j_v,k) \rightarrow (0,v,S_1,a_2,j_v,k)$ with rate $\beta_1 I_{n_v m}$ for $1 \leq v \leq 3$, $0 \leq a_1 \leq s_1$, $0 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(0,0,a_1,a_2,k) \rightarrow (0,0,a_1,S_2,k)$ with rate $\beta_2 I_m$ for $0 \leq v \leq 3$, $0 \leq a_1 \leq S_1$, $0 \leq a_2 \leq s_2$, $1 \leq k \leq m$,
- $(0,v,a_1,a_2,j_v,k) \rightarrow (0,v,a_1,S_2,j_v,k)$ with rate $\beta_2 I_{n_v m}$ for $1 \leq v \leq 3$, $0 \leq a_1 \leq S_1$, $0 \leq a_2 \leq s_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(0,0,0,a_2,k) \rightarrow (0,2,0,a_2-1,j_2,k)$ with rate $\alpha_2 \otimes (c_2+c_3)D_1$ for $1 \leq a_2 \leq S_2$,$1 \leq j_2 \leq n_2$, $1 \leq k \leq m$,
- $(0,0,a_1,0,k) \rightarrow (0,1,a_1-1,0,j_1,k)$ with rate $\alpha_1 \otimes (c_1+c_3)D_1$ for $1 \leq a_1 \leq S_1$, $1 \leq j_1 \leq n_1$, $1 \leq k \leq m$,
- $(0,0,a_1,a_2,k) \rightarrow (0,1,a_1-1,a_2,j_1,k)$ with rate $\alpha_1 \otimes c_1 D_1$ for $1 \leq a_1 \leq S_1$, $1 \leq a_2 \leq S_2$, $1 \leq j_1 \leq n_1$, $1 \leq k \leq m$,
- $(0,0,a_1,a_2,k) \rightarrow (0,2,a_1,a_2-1,j_2,k)$ with rate $\alpha_2 \otimes c_2 D_1$ for $1 \leq a_1 \leq S_1$, $1 \leq a_2 \leq S_2$, $1 \leq j_2 \leq n_2$, $1 \leq k \leq m$,
- $(0,0,a_1,a_2,k) \rightarrow (0,3,a_1-1,a_2-1,j_3,k)$ with rate $\alpha_3 \otimes c_3 D_1$ for $1 \leq a_1 \leq S_1$, $1 \leq a_2 \leq S_2$, $1 \leq j_3 \leq n_3$, $1 \leq k \leq m$.

The matrix $B_{(p+1)0}$, $p \geq 1$, governs

- $(p,v,0,0,j_v,k) \rightarrow (0,0,0,0,k)$ with rate ${T_v}^0 \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(p,v,0,a_2,j_v,k) \rightarrow (0,0,0,a_2,k)$ with rate ${T_v}^0 {c_1}^p \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(p,v,0,a_2,j_v,k) \rightarrow (0,2,0,a_2-1,j_2,k)$ with rate ${T_v}^0 {c_1}^{p-1}(c_2+c_3)\alpha_2 \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(p,v,a_1,0,j_v,k) \rightarrow (0,0,a_1,0,k)$ with rate ${T_v}^0 {c_2}^p \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_1 \leq S_1$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(p,v,a_1,0,j_v,k) \rightarrow (0,2,a_1-1,0,j_1,k)$ with rate ${T_v}^0 {c_2}^{p-1}(c_1+c_3)\alpha_1 \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_1 \leq S_1$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(1,v,a_1,a_2,j_v,k) \rightarrow (0,1,a_1-1,a_2,j_1,k)$ with rate ${T_v}^0 c_1 \alpha_1 \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_1 \leq S_1$, $1 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(1,v,a_1,a_2,j_v,k) \rightarrow (0,2,a_1,a_2-1,j_2,k)$ with rate ${T_v}^0 c_2 \alpha_2 \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_1 \leq S_1$, $1 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(1,v,a_1,a_2,j_v,k) \rightarrow (0,3,a_1-1,a_2-1,j_3,k)$ with rate ${T_v}^0 c_3 \alpha_3 \otimes I_m$ for $1 \leq v \leq 3$, $1 \leq a_1 \leq S_1$, $1 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$.

The matrix A_1, $p \geq 1$, governs

- $(p,v,a_1,a_2,j_v,k) \rightarrow (p,v,S_1,a_2,j_v,k)$ with rate $\beta_1 I_{n_v m}$ for $1 \leq v \leq 3$, $0 \leq a_1 \leq s_1$, $0 \leq a_2 \leq S_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$,
- $(p,v,a_1,a_2,j_v,k) \rightarrow (p,v,a_1,S_2,j_v,k)$ with rate $\beta_2 I_{n_v m}$ for $1 \leq v \leq 3$, $0 \leq a_1 \leq S_1$, $0 \leq a_2 \leq s_2$, $1 \leq j_v \leq n_v$, $1 \leq k \leq m$.

The matrix A_{l+1}, $1 \leq l \leq p-1$, $p \geq 3$, governs

- $(p, v, 0, a_2, j_v, k) \to (p-l, 2, 0, a_2-1, j_2, k)$ with rate ${T_v}^0 {c_1}^{l-1}(c_2+c_3)\alpha_2 \otimes I_m$ for $1 \le v \le 3$, $1 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$,
- $(p, v, a_1, 0, j_v, k) \to (p-l, 1, a_1-1, 0, j_1, k)$ with rate ${T_v}^0 {c_2}^{l-1}(c_1+c_3)\alpha_1 \otimes I_m$ for $1 \le v \le 3$, $1 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$,
- $(p, v, a_1, a_2, j_v, k) \to (p-1, 1, a_1-1, a_2, j_1, k)$ with rate ${T_v}^0 c_1 \alpha_1 \otimes I_m$ for $1 \le v \le 3$, $1 \le a_1 \le S_1$, $1 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$,
- $(p, v, a_1, a_2, j_v, k) \to (p-1, 2, a_1, a_2-1, j_2, k)$ with rate ${T_v}^0 c_2 \alpha_2 \otimes I_m$ for $1 \le v \le 3$, $1 \le a_1 \le S_1$, $1 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$,
- $(p, v, a_1, a_2, j_v, k) \to (p-1, 3, a_1-1, a_2-1, j_3, k)$ with rate ${T_v}^0 c_3 \alpha_3 \otimes I_m$ for $1 \le v \le 3$, $1 \le a_1 \le S_1$, $1 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$.

The matrix B_{01}, governs

- $(0, v, a_1, a_2, j_v, k) \to (1, v, a_1, a_2, j_v, k)$ with rate $I_{n_v} \otimes D_1$ for $1 \le v \le 3$, $0 \le a_1 \le S_1$, $0 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$.

The matrix A_0, $p \ge 1$, governs

- $(p, v, a_1, a_2, j_v, k) \to (p+1, v, a_1, a_2, j_v, k)$ with rate $I_{n_v} \otimes D_1$ for $1 \le v \le 3$, $0 \le a_1 \le S_1$, $0 \le a_2 \le S_2$, $1 \le j_v \le n_v$, $1 \le k \le m$.

4. Analysis of Steady-State

The nonsingularity of B_{00} and A_1 is need for Q to be irreducible. Consider the matrix $A = \sum_{l=0}^{\infty} A_l$. Let the unique stationary distribution of A be ψ. Under the condition (Neuts [15]),

$$\psi A_0 e < \sum_{l=2}^{\infty} (l-1)\psi A_l e,$$

an irreducible Markov chain with generator Q possesses a unique stationary solution vector $Y = (y_0, y_1, y_2, \dots)$ satisfying

$$YQ = 0, \quad Ye = 1.$$

Partitioning Y as $Y = (y_0, y_1, y_2, \dots)$ where
$y_0 = (y_0(v, a_1, a_2, j_v, k) : 0 \le v \le 3,\ 0 \le a_1 \le S_1,\ 0 \le a_2 \le S_2,\ 1 \le j_v \le n_v,\ 1 \le k \le m)$,

$y_p = (y_p(v, a_1, a_2, j_v, k) : 1 \le v \le 3,\ 0 \le a_1 \le S_1,\ 0 \le a_2 \le S_2,\ 1 \le j_v \le n_v,\ 1 \le k \le m)$, $for\ p \ge 1$,

where y_0 is of dimension $1 \times (S_1+1)(S_2+1)m + (S_1+1)(S_2+1)n_1 m + (S_1+1)(S_2+1)n_2 m + (S_1+1)(S_2+1)n_3 m$ and y_p for $p \ge 1$, is of dimension $1 \times (S_1+1)(S_2+1)n_1 m + (S_1+1)(S_2+1)n_2 m + (S_1+1)(S_2+1)n_3 m$. Then Y is obtained from

$$y_p = y_1 R^{p-1}, \quad p \ge 2$$

where R is the minimal nonnegative solution of the matrix equation $\sum_{j=0}^{\infty} Y^j A_j = 0$.
The boundary equations are given by

$$\sum_{p=0}^{\infty} y_p B_{p0} = 0$$

$$y_0 B_{00} + \sum_{p=1}^{\infty} y_p A_p = 0$$

The normalizing condition $Ye = 1$ gives

$$y_0 e + y_1 [I - R]^{-1} e = 1$$

R matrix is obtained using the algorithm:

$$R(0) = 0$$

$$R(p+1) = -A_0 {A_1}^{-1} - R^2(p) A_2 {A_1}^{-1} - R^3(p) A_3 {A_1}^{-1} - \dots, \quad p \ge 0$$

5. Performance Measure

- Expected number of customers in the system, $E_N = \sum_{p=1}^{\infty} py_p$
- Expected number of customers demanding I_1 alone, $E_{I_1} = c_1 E_N$
- Expected number of customers demanding I_2 alone, $E_{I_2} = c_2 E_N$
- Expected number of customers demanding both I_1 and I_2, $E_{I_{12}} = c_3 E_N$
- Expected rate of replenishment for item I_1,

$$E_{RI_1} = \beta_1 \{ \sum_{a_1=0}^{s_1} \sum_{a_2=0}^{S_2} \sum_{k=1}^{m} y_0(0, a_1, a_2, k) + \sum_{p=0}^{\infty} \sum_{v=1}^{3} \sum_{a_1=0}^{s_1} \sum_{a_2=0}^{S_2} \sum_{j_v=1}^{n_v} \sum_{k=1}^{m} y_p(v, a_1, a_2, n_v, k) \}$$

- Expected rate of replenishment for item I_2,

$$E_{RI_2} = \beta_2 \{ \sum_{a_1=0}^{S_1} \sum_{a_2=0}^{s_2} \sum_{k=1}^{m} y_0(0, a_1, a_2, k) + \sum_{p=0}^{\infty} \sum_{v=1}^{3} \sum_{a_1=0}^{S_1} \sum_{a_2=0}^{s_2} \sum_{j_v=1}^{n_v} \sum_{k=1}^{m} y_p(v, a_1, a_2, n_v, k) \}$$

- Expected reorder rate of commodity I_1,

$$E_{R_1} = \mu_1 \sum_{p=0}^{\infty} \sum_{a_2=0}^{S_2} \sum_{j_1=1}^{n_1} \sum_{k=1}^{m} y_p(1, s_1 + 1, a_2, n_1, k)$$

- Expected reorder rate of commodity I_2,

$$E_{R_2} = \mu_2 \sum_{p=0}^{\infty} \sum_{a_1=0}^{S_1} \sum_{j_2=1}^{n_2} \sum_{k=1}^{m} y_p(2, a_1, s_2 + 1, n_2, k)$$

- Expected reorder rate of commodity I_1 and I_2,

$$E_{R_{12}} = \mu_3 \sum_{p=0}^{\infty} \sum_{j_3=1}^{n_3} \sum_{k=1}^{m} y_p(3, s_1 + 1, s_2 + 1, n_3, k)$$

6. Numerical Implementation

In this section, we examine the outcome of our system using numerical and graphical representations. The three different MAP representations are distinct with the following variance and correlation structures and their mean values are 1.

Arrival in Erlang of order 2(ERL-A):

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix} D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

Arrival in Exponential(EXP-A):

$$D_0 = [-1] D_1 = [1]$$

Arrival in Hyper exponential(HYP-EXP-A):

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix} D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}$$

Let us consider PH-distributions for the service process as follows:
ERL-S (Service in Erlang of order 2):

$$\alpha_1 = \alpha_2 = \alpha_3 = [1, 0] \quad T_1 = T_2 = T_3 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

EXP-S(Service in Exponential):

$$\alpha_1 = \alpha_2 = \alpha_3 = [1] \quad T_1 = T_2 = T_3 = [-1]$$

HYP-EXP-S(Service in Hyper exponential):

$$\alpha_1 = \alpha_2 = \alpha_3 = [0.8, 0.2] \quad T_1 = T_2 = T_3 = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

6.1. Illustration

In the following tables 1, 2 and 3, we have examined the impact of the arrival rate λ on the expected system size. Fix $S_1 = 8$, $S_2 = 10$, $s_1 = 2$, $s_2 = 3$, $\mu_1 = 2$, $\mu_2 = 3$, $\mu_3 = 4$, $\beta_1 = 2$, $\beta_2 = 3$, $c_1 = 0.1$, $c_2 = 0.1$, $c_3 = 0.8$.

Table 1: *Arrival rate(λ) vs E_N*

ERL-S			
λ	ERL-A	EXP-A	HYP-EXP-A
1	0.038353086	0.090270325	0.197559861
1.1	0.050688987	0.113781945	0.256893077
1.2	0.065624277	0.141308387	0.329498505
1.3	0.083549669	0.173400817	0.417940725
1.4	0.104926605	0.210715786	0.525249133
1.5	0.130305758	0.254041436	0.654975645
1.6	0.160351157	0.304331913	0.81126147
1.7	0.195872159	0.36275318	0.998927098
1.8	0.23786654	0.430744844	1.223609249
1.9	0.287579538	0.51010493	1.491978161
2.0	0.346586293	0.603108157	1.812077611

Table 2: *Arrival rate(λ) vs E_N*

EXP-S			
λ	ERL-A	EXP-A	HYP-EXP-A
1	0.062592538	0.120376635	0.254416346
1.1	0.08207371	0.151739687	0.328650275
1.2	0.105500217	0.188464919	0.418427098
1.3	0.133449138	0.231291358	0.526393629
1.4	0.166604203	0.281099839	0.655616534
1.5	0.205783572	0.338948174	0.809643341
1.6	0.251976048	0.406117251	0.992584752
1.7	0.306389033	0.484172226	1.20923196
1.8	0.370513054	0.575044848	1.465227137
1.9	0.446210016	0.681145853	1.767309959
2	0.535836027	0.80552094	2.123668949

Table 3: *Arrival rate(λ) vs E_N*

HYP-EXP-S			
λ	ERL-A	EXP-A	HYP-EXP-A
1	0.249031469	0.32018599	0.571239637
1.1	0.319730773	0.403892787	0.721444007
1.2	0.403462978	0.502016937	0.897044873
1.3	0.502065778	0.616543171	1.10107116
1.4	0.617713186	0.749810152	1.336958686
1.5	0.752987938	0.904583472	1.608618699
1.6	0.910971089	1.084145482	1.920517083
1.7	1.095353193	1.292405822	2.27776399
1.8	1.310572349	1.534037132	2.686212657
1.9	1.561985236	1.814640848	3.152564574
2	1.856077719	2.140947812	3.684475578

We observe that from the above tables 1, 2 and 3:

- As arrival rate (λ) increases, the variety of arrangements of arrival and service times then the corresponding E_N also increases.
- Observe the arrival times, E_N rises more quickly in the case of $HYP - EXP - A$ and more slowly in the case of $ERL - A$. Similarly, it rises gradually in the case of $ERL - S$ and rapidly in the case of $HYP - EXP - A$.

6.2. Illustration

We have investigated the consequence of the arrival rate λ against the Expected to reorder rate of commodity I_1 (E_{R_1})in the obeying table 4, 5 and 6. Fix $S_1 = 8$, $S_2 = 10$, $s_1 = 2$, $s_2 = 3$, $\mu_1 = 2$, $\mu_2 = 3$, $\mu_3 = 4$, $\beta_1 = 2$, $\beta_2 = 3$, $c_1 = 0.1$, $c_2 = 0.1$, $c_3 = 0.8$.

Table 4: *Arrival rate(λ) vs E_{R_1}*

ERL-S			
λ	ERL-A	EXP-A	HYP-EXP-A
1.0	0.000025	0.003443	0.000174
1.1	0.000036	0.003781	0.000228
1.2	0.000050	0.004121	0.000290
1.3	0.000069	0.004461	0.000360
1.4	0.000092	0.004804	0.000437
1.5	0.000121	0.005148	0.000522
1.6	0.000155	0.005495	0.000612
1.7	0.000196	0.005846	0.000707
1.8	0.000245	0.006199	0.000806
1.9	0.000302	0.006557	0.000908
2.0	0.000367	0.006918	0.001013

Table 5: *Arrival rate(λ) vs E_{R_1}*

EXP-S			
λ	ERL-A	EXP-A	HYP-EXP-A
1.0	0.003780	0.015642	0.004010
1.1	0.004109	0.017110	0.004381
1.2	0.004435	0.018566	0.004750
1.3	0.004758	0.020010	0.005117
1.4	0.005080	0.021444	0.005481
1.5	0.005402	0.022870	0.005844
1.6	0.005723	0.024289	0.006203
1.7	0.006047	0.025702	0.006560
1.8	0.006372	0.027112	0.006914
1.9	0.006701	0.028519	0.007264
2.0	0.007034	0.029924	0.007610

Table 6: *Arrival rate(λ) vs E_{R_1}*

HYP-EXP-S			
λ	ERL-A	EXP-A	HYP-EXP-A
1	0.000059	0.002861	0.000272
1.1	0.000087	0.003140	0.000356
1.2	0.000124	0.003421	0.000453
1.3	0.000170	0.003705	0.000561
1.4	0.000228	0.003993	0.000679
1.5	0.000297	0.004286	0.000808
1.6	0.000378	0.004585	0.000947
1.7	0.000473	0.004888	0.001093
1.8	0.000583	0.005197	0.001247
1.9	0.000707	0.005512	0.001408
2	0.000845	0.005832	0.001573

We observe that from the above table 4, 5 and 6:

- As arrival rate (λ) increases, the variety of arrangements of arrival and service times then the corresponding E_{R_1} also increases.
- Observe the arrival times, E_{R_1} rises faster in the case of $EXP - A$ and more gradually in the case of $HYP - EXP - A$. Comparably, it rises gradually in the case of $HYP - EXP - S$ and significantly in the case of EXP-S.

6.3. Illustration

In the 2D image, the influence of arrival rate(λ) on the expected number of customers demanding both I_1 and I_{12} has been examined. Fix $S_1 = 8$, $S_2 = 10$, $s_1 = 2$, $s_2 = 3$, $\mu_1 = 2$, $\mu_2 = 3$, $\mu_3 = 4$, $\beta_1 = 2$, $\beta_2 = 3$, $c_1 = 0.1$, $c_2 = 0.1$, $c_3 = 0.8$ so that the stability condition is satisfied.

From Figures 1 to 9,

- we can visualize that as the arrival rate (λ) maximizes, both the value of E_{I_1} and $E_{I_{12}}$ maximizes.
- Furthermore, the rate of an increase of E_{I_1} and $E_{I_{12}}$ for $HYP - EXP - A$ is rapid and slow for $ERL - A$. It is also faster for $HYP - EXP - S$ and shorter for $ERL - S$.

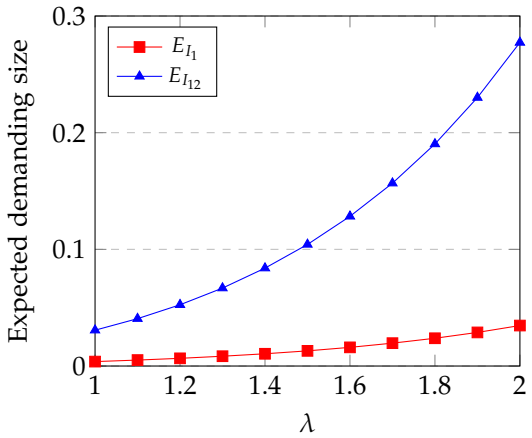


Figure 1: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $E_k/E_k/1$*

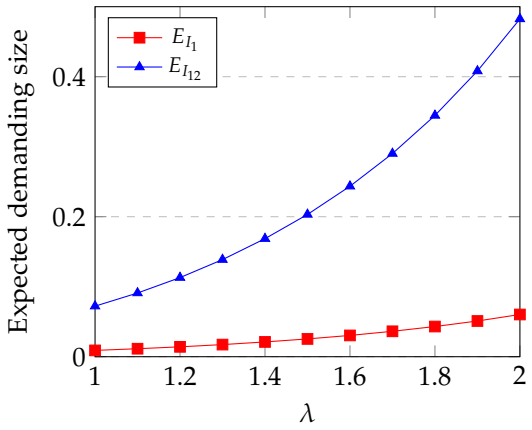


Figure 2: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $M/E_k/1$*

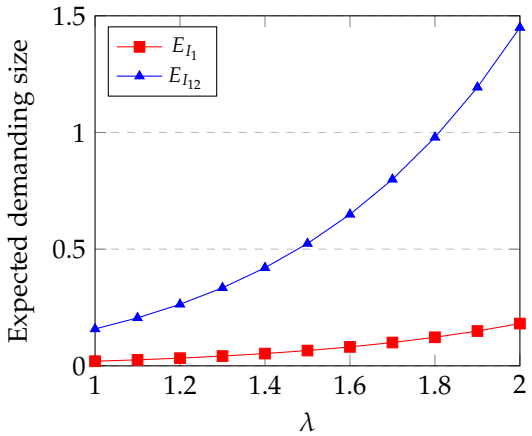


Figure 3: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $H_k/E_k/1$*

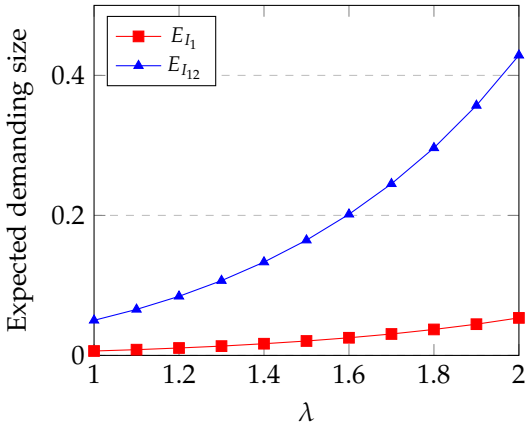


Figure 4: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $E_k/M/1$*

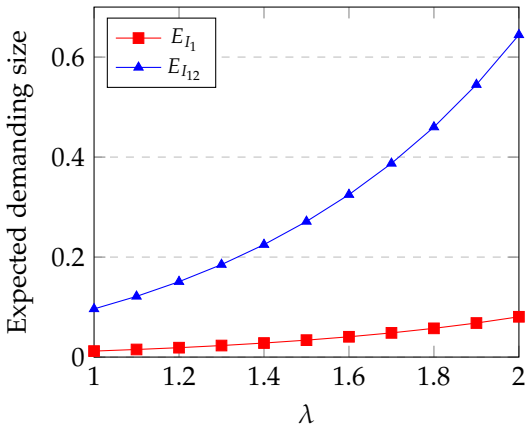


Figure 5: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $M/M/1$*

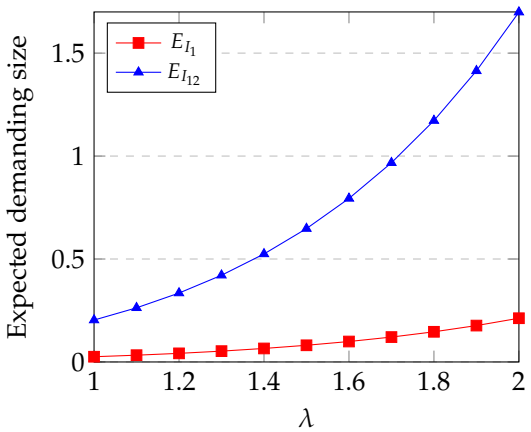


Figure 6: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $H_k/M/1$*

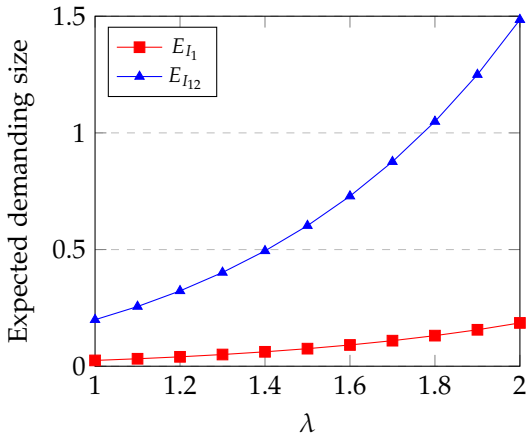


Figure 7: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $E_k/H_k/1$*

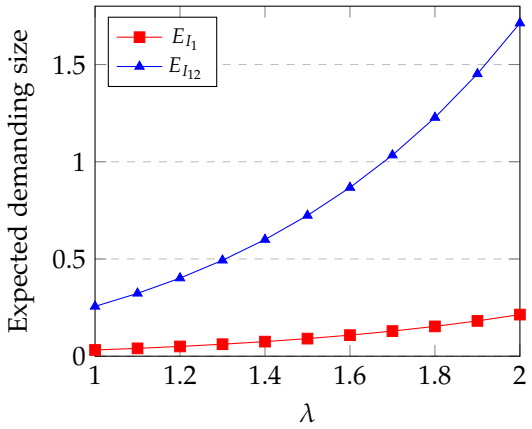


Figure 8: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $M/H_k/1$*

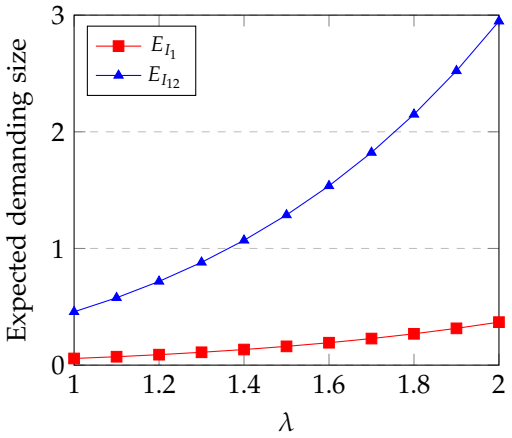


Figure 9: *Arrival rate(λ) vs both E_{I_1} and $E_{I_{12}}$ - $H_k/H_k/1$*

7. Conclusion

We looked at an inventory problem with two commodities and MAP demand arrival. When being taken for service, customers express their needs. If the requested item is unavailable, the customer is permanently removed from the system. When taken for service, if both goods are demanded, and when there is only one thing left, it is served to the customer. Depending on the type of demand, service times are distributed using a phase-type parameter. With parameter β_i for I_i, $i = 1, 2$, the lead times for each commodity are exponentially distributed. It is determined that the continuous-time Markov chain is of type GI/M/1. The stability of the system is demonstrated. Many system performance indices are developed, along with numerical examples and numerical studies.

References

[1] Amirthakodi, M.; Sivakumar, B. (2014). An inventory system with a service facility and finite orbit size for feedback customers, *OPSEARCH*, 52: 225-255.

[2] Anbazhagan, N., Arivarignan, G. (2001). Analysis of two commodity Markovian inventory system with lead time, *Journal of Applied Mathematics and Computing*, 8(2): 427-438.

[3] Anbazhagan, N., Arivarignan, G. and Irle, A. (2012). A two-commodity continuous review inventory system with substitutable items, *Stochastic Analysis and Applications*, 30: 11–19.

[4] Anbazhagan, N. and Jeganathan, K. (2013). Two-commodity Markovian inventory system with a compliment and retrial demand, *British Journal of Mathematics & Computer Science*, 3(2): 115–134.

[5] Balintfy, J.L. (1964). On a basic class of multi-item inventory problems, *Management Science*, 10(2): 287-297.

[6] Benny B., Chakravarthy S.R. and Krishnamoorthy A. (2018). Queueing-inventory system with two commodities, *Journal of the Indian Society for Probability and Statistics*, 19(1): 437–454.

[7] Berman, O., Kaplan, E.H. and Shimshak, D.G. (1993). Deterministic Approximations for Inventory Management at Service Facilities, *IIE Transactions* 25(5): 98–104.

[8] Chakravarthy, S. R. (2010). Markovian Arrival Processes, *Wiley Encyclopedia of Operations Research and Management Science*, 1-17.

[9] Federgruen, A., Groenevelt, H. and Tijms, H.C. (1984). Coordinated replenishments in a multi-item inventory a system with compound Poisson demands, *Management Science*, 30(3): 344–357.

[10] Jeganathan, K., Reiyas, M.A., Selvakumar, S., Anbazhagan, N., Amutha, S., Joshi, G.P., Jeon, D. and Seo, C. (2022). Markovian Demands on Two Commodity Inventory System with Queue-Dependent Services and an Optional Retrial Facility, *Mathematics*, 10(12): 2046.

[11] Kalpakam, S. and Arivarignan, G. (1993). A coordinated multicommodity (S, s) inventory system. *Mathematical and Computer Modelling*, 18(11): 69–73.

[12] Krishnamoorthy, A., Iqbal Basha, R. and Lakshmy, B. (1994). Analysis of a two commodity inventory problem, *International Journal of Information and Management Science*, 5(1): 127-136.

[13] Krishnamoorthy, A. and Varghese, TV. (1994). A two commodity inventory problem, *Information and Management Sciences*, 5(37): 127–138.

[14] Lakshmanan, K., Padmasekaran, S. and Jeganathan, K. (2019). Mathematical analysis of queueing-inventory model with compliments and multiple working vacations, *International Journal of Engineering and Advanced Technology*, 8(6): 4239–4240.

[15] Neuts, M.F. (1981). Matrix-geometric solutions in stochastic models: an algorithmic approach. *The Johns Hopkins University Press*, Baltimore.

[16] Paul Manual, Sivakumar, B. and Arivarignan, G. (2008). A Perishable inventory system with service facilities and retrial customers, *Computers & Industrial Engineering*, 54(3): 484–501.

[17] Schwarz, M., Wichelhaus, C. and Daduna, H. (2007). Product form models for queueing networks with an inventory. *Stochastic Models*, 23(4): 627–663.

[18] Sigman, K. and Simchi-Levi, D. (1992). Light traffic heuristic for an M/G/1 queue with limited inventory, *Annals of Operation Research*, 40: 371-380.

[19] Silver, E.A. (1974). A control system for coordinated inventory replenishment, *International Journal of Production Research*, 12: 647–671.

[20] Sivakumar, B., Anbazhagan, N. and Arivarignan, G. (2007). Two-commodity inventory system with individual and joint ordering policies and renewal demands, *Stochastic Analysis and Applications*, 25(6): 1217–1241.

[21] Sivakumar, B. (2008). Two-commodity inventory system with retrial demand. *European Journal of Operational Research*, 187(1): 70–83.

[22] Yadavalli, V.S.S., Anbazhagan, N. and Arivarignan, G. (2004). A two-commodity continuous review inventory system with lost sales. *Stochastic Analysis and Applications*, 2(2): 479-497.

[23] Yadavalli, V.S.S., Adetunji, O., Sivakumar, B. and Arivarignan, G. (2010). Two-commodity perishable inventory a system with bulk demand for one commodity, *South African Journal of Industrial Engineering*, 21(1): 137-155.

RT&A, No 4 (76)
Volume 18, December 2023

EXPLORING THE ADAPTABILITY OF THE UNIT INVERSE WEIBULL DISTRIBUTION FOR MODELING DATA ON THE UNIT INTERVAL

Shameera T

•

Department of Statistics, MES Mampad College, Malappuram, Kerala, India
shameeranaseerok@gmail.com

Bindu P.P

•

Department of Statistics, Govt. Arts & Science College, Kozhikode, Kerala, India
ppbindukannan@gmail.com

Abstract

This paper derives a new lifetime distribution called the unit inverse Weibull distribution(UIWD) from inverse weibull distribution. Various statistical properties such as the survival function, hazard rate function, revised hazard rate function, cumulative hazard rate function, moments, and quartiles have been discussed. Additionally, we have explored other properties like skewness, kurtosis, order statistics, and the quantile function. Various methods of estimation, including maximum likelihood, moments, percentiles, and the Cramer-Von Mises, have been discussed. Simulation studies were conducted to assess the accuracy and precision of the parameters. Comparative analyses were performed to highlight the effectiveness and utility of the proposed model in comparison to other existing models, using two real-life applications. Finally, real life data analysis reveals that derived distribution can provide a better fit than several well-known distributions.

Keywords: Unit inverse Weibull distribution, statistical properties, estimation, simulation study

1. Introduction

Over the last couple of decades, multiple authors have introduced a range of fresh methodologies for creating novel sets of distributions. This has significantly expanded the potential for accurately modeling real-world data across a variety of fields. This concept of devising new models and families has garnered notable attention recently, often termed as "parameter addition" and "parameter induction." The core objective of these efforts is to formulate models capable of effectively capturing real-life phenomena by utilizing available data sets from diverse domains.

In applied statistics, a prevalent hurdle involves addressing uncertain occurrences that exist within the confined range of (0,1). Consider instances from the real world, where measurements frequently involve proportions, fractions representing certain attributes, scores obtained from aptitude assessments, assorted indices, rates, and various other data points that inherently lie within the interval (0,1). In such scenarios, continuous distributions characterized within the domain of (0,1) prove indispensable for the probabilistic representation of these phenomena. The

distribution that holds sway over the unit interval finds application across numerous sectors, encompassing economics and biology.

Distributions that find definition over the (0,1) range are conventionally harnessed to model random variables that are inherently limited within the confines of (0,1), such as percentages and proportions. The Beta distribution, renowned for its convenience and utility across a plethora of statistical domains, is a standard choice for tackling such scenarios. Nevertheless, there are situations where the Beta distribution might fall short in adequately elucidating the data, thus prompting the quest for alternative distributions defined within the unit interval.

Various distributions defined on the unit interval have been proposed in the literature, including Topp-Leone [1], Johnson SB [2], unit Gamma [3], Kumaraswamy [4], Arcsine [5], unit Logistic [6], generalized Beta type I [7], Simplex [8], standard two-sided Power[9], Mc Arcsine [10], Log-Lindley [11], two-sided generalized Kumaraswamy [12], and Log-Xgamma [13].

More recently, researchers have proposed new families of transformed distributions on the unit interval. Examples include the unit Birnbaum-Saunders [14], unit Lindley [15], unit inverse Gaussian [16], unit Gompertz [17], unit improved second-degree Lindley [18], Log-weighted exponential [19] , Logit Slash [20], and unit generalized Half Normal [21] distributions.

In this study, we propose a probability distribution called the Unit Inverse Weibull Distribution (UIWD) specifically designed for modeling data on the interval (0,1). The UIWD is derived from a type transformation of the Inverse Weibull Distribution (IWD), and we provide various methods of estimation approach for estimating its parameters.

The motivation behind the development of the UIWD arises from several factors. Firstly, the UIWD exhibits simple and closed-form expressions for its distribution function and quantile function. Moreover, it demonstrates superior fitting performance compared to other commonly used distributions on the unit interval. The UIWD allows for the derivation of various statistical properties, and we utilize the various estimators method to estimate its parameters. Through simulation studies, we assess the accuracy and precision of different estimators and compare the UIWD model with existing models to showcase its utility and effectiveness. Overall, this study aims to contribute to the understanding and application of the UIWD as a flexible probability distribution for modeling data on the unit interval (0,1).

2. The Unit Inverse Weibull Distribution: Derivation of Pdf and Cdf

The two-parameter Weibull distribution defines the probability density function (PDF) of a random variable U as

$$f(u;\alpha,\beta) = \alpha\beta u^{\beta-1}e^{-\alpha u^{\beta}}, \quad u > 0; \alpha, \beta > 0. \tag{1}$$

To explore a related distribution, we introduce a transformation by defining $V = 1/U$. Consequently, V follows the two-parameter Inverse Weibull Distribution (IWD), and its PDF is given by

$$f(v;\alpha,\beta) = \alpha\beta v^{-\beta-1}e^{-\alpha v^{-\beta}}, v > 0; \alpha, \beta > 0. \tag{2}$$

Now, we propose the Unit Inverse Weibull Distribution (UIWD) by introducing a further transformation, $X = 1/1 + V$. As a result, X follows the UIWD, and its PDF is expressed as

$$f(x;\alpha,\beta) = \alpha\beta\left(\frac{1}{x^2}\right)\left(\frac{1}{x}-1\right)^{-\beta-1} e^{-\alpha\left(\frac{1}{x}-1\right)^{-\beta}}, 0 < x < 1; \alpha, \beta > 0. \tag{3}$$

Here, α and β serve as shape parameters in the UIWD, with the constraint that $\alpha > 0$ and $\beta > 0$.

The cumulative distribution function (CDF) of the Unit Inverse Weibull Distribution (UIWD) is defined as follows:

$$F(x|\alpha,\beta) = P(X \leq x) = 1 - e^{-\alpha(1/x-1)^{-\beta}}, \quad 0 < x < 1, \quad \alpha, \beta > 0. \tag{4}$$

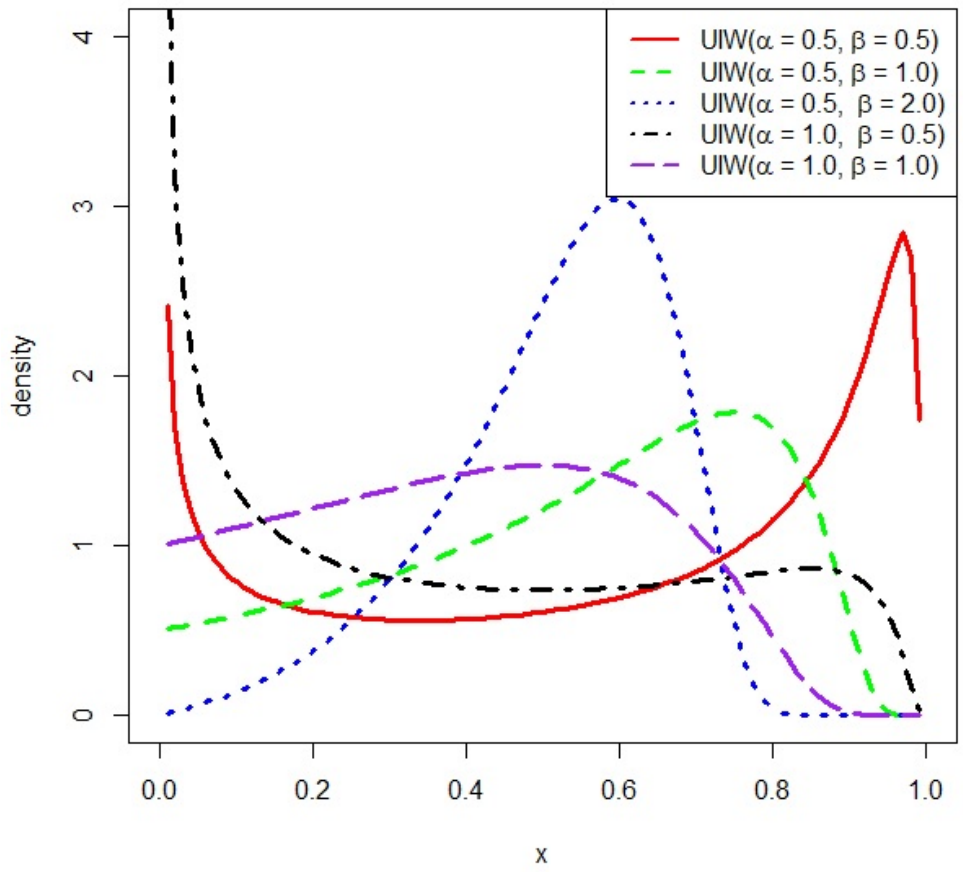


Figure 1: *Pdf of UIWD with different values of alpha and beta*

3. Reliability Properties

3.1. Survival Function

The survival function of the Unit Inverse Weibull Distribution (UIWD) is given by:

$$S(x|\alpha,\beta) = 1 - F(x) = e^{-\alpha(1/x-1)^{-\beta}}, \quad 0 < x < 1, \quad \alpha,\beta > 0. \tag{5}$$

Here the parameters α and β of the UIWD control the shape and scale of the survival function.

The survival function of the Unit Inverse Weibull Distribution (UIWD) exhibits the following characteristics:

1.Monotonic Decrease: The survival function is a monotonically decreasing function. As the value of x increases within the (0, 1) interval, the probability of the random variable exceeding x decreases. This is evident from Figure 3.

2.Asymptotic Behavior: As x approaches 0, the survival function approaches 1. This indicates that the probability of the random variable being greater than a value close to 0 tends to 1. In other words, the UIWD has a high probability of taking on values very close to 0.

3. As x approaches 1, the survival function approaches 0. This implies that the probability of the random variable exceeding a value close to 1 tends to 0. Consequently, the UIWD has a low probability of taking on values very close to 1.

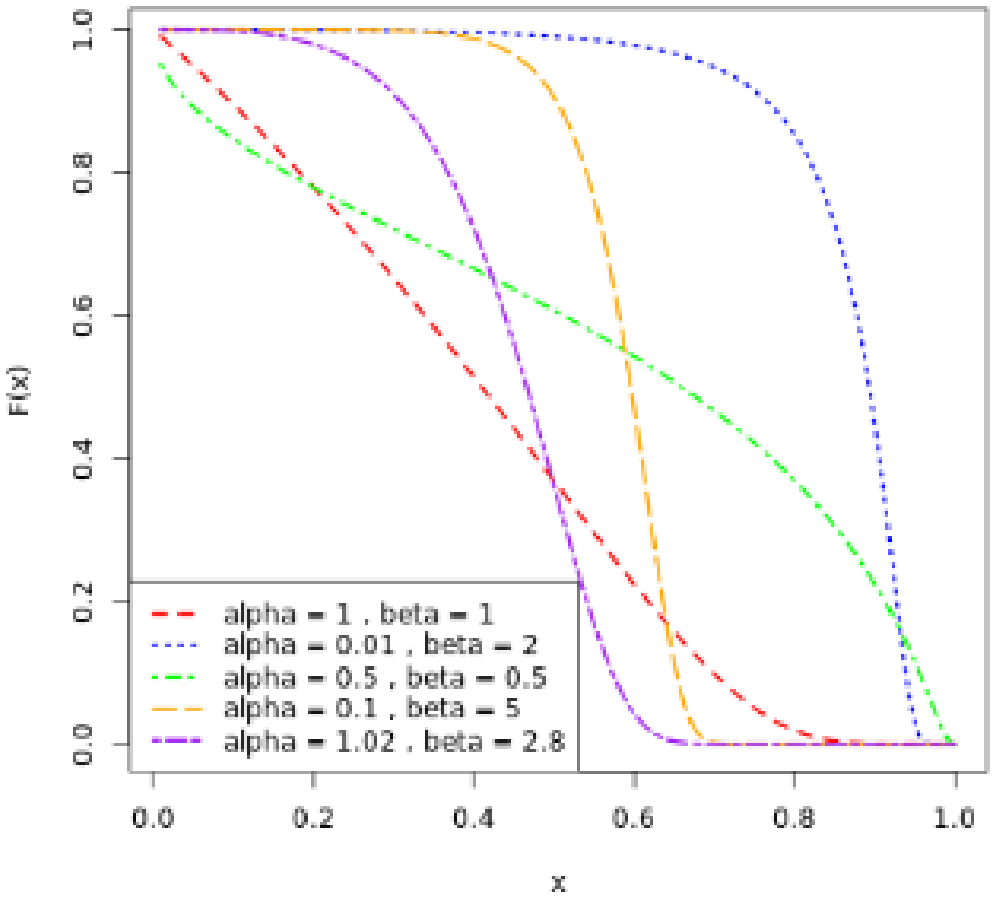


Figure 2: *Distribution function of UIWD with different values of alpha and beta*

3.2. Hazard Rate,Reversed Hazard Rate, Cumulative Hazard Rate Functions

The hazard rate function of the Unit Inverse Weibull Distribution (UIWD) is given by:

$$h(x|\alpha,\beta) = \frac{f(x)}{1-F(x)}$$

$$= \alpha\beta\left(\frac{1}{x^2}\right)\left(\frac{1}{x}-1\right)^{-\beta-1}, \quad 0<x<1; \quad \alpha,\beta>0. \tag{6}$$

Hazard Rate function has the following characteristics:

1. Monotonic Increase: The hazard function is a monotonically increasing function. As the value of x increases within the (0, 1) interval, the hazard rate, which measures the instantaneous rate of occurrence of an event, increases. This is evident from Figure 4.

2. Asymptotic Behavior: As x approaches 1, the hazard rate approaches infinity. This indicates that the event becomes more likely to occur as time approaches the maximum value of 1.

The reversed hazard rate function of the Unit Inverse Weibull Distribution (UIWD) is given by:

$$h_{\text{rev}}(x|\alpha,\beta) = \frac{f(x)}{F(x)}$$

$$= \frac{\alpha\beta\left(\frac{1}{x^2}\right)\left(\frac{1}{x}-1\right)^{-\beta-1} e^{\left(-\alpha\left(\frac{1}{x}-1\right)^{-\beta}\right)}}{1-e^{\left(-\alpha\left(\frac{1}{x}-1\right)^{-\beta}\right)}}, 0<x<1, \quad \alpha,\beta>0. \tag{7}$$

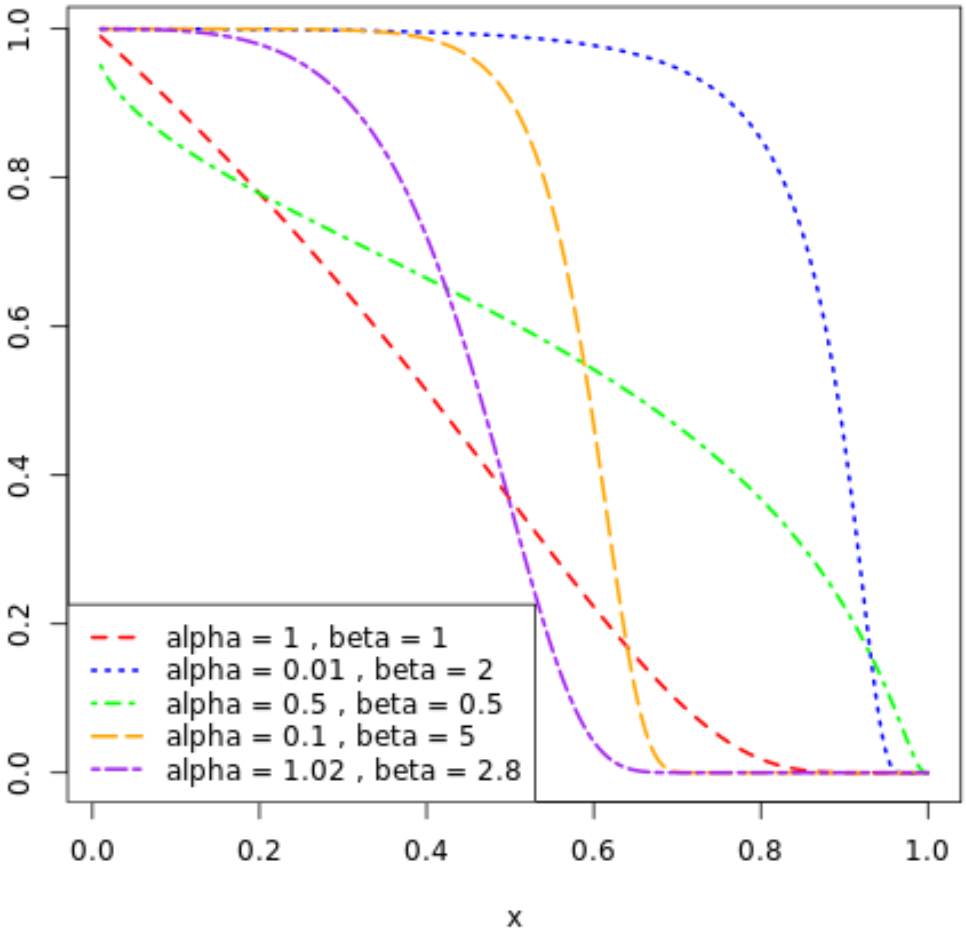


Figure 3: *Survival function of UIWD with different values of alpha and beta*

The cumulative hazard rate function of the Unit Inverse Weibull Distribution (UIWD) is given by:

$$C(x;\alpha,\beta) = -lns(x)$$

$$= \alpha\left(\frac{1}{x}-1\right)^{-\beta}, \quad 0<x<1, \quad \alpha,\beta>0. \tag{8}$$

4. Moments and Related Properties

4.1. Raw Moments and central moments

The r^{th} raw moment about the origin of the random variable X is:

$$\mu'_r = \int_0^1 x^r \alpha\beta\left(\frac{1}{x^2}\right)\left(\frac{1}{x}-1\right)^{-\beta-1} e^{-\alpha\left(\frac{1}{x}-1\right)^{-\beta}} dx$$

Substituting u = $\left(\frac{1}{x}-1\right)^{-\beta}$, we get

$$\mu'_r = \int_0^\infty \alpha \frac{1}{(1+u^{\frac{-1}{\beta}})} e^{-\alpha u}\, du$$

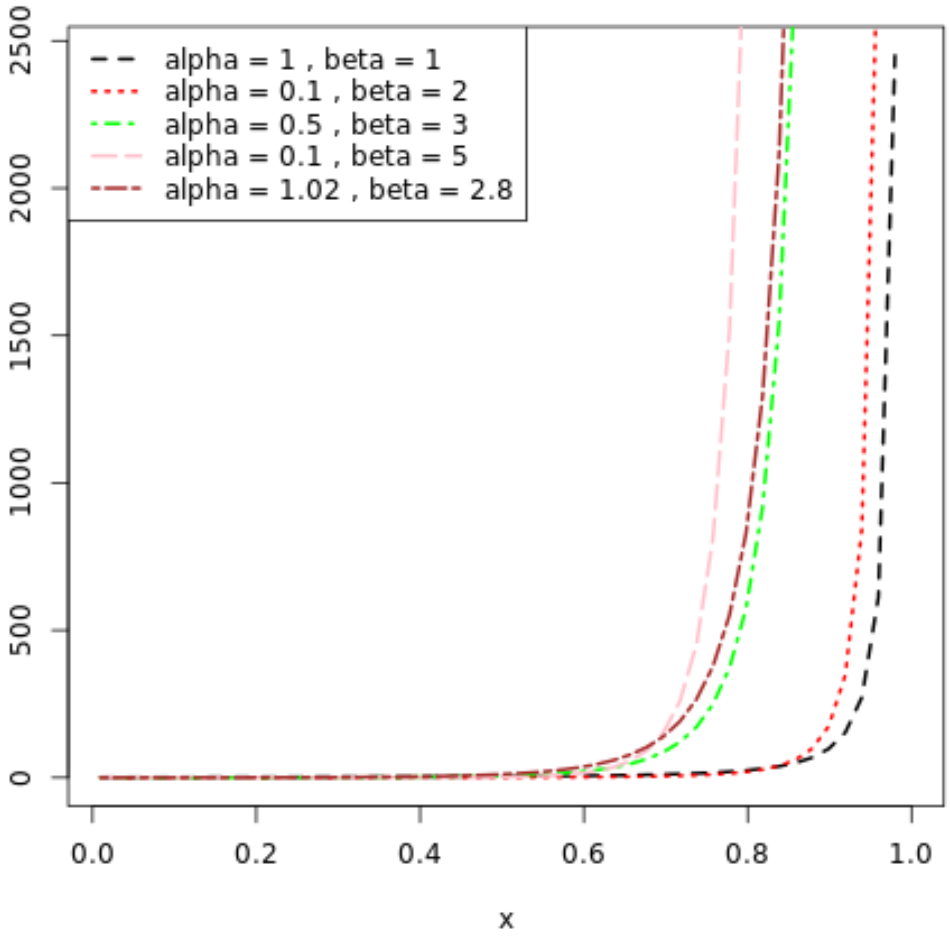


Figure 4: *Hazard Rate function of UIWD with different values of alpha and beta*

$$= 1 + \sum_{k=1}^{\infty} (-1)^k r(r+1).....(r+k-1) \frac{\Gamma_{1-\frac{k}{\beta}}}{\alpha^{\frac{-k}{\beta}}}, \beta > k. \tag{9}$$

The first four raw moments are obtained by putting r = 1, 2, 3,4 in (9) .

The central moments are obtained from the raw moments with the help of the recurrence relationship between raw moments and central moments.

$\mu_1 = 0$
$\mu_2 = \mu_2' - {\mu_1'}^2$
$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$
$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$

4.2. Quartiles

Consider the function $f(x)$ of UIWD. The first quartile (Q_1) is given by:

$$\int_0^{Q_1} f(x)\, dx = \frac{1}{4}$$

which implies

$$Q_1 = \frac{1}{1 + (0.7213\alpha)^{-\beta}}. \tag{10}$$

The median (m, Q_2) is given by:

$$\int_0^m f(x)\,dx = \frac{1}{2}$$

which implies

$$m = \frac{1}{1+(1.4423\alpha)^{-\beta}}. \tag{11}$$

The third quartile (Q_3) is given by:

$$\int_0^{Q_3} f(x)\,dx = \frac{3}{4}$$

which implies

$$Q_3 = \frac{1}{1+(3.4761\alpha)^{-\beta}}. \tag{12}$$

4.3. Bowley's Coefficient of Skewness

The Bowley's coefficient of skewness (S_{kp}) of UIWD is given by:

$$S_{kp} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

.

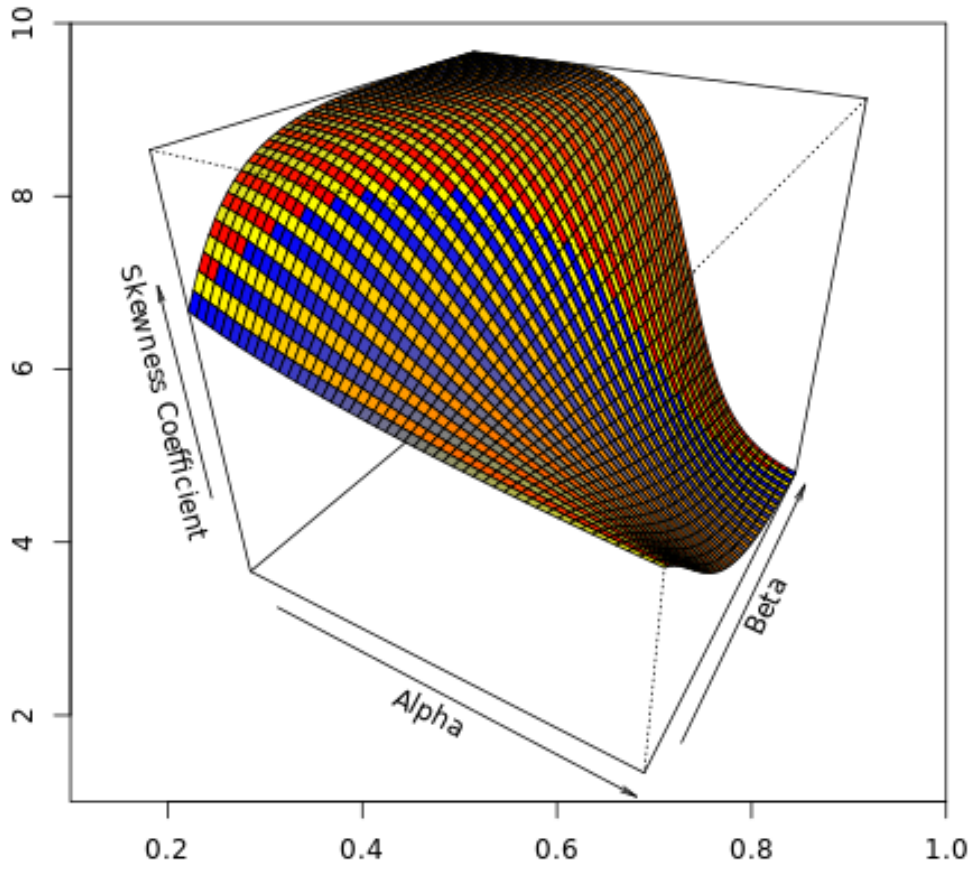


Figure 5: *skewness of UIWD*

4.4. Percentile Coefficient of Kurtosis

The percentile coefficient of kurtosis of UIWD is given by:

$$K = \frac{Q.D}{P_{90} - P_{10}}$$

where

$$Q.D = \frac{Q_3 - Q_1}{2}$$

10^{th} percentile, P_{10} is given by

$$\int_0^{P_{10}} f(x)\,dx = \frac{1}{10}$$

implies

$$P_{10} = \frac{1}{1 + (0.4343\alpha)^{-\beta}}. \tag{13}$$

90^{th} percentile, P_{90} is given by

$$\int_0^{P_{90}} f(x)\,dx = \frac{9}{10}$$

implies

$$P_{90} = \frac{1}{1 + (9.4912\alpha)^{-\beta}}. \tag{14}$$

Table 1: *Skewness, Kurtosis, Mean, and Variance for Different α and β*

α	β	Skewness	Kurtosis	Mean	Variance
0.5	0.2	0.1208	0.2563	0.5975	0.0712
0.5	1	0.1673	0.2855	0.5385	0.0164
0.5	2.6	0.3706	0.3848	0.5127	0.0028
0.5	7.8	0.8538	0.4932	0.5038	0.0004
1	0.2	0.1154	0.2571	0.3687	0.0880
1	1	0.0407	0.2970	0.4037	0.0122
1	2.6	-0.2735	0.3704	0.4496	0.0022
1	7.8	-0.8829	0.4644	0.4817	0.0003
2	0.2	0.1101	0.2579	0.1511	0.0448
2	1	-0.0716	0.2924	0.2773	0.0063
2	2.6	-0.6071	0.2304	0.3879	0.0015
2	7.8	-0.9905	0.0362	0.4597	0.0003
10	0.2	0.0981	0.2596	0.0010	0.0000
10	1	-0.2137	0.2654	0.0844	0.0003
10	2.6	-0.6969	0.1336	0.2592	0.0005
10	7.8	-0.9910	0.0096	0.4094	0.0002

4.5. Order Statistics

Let's consider a random sample of size n from the UIWD, denoted as $X_1, X_2, \ldots, X_n$. The order statistics of the UIWD are defined as $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, where $X_{(1)}$ represents the smallest observed value, $X_{(2)}$ represents the second smallest value, and so on, up to $X_{(n)}$, which represents the largest observed value.

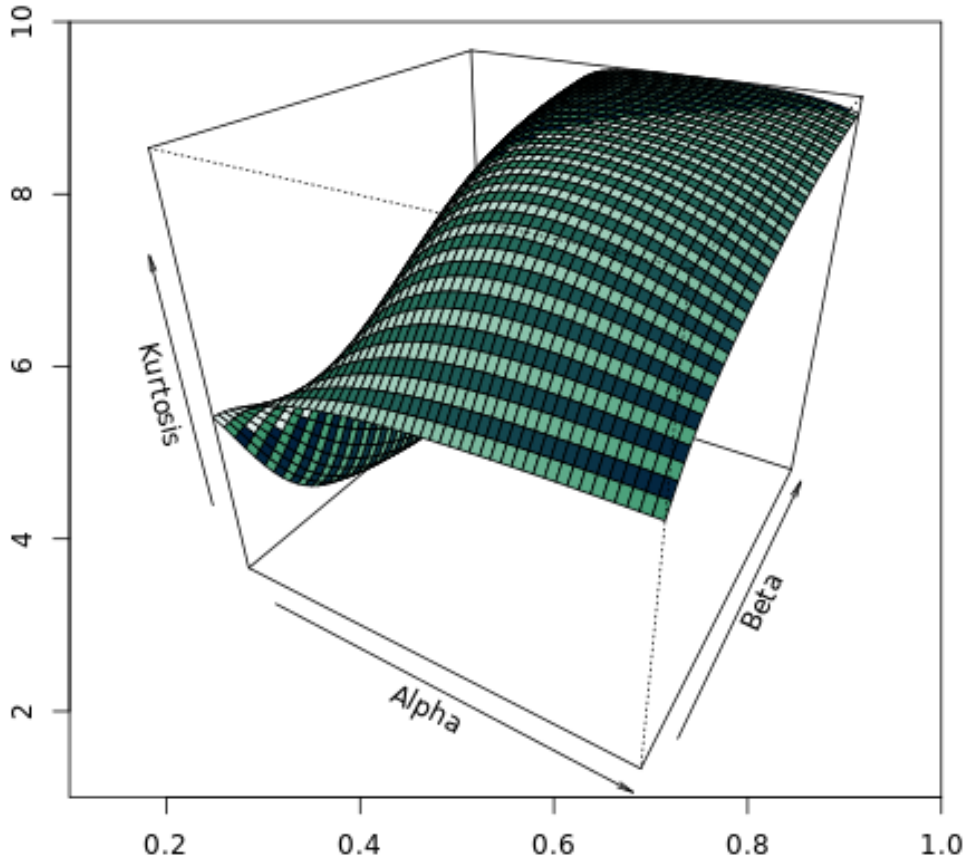


Figure 6: *Kurtosis of UIWD*

The r^{th} order statistics is given by:

$$f_{r:n}(x|\alpha,\beta) = C_{r:n}F(x|\alpha,\beta)^{r-1}(1-F(x|\alpha,\beta))^{n-r}f(x|\alpha,\beta)$$

$$= C_{r:n}[1-e^{-\alpha(1/x-1)^{-\beta}}]^{r-1}\alpha\beta\left(\frac{1}{x^2}\right)\left(\frac{1}{x}-1\right)^{-\beta-1}e^{-\alpha\left(\frac{1}{x}-1\right)^{-\beta(n-r+1)}}, \tag{15}$$

where,

$$C_{r:n} = \frac{n!}{(r-1)!(n-r)!}; 0 < x < 1, \quad \alpha,\beta > 0.$$

4.6. Quantile Function

The quantile function of a distribution gives the inverse mapping of the cumulative distribution function (CDF), allowing you to find the value corresponding to a given probability.Let's denote the quantile function as Q(p), where p is the probability for which we want to find the corresponding value. Then quantile function of UIWD is given by:

$$p = F(x)$$

Then,

$$Q(p) = \frac{1}{1+\left(-\frac{\ln(1-p)}{\alpha}\right)^{\frac{-1}{\beta}}}. \tag{16}$$

5. Simulation study

In this section, we present a Monte Carlo simulation study conducted for the purpose of evaluation of the finite-sample behavior of the maximum likelihood estimates of the UIWD. Generate $n = 50, 100, 200, 400$ and 800 as sample size and considered 1000 replications for each sample size. To simulate n observations from the proposed distribution, we implemented the following algorithm:

1. Generate n random numbers from the uniform distribution $U(0,1)$, denoted as U_i, where $i = 1, 2, \ldots, n$.
2. For each i, solve the equation $G(y_i)$ by finding the inverse of the cumulative distribution function:

$$y_i = G^{-1}(U_i) = \frac{1}{1 - \left(\frac{-\ln(1-U_i)}{\alpha}\right)^{\frac{1}{\beta}}}$$

.

Where $G^{-1}(.)$ is the quantile function of UIWD.

This algorithm allows us to generate a sample of n observations that follow the desired distribution.

Table 2: *Estimates, Bias, and MSE for Different n*

n	Estimates	Bias	MSE
50	$\hat{\alpha}$: 3.0766	-0.3234	0.1046
	$\hat{\beta}$: 1.1433	-0.0567	0.0092
100	$\hat{\alpha}$: 3.5584	0.1584	0.0551
	$\hat{\beta}$: 1.2334	0.0334	0.0071
250	$\hat{\alpha}$: 3.6005	0.2005	0.0402
	$\hat{\beta}$: 1.2855	0.0855	0.0053
500	$\hat{\alpha}$: 3.6432	0.2432	0.0292
	$\hat{\beta}$: 1.2977	0.0977	0.0036
1000	$\hat{\alpha}$: 3.4555	0.0555	0.0031
	$\hat{\beta}$: 1.2490	0.0490	0.0024

6. Estimation

6.1. Maximum Likelihood Estimator

Let's denote the sample as $\{x_1, x_2, \ldots, x_n\}$ taken from the population which follows UIWD and the parameters are α and β.

The likelihood function,

$$L(\alpha, \beta) = \prod_{i=1}^{n} \alpha\beta \left(\frac{1}{x_i^2}\right) \left(\frac{1}{x_i} - 1\right)^{-\beta-1} e^{-\alpha\left(\frac{1}{x_i}-1\right)^{-\beta}}.$$

Taking the natural logarithm of both sides, we get:

$$\ln L(\alpha, \beta) = \sum_{i=1}^{n} \ln(\alpha\beta) - 2\ln(x_i) - (\beta+1)\ln\left(\frac{1}{x_i} - 1\right) - \alpha\left(\frac{1}{x_i} - 1\right)^{-\beta}.$$

Now, let's differentiate $\ln(L(\alpha,\beta))$ with respect to α and β and set the derivatives to zero:

$$\frac{\partial}{\partial\alpha}\ln(L(\alpha,\beta)) = 0$$

and

$$\frac{\partial}{\partial\beta}\ln(L(\alpha,\beta)) = 0$$

Taking the derivative with respect to α:

$$\frac{\partial}{\partial\alpha}\ln(L(\alpha,\beta)) = \sum_{i=1}^{n}\left[\frac{1}{\alpha} - (x_i-1)^{-\beta}\right].$$

Setting it to zero:

$$\sum_{i=1}^{n}\left[\frac{1}{\alpha} - (x_i-1)^{-\beta}\right] = 0.$$

Rearranging, we get:

$$\frac{1}{\alpha} = \sum_{i=1}^{n}(x_i-1)^{-\beta}.$$

Solving for α:

$$\hat{\alpha} = \frac{1}{\sum_{i=1}^{n}(x_i-1)^{-\beta}}. \tag{17}$$

Now, taking the derivative with respect to β:

$$\frac{\partial}{\partial\beta}\ln(L(\alpha,\beta)) = \sum_{i=1}^{n}\left[\ln(x_i-1) + \alpha(x_i-1)^{-\beta}\ln(x_i-1) - \alpha(x_i-1)^{-\beta-1}(x_i-1)\right].$$

Setting it to zero:

$$\sum_{i=1}^{n}\left[\ln(x_i-1) + \alpha(x_i-1)^{-\beta}\ln(x_i-1) - \alpha(x_i-1)^{-\beta-1}(x_i-1)\right] = 0. \tag{18}$$

It is non linear function and it can be found numerically.

These estimates of α and β obtained through maximizing the log-likelihood function are the maximum likelihood estimators for the given model.

6.2. Method of moments

Let ${m_r}'$ and ${\mu_r}'$ represent the sample and population raw moments, respectively, for a given data set. These moments are defined for different orders of r, where r ranges from 1 to n.

The r^{th} raw moment about the origin of the random variable X is given by:

$$\mu_r' = 1 + \sum_{k=1}^{\infty}(-1)^k r(r+1)\dots(r+k-1)\frac{\Gamma_{1-\frac{k}{\beta}}}{\alpha^{\frac{-k}{\beta}}}, \quad \beta > k.$$

The r^{th} sample raw moment, denoted as $m_r^{'}$, is calculated as:

$$m_r^{'} = \frac{1}{n}\sum_{i=1}^{n}x_i^r, \quad r = 1,2,3,\dots,n.$$

where x_i represents the individual data points in the data set.

By equating ${m_r}'$ to ${\mu_r}'$, we can establish the following set of equations:

$$m_1^{'} = \mu_1^{'}, \quad m_2^{'} = \mu_2^{'}, \quad m_3^{'} = \mu_3^{'}, \quad \dots, \quad m_n^{'} = \mu_n^{'}.$$

Solving these equations will allow us to determine the values of α and β.

6.3. Percentile estimation method

consider 25^{th} percentile (P_{25}) and 75^{th} percentile (P_{75}) of UIWD.

Then,

$$P_{25} = \frac{1}{1 + (0.7213\alpha)^{-\beta}}$$

and

$$P_{75} = \frac{1}{1 + (3.4761\alpha)^{-\beta}}$$

solving these two equations, we get,

$$\hat{\alpha} = -0.1419 + e^{\frac{\log\left(\frac{1}{p_{25}} - 1\right)}{-1.4641 \log\left(\frac{(1-p_{25})p_{75}}{(1-p_{75})p_{25}}\right)}} \tag{19}$$

and

$$\hat{\beta} = 1.4641 \log\left(\frac{(1-p_{25})p_{75}}{(1-p_{75})p_{25}}\right). \tag{20}$$

6.4. Cramer-von mises method

The Cramer-von Mises estimation equation is used to estimate the parameters of a distribution by minimizing the Cramer-von Mises objective function. For the Unit Inverse Weibull Distribution (UIWD), the Cramer-von Mises estimation equation can be defined as follows:

$$C(\alpha, \beta) = \frac{1}{12n} + \sum \left[F(t_i|\alpha, \beta) - \frac{2i-1}{2n}\right]^2.$$

In this equation, $C(\alpha, \beta)$ represents the Cramer-von Mises objective function, α and β are the parameters of the UIWD that need to be estimated, n is the number of data points, t_i represents the observed data points, and $F(t_i|\alpha, \beta)$ is the cumulative distribution function (CDF) of the UIWD for each data point.

To estimate the parameters α and β, the Cramer-von Mises objective function is minimized by finding the values of α and β that result in the smallest value of $C(\alpha, \beta)$. This can be done using partial differentiation of α and β. Equating these to zero, we get normal equations. since these equations are non-linear, we can use iterative method to find the estimates.

7. Applications

Data set 1: Here, we apply real data to show how adaptable and applicable the suggested distribution is in comparison to a variety of other well-known distributions on the unit interval. The data set is used for fitting is COVID-19 of Britain: This data set covered a period of 47 days, from 1 May 2021 to 17 June 2021 [22]. The following information is created using daily new deaths (DNDs), daily cumulative cases (DCCs), and daily cumulative deaths (DCDs): 0.0023, 0.0023, 0.0023, 0.0046, 0.0065, 0.0067, 0.0069, 0.0069, 0.0091, 0.0093, 0.0093, 0.0093, 0.0111, 0.0115, 0.0116, 0.0116, 0.0119, 0.0133, 0.0136, 0.0138, 0.0138, 0.0159, 0.0161, 0.0162, 0.0162, 0.0162, 0.0163, 0.0180, 0.0187, 0.0202, 0.0207, 0.0208, 0.0225, 0.0230, 0.0230, 0.0239, 0.0245, 0.0251, 0.0255, 0.0255, 0.0271, 0.0275, 0.0295, 0.0297, 0.0300, 0.0302, 0.0312, 0.0314, 0.0326

Data set 2:It is the recovery rates of COVID-19 patients in Spain from 3 March to 7 May 2020 [23] .The data set are: 0.6670, 0.5000, 0.5000, 0.4286, 0.7500, 0.6531, 0.5161, 0.7895, 0.7689, 0.6873, 0.5200, 0.7251, 0.6375, 0.6078, 0.6289, 0.5712, 0.5923, 0.6061, 0.5924, 0.5921, 0.5592, 0.5954, 0.6164, 0.6455, 0.6725, 0.6838, 0.6850, 0.6947, 0.7210, 0.7315, 0.7412, 0.7508, 0.7519, 0.7547, 0.7645, 0.7715, 0.7759, 0.7807, 0.7838, 0.7847, 0.7871, 0.7902, 0.7934, 0.7913, 0.7962, 0.7971, 0.7977, 0.8007, 0.8038,

0.8289, 0.8322, 0.8354, 0.8371, 0.8387, 0.8456, 0.8490,0.8535, 0.8547, 0.8564, 0.8580, 0.8604, 0.8628, 0.6586, 0.7070, 0.7963, 0.8516.

To compare adaptability of unit inverse weibull distribution, we use the following two parameter distributions on the unit interval.

1.**Beta distribution of first kind (BD)**:

The probability density function (PDF) of the beta distribution is given by,

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

where $0 \leq x \leq 1$, $\alpha > 0$, $\beta > 0$, and $B(\alpha,\beta)$ is the beta function.

2.**Unit Weibull distribution(UWD)**:

The probability density function (PDF) UWD is given by,

$$f(x;\alpha,\beta) = \frac{1}{x}\alpha\beta\left(-\log x\right)^{(\beta-1)}\exp\left[-\alpha\left(-\log x\right)^{\beta}\right]$$

where $0 < x < 1$, $\alpha > 0$, and $\beta > 0$.

3. **Kumaraswamy Distribution(KD)**:

The probability density function (PDF) of the Kumaraswamy distribution is given by:

$$f(x;\alpha,\beta) = \alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}$$

where $0 < x < 1$, $\alpha > 0$, and $\beta > 0$.

4. **Unit Birnbaum-Saunders distribution (UBSD)**:

The probability density function (PDF) of the UBSD is given by:

$$f\left(x;\alpha,\beta\right) = \frac{1}{2x\alpha\beta\sqrt{2\pi}}\left[\left(\frac{-\beta}{\log x}\right)^{\frac{1}{2}} + \left(\frac{-\beta}{\log x}\right)^{\frac{3}{2}}\right]\exp\left\{\frac{1}{2\alpha^2}\left(\frac{\log x}{\beta} + \frac{\beta}{\log x} + 2\right)\right\}$$

where $0 < x < 1$, $\beta > 0$, and $\alpha > 0$.

5. **Unit Gompertz distribution(UGD)**:

The probability density function (PDF) of the UGD is given by:

$$f(x;\alpha,\beta) = \alpha\beta x^{-(\beta+1)}\exp\left(-\alpha\left(x^{-\beta}-1\right)\right)$$

where $0 < x < 1$, $\alpha > 0$, and $\beta > 0$.

6. **Kumaraswamy distribution(KD)**:

The pdf of KD is given by,

$$f(x) = abx^{a-1}(1-x^{a})^{b-1}, 0 < x < 1, a > 0, b > 0.$$

7. **Unit inverse weibull distribution (UIWD)**:

The pdf of UIWD is given by,

$$f(x|\alpha,\beta) = \alpha\beta\left(\frac{1}{x^2}\right)\left(\frac{1}{x}-1\right)^{-\beta-1}e^{-\alpha\left(\frac{1}{x}-1\right)^{-\beta}}$$

where $0 < x < 1$ and $\alpha, \beta > 0$.

From the table 2, we can conclude that the unit inverse weibull model has the lowest AIC, AICC, and BIC values among the listed models, indicating that it provides the best trade-off between goodness of fit and model complexity based on all three criteria. Lower values of AIC, AICC, and BIC indicate better fitting models with lower complexity.

From table 3 , we can conclude that , if the unit inverse weibull distribution model has the least values for AIC, AICC, and BIC, it indicates that the unit inverse weibull distribution model is the best-fitting model among the listed models. The lower values of these criteria suggest that the unit inverse weibull distribution model provides a better trade-off between goodness of fit and model complexity compared to the other models. Therefore, based on the provided information, the unit inverse weibull distribution model is the most favorable choice.

Table 3: *Description of Models with AIC, AICC, and BIC Values*

Model	Log Likelihood	AIC	AICC	BIC
UIWD	164.4482	-324.8965	-324.6356	-321.1128
UIGD	161.1174	-318.2347	-317.9739	-314.4511
BD	162.1896	-320.3791	-320.1182	-316.5955
UBSD	161.1062	-318.2125	-317.9516	-314.4288
UWD	152.3815	-300.763	-300.5021	-296.9793
UGD	146.5113	-289.0226	-288.7617	-285.239
KD	164.3392	-324.6785	-324.4176	-320.8948

Table 4: *Model Comparison with AIC,AICC and BIC values*

Model	Log Likelihood	AIC	AICC	BIC
UIWD	60.5479	-117.0958	-116.9053	-112.7165
UIGD	60.0268	-116.0535	-115.8631	-111.6742
BD	57.57423	-111.1486	-110.9581	-106.7692
UBSD	59.9357	-115.8715	-115.681	-111.4921
UWD	53.9658	-103.9316	-103.7411	-99.5523
UGD	46.02843	-116.0535	-115.8631	-111.6742
KD	58.8343	-113.6686	-113.4782	-109.2893

8. Conclusion

In this study, we introduced a probability distribution called the Unit Inverse Weibull Distribution (UIWD) for modeling data on the interval $(0,1)$. The UIWD was derived through a type transformation involving the Inverse Weibull Distribution (IWD). We provided the probability density function (PDF) and cumulative distribution function (CDF) of the UIWD, along with their respective mathematical expressions.We highlighted the key features and properties of the UIWD, including its simple and closed-form expressions for the distribution and quantile functions, superior fitting performance compared to other unit interval distributions, and its ability to derive various statistical properties such as the survival function, hazard rate function, revised hazard rate function, cumulative hazard rate function, moments, mode, and order statistics.

For parameter estimation, we employed different estimation methods and conducted simulation studies to assess the accuracy and precision of different estimators. The results demonstrated the effectiveness and utility of the UIWD model in capturing and analyzing data on the unit interval $(0,1)$. Comparative analyses were performed to highlight the advantages of the UIWD model over other existing models. The UIWD showed superior performance in terms of fitting data and providing a flexible framework for statistical modeling.

Overall, this research contributes to the understanding and application of the UIWD as a flexible probability distribution for modeling data on the unit interval $(0,1)$. The UIWD offers a reliable alternative to existing models and can be effectively used in various fields requiring modeling and analysis of data on the unit interval.

References

[1] Topp, C. W., and Leone, F. C. (1955). A family of J-shaped frequency functions. *Journal of the American Statistical Association*,50: 209–219.

[2] Johnson, N. L. (1955). Systems of frequency curves derived from the first law of Laplace. *Trabajos de estadistica*, 5:283–291.

[3] Consul, P. C., and Jain, G. C. (1971). On the log-gamma distribution and its properties. *Statistische Hefte*, 12: 100–106.

[4] Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of hydrology*, 46:79–88.
[5] Arnold, B. C., and Groeneveld, R. A. (1980). Some properties of the arcsine distribution. *Journal of the American Statistical Association*, 75:173–175.
[6] Tadikamalla, P. R., and Johnson, N. L. (1982). Systems of frequency curves generated by transformations of logistic variables. *Biometrika*, 69:461–465.
[7] McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrika*, 52:647–663.
[8] Barndorff-Nielsen, O. E., and Jorgensen, B. (1991). Some parametric models on the simplex. *Journal of multivariate analysis*, 39:106–116.
[9] Van Dorp, J. R., and Kotz, S.(2002). The standard two-sided power distribution and its properties: with applications in financial engineering. *The American Statistician*, 56:90–99.
[10] Cordeiro, G. M.,and Lemonte, A. J. (2014). The McDonald arcsine distribution: A new model for proportional data. *Statistica*, 48:182–199.
[11] Gomez-Deniz, E., Sordo, M. A., and Calderin-Ojeda, E. (2014). The log–Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics*, 54:49–57.
[12] Korkmaz, M. C. (2020). A new heavy-tailed distribution defined on the bounded interval: the logit slash distribution and its application. *Journal of Applied Statistics*, 47:2097–2119.
[13] Altun, E., and Hamedani, G. (2018). The log-xgamma distribution with inference and application. *Journal de la Societe Francaise de Statistique*, 159:40–55.
[14] Mazucheli, J., Menezes, A. F., and Dey, S. (2018). The unit Birnbaum-Saunders distribution with applications. *Chilean Journal of Statistics*, 9:47–57.
[15] Mazucheli, J., Menezes, A. F. B.,and Chakraborty, S. (2019). On the one-parameter unit Lindley distribution and its associated regression model for proportion data. *Journal of Applied Statistics*, 46:700–714.
[16] Ghitany, M. E., Mazucheli, J., Menezes, A. F. B., and Alqallaf, F. (2018). The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval. *Communications in Statistics - Theory and Methods*, 48:3423–3438.
[17] Mazucheli, J., Menezes, A. F. B., and Dey, S. (2019). Unit-Gompertz distribution with applications. *STATISTICA*, 79:25–43.
[18] Altun, E. (2020). The log-weighted exponential regression model: An alternative to the beta regression model. *Communications in Statistics-Theory and Methods*, 50:2306-2321.
[19] Altun, E., and Cordeiro, G. M. (2020). The unit-improved second-degree Lindley distribution: Inference and regression modeling. *Computational Statistics*, 35:259–279.
[20] Korkmaz, M. C. (2020). A new heavy-tailed distribution defined on the bounded interval: The logit slash distribution and its application. *Journal of Applied Statistics*, 47:2097–2119.
[21] Korkmaz, M. C. (2020). The unit generalized half normal distribution: A new bounded distribution with inference and application. *University Politehnica of Bucharest Scientific Bulletin-Series A-Applied Mathematics and Physics*, 82:133–140.
[22] Abu El Azm, W. S., Almetwally, E. M., Naji AL-Aziz, S., El-Bagoury, A. A. A. H., Alharbi, R., and Abo-Kasem, O. E. (2021). A new transmuted generalized Lomax distribution: Properties and applications to COVID-19 data. *Computational Intelligence and Neuroscience*,Article ID 5918511,https://doi.org/10.1155/2021/5918511
[23] Nasiru, S., Abubakari, A. G., and Chesneau, C. (2022). New Lifetime Distribution for Modeling Data on the Unit Interval: Properties, Applications and Quantile Regression. *Mathematics and Computers in Applied Science*,27,105, https://doi.org/10.3390/mca27060105

RT&A, No 4 (76)
Volume 18, December 2023

On MIXTURE OF BURR XII AND NAKAGAMI DISTRIBUTIONS: PROPERTIES AND APPLICATIONS

Hemani Sharma and Parmil Kumar

•

Department of Statistics, University of Jammu, J&K.
hemanisharma124@gmail.com, parmil@yahoo.com

Abstract

The Burr XII and Nakagami distributions hold significant importance in both lifetime distribution and wealth distribution analyses. The Burr XII distribution serves as a valuable tool for understanding the distribution of wealth and wages within specific societies, while the Nakagami distribution finds its application in the realm of communication engineering. The incorporation of finite mixture distributions, aimed at accounting for unobserved variations, has gained substantial traction, particularly in the estimation of dynamic discrete choice models. This research delves into the fundamental characteristics of the mixture Burr XII and Nakagami distributions. The study introduces parameter estimation techniques and explores various aspects, including the cumulative distribution function, hazard rate, failure rate, inverse hazard function, odd function, cumulative hazard function, r^{th} moment, moment generating function, characteristic function, moments, mean and variance, Renyi and Beta entropies, mean deviation from mean, and mean time between failures (MTBF). The paper also addresses the estimation of the mixing parameter through a Bayesian approach. To illustrate the effectiveness of the proposed model, two real-life datasets are examined.

Keywords: Nakagami distribution, Burr XII, mixture distribution, Renyi and beta entropy, MTBF

1. Introduction

Mixture distributions play a pivotal role in modeling complex phenomena where a single underlying distribution cannot fully capture the observed data. These distributions offer a flexible framework by combining multiple component distributions, each representing a distinct pattern or source of variation. In various fields, ranging from statistics and economics to engineering and biology, mixture distributions have gained prominence due to their ability to accommodate diverse data characteristics. By utilizing mixture distributions, researchers can uncover latent subpopulations, account for unobserved factors, and enhance the accuracy of modeling real-world phenomena. The exploration of mixture distributions was instigated by Pearson [14], who initially investigated the blending of two normal distributions. Following Pearson's pioneering work [14], there emerged a significant hiatus in the advancement of mixture distributions. However, this period of inactivity was later reinvigorated by subsequent scholars. Decay [7] built upon Pearson's foundation, while Hasselblad [8] delved extensively into the intricate realm of finite mixture distributions. Numerous practical scenarios involve interpreting observed data as a blend originating from two or more distinct distributions. By leveraging this notion, we have the capability to merge different statistical distributions, thereby forging a novel distribution that inherits the traits of its individual components. This approach holds true particularly when dealing with continuous underlying random variables. An essential characteristic of a mixture distribution is that its cumulative distribution function (CDF) is formed by skillfully combining the CDFs of its constituent distributions through a convex amalgamation. Similarly, the probability density function (PDF) of the mixture can be elegantly expressed as a convex blend, seamlessly

incorporating the individual PDFs of its constituent parts. This versatile mathematical framework not only allows us to effectively handle intricate data but also enables the extraction of meaningful insights by synergistically harnessing the statistical attributes of diverse distributions. Assume T is a continuous random variable with the following probability density function:

$$f(t_i) = \sum_{i=1}^{k} p_j f_j(t); t > 0, k > 0 \tag{1}$$

where $\leq p_j \leq 1$, j=1,2,3...,k and$\sum_{i=1}^{k} p_i = 1$. The corresponding c.d.f. is given by

$$F(t_i) = \sum_{i=1}^{k} p_j F_j(t); t > 0, k > 0 \tag{2}$$

where k is the number of components, the parameters $p_1, p_2, ..., p_n$are called mixing parameters, where p_i, represent the probability that a given observation comes from population "i" with density $f_i(.)$, and $f_1(.), f_2(.), ..., f_3(.)$, are the component densities of the mixture. When the number of components k=2, a two component mixture and can be written as:

$$f(t) = p_1 f_1(t) + p_2 f_2(t) \tag{3}$$

Different research studies have explored various mixture distribution models and their applications. For example, Zaman and colleagues [19] introduced a mixture of chi-square distributions using Poisson elements. Chen and co-authors [4] discussed likelihood inference in finite mixture models, explaining different scenarios for these models. Martin and collaborators [11] utilized a mixture of chi-square distributions to describe the distribution of light in imaging, and they also introduced this distribution to model unstructured light distribution. Daniyel et al. [6] proposed a mixture of Burr II and Weibull distributions and examined their properties. Jaspers and co-workers [10] used a Bayesian approach to estimate mixing weights for multivariate mixture models. Nasiri and Azarian [13] explored estimating mixing proportions for a mixture of two chi-square distributions using various methods and compared their performance. Singh and colleagues [16] introduced a generalized distribution for modeling lifetime data and discussed its statistical properties with real-world applications. Daghestani and team [5] proposed a mixture of Lindley and Weibull distributions, investigated their statistical properties, and estimated their unknown parameters. In this context, this study focuses on the utilization of mixture distributions, specifically exploring the properties and applications of the mixture Burr XII and Nakagami distributions. By delving into their mathematical foundations, parameter estimation methods, and various statistical measures, we aim to provide a comprehensive understanding of how mixture distributions can enhance our ability to capture the intricate nuances present in real-world datasets.

The three-parameter Burr XII distribution was introduced by Titterington et al. [17]. For values of x greater than zero, its cumulative distribution function and probability density function are as follows:

$$F(x; q, d, b) = 1 - \left[1 + \left(\frac{x}{q}\right)^{b}\right]^{-d}; d, b, q > 0 \tag{4}$$

And

$$f(x; q, d, b) = bdq^{-b}x^{b-1}\left[1 + \left(\frac{x}{q}\right)^{b}\right]^{-d-1}; d, b, q > 0 \tag{5}$$

respectively, where d and b are the shape parameters and q is the scale parameter. It is frequently employed to model income data.

Nakagami [12] introduced the Nakagami distribution, also referred to as the Nakagami-m distribution, to describe radio signal fading. It involves a shape parameter denoted as 'm.' If a random variable 'X' conforms to the Nakagami distribution with a shape parameter $\alpha > 0.5$ and

a scale parameter $\lambda > 0$, its probability density function (p.d.f) can be expressed in the following manner:

$$f(x;\alpha,\lambda) = \frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha} x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right); x > 0, \alpha \geq 0.5, \lambda > 0 \tag{6}$$

The corresponding commulative distribution function is given by

$$F(x) = \frac{1}{\Gamma(\alpha)}\Gamma(\frac{\alpha}{\lambda}x^2,\alpha); x > 0, \alpha \geq 0.5, \lambda > 0 \tag{7}$$

where $\Gamma(x,a) = \int_0^x t^{a-1}e^{-t}dt$ is the lower incomplete gamma function.

This particular distribution has found applications spanning a diverse range of fields, encompassing hydrology, the transmission of multimedia data across networks, investigations into medical imaging, and the modeling of high-frequency seismogram envelopes.

The primary objective of this research paper is to introduce and explore a novel distribution, which emerges as a fusion of the Burr XII and Nakagami distributions, termed as the Mixture of Burr II and Nakagami Distribution (MBND). The study delves into the statistical properties inherent to this mixture distribution, in addition to the techniques for parameter estimation through Maximum Likelihood Estimation. Furthermore, the paper extends its investigation to include the estimation of the mixing parameter using a Bayesian approach. To validate the efficacy of the proposed model, real-world data is employed, allowing for a comparative assessment against other optimally fitting models.

2. Probability Density Function of MBND

The PDF of the mixture of Burr XII Nakagami distribution (MBND) has the following form

$$f(x;q,d,b,\alpha,\lambda) = p_1 f_1(x;q,d,b) + p_2 f_2(x;\alpha,\lambda) \tag{8}$$

Here, with p_1 and p_2 representing the mixing proportions where their sum equals one ($p_1 + p_2 = 1$), $f_1(x;q,d,b)$ signifies the PDF of the Burr XII distribution, and $f_2(x;\alpha,\lambda)$ denotes the PDF of the Nakagami distribution. Combining these individual probability densities results in the mixture density as follows:

$$f(x;q,d,b,\alpha,\lambda) = p_1 bdq^{-b}x^{b-1}\left[1+(\frac{x}{q})^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right) \tag{9}$$

where $x > 0$ and $d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

3. Cumulative Probability distribution of MBND

The cumulative distribution function is given by

$$F(x;q,d,b,\alpha,\lambda) = p_1 F_1(x;q,d,b) + p_2 F_2(x;\alpha,\lambda) \tag{10}$$

Here, with p_1 and p_2 representing the mixing proportions where their sum equals one ($p_1 + p_2 = 1$), $f_1(x;q,d,b)$ signifies CDF of the Burr XII distribution, and $f_2(x;\alpha,\lambda)$ denotes the CDF of the Nakagami distribution. Combining these individual probability densities results in the mixture density as follows:

$$F(x;q,d,b,\alpha,\lambda) = p_1\left(1-\left[1+(\frac{x}{q})^b\right]^{-d}\right) + p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma(\frac{\alpha}{\lambda}x^2,\alpha)\right) \tag{11}$$

where $x > 0$, $d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

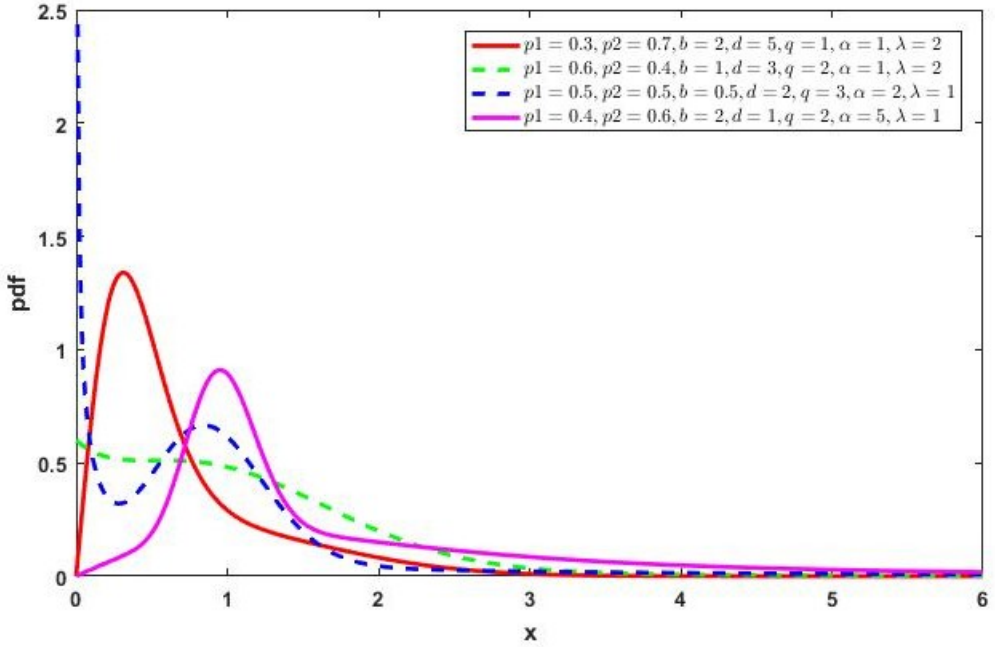


Figure 1: *Plot of the PDF for the MBND*

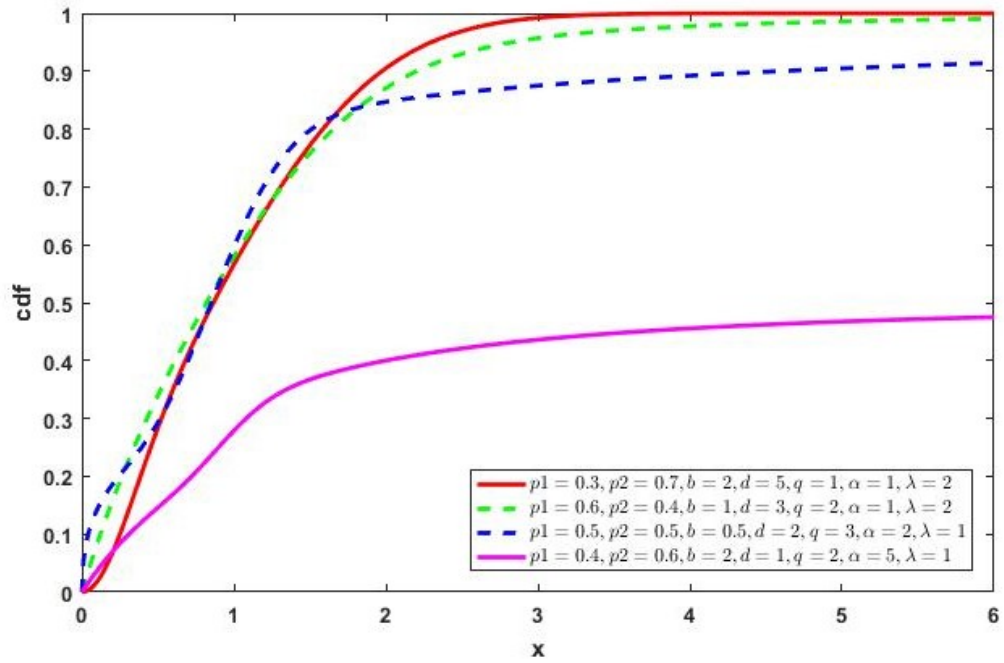


Figure 2: *Plot of the CDF for the MBND*

4. Properties of Bunami distribution

4.1. Area under the curve

Both the Burr XII and Nakagami distributions constitute complete probability density functions individually, their amalgamation as a mixture will similarly result in a comprehensive probability

density function.

4.2. Reliability function

The reliability function, also called the survival function, describes a feature of a random variable related to how a system might stop working within a certain period.It is defined as the probability that the system will keep working beyond a particular point in time. Mathematically,

$$R(x) = 1 - F(x) \tag{12}$$

By inserting equation (11) into equation (12), the reliability or survival function for the MBND can be expressed in the following manner

$$R(x) = 1 - \left(p_1 \left(1 - \left[1 + \left(\frac{x}{q} \right)^b \right]^{-d} \right) + p_2 \left(\frac{1}{\Gamma(\alpha)} \Gamma\left(\frac{\alpha}{\lambda} x^2, \alpha \right) \right) \right) \tag{13}$$

where $x > 0, d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

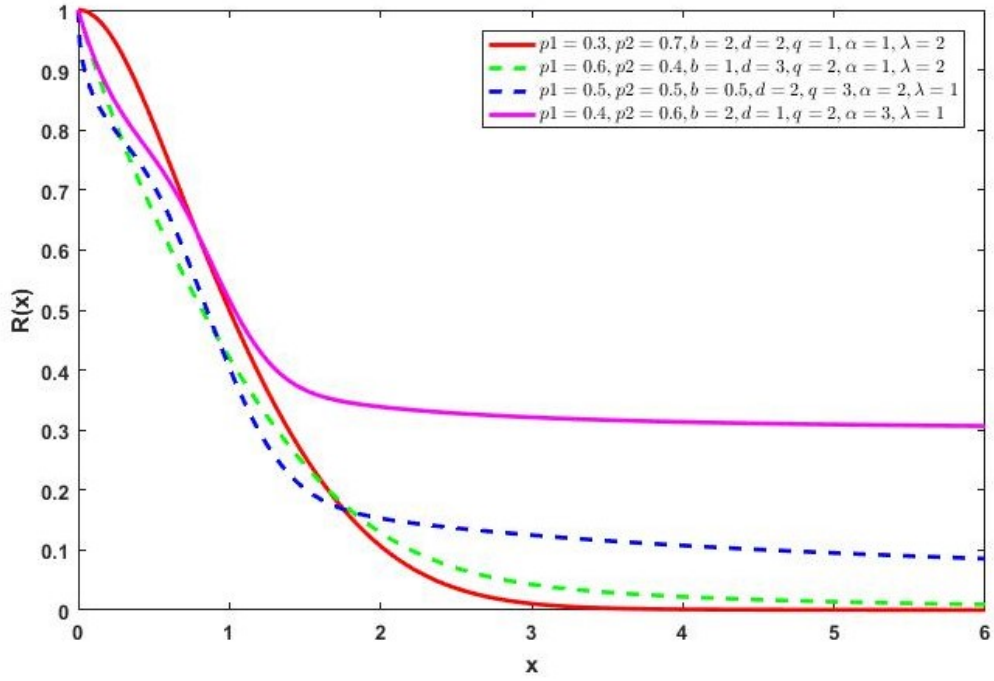


Figure 3: *Plot of the Reliability function for the MBND*

4.3. Hazard Function

The hazard function is mathematically defined as the quotient of the probability density function divided by the reliability function, and its expression takes the following form:

$$h(x) = \frac{f(x)}{R(x)} \tag{14}$$

By substituting equations (9) and (13) into equation (14), the hazard rate for the MBND can be derived, leading to the subsequent definition.

$$h(x) = \frac{p_1 bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)}{1-\left(p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right)+p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma\left(\frac{\alpha}{\lambda}x^2,\alpha\right)\right)\right)} \tag{15}$$

where $x > 0, d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

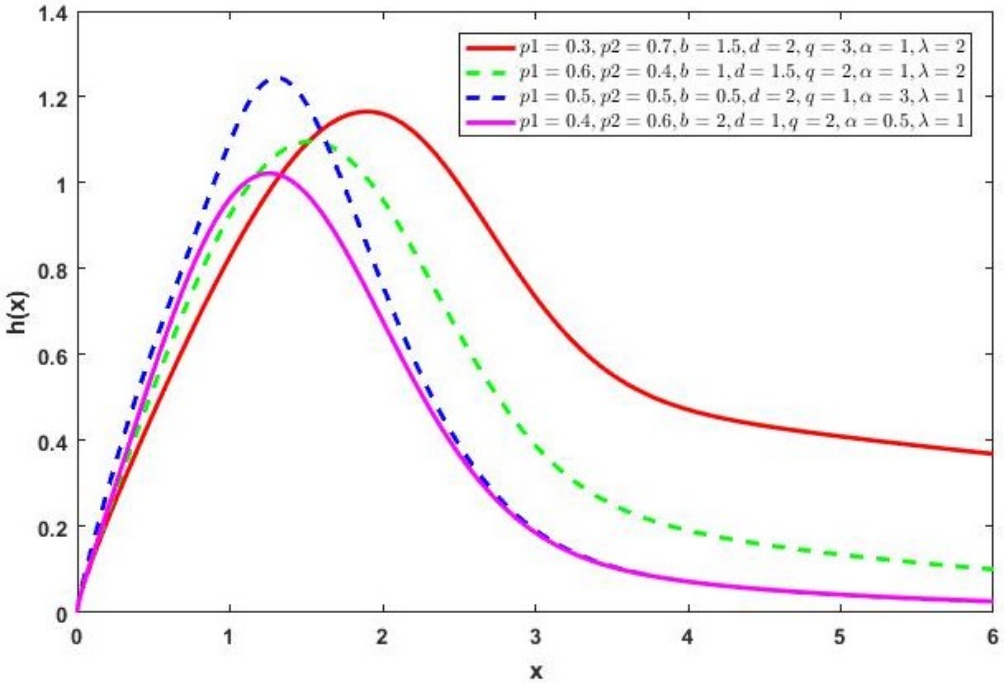


Figure 4: *Plot of the Hazard function of the MBND*

4.4. Cumulative Hazard Function

The cumulative hazard function can be defined as

$$\Lambda(x) = -\log R(x) \tag{16}$$

By substituting equation (13) into equation (16), the cumulative hazard function for the MBND can be derived, yielding the subsequent expression.

$$\Lambda(x) = -\log\left[1-\left(p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right)+p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma\left(\frac{\alpha}{\lambda}x^2,\alpha\right)\right)\right)\right] \tag{17}$$

where $x > 0, d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

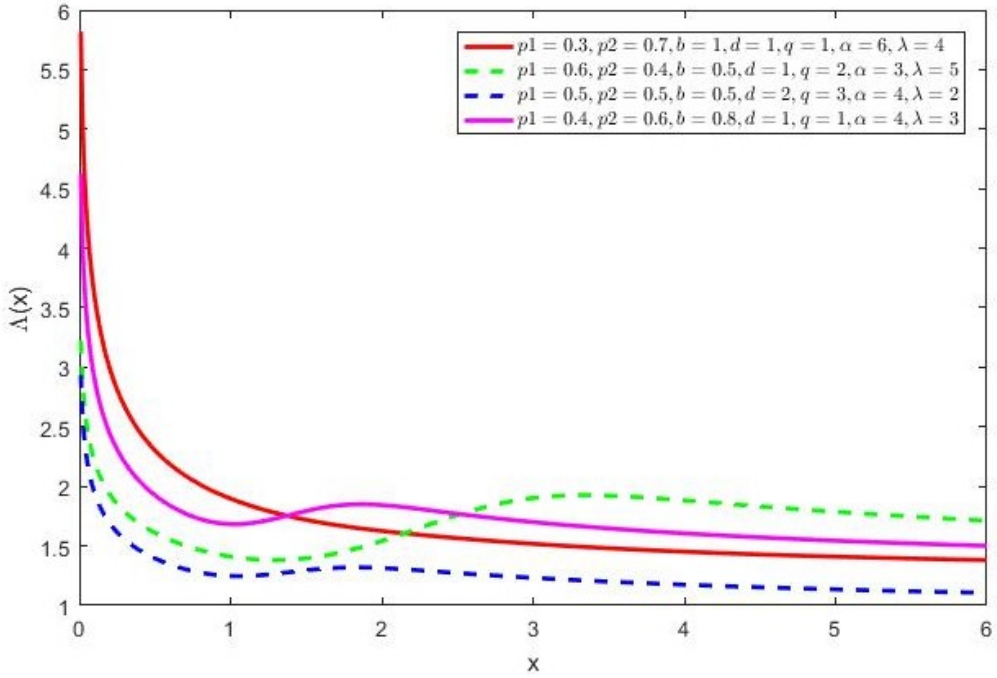


Figure 5: *Plot of Cumulative Hazard function for the MBND*

4.5. Reversed Hazard Rate

The reversed hazard rate can be conceptually understood as the result of dividing the probability density function by the cumulative distribution function, as demonstrated by the following formulation:

$$r(x) = \frac{f(x)}{F(x)} \tag{18}$$

The reversed hazard rate corresponding to the MBND can be derived by substituting equations (9) and (11) into equation (18), resulting in the following expression.

$$r(x) = \frac{p_1 bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)}{p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right)+p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma\left(\frac{\alpha}{\lambda}x^2,\alpha\right)\right)} \tag{19}$$

where $x > 0$, $d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

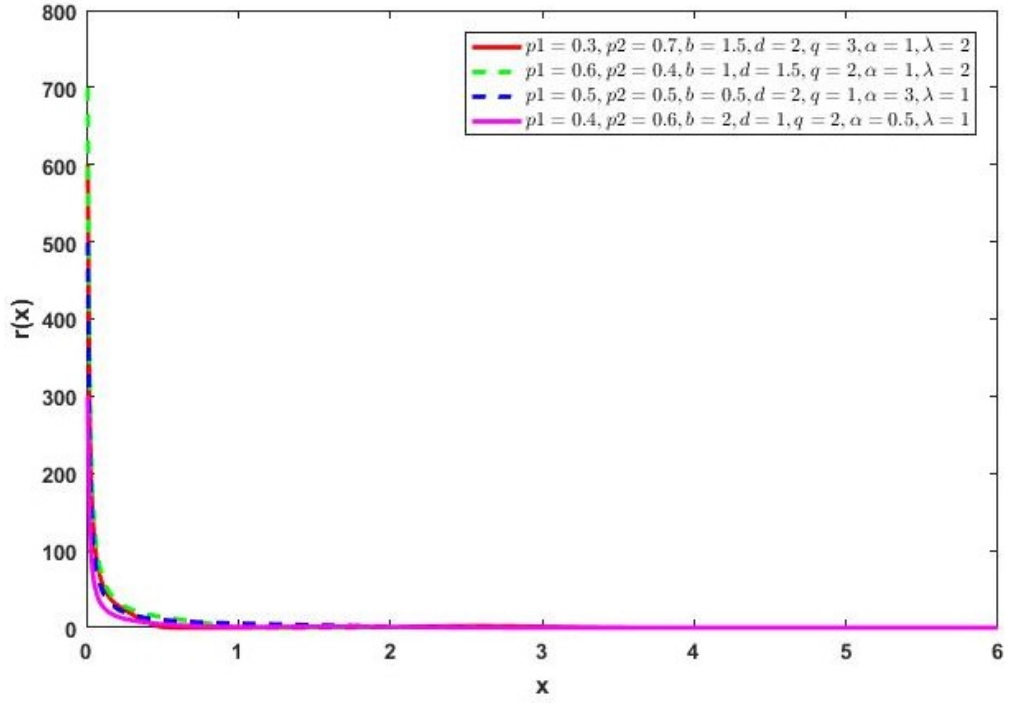


Figure 6: *Plot of the Reversed Hazard function for the MBND*

4.6. Odds Function

The odds function, represented as O(x), is characterized as the proportion of the cumulative distribution function to the reliability function, and it is expressed through the following equation:

$$O(x) = \frac{F(x)}{R(x)} \tag{20}$$

The odds function corresponding to the MBND can be derived by substituting equations (11) and (13) into equation (20), resulting in the following expression.

$$O(x) = \frac{p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right)+p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma\left(\frac{\alpha}{\lambda}x^2,\alpha\right)\right)}{1-\left(p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right)+p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma\left(\frac{\alpha}{\lambda}x^2,\alpha\right)\right)\right)} \tag{21}$$

where $x > 0$, $d, b, q >, \lambda > 0, \alpha \geq 0.5$, $p_1 + p_2 = 1$.

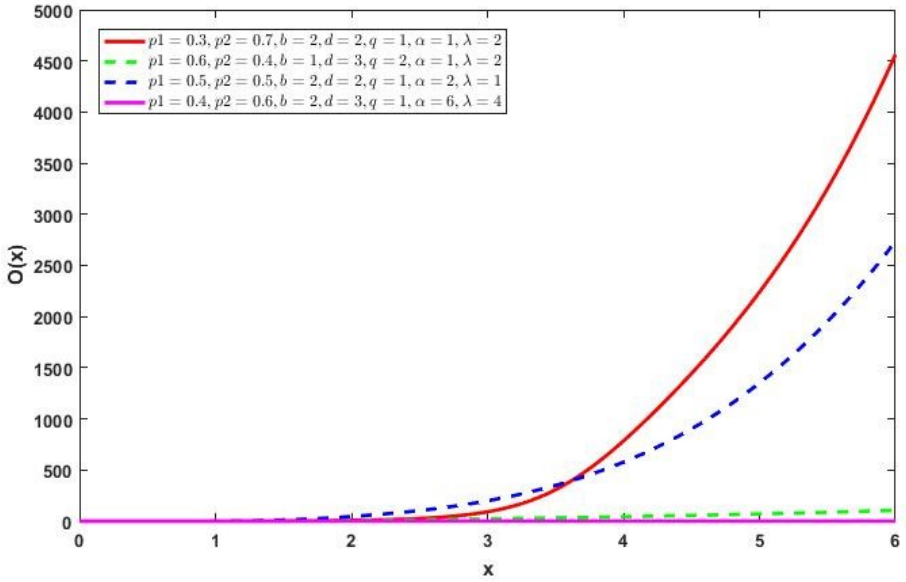


Figure 7: *Plot of the Odd Function for the MBND*

4.7. r^{th} Moment about Origin

The r^{th} moment of the real-valued function can be acquired by applying

$$\mu_r^{'} = \int x^r f(x) dx \tag{22}$$

$$\mu_r^{'} = \int_0^{\infty} x^r \left(p_1 bdq^{-b} x^{b-1} \left[1 + (\frac{x}{q})^c\right]^{-k-1} + p_2 \frac{2}{\Gamma(\alpha)} \left(\frac{\alpha}{\lambda}\right)^{\alpha} x^{2\alpha-1} \exp\left(\frac{-\alpha}{\lambda} x^2\right)\right) dx \tag{23}$$

$$\mu_r^{'} = p_1 \int_0^{\infty} x^r bdq^{-c} x^{c-1} \left[1 + (\frac{x}{q})^b\right]^{-d-1} dx + p_2 \int_0^{\infty} x^r \frac{2}{\Gamma(\alpha)} \left(\frac{\alpha}{\lambda}\right)^{\alpha} x^{2\alpha-1} \exp\left(\frac{-\alpha}{\lambda} x^2\right) dx \tag{24}$$

$$\mu_r^{'} = p_1 q^r \frac{\Gamma(\frac{r}{b}+1)\Gamma(k-\frac{r}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{r}{2})}{\Gamma\alpha} \left(\frac{\lambda}{\alpha}\right)^{\frac{r}{2}} \tag{25}$$

4.8. Raw moments about Origin

Putting r=1, 2, 3 and 4 in (25),first four raw moments are

$$\mu_1^{'} = p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha} \left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}} = Mean \tag{26}$$

$$\mu_2^{'} = p_1 q^2 \frac{\Gamma(\frac{2}{b}+1)\Gamma(k-\frac{2}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+1)}{\Gamma\alpha} \left(\frac{\lambda}{\alpha}\right) \tag{27}$$

$$\mu_3^{'} = p_1 q^3 \frac{\Gamma(\frac{3}{b}+1)\Gamma(k-\frac{3}{b})}{\Gamma(d)} p_2 \frac{\Gamma(\alpha+\frac{3}{2})}{\Gamma\alpha} \left(\frac{\lambda}{\alpha}\right)^{\frac{3}{2}} \tag{28}$$

$$\mu_4^{'} = p_1 q^4 \frac{\Gamma(\frac{4}{b}+1)\Gamma(k-\frac{4}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+2)}{\Gamma\alpha} \left(\frac{\lambda}{\alpha}\right)^{2} \tag{29}$$

4.9. Moments about Mean

$$\mu_1 = 0 \tag{30}$$

$$\mu_2 = \mu_2^{'} - (\mu_1^{'})^2 = Variance \tag{31}$$

$$\mu_2 = p_1 q^2 \frac{\Gamma(\frac{2}{b}+1)\Gamma(k-\frac{2}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+1)}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right) - \left(p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right)^2 \tag{32}$$

$$\mu_3 = \mu_3^{'} - 3\mu_2^{'}\mu_1^{'} + (\mu_1^{'})^3$$

$$\mu_3 = p_1 q^3 \frac{\Gamma(\frac{3}{b}+1)\Gamma(k-\frac{3}{b})}{\Gamma(d)} p_2 \frac{\Gamma(\alpha+\frac{3}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{3}{2}} - 3\left[p_1 q^2 \frac{\Gamma(\frac{2}{b}+1)\Gamma(k-\frac{2}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+1)}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)\right]$$
$$\left[p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right] + \left[p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right]^3 \tag{33}$$

$$\mu_4 = \mu_3^{'} - 4\mu_1^{'}\mu_3^{'} + 6(\mu_1^{'})^2(\mu_2^{'}) - 3(\mu_1^{'})^4$$

$$\mu_4 = p_1 q^3 \frac{\Gamma(\frac{3}{b}+1)\Gamma(k-\frac{3}{b})}{\Gamma(d)} p_2 \frac{\Gamma(\alpha+\frac{3}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{3}{2}} - 4\left(p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right)$$
$$p_1 q^3 \frac{\Gamma(\frac{3}{b}+1)\Gamma(k-\frac{3}{b})}{\Gamma(d)} p_2 \frac{\Gamma(\alpha+\frac{3}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{3}{2}} + 6\left(p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right)^2$$
$$p_1 q^2 \frac{\Gamma(\frac{2}{b}+1)\Gamma(k-\frac{2}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+1)}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right) - 3\left(p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right)^4 \tag{34}$$

4.10. Measure of Skewness and Kurtosis

Skewness refers to the measure of asymmetry exhibited by a probability distribution. It indicates the extent to which the distribution leans or deviates from being symmetric around its mean. Kurtosis helps characterize the peakedness or flatness of the distribution's curve. Skewness and kurtosis are symbolized as β_1 and β_2 respectively, and their mathematical representation is given as:

$$\beta_1 = \frac{(\mu_3)^2}{(\mu_2)^3} \tag{35}$$

Using eq.(32) and (33) in the above equation, we can get the value of skewness.

And,

$$\beta_2 = \frac{\mu_4}{(\mu_2)^2} \tag{36}$$

Using eq.(32) and (34) in the equation (36), the expression for the β_2 can be obtained.

4.11. Moment Generating Function

The moment generating function is a mathematical function that can be defined as follows: $M_x(t) = E(e^{tx})$, provided it exists.

where $E(e^{tx}) = \int e^{tx} f(x)dx$ and is given by

$$E(e^{tx}) = \int_0^\infty e^{tx}\left(p_1 bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^\alpha x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right)dx$$

$$E(e^{tx}) = p_1\int_0^\infty e^{tx}bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1}dx + p_2\int_0^\infty e^{tx}\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^\alpha x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)dx \tag{37}$$

$$E(e^{tx}) = p_1\sum_{j=0}^{\infty}\frac{(qt)^j}{j!}\frac{\Gamma(\frac{j}{b}+1)\Gamma(k-\frac{j}{b})}{\Gamma(d)} + p_2\left[1-t\sqrt{\frac{\lambda}{\alpha}}\right]^{-\alpha} \tag{38}$$

4.12. Characteristic Function

The characteristic function is derived from the moment generating function by substituting 't' with 'it', and its expression is given as follows:

$$\phi_x(t) = E(e^{itx}) \tag{39}$$

$$E(e^{itx}) = p_1\sum_{j=0}^{\infty}\frac{(qit)^j}{j!}\frac{\Gamma(\frac{j}{b}+1)\Gamma(k-\frac{j}{b})}{\Gamma(d)} + p_2\left[1-it\sqrt{\frac{\lambda}{\alpha}}\right]^{-\alpha} \tag{40}$$

4.13. Mean deviation about Mean

The mean deviation represents the average of the absolute differences between data points and the mean of a dataset. It offers a way to estimate the level of variability in a population. The mean deviation about mean can be obtained as follows:

$$\delta_1(x) = \int_0^\infty |x-\mu| f(x)dx \tag{41}$$

where $\mu = E(x)$. It can be calculated by the following relationship

$$\delta_1(x) = \int_0^\mu (\mu - x)f(x)dx + 2\int_\mu^\infty (\mu - x)f(x)dx$$

$$\delta_1(x) = 2\int_0^\mu (\mu - x)f(x)dx$$

$$\delta_1(x) = 2\left[\mu F(\mu) + \int_0^\mu xf(x)dx\right] \tag{42}$$

$$\delta_1(x) = 2\Bigg[\left(p_1 q^1\frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2\frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right)\left(p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right) + p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma(\frac{\alpha}{\lambda}\mu^2,\alpha)\right)\right)$$
$$-\left(p_1 bd\sum_{j=0}^{\infty}(-1)^j\binom{d+1}{j}\frac{\left(\frac{\mu}{q}\right)^{\frac{j+1}{b}+1}}{j+\frac{1}{b}+1} + p_2\frac{1}{\Gamma(\alpha)}\sqrt{\frac{\lambda}{\alpha}}\sum_{i=0}^{\infty}\frac{(-i)}{i!}\frac{\left(\frac{\alpha}{\lambda}\mu^2\right)^{i+\alpha-\frac{1}{2}}}{i+\alpha+\frac{1}{2}}\right)\Bigg] \tag{43}$$

4.14. Renyi Entropy

The quantification of uncertainty inherent in random variable X can be computed using Renyi entropy [15]. This can be articulated using the subsequent equation:

$$\delta_r(x) = \frac{1}{1-r}\log\left(\int f^r(x)dx\right) \tag{44}$$

where

$$f^r(x) = \left[p_1 bdq^{-b}x^{b-1}\left[1+(\frac{x}{q})^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right]^r \tag{45}$$

By using binomial expansion

$$(a+b)^n = \sum_{k=0}^{\infty}\binom{n}{k}a^{n-k}b^k$$

The integral will take the form

$$= \sum_{k=0}^{\infty}\binom{r}{n}\int_0^{\infty}\left[p_1 bdq^{-b}x^{b-1}\left[1+(\frac{x}{q})^b\right]^{-d-1}\right]^{r-n}\left[p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right]^n dx$$

Due to their independence, the integral can be treated as the multiplication of the two integrals, leading to the resulting expression for the Renyi entropy as:

$$\delta_r(x) = \frac{1}{1-r}\log\left[\frac{(p_1 bdq^{-b})^{r-n}q^{(b-1)(r-n)+1}}{b}\frac{\Gamma\left(\frac{(r-n)(b-1)+1}{b}\right)\Gamma\left(\frac{brd-bnd+r-n-1}{b}\right)}{\Gamma\left(\frac{(r-n)(b-1)+1}{b}+\frac{brd-bnd+r-n-1}{b}\right)}\right.$$
$$\left. + p_2\frac{2^{n-1}}{(\Gamma(\alpha))^n}\left(\frac{\alpha}{\lambda}\right)^{1-\frac{n}{2}}\left(\frac{1}{n}\right)^{\left(\alpha n-\frac{n}{2}+1\right)}\Gamma(\alpha n-\frac{n}{2}+\frac{1}{2})\right] \tag{46}$$

4.15. β-Entropy

The β-entropy is formulated as an extension of the Shannon entropy with a single parameter. This concept was initially introduced by Havrda and Charvat [9], and subsequently adapted for use in physical contexts by Tsallis [18]. The β-entropy can be defined as:

$$H_{\overline{\beta}} = \frac{1}{\overline{\beta}-1}\left[1-\int_0^{\infty}f^{\overline{\beta}}(x)\right]dx, for\overline{\beta}\neq 1. \tag{47}$$

β-entropy for the MBND is given by

$$H_{\overline{\beta}} = \frac{1}{\overline{\beta}-1}\int_0^{\infty}\left[p_1 bdq^{-b}x^{b-1}\left[1+(\frac{x}{q})^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right]^{\overline{\beta}}dx, for\overline{\beta}\neq 1. \tag{48}$$

By using binomial expansion

$$(a+b)^n = \sum_{k=0}^{\infty}\binom{n}{k}a^{n-k}b^k$$

The integral will take the form

$$= \sum_{n=0}^{\overline{\beta}}\overline{\beta}n\int_0^{\infty}\left[p_1 bdq^{-b}x^{b-1}\left[1+(\frac{x}{q})^b\right]^{-d-1}\right]^{\overline{\beta}-n}\left[p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right]^n dx$$

Due to their independence, the integral can be treated as the multiplication of the two integrals, leading to the resulting expression for the β-entropy as:

$$H_{\overline{\beta}} = \frac{1}{1-\overline{\beta}}\Bigg[\frac{(p_1 bdq^{-b})^{\overline{\beta}-n} q^{(b-1)(\overline{\beta}-n)+1}}{b}\frac{\Gamma\left(\frac{(\overline{\beta}-n)(b-1)+1}{b}\right)\Gamma\left(\frac{b\overline{\beta}d-bnd+\overline{\beta}-n-1}{b}\right)}{\Gamma\left(\frac{(\overline{\beta}-n)(b-1)+1}{b}+\frac{b\overline{\beta}d-bnd+\overline{\beta}-n-1}{b}\right)} + p_2\frac{2^{n-1}}{(\Gamma(\alpha))^n}\left(\frac{\alpha}{\lambda}\right)^{1-\frac{n}{2}}\left(\frac{1}{n}\right)^{\left(\alpha n-\frac{n}{2}+1\right)}\Gamma\left(\alpha n-\frac{n}{2}+\frac{1}{2}\right)\Bigg] \tag{49}$$

4.16. Mean Time between failures

The Mean Time Between Failures, often abbreviated as MTBF, serves as an indicator of a product's or component's reliability. This metric represents the average interval between successive failures in a process. Its calculation involves determining:

$$MTBF = \int_0^\infty t f(t) dt \tag{50}$$

It can be calculated as

$$MTBF = \int_0^\infty t\left(p_1 bdq^{-b}t^{b-1}\left[1+\left(\frac{t}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^\alpha t^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}t^2\right)\right)dt \tag{51}$$

$$= p_1\int_0^\infty tbdq^{-b}t^{b-1}\left[1+\left(\frac{t}{q}\right)^b\right]^{-d-1}dx + p_2\int_0^\infty x^r\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^\alpha t^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}t^2\right)dt \tag{52}$$

$$MTSF = p_1 q\frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2\frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}} \tag{53}$$

4.17. Bonferoni and Lorenz curve

The Bonferroni Curve, introduced by Bonferroni [3], serves as a method for analyzing income inequality. It finds utility across various domains, including the examination of economic factors like income and poverty, as well as applications in reliability, demographics, insurance, and medicine.

On the other hand, the Lorenz curve provides a visual representation of the distribution of income or wealth. Though commonly associated with depicting economic inequality, it holds the ability to illustrate uneven distributions in diverse systems. The curvature of the Lorenz curve relative to a baseline, often represented as a straight line, signifies the extent of inequality – the greater the inequality, the farther the curve deviates from this baseline. The following expression is used to calculate the Bonferroni curve:

$$B_F[F(x)] = \frac{1}{\mu F(x)}\int_0^x u f(u) du \tag{54}$$

where

$$\int_0^x u f(u) du = \int_0^x u\left(p_1 bdq^{-b}u^{b-1}\left[1+\left(\frac{u}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^\alpha u^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}u^2\right)\right)du \tag{55}$$

$$= \int_0^x p_1 ubdq^{-b}u^{b-1}\left[1+\left(\frac{u}{q}\right)^b\right]^{-d-1}dx + \int_0^x p_2 u\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^\alpha u^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}u^2\right)du \tag{56}$$

$$\int_0^x uf(u)du = p_1 dq \sum_{j=0}^{\infty}(-1)^j \binom{d+1}{j} \frac{\left(\frac{x}{q}\right)^{\frac{j+1}{b}+1}}{j+\frac{1}{b}+1} + p_2 \frac{1}{\Gamma(\alpha)} \sqrt{\frac{\lambda}{\alpha}} \sum_{i=0}^{\infty} \frac{(-i)}{i!} \frac{\left(\frac{\alpha}{\lambda}x^2\right)^{i+\alpha-\frac{1}{2}}}{i+\alpha+\frac{1}{2}} \tag{57}$$

$$B_F[F(x)] = \frac{p_1 dq \sum_{j=0}^{\infty}(-1)^j \binom{d+1}{j} \frac{\left(\frac{x}{q}\right)^{\frac{j+1}{b}+1}}{j+\frac{1}{b}+1} + p_2 \frac{1}{\Gamma(\alpha)} \sqrt{\frac{\lambda}{\alpha}} \sum_{i=0}^{\infty} \frac{(-i)}{i!} \frac{\left(\frac{\alpha}{\lambda}x^2\right)^{i+\alpha-\frac{1}{2}}}{i+\alpha+\frac{1}{2}}}{\left(p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}\right)\left(p_1\left(1-\left[1+\left(\frac{x}{q}\right)^b\right]^{-d}\right) + p_2\left(\frac{1}{\Gamma(\alpha)}\Gamma\left(\frac{\alpha}{\lambda}\mu^2,\alpha\right)\right)\right)} \tag{58}$$

Now, Lorenz curve is given as

$$L(z) = \frac{\int_0^z xf(x)dx}{\mu} \tag{59}$$

where

$$\int_0^z xf(x)du = \int_0^z x\left[p_1 bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right]dx \tag{60}$$

The final expression for Lorenz curve is

$$L(z) = \frac{p_1 dq \sum_{j=0}^{\infty}(-1)^j \binom{d+1}{j} \frac{\left(\frac{z}{q}\right)^{\frac{j+1}{b}+1}}{j+\frac{1}{b}+1} + p_2 \frac{1}{\Gamma(\alpha)} \sqrt{\frac{\lambda}{\alpha}} \sum_{i=0}^{\infty} \frac{(-i)}{i!} \frac{\left(\frac{\alpha}{\lambda}z^2\right)^{i+\alpha-\frac{1}{2}}}{i+\alpha+\frac{1}{2}}}{p_1 q^1 \frac{\Gamma(\frac{1}{b}+1)\Gamma(k-\frac{1}{b})}{\Gamma(d)} + p_2 \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma\alpha}\left(\frac{\lambda}{\alpha}\right)^{\frac{1}{2}}} \tag{61}$$

5. Maximum Likelihood Estimation

Suppose $x_1, x_2, ...x_n$ be the random sample derived from the MBND having the probability density function defined in (9). Therefore, for n observations , the logarithm of the likelihood function is expressed as below

$$L(x) = \prod_{i=1}^{n}\left(p_1 bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1} + p_2\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right) \tag{62}$$

The log likelihood function is given as

$$logL(x) = \sum_{i=0}^{n} log\left(p_1 bdq^{-b}x^{b-1}\left[1+\left(\frac{x}{q}\right)^b\right]^{-d-1} + (1-p_1)\frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{\lambda}\right)^{\alpha}x^{2\alpha-1}\exp\left(\frac{-\alpha}{\lambda}x^2\right)\right)$$

$$= nlog(p_1) + nlog(b) + nlog(d) + -nblog(q) + (b-1)log\sum_{i=1}^{n}x_i + (-d-1)log\sum_{i=1}^{n}\left[1+\left(\frac{x_i}{q}\right)^b\right] + nlog(1-p_1)$$
$$+nlog\left(\frac{2}{\Gamma(\alpha)}\right) + \alpha nlog\left(\frac{\alpha}{\lambda}\right) + (2\alpha-1)log\sum_{i=1}^{n}x_i + log\sum_{i=1}^{n}\left(\frac{-\alpha}{\lambda}x_i^2\right) \tag{63}$$

The equation are in implicit form, so it can be solved using numerical iteration method, such as the Newton–Raphson method via R to obtain the estimates of b, d, q, α, λ and p.

6. Bayesian estimation for mixing parameter

To estimate the mixing parameter p_1 within the context of MBND, we can adopt the Beta distribution as a suitable prior due to the range constraint $p_1 \in (\theta, 1)$. Let's assume that the prior distribution for p_1 is represented by $\beta(\theta, 1)$.

$$\pi(p) = \theta p_1^{\theta-1} \tag{64}$$

The likelihood function is given by

$$L(x) = \prod_{i=1}^{n} \left(p_1 bdq^{-b} x^{b-1} [1 + (\frac{x}{q})^b]^{-d-1} + p_2 \frac{2}{\Gamma(\alpha)} \left(\frac{\alpha}{\lambda}\right)^{\alpha} x^{2\alpha-1} \exp\left(\frac{-\alpha}{\lambda} x^2\right)\right) \tag{65}$$

And, the posterior distribution of p_1 is obtained by

$$\pi(p|x) = \frac{L(x)\pi(p)}{\int_0^1 L(x)\pi(p)dp} \tag{66}$$

The marginal distribution is given as

$$m(x) = \int_0^1 L(x)\pi(p)dp \tag{67}$$

$$m(x) = \int_0^1 p_1^{\theta-1} \left[(p_1 bdq^{-b})^n \sum_{i=1}^{n} x_i^{c-1} \prod_{i=1}^{n} \left(1 + \left(\frac{x_i}{b}\right)^c\right)^{-k-1} + (1-p_1)^n \left(\frac{2}{\Gamma(\alpha)}\right)^n \left(\frac{\alpha}{\lambda}\right)^{\alpha n} \sum_{i=1}^{n} x_i^{2\alpha-1} \prod_{i=1}^{n} \exp\left(\frac{-\alpha}{\lambda} x_i^2\right)\right) \right] dp$$

$$m(x) = f_1^n(x; b, d, q)\left(\frac{\theta}{\theta+n}\right) + \theta f_2^n(x; \alpha, \lambda)\beta(\theta, n+1) \tag{68}$$

Eq.(66) becomes

$$\pi(p|x) = \frac{\theta p_1^{\theta-1}\left(p_1^n f_1^n(x; b, d, q) + (1-p_1)^n f_2^n(x; \alpha, \lambda)\right)}{f_1^n(x; b, d, q)\left(\frac{\theta}{\theta+n}\right) + \theta f_2^n(x; \alpha, \lambda)\beta(\theta, n+1)} \tag{69}$$

$$= \frac{\theta p_1^{n+\theta-1} f_1^n(x; b, d, q)}{f_1^n(x; b, d, q)\left(\frac{\theta}{\theta+n}\right) + \theta f_2^n(x; \alpha, \lambda)\beta(\theta, n+1)} + \frac{(1-p_1)^n \theta p_1^{\theta-1} f_2^n(x; \alpha, \lambda)}{f_1^n(x; b, d, q)\left(\frac{\theta}{\theta+n}\right) + \theta f_2^n(x; \alpha, \lambda)\beta(\theta, n+1)}$$

The Bayesian estimate of p_1 under Squared Error Loss Function (SELF) is the mean of the posterior distribution denoted as $\hat{p_{B1}}$.

7. Simulation Study

In this part, a Monte Carlo simulation is performed using R for estimating the parameters of mixture of Burr XII and Nakagami distribution (MBND) . The values of n and the parameters are fixed as $n = (20, 40, 70, 100)$, $(b, d, q, \alpha, \lambda) = (1.5, 1, 1.5, 1.5, 1)$ respectively, the mixing parameter p_1 is fixed at 0.8 which is estimated using the classical and bayesian approach. The performance of the estimates was evaluated using the Mean Squared Error (MSE). The procedure is replicated 1000 times and obtained results are shown in table 1.

$$MSE = \frac{1}{N} \sum_{i=1}^{n} (\hat{\alpha} - \alpha)^2$$

Table 1: *Average values of estimates and their MSE in parentheses*

n	$\hat{b}$	$\hat{d}$	$\hat{q}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{p}_1$	$\hat{p}_{B1}$
20	1.9703	1.2695	0.8140	2.4148	0.6744	0.3404	0.4231
	(0.8963)	(0.5409)	(0.7580)	(0.7568)	(0.7631)	(0.4065)	(0.3932)
40	1.7349	0.8504	0.9948	2.3071	0.7112	0.5763	0.5892
	(0.7130)	(0.5140)	(0.4177)	(0.7200)	(0.6878)	(0.3896)	(0.3624)
70	1.6313	0.9631	1.4108	2.1647	0.8132	0.6296	0.6742
	(0.5990)	(0.3167)	(0.3139)	(0.4988)	(0.6157)	(0.2329)	(0.2004)
100	1.3119	0.9726	1.4372	1.9322	0.9542	0.7608	0.7814
	(0.2175)	(0.1400)	(0.1880)	(0.2604)	(0.2332)	(0.1498)	(0.1413)

From the table 1, it is noted that

• As the sample size increases, the MSE of the estimates comes down.

• With increase of n, the estimates comes closer to their true values, shows the consistency of the estimates.

• The mixing parameter performs better under the bayesian approach.

8. Real Data Analysis

In this section, we will use two real-world datasets to demonstrate the significance and versatility of the Mixture of Burr XII and Nakagami Distribution (MBND). We will compare the suitability of our proposed model (MBND) with other competing models, namely the Burr XII Lomax Distribution (BLD), Burr II Distribution (BD), and the Length Based Nakagami Distribution (LBND).
To conduct this comparison, we will employ various goodness-of-fit metrics, including the $-2\hat{l}$ statistic, the Akaike Information Criterion (AIC), the Corrected Akaike Information Criterion (AICc), and the Bayesian Information Criterion (BIC). These criteria aid in selecting the most suitable distribution by identifying the one with the lowest values of AIC, AICc, and BIC. Additionally, we will calculate the Kolmogorov-Smirnov (KS) distance and its associated p-value, providing a further assessment of the goodness of fit.

$$AIC = 2k - 2\hat{l} \tag{70}$$

$$BIC = kln(n) - 2\hat{l} \tag{71}$$

and

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \tag{72}$$

where $\hat{l}$ is the maximized log-likelihood, k is the number of parameters, and n is the sample size. Through this comprehensive analysis, we aim to establish the effectiveness of the MBND model in capturing the characteristics of the given datasets and to ascertain its superiority over the alternative distribution models.

Data Set I: The dataset pertains to the COVID-19 vaccination rate across 46 distinct countries situated in southern Africa. This dataset has been previously subjected to analysis by [2]. The data is as follows: 0.042, 0.205, 0.285, 0.319, 0.464, 0.550, 0.889, 0.895, 0.939, 0.986, 1.000, 1.088, 1.212, 1.244, 1.450, 1.593, 1.844, 2.039, 2.157, 2.167, 2.334, 2.440, 2.657, 3.685, 3.879, 4.493, 4.800, 4.944, 5.155, 5.674, 7.602, 10.004, 12.238, 12.520, 12.553, 13.063, 15.105, 15.229, 15.629, 15.848, 18.641, 18.940, 29.885, 58.162, 61.838, 72.286

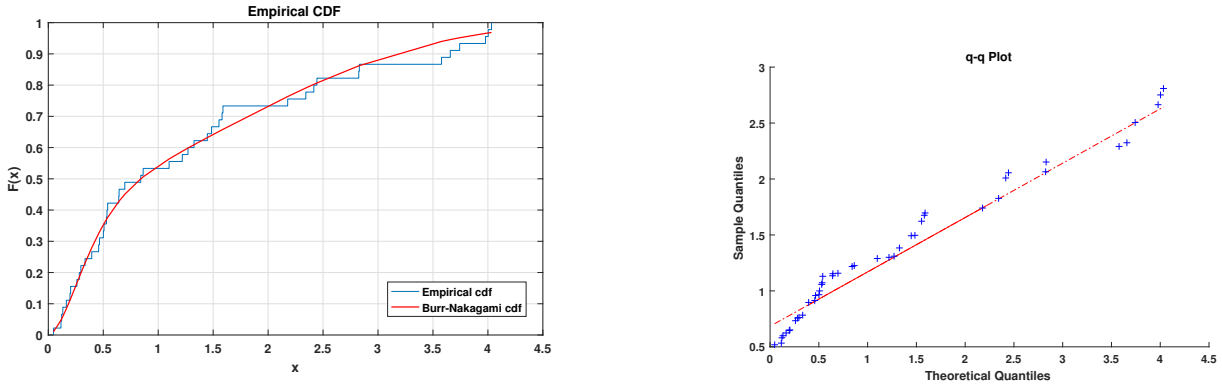


Figure 8: *(a) Ecdf plot for the dataset I (b) q-q plot for the dataset I*

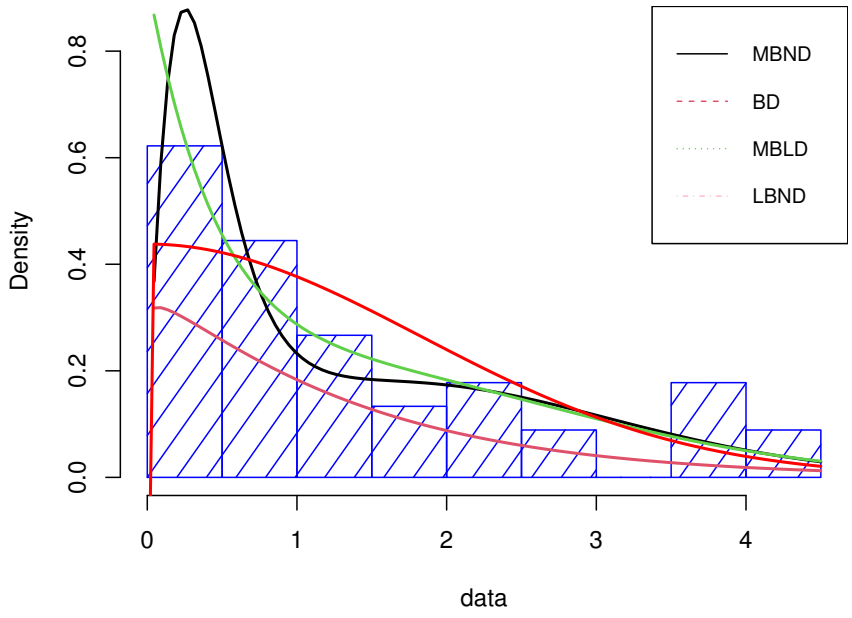


Figure 9: *Plot of the fitted densities for dataset I*

Table 2: *Maximum Likelihood Estimates of the parameters of the different models.*

Model	$\hat{b}$	$\hat{q}$	$\hat{d}$	$\hat{\alpha}$	$\hat{\lambda}$
MBND	1.7543	0.8746	3.4638	1.0266	6.3816
BLD	1.8732	4.1136	19.3219	0.03191	13.995
BD	1.0461	32.1767	28.9722	-	-
LBND	-	-	-	0.4712	1.6594

Table 3: *Goodness of fit measures for the different models.*

Model	$-2\hat{l}$	AIC	AICC	BIC	K-S	p-value
MBND	**110.3650**	**120.3650**	**121.9034**	**129.3983**	**0.0773**	**0.9313**
MBLD	114.4867	124.4867	126.0251	133.5200	0.0835	0.8864
BD	178.6642	184.6642	185.2496	190.0842	0.5233	$5.279e^{-}12$
LBND	155.8333	159.8333	160.1190	163.4466	0.1757	0.1096

Data Set II: Almongy et al. [1] analyzed this dataset to showcase the utility of an innovative extended Rayleigh distribution. This distribution was employed to model COVID-19 mortality rate data from Mexico spanning a period of 108 days, starting from March 4th and concluding on July 20th, 2020. The dataset pertains to a raw mortality rate and is characterized by its relatively unrefined nature. The data are as follows: 8.826, 6.105, 10.383, 7.267, 13.220, 6.015, 10.855, 6.122, 10.685, 10.035, 5.242, 7.630, 14.604, 7.903, 6.327, 9.391, 14.962, 4.730, 3.215, 16.498, 11.665, 9.284, 12.878, 6.656, 3.440, 5.854, 8.813, 10.043, 7.260, 5.985, 4.424, 4.344, 5.143, 9.935, 7.840, 9.550, 6.968, 6.370, 3.537, 3.286, 10.158, 8.108, 6.697, 7.151, 6.560, 2.988, 3.336, 6.814, 8.325, 7.854, 8.551, 3.228, 3.499, 3.751, 7.486, 6.625, 6.140, 4.909, 4.661, 1.867, 2.838, 5.392, 12.042, 8.696, 6.412, 3.395, 1.815, 3.327, 5.406, 6.182, 4.949, 4.089, 3.359, 2.070, 3.298, 5.317, 5.442, 4.557, 4.292, 2.500, 6.535, 4.648, 4.697, 5.459, 4.120, 3.922, 3.219, 1.402, 2.438, 3.257, 3.632, 3.233, 3.027, 2.352, 1.205, 2.077, 3.778, 3.218, 2.926, 2.601, 2.065, 1.041, 1.800, 3.029, 2.058, 2.326, 2.506, 1.923.

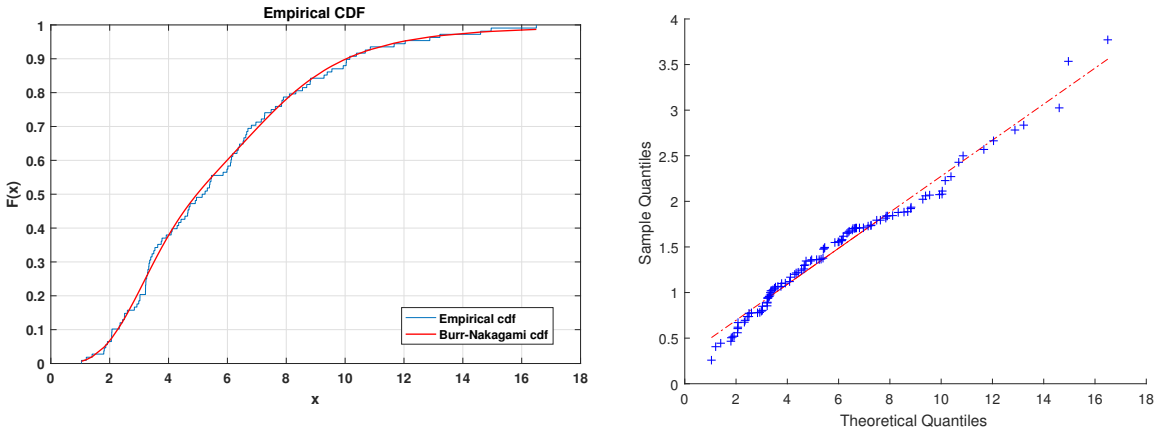


Figure 10: *(a) Ecdf plot for the dataset II (b) q-q plot for the dataset II*

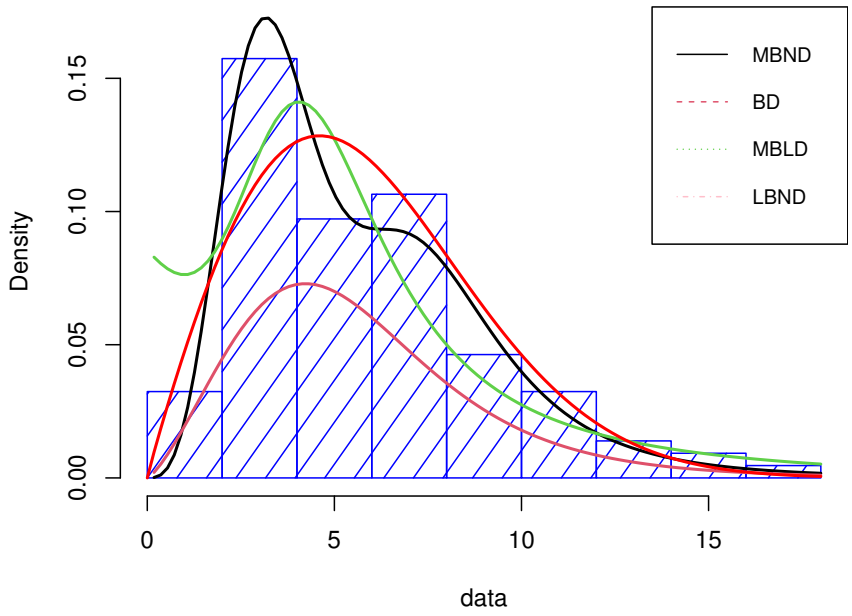


Figure 11: *Plot of the fitted densities for dataset II*

Table 4: *Maximum Likelihood Estimates of the parameters of the different models.*

Model	$\hat{b}$	$\hat{q}$	$\hat{d}$	$\hat{\alpha}$	$\hat{\lambda}$
MBND	5.2318	7.4680	0.9442	2.0159	12.4548
MBLD	3.7382	4.8573	0.9421	0.0090	18.9176
BD	2.3266	8.7808	2.6983	-	-
LBND	-	-	-	0.4719	21.1924

Table 5: *Goodness of fit measures for the different models.*

Model	$-2\hat{l}$	AIC	AICC	BIC	K-S	p-value
MBND	**530.8075**	**540.8075**	**541.3957**	**554.2181**	**0.0486**	**0.9602**
MBLD	558.2015	568.2015	568.7097	581.6121	0.1151	0.1144
BD	685.1428	691.1428	691.3735	699.1892	0.5054	$2.20e^{-}16$
LBND	538.4605	542.4605	542.5748	547.8247	0.3049	$3.676e^{-}09$

Table 2 and Table 3 provide the Maximum Likelihood Estimates (MLEs) and offer a comparison of the performance between the Mixture of Burr XII and Nakagami Distribution (MBND) and the alternative distributions for the first dataset. Similarly, Table 4 and Table 5 present the MLEs and compare the performance of MBND with other distributions for the second dataset. The findings displayed in Table 3 and 5 demonstrate that MBND achieves the lowest values for AIC, AICc, and BIC, indicating its superiority over several well-established competing models. The claim is

further supported by Figures 8 and 10. Additionally, the results presented in Table 3 and 5 are corroborated by examining the Empirical Cumulative Distribution Function (ECDF) plots and Quantile-Quantile (Q-Q) plots for the MBND model in the context of both datasets, showcased in Figure 7 and 9. These visual representations lend support to the outcomes highlighted in Table 3 and 5.

9. Conclusion

In this paper, we presented a mixture of the Burr XII and Nakagami distribution (MBND). We investigate its statistical properties and demonstrate that this new mixture distribution can serve as an alternative model to some existing distributions. For MLEs, simulation schemes are developed that yield less Mean Squared Error (MSE) as sample size increases. Additionally, Bayesian estimates of the mixing parameter are computed, demonstrating their superiority over classical estimates. We utilize real-world datasets to illustrate the efficacy of the MBND model and compare it with other competing distributions. The results of this comparison reveal that the MBND model offers a better fit than the considered distributions. Through this comprehensive analysis, we establish the viability of MBND as a valuable tool for modeling and analyzing various data scenarios.

References

[1] Almetwally E. M., Aljohani H.M., Alghamdi A.S, Almongy, H. M. and Hafez, E.H. (2021). A new extended rayleigh distribution with applications of covid-19 data. *Results in Physics*, 23, 104012.

[2] Almetwally E. M., Haj Ahmad, H., Almongy, H. M. and Al-nefaie, A. H. (2022). Modeling of covid 19 vaccination rate using odd lomax inverted nadarajah- haghighi distribution. *Plos one*, 17(10), e0276181.

[3] Berferroni, C. E. Elmenti di statitica generale. Libreria Seber, Firenze., 1930.

[4] Chen, X., Ponomareva, M. and Tamer, E. Likelihood inference in some finite mixture models. *Journal of Econometrics*, 182:87–99.

[5] Daghestani, A. F., Sultan, K. S. and Al-Moisheer, A. S. (2021). Mixture of lindley and weibull distributions: Properties and estimation. *Journal of Statistics Applications and Probability*, 10(2):301—313.

[6] Daniyal, M. and Aleem, M. (2014). On the mixture of burr xii and weibull distributions. journal of statistics applications and probability, 3:251-–267.

[7] Decay, M. F. (1964). Modified poisson probability law for a point pattern more than random. (1964). *Association of American Geographers Annals*, 54:559—565.

[8] Hasselblad, V. (1968). Estimation of mixtures of distribution from exponential family. *Journal of American Statistical Association*, 64:300—304.

[9] Havrda, J. and Charvat, F. (1967). Quantification method in classification processes: concept of structural α-entropy. *Kybernetika*, 3:30-–35.

[10] Jaspers, S., Komrek, A. and Aerts, M. (2018). Bayesian estimationof multivariate normal mixtures with covariate-dependent mixing weights, with an application in antimicrobial resistance monitoring. Biometrical JournaL, 60:7—19.

[11] Martin, M., Hory, C. and Huchard, M. Chi-square law mixture for an unsupervised time frequency pattern extraction. Proceedings of the Tenth International Congress on Sound and Vibration, Stockholm, Sweden, 7-10 July, Code 62825, 2003.

[12] Nakagami, M. (1960). The m-distribution-a general formulation of intensity distribution of rapid fading. in W. C. Hoffman (ed.) *Statistical Methods in Radio Wave Propagations*, 3-–36.

[13] Nasiri, P. and Azarian, A. A. (2018). Estimating the mixing proportion of mixture of two chi square distributions. *International Journal of Applied Mathematics and Information Sciences*, 12(4):753-–759.

[14] Pearson, k. (1894). Contribution to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London*, 185:71–110, 1894.
[15] Renyi, A. On the measure of entropy and information. Proceedings of the fourth Berkely symposium on mathematical statistics and probability, 1:547–561, 1961.
[16] Singh, B. P., Singh, S. and Das, U. D. (2021). A general class of new continuous mixture distribution and applications. *J. Math. Comput. Sci.*, 11(1):585—602.
[17] Titterington, D.M., Smith, A.F.M. and Makov, U.E. Statistical Analysis of Finite Mixture Distributions. John Wiley and Sons, New York, 1985.
[18] Tsallis, C. (1968). Possible generalization of boltzmann–gibbs statistics. *Journal of Statistical Physics*, 52:479-–487.
[19] Zaman, M. R., Paul, D. N. R., Akhter, N., Howlader, M. H. and Kabir, M. S. (2006). Chi-square mixture of chi-square distributions. *Journal of Applied Sciences*, 6:243-–246.

STARTING MODE OF SYNCHRONOUS MACHINES WITH MASSIVE ROTORS

Laman Hasanova[1,2], Nurali Yusifbayli[1]

[1]Azerbaijan Technical University,
[2]Azerbaijan Scientific-Research and Designed-Prospecting Institute of Energetics
gasanovalg@mail.ru, yusifbayli.n@gmail.com

Abstract

It is known that in synchronous machines with massive rotors, it is required to take into account the change in the equivalent rotor of active resistance depending on the frequency of the current in the rotor at starting of these machines. A three-phase mathematical model of these machines has been compiled, the equations of which are written in axes rotating at the speed of the machine rotor. Study of the start-up modes of these machines and operation in synchronous mode with a load surge (dynamic mode) has been carried out on this model. The studies have allowed for making the following conclusions. When starting synchronous machines with massive rotors, it is most preferable to take into account the change in the equivalent resistance in the form of a linear sliding function. It was found out that in the synchronous mode of operation of these machines, including in stable dynamic modes, there is no need to take into account changes in the rotor resistance-sliding, since the sliding in these modes oscillatory damps around zero, thereby not affecting changes in the value of the equivalent resistance. Therefore, in these modes it is recommended to consider the value of this resistance constant and definite at slip equal to zero.

Keywords: synchronous machines with massive rotor, start mode, dynamic mode, three-coordinate system

I. Introduction

Synchronous machines with massive rotors are widely used both as generators (turbo generators) and as motors [1,2,3]. For the studying of static and, especially, dynamic modes of their operation, the well-known Park equations are widely used [4,5].

However, the presence of a massive rotor in them requires to take into account the change in the active resistance of the rotor depending on the frequency of the current in the rotor. This question was most completely solved analytically in [6]. Although there we are talking about synchronous motors with massive rotor, but this can, of course, be extended to low-power synchronous generators, since their asynchronous start is identical.

II. Methods

In [6], it is proposed to change the active resistance of a massive rotor according to the following expression:

$$r_r = r_{r_{s=0}} + (r_{r_{s=1}} - r_{r_{s=0}})\sqrt{s} \quad (1)$$

where $r_{rs=0}$ – active resistance of the rotor at sliding s=0 and short-circuited excitation winding; $r_{rs=1}$ – active resistance of the rotor at s=1, also at short-circuited excitation winding. Further, in [6], as in other works [7], the static characteristic of the asynchronous moment is determined, i.e. on the basis of this characteristic, the start of the electric machine is defined.

For a more accurate idea of the starting modes of synchronous machines with massive rotor, it is necessary to take into account the effect of transient processes during start-up.

On this basis, the purpose of this paper is the issue of mathematical modeling of synchronous machines with massive rotor and the study of the dynamic modes of their operation during direct asynchronous start.

As the basis of the mathematical model, it is proposed to take the equations of synchronous machines written in a three-phase coordinate system, in the axes α_r, β_r, γ_r, rotating with the speed of the rotor ω_r of the synchronous machine. It should be noted that it is relatively easy to obtain these equations from the equation of a three-phase model of a double-fed machine, also written in the axes α_r, β_r, γ_r, rotating at the speed of the machine rotor [8,9] and this can be done since a synchronous machine with a massive rotor is magnetically symmetrical electric machine, the rotor of which is non-salient pole, i.e. the air gap remains practically unchanged along the entire perimeter of the machine rotor.

On this basis, equations of synchronous machines with massive rotor in a three-coordinate (three-phase) system in extensive form will appear as follows:

$$\left.\begin{aligned}
p\Psi_{s\alpha} &= U_{s\alpha}\cdot\sin\theta+\frac{1}{\sqrt{3}}\omega_r(\Psi_{s\beta}-\Psi_{s\gamma})-r_s\cdot i_{s\alpha}\\
p\Psi_{s\beta} &= U_{s\beta}\cdot\sin\left(\theta-\frac{2\pi}{3}\right)+\frac{1}{\sqrt{3}}\omega_r(\Psi_{s\gamma}-\Psi_{s\alpha})-r_s\cdot i_{s\beta}\\
p\Psi_{s\gamma} &= U_{s\gamma}\cdot\sin\left(\theta+\frac{2\pi}{3}\right)+\frac{1}{\sqrt{3}}\omega_r(\Psi_{s\alpha}-\Psi_{s\beta})-r_s\cdot i_{s\gamma}\\
p\Psi_{f} &= U_{f}-r_f\cdot i_f\\
p\Psi_{r\alpha} &= 0-r_r\cdot i_{r\alpha}\\
p\Psi_{r\beta} &= 0-r_r\cdot i_{r\beta}\\
p\Psi_{r\gamma} &= 0-r_r\cdot i_{r\gamma}
\end{aligned}\right\} \qquad (2)$$

The system of equations (2) describes the balance of voltages in the stator and rotor circuits of the machine. At that, it is meant that the stator has three symmetrical windings shifted in space by 120 electric degrees, and four windings are placed on the rotor – one excitation winding and three damper (starting) windings equivalent to the massive rotor of the machine, moreover, we agree to assume that the axis of excitation winding coincides with the axis of the damper winding located along the α axis.

Then the equations of connection between flux linkages and currents will appear in the form:

$$\left.\begin{aligned}
\Psi_{s\alpha} &= x_s\cdot i_{s\alpha}-0.5\cdot x_m\cdot i_{s\beta}-0.5\cdot x_m i_{s\gamma}+x_m\cdot i_f+x_m\cdot i_{r\alpha}-0.5\cdot x_m\cdot i_{r\beta}-0.5\cdot x_m\cdot i_{r\gamma}\\
\Psi_{s\beta} &= -0.5\cdot x_m\cdot i_{s\alpha}+x_s\cdot i_{s\beta}-0.5\cdot x_m\cdot i_{s\gamma}-0.5\cdot x_m\cdot i_f-0.5\cdot x_m\cdot i_{r\alpha}+x_m\cdot i_{r\beta}-0.5\cdot x_m\cdot i_{r\gamma}\\
\Psi_{s\gamma} &= -0.5\cdot x_m\cdot i_{s\alpha}-0.5\cdot x_m\cdot i_{s\beta}+x_s\cdot i_{s\gamma}-0.5\cdot x_m\cdot i_f-0.5\cdot x_m\cdot i_{r\alpha}-0.5\cdot x_m\cdot i_{r\beta}+x_m\cdot i_{r\gamma}\\
\Psi_{f} &= x_m\cdot i_{s\alpha}-0.5\cdot x_m\cdot i_{s\beta}-0.5\cdot x_m\cdot i_{s\gamma}+x_f\cdot i_f+x_m\cdot i_{r\alpha}-0.5\cdot x_m\cdot i_{r\beta}-0.5\cdot x_m\cdot i_{r\gamma}\\
\Psi_{r\alpha} &= x_m\cdot i_{s\alpha}-0.5\cdot x_m\cdot i_{s\beta}-0.5\cdot x_m\cdot i_{s\gamma}+x_m\cdot i_f+x_r\cdot i_{r\alpha}-0.5\cdot x_m\cdot i_{r\beta}-0.5\cdot x_m\cdot i_{r\gamma}\\
\Psi_{r\beta} &= -0.5\cdot x_m\cdot i_{s\alpha}+x_m\cdot i_{s\beta}-0.5\cdot x_m\cdot i_{s\gamma}-0.5\cdot x_m\cdot i_f-0.5\cdot x_m\cdot i_{r\alpha}+x_r\cdot i_{r\beta}-0.5\cdot x_m\cdot i_{r\gamma}\\
\Psi_{r\gamma} &= -0.5\cdot x_m\cdot i_{s\alpha}-0.5\cdot x_m\cdot i_{s\beta}+x_m\cdot i_{s\gamma}-0.5\cdot x_m\cdot i_f-0.5\cdot x_m\cdot i_{r\alpha}-0.5\cdot x_m\cdot i_{r\beta}+x_r\cdot i_{r\gamma}
\end{aligned}\right\} \qquad (3)$$

In equations (2), (3): $\Psi_{s\alpha}$, $\Psi_{s\beta}$, $\Psi_{s\gamma}$, $i_{s\alpha}$, $i_{s\beta}$, $i_{s\gamma}$ are flux linkages and currents of stator windings, Ψ_f, $\Psi_{r\alpha}$, $\Psi_{r\beta}$, $\Psi_{r\gamma}$, i_f, $i_{r\alpha}$, $i_{r\beta}$, $i_{r\gamma}$ are flux linkages and currents of rotor circuits, and Ψ_f and $\Psi_{r\alpha}$ are located along the rotor axis α_r; r_r is active resistance equivalent to the rotor circuits; x_s, x_r, x_f are inductive resistances (in relative units they are equal to inductances) of the stator windings, damper and rotor

excitation windings, x_m is mutual induction resistance between the stator and rotor circuits (i.e. in relative units mutual inductance).

The relationship between currents and flux linkages is determined using the inverse matrix according to the relations:

$$\begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \\ i_{s\gamma} \\ i_f \\ i_{r\alpha} \\ i_{r\beta} \\ i_{r\gamma} \end{bmatrix} = \begin{bmatrix} x_s - 0.5\cdot x_m - 0.5\cdot x_m + x_m + x_m - 0.5\cdot x_m - 0.5\cdot x_m \\ -0.5\cdot x_m + x_s - 0.5\cdot x_m - 0.5\cdot x_m - 0.5\cdot x_m + x_m - 0.5\cdot x_m \\ -0.5\cdot x_m - 0.5\cdot x_m + x_s - 0.5\cdot x_m - 0.5\cdot x_m + 0.5\cdot x_m + x_m \\ x_m - 0.5\cdot x_m - 0.5\cdot x_m + x_f + x_m - 0.5\cdot x_m - 0.5\cdot x_m \\ x_m - 0.5\cdot x_m - 0.5\cdot x_m + x_m + x_r - 0.5\cdot x_m - 0.5\cdot x_m \\ -0.5\cdot x_m + x_m - 0.5\cdot x_m - 0.5\cdot x_m - 0.5\cdot x_m + x_r - 0.5\cdot x_m \\ -0.5\cdot x_m - 0.5\cdot x_m + x_m - 0.5\cdot x_m - 0.5\cdot x_m - 0.5\cdot x_m + x_r \end{bmatrix}^{-1} \begin{bmatrix} \Psi_{s\alpha} \\ \Psi_{s\beta} \\ \Psi_{s\gamma} \\ \Psi_f \\ \Psi_{r\alpha} \\ \Psi_{r\beta} \\ \Psi_{s\gamma} \end{bmatrix} \quad (4)$$

Equations (2), (3) and (4) are supplemented with the equations of electromagnetic torque [10], motion and sliding:

$$\left.\begin{aligned} m_{em} &= p_m \frac{\sqrt{3}}{2} x_m \left[\left(i_{s\alpha}\cdot i_{r\gamma} + i_{s\beta}\cdot i_{r\alpha} + i_{s\beta}\cdot i_f + i_{s\gamma}\cdot i_{r\beta}\right) - \left(i_{s\alpha}\cdot i_{r\beta} + i_{s\beta}\cdot i_{r\gamma} + i_{s\gamma}\cdot i_{r\alpha} + i_{s\gamma}\cdot i_f\right)\right] \\ p\theta &= 1 - \omega_r = s \\ p\omega_r &= \frac{p_m}{J}\left(m_{em} - m_T\right) \end{aligned}\right\} \quad (5)$$

here – p_m – number of pole pairs; $s=p\theta$ – sliding; J – moment of inertia of moving parts.

Equations (4) in extensive form will appear in the form:

$$\left.\begin{aligned} i_{s\alpha} &= k_{s\alpha 1}\cdot\psi_{s\alpha} + k_{s\alpha 2}\cdot\psi_{s\beta} + k_{s\alpha 3}\cdot\psi_{s\gamma} + k_{s\alpha 4}\cdot\psi_f + k_{s\alpha 5}\cdot\psi_{r\alpha} + k_{s\alpha 6}\cdot\psi_{r\beta} + k_{s\alpha 7}\cdot\psi_{r\gamma} \\ i_{s\beta} &= k_{s\beta 1}\cdot\psi_{s\alpha} + k_{s\beta 2}\cdot\psi_{s\beta} + k_{s\beta 3}\cdot\psi_{s\gamma} + k_{s\beta 4}\cdot\psi_f + k_{s\beta 5}\cdot\psi_{r\alpha} + k_{s\beta 6}\cdot\psi_{r\beta} + k_{s\beta 7}\cdot\psi_{r\gamma} \\ i_{s\gamma} &= k_{s\gamma 1}\cdot\psi_{s\alpha} + k_{s\gamma 2}\cdot\psi_{s\beta} + k_{s\gamma 3}\cdot\psi_{s\gamma} + k_{s\gamma 4}\cdot\psi_f + k_{s\gamma 5}\cdot\psi_{r\alpha} + k_{s\gamma 6}\cdot\psi_{r\beta} + k_{s\gamma 7}\cdot\psi_{r\gamma} \\ i_f &= k_{f1}\cdot\psi_{s\alpha} + k_{f2}\cdot\psi_{s\beta} + k_{f3}\cdot\psi_{s\gamma} + k_{f4}\cdot\psi_f + k_{f5}\cdot\psi_{r\alpha} + k_{f6}\cdot\psi_{r\beta} + k_{f7}\cdot\psi_{r\gamma} \\ i_{r\alpha} &= k_{r\alpha 1}\cdot\psi_{s\alpha} + k_{r\alpha 2}\cdot\psi_{s\beta} + k_{r\alpha 3}\cdot\psi_{s\gamma} + k_{r\alpha 4}\cdot\psi_f + k_{r\alpha 5}\cdot\psi_{r\alpha} + k_{r\alpha 6}\cdot\psi_{r\beta} + k_{r\alpha 7}\cdot\psi_{r\gamma} \\ i_{r\beta} &= k_{r\beta 1}\cdot\psi_{s\alpha} + k_{r\beta 2}\cdot\psi_{s\beta} + k_{r\beta 3}\cdot\psi_{s\gamma} + k_{r\beta 4}\cdot\psi_f + k_{r\beta 5}\cdot\psi_{r\alpha} + k_{r\beta 6}\cdot\psi_{r\beta} + k_{r\beta 7}\cdot\psi_{r\gamma} \\ i_{r\gamma} &= k_{r\gamma 1}\cdot\psi_{s\alpha} + k_{r\gamma 2}\cdot\psi_{s\beta} + k_{r\gamma 3}\cdot\psi_{s\gamma} + k_{r\gamma 4}\cdot\psi_f + k_{r\gamma 5}\cdot\psi_{r\alpha} + k_{r\gamma 6}\cdot\psi_{r\beta} + k_{r\gamma 7}\cdot\psi_{r\gamma} \end{aligned}\right\} \quad (6)$$

The coefficient $k_{s\alpha 1} \div k_{r\gamma 2}$ is determined from the inverse matrix of machine parameters (4).

Thus, comparing these equations with the equations of double-fed machine given in [8,9,10], it can be noted that their structure remains practically unchanged, but the action of the excitation winding, which is proposed to be placed on the same axis as the damper (starting) winding located on the axis, is additionally taken into account. Naturally, it is also taken into account in the equations of flux linkages of rotor and electromagnetic torque.

Once again it is necessary to emphasize, as it was noted in [6], it is necessary to use the three-phase model only in extreme cases, when it is necessary to bring additional clarity to issues that are impossible (or difficult to implement) to study in two-coordinate models. For example, such as the study of asymmetric modes in rotary circuits, etc.

III. Results

In this case, the proposed model is used for studying the starting modes and modes of involvement into synchronism of synchronous machines with massive rotor – this is the formation of asynchronous starting torques with the equivalent massive rotors of these machines, as well as load surge in synchronous mode.

On Fig. 1 (*a, b, c*) are shown, respectively, the fluctograms of the change in the rotational speed ω_r of the rotor, the electromagnetic torque m_{em} and the excitation current i_f of the model synchronous

generator, the parameters of which are given in Appendix 1. In Fig. 1 (*d*) the extensive fluctogram of the starting torque m_{em} is shown. These fluctograms were obtained at rotor resistance equivalent to massive rotor according to the expression [6]:

$$r_r = r_{r_{S=0}} + (r_{r_{S=1}} - r_{r_{S=0}}) \cdot \sqrt{s} \tag{7}$$

For the generator under study $r_{rs=0}$=0.01 – the active resistance of the rotor at s=0; $r_{rs=1}$=0.05 – active resistance of the rotor at s=1; s=(1–ω_r) – sliding, i.e.

$$r_r = 0.01 + 0.04 \cdot \sqrt{1-\omega_r} \ .$$

On fluctograms after asynchronous start at $2 \cdot 10^3$ rad. the voltage is applied to the excitation winding, the machine is involved into synchronism, and at 3000 rad. the moment of resistance equal to the rated moment of the machine is applied abruptly (m_n=1.596).

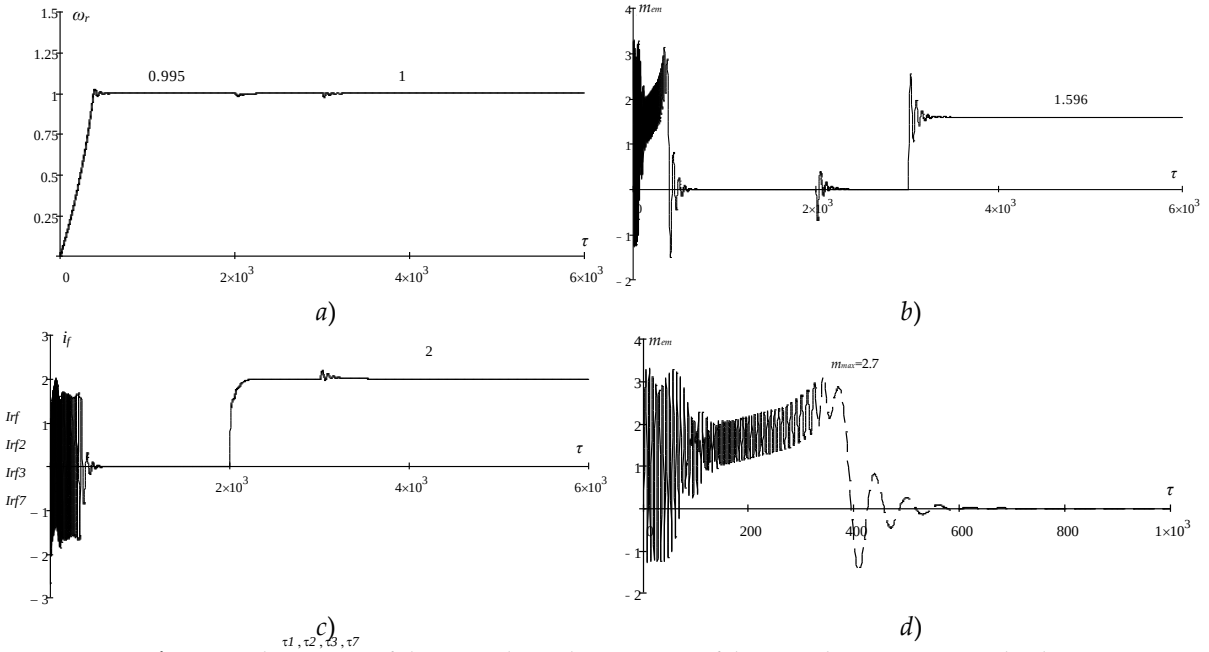
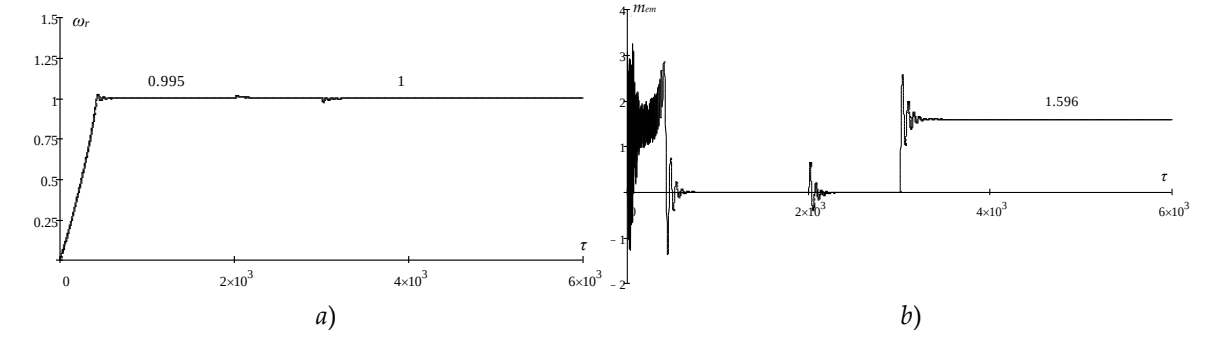


Figure 1: *Fluctograms of change in the mode parameters of the equivalent massive rotor by the expression*

$$r_r = 0.01 + 0.04 \cdot \sqrt{1-\omega_r}$$

On Fig. 2 (*a, b, c, d*) the fluctograms of these mode parameters are presented in the same sequence when the massive rotor is equivalent to active resistance, which varies depending on the sliding according to the linear law:

$$r_r = 0.01 + 0.04 \cdot (1-\omega_r) \tag{8}$$



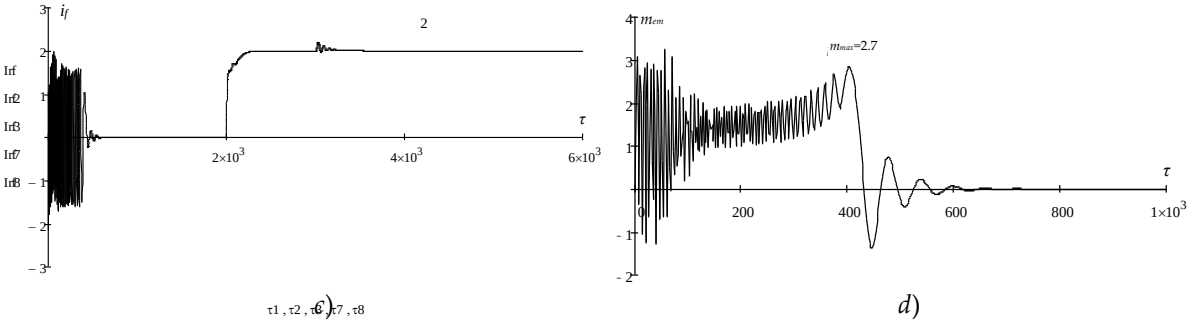


Figure 2: *Fluctograms of change in mode parameters when massive rotor is equivalent to active resistance, which varies depending on sliding according to linear law*

And finally, in Fig. 3 (*a, b, c, d*) and Fig. 4 (*a, b, c, d*) the fluctograms are shown at constant values of r_r, equal to $r_{rs=1}$=0.05 and $r_{rs=0}$=0.01 respectively.

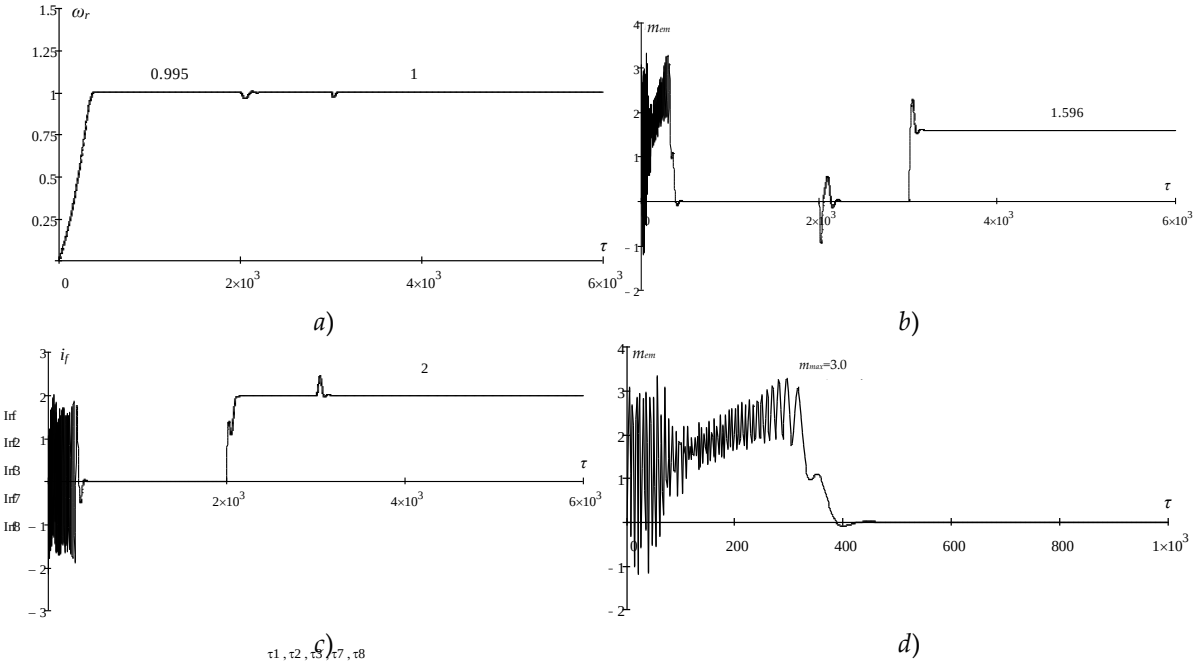


Figure 3: *Fluctograms of change in mode parameters, when the massive rotor is equivalent to active resistance equal to* $r_{rs=1}$=0.05

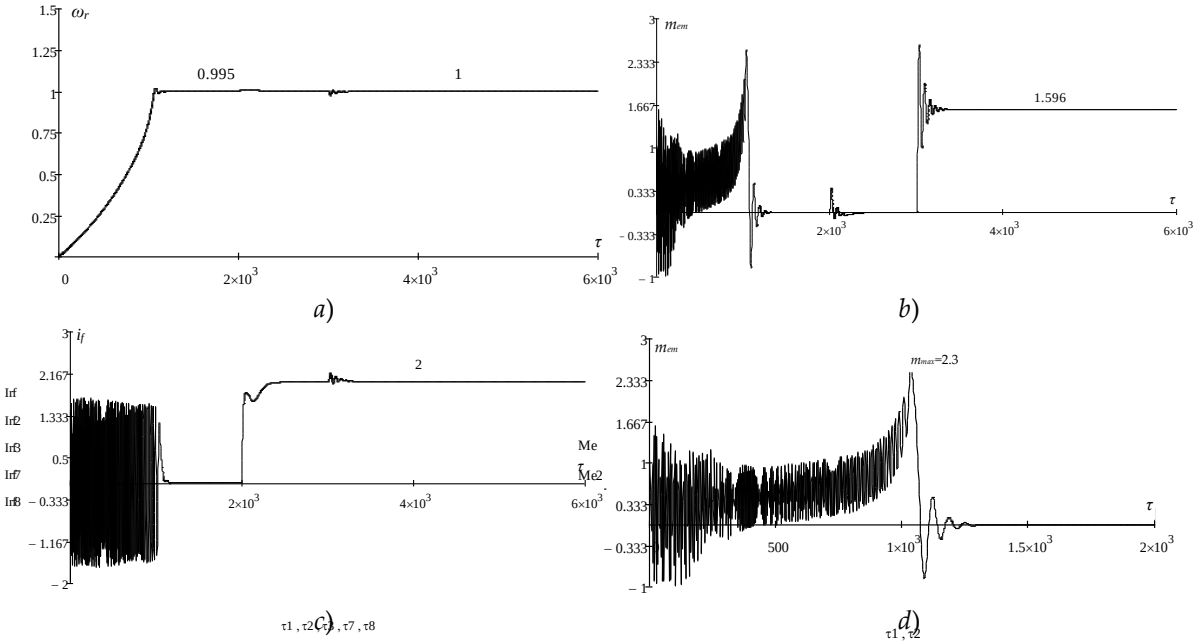


Figure 4: *Fluctograms of change in mode parameters, when the massive rotor is equivalent to active resistance equal to* $r_{rs=0}$=0.01

Before proceeding to the analysis of fluctograms, it is necessary to note the following. The implementation of expression (7) on the computer will introduce a significant error, since due to the presence of even a small sliding after the completion of the asynchronous start (in the example given, $s_{st\text{-}st}$=0.005, the radicand from it will be equal to $\sqrt{s} = 0.071$, i.e. expression (7) after the start will be equal not to $r_{rs=0}$=0.01, but to $r_{rs=0}$=0.0128 (i.e. the error will be 28%). When using expression (8), there will also be an error, but it will not exceed 2%, which is quite acceptable. On this basis, it is proposed to simulate expression (7) by a piecewise linear approximation, i.e. in the range of change of ω_r from 0 to 0.8 – to represent r_r with linear dependence of the form r_{r1}=0.05 – 0.0275·ω_r, and in the range of change of ω_r from 0.8 to 1 – r_{r1} = 0.1 – 0.09·ω_r. At that, the maximum error will not exceed 12%. Thus, the fluctograms in Fig. 1 (*a, b, c, d*) are obtained just on the basis of the piecewise linear approximation of expression (7).

Comparative analysis of fluctograms shows the following. The start-up process is a dynamic process, i.e. the acceleration time cannot be determined only by the static characteristic of the asynchronous moment of the synchronous machine, since during the start-up the kinetic energy of the rotating masses of the unit together with the rotor of the synchronous machine takes an active part. However, this characteristic is necessary, since with its help two important parameters of the starting characteristic are specified – the equivalent active resistances of the rotor during s=0 and s=1 slidings.

The extensive fluctograms of the electromagnetic moment m_{em}, presented in Fig. 1 and Fig. 2, according to the average values of starting moment m_{s1} and maximum moment m_{max} practically coincide m_{s1}=m_{s2}≈1.1 (~0.7·m_n)·m_{max1}=m_{max2}=2.7(~1.7·m_n), the start time at change of r_r according to expressions (7) and (8) is also the same and equal to τ_p~600 rad. On the fluctograms in Fig. 3 and 4 at constant values r_r=$r_{rs=1}$=0.05 and r_r=$r_{rs=0}$=0.01, the average values of the starting torque are respectively equal to m_{s3}≈1.25 and m_{s4}≈0.4, and the maximum moments m_{max3}=3 and m_{max4}=2.3, the start time, respectively, is equal to τ_{s3}=400 rad. and τ_{s4}=1250 rad. That is, the options of Fig. 3 and 4 cannot reflect the real dynamics of the starting mode of a synchronous machine.

Thus, a comparative analysis of the fluctograms shows that since the options in Fig. 1 and 2 in terms of the dynamic starting torque practically coincide with each other, then, based on the simplicity of implementation, it is necessary to stop at the 2nd option, when the resistance r_r changes linearly as a function of sliding (8).

Analysis of the synchronous mode of fluctograms (we remind that at the 2000th rad. the excitation winding opens and connects to the power source – the machine enters synchronism ω_r=1, and at the 3000th rad. the moment of resistance equal to the rated moment of the machine m_r=1.596 is supplied abruptly) shows, that during the rated load surge on the synchronous motor in Fig. 1, 2, and 4, the process of steadiness of the moment after the transient mode is almost the same on all the above fluctograms (i.e. the maximum radius and the number of oscillations are the same). This, of course, also confirms the adequacy of the model, since in this mode the synchronous machine "operates" with a steady value of the equivalent rotor active resistance equal to r_r=$r_{rs=0}$ (for the simulative motor under study $r_{rs=0}$=0.01).

In Fig. 3, the load surge process is almost aperiodic. This mode does not exist in reality, since the equivalent rotor resistance r_r is assumed to be equal to r_r=$r_{rs=1}$ (which leads to incorrect results in the calculations both in the asynchronous mode and in the synchronous mode of operation of the synchronous machine).

IV. Discussion

I. Subsection One

A three-coordinate (three-phase) digital model of synchronous machines with massive rotors is

proposed, which makes it possible to study in one structure asynchronous and synchronous modes of operation of these machines.

II. Subsection Two

In the starting mode and synchronous mode of operation, it is possible to represent practically without error the formula for the equivalent active resistance of the rotor, which varies as a function of sliding by a linear expression (8).

III. Subsection Three

When studying only synchronous modes of operation of these machines, it is sufficient to equivalent the active resistance of the massive rotor with a constant value equal to $r_r=r_{r_s=0}$, i.e. oscillations in sliding in this mode does not affect the average value of this resistance

References

[1] Ewald Fuchs and Mohammad Masoum (2015). Power Quality in Power Systems and Electrical Machines. eBook ISBN: 9780128009888: Hardback ISBN: 9780128007822.

[2] V. del Toro (1985). Electric Machines and Power Systems. Prentice Hall, 1985.

[3] D. Zorbas (1989). Electric Machines-Principles, Applications, and Control Schematics. West, 1989.

[4] Jerom Cros, Stephanie Rakotovolona, Maxim Bergeron, Philippe Viarouge and etc. (2016). Simulation Methods for the Transient Analysis of Synchronous Alternators. In book: Renewable Energy – Utilisation and System Integration DOI: 10.5772/61604 Publisher: INTECH May 2016.

[5] Roberto Felicetti (2021). Voltage Transients in the Field Winding of Salient Pole Wound Synchronous Machines: Implications from fast switching power electronics. 91 pp. Uppsala.

[6] Andre, E. Blondel (1913) Synchronous motors and converters. MCG RAW–HILL Book Company 239 West 39th Street, New York 6 Bouverie Street, London. E. C.

[7] Zhang, L. Z, Chen L. F (1991) Analysis of subsynchronous resonance in a turbogenerator and its shstem using compensation capacitors. Proc. CICEM, Huazhong University of Science and Technology Press, Wuhan, pp 104–108.

[8] Franklin, P. W (1972) Theory of the 3-phase salient pole type generator with bridge rectifier output. IEEE Trans. PAS-91(5): Part I, 1960–1967, Part II, 1968–1975

[9] Vazquez, S. A model-based direct power control for three-phase power converters /S.Vazquez, J. A.Sanchez, J. M.Carrasco [et al.] // IEEE Trans. Ind. Electron., – 2008, v.55, No. 4, – p. 1647-1657.

[10] Kopylov, I. P. (2012). Mathematical Models of Electric Machines. Moscow "High School", 2012.

Appendices

Appendix 1. Model generator parameters

$x_s = x_d = x_q = 2.78(\text{r.u.})$ $U_{\text{baz}} = 310$ B $r_f = 0.03(\text{r.u.})$

$x_m = 2.69(\text{r.u.})$ $I_{\text{baz}} = 30.44$ A $r_s = 0.045\,(\text{r.u.})$

$x_r = x_{dr} = x_{qr} = 2.82(\text{r.u.})$ $R_{\text{baz}} = 10.2$ Ohm $r_{r_{s=0}} = 0.01\,(\text{r.u.})$

$x_f = 3(\text{o.e})$ $\Psi_{\text{baz}} = 0.987\text{Bc}$ $r_{r_{s=1}} = 0.05\,(\text{r.u.})$

$J_{\text{об}} = 0.3 kgm^2$ $p_{\text{baz}} = 14.15$ kW $r_r = 0.01 + 0.04 \cdot \sqrt{1-\omega_r}$ (r.u.)

$J_{\text{baz}} = \dfrac{\text{M}_{\text{baz}}}{\omega_{\text{baz}}^2} = 0.00046$ $\omega_{\text{baz}} = 314$ $r_{r2} = 0.01 + (0.05 - 0.01)\cdot(1-\omega)$ (r.u.)

$J^* = \dfrac{J}{J_{baz}} = \dfrac{0.3}{0.0006} \approx 656.5$ $M_{\text{baz}} = 45.1$ Hm $r_{r3} = r_{rs=1} = 0.05(\text{r.u.})$

$U_{fn} = 0.06(\text{r.u.})$ $p_m = 1$ $r_{r4} = r_{rs=0} = 0.01(\text{r.u.})$

Appendix 2. Equations of three-phase model of simulative synchronous generator

$$p\psi_{s\alpha} = 1\cdot\sin\theta + 0.577\cdot\omega_r(\Psi_{s\beta} - \Psi_{s\gamma}) - 0.045\cdot i_{s\alpha}$$

$$p\psi_{s\beta} = 1\cdot\sin(\theta - 2.09) + 0.577\cdot\omega_r(\Psi_{s\gamma} - \Psi_{s\alpha}) - 0.045\cdot i_{s\beta}$$

$$p\psi_{s\gamma} = 1\cdot\sin(\theta + 2.09) + 0.577\cdot\omega_r(\Psi_{s\alpha} - \Psi_{s\beta}) - 0.045\cdot i_{s\gamma}$$

$$p\psi_f = U_f - 0.03\cdot i_f$$

$$p\psi_{r\alpha} = 0 - r_{rn}\cdot i_{r\alpha}$$

$$p\psi_{r\beta} = 0 - r_{rn}\cdot i_{r\beta}$$

$$p\psi_{r\gamma} = 0 - r_{rn}\cdot i_{r\gamma}$$

$$p\omega_r = \frac{1}{656.46}(m_{em} - m_T) = 0.0015(m_{em} - m_T)$$

$$p\theta = s = (1-\omega_r)$$

$$i_{s\alpha} = 7.689\cdot\Psi_{s\alpha} + 2.108\cdot\Psi_{s\beta} + 2.108\cdot\Psi_{s\gamma} - 1.165\cdot\Psi_f - 2.856\cdot\Psi_{r\alpha} + 1.428\cdot\Psi_{r\beta} + 1.428\cdot\Psi_{r\gamma}$$

$$i_{s\beta} = 2.108\cdot\Psi_{s\alpha} + 7.346\cdot\Psi_{s\beta} + 2.451\cdot\Psi_{s\gamma} + 0.582\cdot\Psi_f + 1.428\cdot\Psi_{r\alpha} - 3.088\cdot\Psi_{r\beta} + 1.66\cdot\Psi_{r\gamma}$$

$$i_{s\gamma} = 2.108\cdot\Psi_{s\alpha} + 2.451\cdot\Psi_{s\beta} + 7.346\cdot\Psi_{s\gamma} + 0.582\cdot\Psi_f + 1.428\cdot\Psi_{r\alpha} + 1.66\cdot\Psi_{r\beta} - 3.088\cdot\Psi_{r\gamma}$$

$$i_f = -1.165\cdot\Psi_{s\alpha} + 0.582\cdot\Psi_{s\beta} + 0.582\cdot\Psi_{s\gamma} + 2.968\cdot\Psi_f - 0.789\cdot\Psi_{r\alpha} + 0.395\cdot\Psi_{r\beta} + 0.395\cdot\Psi_{r\gamma}$$

$$i_{r\alpha} = -2.856\cdot\Psi_{s\alpha} + 1.428\cdot\Psi_{s\beta} + 1.428\cdot\Psi_{s\gamma} - 0.789\cdot\Psi_f + 6.13\cdot\Psi_{r\alpha} + 0.967\cdot\Psi_{r\beta} + 0.967\cdot\Psi_{r\gamma}$$

$$i_{r\beta} = 1.428\cdot\Psi_{s\alpha} - 3.088\cdot\Psi_{s\beta} + 1.66\cdot\Psi_{s\gamma} + 0.395\cdot\Psi_f + 0.967\cdot\Psi_{r\alpha} + 5.972\cdot\Psi_{r\beta} + 1.125\cdot\Psi_{r\gamma}$$

$$i_{r\gamma} = 1.428\cdot\Psi_{s\alpha} + 1.66\cdot\Psi_{s\beta} - 3.088\cdot\Psi_{s\gamma} + 0.395\cdot\Psi_f + 0.967\cdot\Psi_{r\alpha} + 1.125\cdot\Psi_{r\beta} + 5.972\cdot\Psi_{r\gamma}$$

$$m_{em} = 2.3\left[(i_{s\alpha}\cdot i_{r\gamma} + i_{s\beta}\cdot i_{r\alpha} + i_{s\beta}\cdot i_f + i_{s\gamma}\cdot i_{r\beta}) - (i_{s\alpha}\cdot i_{r\beta} + i_{s\beta}\cdot i_{r\gamma} + i_{s\gamma}\cdot i_{r\alpha} + i_{s\gamma}\cdot i_f)\right]$$

where r_{rn} – equivalent resistance of rotor circuits taking the value r_{r1}, (Figure 1), r_{r2} (Figure 2), r_{r3} (Figure 3) и r_{r4} (Figure 4).

DESIGNING OF PHASE II ARL-UNBIASED S^2-CHART WITH IMPROVED PERFORMANCE USING REPETITIVE SAMPLING

Sonam Jaiswal[1] and Nirpeksh Kumar[2*]

•

[1]DST-Center for Interdisciplinary Mathematical Sciences (DST-CIMS), Banaras Hindu University, Uttar Pradesh, India

[2*]Department of Statistics, Banaras Hindu University, Uttar Pradesh, India

[1]jsonam572@gmail.com

[2*]nirpeksh@gmail.com

Abstract

In this paper, we consider a two-sided Phase II S^2-chart with probability limits because the surveillance of both an increase and decrease in the process variance plays a decisive role in a continuous quality improvement program. We propose a two-sided average run length unbiased S^2-chart under the repetitive sampling with probability limits for a fixed in-control average run length and average sample size to eliminate the average run length biasedness. It is well established that the Shewhart-type charts are less sensitive to detect small to moderate changes in the process parameters. Therefore, a repetitive sampling scheme is taken into consideration to improve the S^2-chart's ability to detect changes in the process variance. Under the repetitive sampling methodology, the detection ability of the chart is improved by using more than one samples if the first sample does not provide sufficient evidence to decide whether the process is in-control or out-of-control. The proposed chart is compared with the existing charts, such as the equal tailed standard Shewhart S^2-chart and unequal tailed S^2-chart under repetitive sampling. Results show that the proposed chart is more efficient than the existing chart. Finally, an illustration has been provided in the favor of the proposed chart with the help of a published dataset.

Keywords: Average run length, average sample number, average run length-unbiased, Control chart, in-control and out-of-control performances, Process variability.

1. Introduction

Recently, the S^2-chart has gained popularity to monitor a decrease and increase in the process variability when the quality characteristic is normally distributed. On the other hand, the traditional control charts, such as R- and S-charts with symmetric control limits, have some shortcomings. For example, the negative lower control limits are found for small sample sizes n ($n \leq 5$ for the R-chart and $n \leq 6$ for the S-chart), and their in-control (IC) performance significantly differs from the IC

* Corresponding author. Nirpeksh Kumar. Email: nirpeksh@gmail.com

performance expected one. See Knoth and Morais [13] and references therein for a detailed discussion on the limitations of the R- and S-charts with symmetric limits. In fact, the R and S-charts with symmetric control limits are less sensitive to detect a decrease in the process variance which refers to an improvement case whereas perform satisfactorily to detect an increase in the process variance. Nevertheless, in many situations, it is of interest to know the factors which are responsible for a low process variability so that a new standard can be set for the forthcoming production. Sarmiento, Chakraborti and Epprecht [17] have discussed, however, in the context of statistical tolerance limits to assess the product quality, that lowering the process variability is one of the most important objectives in the quality control studies.

For the S^2-chart, several efforts have been made to increase the efficacy of detecting the upward and downward changes in the process variance. For example, using memory-type charts such as exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) charts (see Chang and Gan [8, 9], Castagliola, Celano and Fichera [7], Lawson [14], and references therein). Besides these, the ability of the charts to detect changes can also be improved by using more than one sample if the first sample does not provide sufficient evidence to decide whether the process is IC or out-of-control (OOC). The idea is derived from the acceptance sampling plan, where the following sample is considered when the decision of rejecting or accepting a lot cannot be made based on the first sample. The different strategies are considered to use the information from the next sample. For example, in a double sampling plan, a second sample is used if the decision is not made on the first sample, and combined information of both samples is used to take a decision (see Montgomery [15]). Various attempts to design the charts based on double sampling are made to improve the chart's ability to detect OOC signals. See, for example, S-chart for agile manufacturing (He and Grigoryan [11]), S^2-chart (Khoo [12]).

In recent years, the repetitive sampling (RS) has been used to design charts by several authors, to name a few; see Ahmad, Aslam and Jun [1], Aslam, Azam, and Jun [4], Aslam, Khan, Azam et al. [6]. The idea of the RS was firstly used by Sherman [18] to develop an attribute acceptance sampling plan. In order to implement the RS to the control chart, the region of the control chart is divided into three sub-regions: the acceptance region (IC region), the indecision region (repetitive region), and the rejection region (OOC region). The quality control personnel use the next sample when the first plotted sample falls in the indecision region and repeats the process until the sample lies either in the acceptance or rejection region. Please note here that in the RS scheme, we do not keep track of how many times we have repeated our sample. Thus, under the RS scheme, on average, more than one charting point is required to make a final decision on IC or OOC status of the process. This is usually measured by the average sample size (ASS) (the expected number of sample size needed until the final decision is made) that can be fixed in advance. In addition, this scheme uses fewer number of parameters to design a control chart based on double sampling (Aslam, Azam and Jun [4], Aslam, Khan, Azam et al. [6]). Moreover, in comparison to other sampling schemes, the RS scheme throws information away by discarding the samples lying in the indecision region. Still, on the other hand, we gain in the simplicity of design and execution. Ahamd, Aslam, and Jun [1] implemented the RS in constructing the $\bar{X}$ control limits based on the capability index. Aslam, Azam and Jun [4] developed the attribute chart and np chart under the RS. Aslam, Khan, Azam et al. [6] used RS to design the T-chart based on the transformed exponential variables. Aslam [3] used the RS to design the S^2-chart for the neutrosophic statistic. Recently, Alevizakos, Chatterjee, Koukouvinos et al. [2] examined the effect of parameter estimation on the performance of the variable control charts under the RS and Singh and Kumar [19] designed the average run length (ARL)-unbiased exponential chart under the RS to monitor the inter-arrival times in high-yield processes. Performance of different chart under RS shows better detection ability in the OOC case than the standard one's.

Note that Aslam, Khan and Jun [5] considered the RS to propose a new S^2-chart with equal

distance (3-sigma) limits (thereafter AKJ chart) and showed superior to the existing standard S^2-chart. Note that they constructed the two-sided chart, but discussed only the case of process deterioration (increased variance). But to have an idea about their chart's performance in the improvement case, we calculated the ARL values using their control limits for shifts representing the improvement case (decreased variance). It is worth to mention here that the chart with equal-tailed limits and using a skewed charting statistic possesses an undesirable property which is known as ARL-biased in the SPC literature. For such (ARL-biased) charts, the ARL function does not achieve its maximum at the IC state. As a result, the chart gives a delayed OOC signal than the false alarm. For more details, please refer to Knoth and Morais [13], Zhang, Bebbington, Lai et al. [20]. To overcome this property, the ARL-unbiased charts are proposed to ensure the detection of shifts in the process parameter more quickly than a false alarm. Note that Aslam, Khan and Jun [5] designed chart is ARL-biased and triggers an OOC signal lately for a decrease in the process variance than it raises a false alarm. However, Knoth and Morais [13] point out that a decrease in the process variance should be followed seriously as it is a synonym for a quality improvement. Thus, the adoption of ARL-unbiased charts can play an absolutely deciding role in achieving the final goal of producing better-quality items (Pignatiello et al. [16]), while using charts to signal to both decreases and increases in a parameter.

The objective of this study is to improve the ability of the Phase II two-sided ARL-biased and ARL-unbiased S^2-chart to detect an increase and decrease in the process variance by implementing the RS scheme and to compare their performances with the existing the AKJ chart and the standard ARL-biased and -unbiased S^2- charts.

The rest of the article is organized as follows. Section 2 discusses a general charting procedure to construct the equal-tailed S^2-chart under the RS. In section 3, the performances of the proposed S^2-chart under the RS are examined and compared with the AKJ and standard S^2-charts. In section 4, we design the ARL-unbiased S^2-chart under the RS and its performance is evaluated in Section 5. An example is provided in Section 6. Finally, the concluding remarks are offered in Section 7.

2. Phase II S^2-chart with equal-tailed probability limits under repetitive sampling

Let $X_1, X_2, \dots, X_n$ be independently and identically distributed $N(\mu, \sigma_0^2)$ random variables where μ and σ_0^2 are the IC process mean and variance, respectively. The charting statistic for the S^2-chart is the sample variance given by $S^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2, \bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$. Let LCL and UCL be the lower and upper control limits, respectively, of the standard S^2-chart (with the single sampling (SS) scheme). Thus, the equal-tailed limits can be obtained from $P[S^2 < \text{LCL}|\sigma_0^2] = P[S^2 > \text{UCL}|\sigma_0^2] = \frac{\alpha}{2}$, where α is the nominal false alarm rate (FAR). It is known that $(n-1)S^2/\sigma_0^2$ follows a χ^2-distribution with $(n-1)$ degrees of freedom (df). Thus, the control limits LCL and UCL can be expressed as follows.

$$\text{LCL} = \frac{\sigma_0^2}{n-1}\chi^2_{\frac{\alpha}{2},n-1} = \frac{\sigma_0^2}{n-1}A_1 \quad \text{and} \quad \text{UCL} = \frac{\sigma_0^2}{n-1}\chi^2_{1-\frac{\alpha}{2},n-1} = \frac{\sigma_0^2}{n-1}A_2 \tag{1}$$

where $A_1 = \chi^2_{\frac{\alpha}{2},n-1}$, $A_2 = \chi^2_{1-\frac{\alpha}{2},n-1}$. The $\chi^2_{\zeta,n-1}$ denotes the 100ζ-percentile of the χ^2-distribution with $(n-1)$ df. The center line (CL) of the S^2-chart is given by

$$\text{CL} = \frac{\sigma_0^2}{n-1}\chi^2_{0.5,n-1} = \frac{\sigma_0^2}{n-1}A_3 \tag{2}$$

where $\text{A}_3 = \chi^2_{0.5,n-1}$. Define $\delta = \frac{\sigma_1^2}{\sigma_0^2} > 0$, which quantifies the magnitude of the process variance shift from IC variance σ_0^2 to the shifted variance $\sigma_1^2 = \delta\sigma_0^2$. Clearly, for $\delta = 1$, the process is IC, otherwise it is OOC. Further note that $\delta > 1$ ($\delta < 1$) represents the upward (downward) shift in the process variance reflecting the increased (decreased) process variability which refers to the deterioration (improvement) case. Let us define a signalling event E that the charting statistic lies outside the control limits. Thus, the probability of signal (PS), when the process variance is σ_1^2, is given by

$$\beta(\delta) = \mathrm{P}(E|\sigma_1^2) = P[S^2 < \mathrm{LCL} \text{ or } S^2 > \mathrm{UCL}|\sigma_1^2 = \delta\sigma_0^2]$$
$$= 1 + F_{\chi^2_{n-1}}\left(\frac{A_1}{\delta}\right) - F_{\chi^2_{n-1}}\left(\frac{A_2}{\delta}\right)$$

where $F_{\chi^2_{n-1}}(\cdot)$ denotes the cumulative distribution function (CDF) of the χ^2-distribution with $(n-1)$ df. Thus, the ARL, the reciprocal of the PS, for a Shewhart chart, is given by

$$\mathrm{ARL}(\delta) = \frac{1}{\beta((\delta)} \tag{3}$$

To implement the RS, in addition to the outer control limits, say, $\mathrm{UCL_{RS}}, \mathrm{LCL_{RS}}$, two additional inner (repetitive) control limits $\mathrm{URL_{RS}}$ and $\mathrm{LRL_{RS}}$ are introduced. Consequently, the whole area of the control chart is divided into six regions which are as follows- (i) one extending above the $\mathrm{UCL_{RS}}$ (region a), (ii) one extending between $\mathrm{UCL_{RS}}$ and $\mathrm{URL_{RS}}$ (region b), (iii) one extending between $\mathrm{URL_{RS}}$ and CL (region c), (iv) one extending between the CL and $\mathrm{LRL_{RS}}$ (region d), (v) one extending between $\mathrm{LRL_{RS}}$ and $\mathrm{LCL_{RS}}$ (region e), (vi) one extending below $\mathrm{LCL_{RS}}$ (region f). These six regions can further be broadly classified as OOC region (regions (a) and (f)), indecision region (regions (b) and (e)) and the IC region (regions (c) and (d)), respectively. The control limits are such that $\mathrm{UCL_{RS}} \geq \mathrm{URL_{RS}} > \mathrm{CL} > \mathrm{LRL_{RS}} \geq \mathrm{LCL_{RS}}$. Under the RS, the process is declared to be OOC (IC), when the charting point lies in the OOC region (IC region). Otherwise, the process is repeated for the resampling when the charting point lies in the indecision region. Let α_1 and α_2 be the probabilities of a charting point lying beyond the $\mathrm{UCL_{RS}}$ and $\mathrm{LCL_{RS}}$, and $\mathrm{URL_{RS}}$ and $\mathrm{LRL_{RS}}$, respectively. Thus, the equal-tailed outer and inner control limits can be obtained from $P[S^2 < \mathrm{LCL_{RS}}|\sigma_0^2] = P[S^2 > \mathrm{UCL_{RS}}|\sigma_0^2] = \alpha_1/2$ and $P[S^2 < \mathrm{LRL_{RS}}|\sigma_0^2] = P[S^2 > \mathrm{URL_{RS}}|\sigma_0^2] = \alpha_2/2$, which can be obtained as follows.

$$\mathrm{LCL_{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{\alpha_1/2,n-1} = \frac{\sigma_0^2}{n-1}B_1 \quad \text{and} \quad \mathrm{UCL_{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha_1/2,n-1} = \frac{\sigma_0^2}{n-1}B_2 \tag{4a}$$

$$\mathrm{LRL_{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{\alpha_2/2,n-1} = \frac{\sigma_0^2}{n-1}R_1 \quad \text{and} \quad \mathrm{URL_{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha_2/2,n-1} = \frac{\sigma_0^2}{n-1}R_2 \tag{4b}$$

The constants α_1 and α_2 $(0 < \alpha_1 \leq \alpha_2 < 1)$ are known as design constants for the S^2-chart under the RS. Note that for the RS scheme, we must have $\alpha_1 < \alpha_2$, otherwise, the $\alpha_1 = \alpha_2$ will result in $\mathrm{UCL_{RS}} = \mathrm{URL_{RS}}$ and $\mathrm{LCL_{RS}} = \mathrm{LRL_{RS}}$. As a result, the control chart under the RS reduces to the standard S^2-chart in Equation (1). Let p_a, p_b, p_c, p_d, p_e and p_f denote the probabilities of a charting point lying in regions a, b, c, d, e, and f, respectively. Then, for a shift $\delta > 0$, and using Equations (2) and (4), these probabilities can be calculated as follows.

$$p_a(\delta) = P[S^2 \in a|\sigma_1^2] = P[S^2 > \mathrm{UCL_{RS}}|\sigma_1^2] = 1 - F_{\chi^2_{n-1}}\left(\frac{B_1}{\delta}\right), \tag{5a}$$

$$p_b(\delta) = P[S^2 \in b|\sigma_1^2] = P[\mathrm{URL_{RS}} < S^2 \leq \mathrm{UCL_{RS}}|\sigma_1^2] = F_{\chi^2_{n-1}}\left(\frac{B_1}{\delta}\right) - F_{\chi^2_{n-1}}\left(\frac{R_1}{\delta}\right), \tag{5b}$$

$$p_c(\delta) = P[S^2 \in c|\sigma_1^2] = P[\mathrm{CL} < S^2 \leq \mathrm{URL_{RS}}|\sigma_1^2] = F_{\chi^2_{n-1}}\left(\frac{R_1}{\delta}\right) - F_{\chi^2_{n-1}}\left(\frac{A_3}{\delta}\right), \tag{5c}$$

$$p_d(\delta) = P[S^2 \in d|\sigma_1^2] = P[\mathrm{LRL_{RS}} < S^2 \leq \mathrm{CL}|\sigma_1^2] = F_{\chi^2_{n-1}}\left(\frac{A_3}{\delta}\right) - F_{\chi^2_{n-1}}\left(\frac{R_2}{\delta}\right), \tag{5d}$$

$$p_e(\delta) = P[S^2 \in e|\sigma_1^2] = P[\mathrm{LCL_{RS}} < S^2 \leq \mathrm{LRL_{RS}}|\sigma_1^2] = F_{\chi^2_{n-1}}\left(\frac{R_2}{\delta}\right) - F_{\chi^2_{n-1}}\left(\frac{B_2}{\delta}\right), \tag{5e}$$

$$p_f(\delta) = P[S^2 \in f|\sigma_1^2] = P[S^2 < \mathrm{LCL_{RS}}|\sigma_1^2] = F_{\chi^2_{n-1}}\left(\frac{B_2}{\delta}\right). \tag{5f}$$

The probability of declaring the process OOC on a single sample is given by

$$\beta_{\mathrm{out}}(\delta) = p_a(\delta) + p_f(\delta),$$

and the probability of repetition (resampling) until the decision made is

$$\beta_{rep}(\delta) = p_b(\delta) + p_e(\delta)$$

Thus, under the RS, the PS is given by (Aslam et al. (2015))

$$\beta_{\mathrm{RS}}(\delta) = \frac{\beta_{\mathrm{out}}(\delta)}{1-\beta_{\mathrm{rep}}(\delta)} = \frac{p_a(\delta)+p_f(\delta)}{1-p_b(\delta)-p_e(\delta)} \tag{6}$$

Clearly, under the RS, the RL will also follow a geometric distribution with parameter, $\beta_{\mathrm{RS}}(\delta)$. Thus, the ARL function of the S^2-chart under the RS is given by

$$\mathrm{ARL_{RS}}(\delta) = \frac{1}{\beta_{\mathrm{RS}}(\delta)} = \frac{1-p_b(\delta)-p_e(\delta)}{p_a(\delta)+p_f(\delta)} \tag{7}$$

and the ASS for a given shift (δ) is given by (Aslam, Khan and Jun [5])

$$\mathrm{ASS}_{\mathrm{RS}}(\delta) = \frac{n}{1-\beta_{\mathrm{rep}}(\delta)} = \frac{n}{1-p_b(\delta)-p_e(\delta)} \quad (8)$$

It follows from Equation (8) that the ASS $= n$ for the chart using the SS whereas $\mathrm{ASS}_{\mathrm{RS}}(1) > n$, for the chart using the RS. Most of the existing works design the control charts under the RS for a fixed nominal IC ARL value, but a possible minimum value of the $\mathrm{ASS}_{\mathrm{RS}}(1)$ under some variations. This provides a subjective selection of the control limits, and each practitioner will have his own control limits. To avoid ambiguity, we fix IC ASS i.e., $\mathrm{ASS}_{\mathrm{RS}}(1)$ in advance and then construct the control limits. The choice of the $\mathrm{ASS}_{\mathrm{RS}}(1)$ depends on the user's own choice that how much he is willing to wait for a decision (large $\mathrm{ASS}_{\mathrm{RS}}(1)$) for a better OOC performance. See Singh and Kumar [35]. Thus, to obtain the design constants α_1 and α_2 for a fixed IC ARL i.e., $\mathrm{ARL}_{\mathrm{RS}}(1)$ value, say, ARL_0 and IC ASS value, $\mathrm{ASS}_{\mathrm{RS}}(1)$ value, say, ASS_0, we need to solve the following pair of Equations

$$\mathrm{ARL}_{\mathrm{RS}}(1) = \frac{1}{\beta_{\mathrm{RS}(1)}} = \mathrm{ARL}_0 \quad (9a)$$

$$\mathrm{ASS}_{\mathrm{RS}}(1) = \frac{n}{1-\beta_{\mathrm{rep}}(1)} = \mathrm{ASS}_0 \quad (9b)$$

Using the design constants (α_1, α_2), we can obtain the control limits for the S^2-chart under the RS. Once we obtain the design constants, we can calculate the control limits of the proposed S^2-chart using Equation (4).

3. Comparative Study

In this section, we compare the performance of the proposed control chart with the existing standard S^2 and the AKJ chart. The performance of the charts are evaluated in terms of the ARL, and lower OOC ARL values are desirable for an efficient chart.

3.1 Proposed equal-tailed S^2-chart versus Aslam et al. (2015) chart

In order to compare the performances, the ARL values of the AKJ chart (Aslam et al. [5] chart) for $n = 4, \mathrm{ARL}_0 = 370, ASS_0 = 4.11$ and $n = 7, \mathrm{ARL}_0 = 370, ASS_0 = 7.36$ are used. Further, we obtained the ARL values for the proposed chart with keeping similar IC metrics for $n = 4, 7$ i.e., $\mathrm{ARL}_0 = 370$ and $\mathrm{ASS}_0 = 4.11, 7.36$, respectively. These values are presented in Table 1. It can be observed that the AKJ chart outperforms the proposed chart in the deterioration case, however, it is very less sensitive to detect a decrease in the process variance. It is expected that with an increase in the shift size in any direction, the chart's ability to detect the changes must increase. However, the AKJ chart's sensitivity decreases to detect the larger downward shifts which contradicts the philosophy of using the control chart. The AKJ chart is insensitive to detect the improvement in the process and hence the decision of using two-sided chart becomes questionable. Thus, we do not recommend the AKJ chart when the interest lies in both improvement and deterioration cases and the direction of shifts is not known. On the other hand, for the proposed chart, the OOC ARL values decreases as δ goes far away from $\delta = 1$, except for some $\delta < 1$ values close to one. As we have discussed earlier, it is undesirable and in the following section, we eliminate it.

Table 1: *Comparison of proposed chart with the chart by Aslam et al. (2015)*

	$n = 4$						$n = 7$					
	AKJ chart		Proposed chart		S^2 chart		AKJ chart		Proposed chart		S^2 chart	
δ	ARL	ASS	ARL	ASS	ARL	ASS	ARL	ASS	ARL	ASS	ARL	ASS
0.1	4.25E+17	4.00	18.77	5.54	25.34	4.00	9.82E+17	7.00	1.15	17.65	2.82	7.00
0.3	3.46E+09	4.00	118.22	4.31	124.14	4.00	6.76E+11	7.00	20.59	10.19	28.69	7.00

0.5	338135.52	4.00	261.32	4.15	263.92	4.00	2589301.43	7.00	99.43	7.98	108.14	7.00
0.7	6778.54	4.02	425.98	4.10	424.74	4.00	15179.88	7.05	266.07	7.45	269.64	7.00
0.9	785.32	4.07	461.51	4.10	459.62	4.00	956.87	7.22	434.79	7.33	432.65	7.00
1	370.00	4.11	370.00	4.11	370.00	4.00	370.00	7.36	370.00	7.36	370.00	7.00
1.1	199.90	4.16	260.76	4.13	262.60	4.00	171.14	7.55	241.47	7.43	244.83	7.00
1.3	77.45	4.27	118.48	4.19	121.62	4.00	52.86	8.01	85.92	7.68	90.87	7.00
1.5	38.59	4.39	59.34	4.27	62.27	4.00	22.59	8.52	35.93	8.03	39.95	7.00
1.7	22.65	4.51	33.95	4.36	36.44	4.00	11.94	9.03	18.24	8.41	21.34	7.00
3	4.16	5.01	5.34	4.80	6.36	4.00	2.00	10.11	2.45	9.68	3.35	7.00
4	2.52	5.07	3.02	4.90	3.68	4.00	1.38	9.40	1.55	9.29	2.04	7.00
$\boldsymbol{k_1}$	4.557690						4.094190					
$\boldsymbol{k_2}$	2.432020						1.873700					
$\boldsymbol{\alpha_1}$			0.002630		0.001351				0.002570		0.001351	
$\boldsymbol{\alpha_2}$			0.029400		0.001351				0.051490		0.001351	

3.2 Proposed equal tails S^2-chart versus equal-tailed (standard) S^2-chart

Now, the ARL comparison of the proposed chart and the standard S^2-chart is presented. The OOC ARL values for the standard S^2-chart are also presented in Table 1. It can be observed that the use of the RS improves the performance of the chart. For example, when $\delta = 1.3$, the ARL values for the proposed chart are 118.48 for $n = 4$ and 85.92 for $n = 7$, respectively whereas these values for the standard S^2-chart are 121.62 and 90.87, respectively. In the improvement case ($\delta < 1$), except small shifts sizes, for example, for $0.7 < \delta < 1$ when $n = 4$ and for $0.9 \leq \delta < 1$ when $n = 7$, the proposed chart also performs better than the standard chart. Table 1 points out that the larger n results in a better performance of the chart in both improvement and deterioration cases. Moreover, with an increase in n, the range of the downward shifts becomes shorter in which the proposed chart is inferior to the standard chart. The study clearly indicates that the performance of the chart using the RS can be improved with choosing larger ASS_0 value which supports the existing research's findings.

4. Design of the Phase II ARL-unbiased S^2-chart under repetitive sampling

In this section, we modify the control limits of the proposed ARL-biased S^2-chart under the RS so that the chart is ARL-unbiased and attains the desired IC performance. To construct the ARL-

unbiased S^2-chart under the RS, we introduce the design constants γ, α_1' and α_2', so that the modified outer and inner control limits under the RS are obtained as follows.

$$\mathrm{LCL}'_{\mathrm{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{\gamma\alpha_1',n-1} = \frac{\sigma_0^2}{n-1}B_1' \quad \text{and} \quad \mathrm{UCL}'_{\mathrm{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha_1',n-1} = \frac{\sigma_0^2}{n-1}B_2' \tag{10a}$$

$$\mathrm{LRL}'_{\mathrm{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{\gamma\alpha_2',n-1} = \frac{\sigma_0^2}{n-1}R_1' \quad \text{and} \quad \mathrm{URL}'_{\mathrm{RS}} = \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha_2',n-1} = \frac{\sigma_0^2}{n-1}R_2' \tag{10b}$$

where the constants α_1', α_2' $(0 < \alpha_1' \le \alpha_2' < 1)$ are determined to keep the IC performance at a desired level and γ is obtained to eliminate the bias in the ARL function. For $\gamma = 1$, the chart reduces to the ARL-biased S^2-chart under the RS. Also, $B_1' = \chi^2_{\gamma\alpha_1',n-1}, B_2' = \chi^2_{1-\alpha_1',n-1}$, $R_1' = \chi^2_{\gamma\alpha_2',n-1}$ and $R_2' = \chi^2_{1-\alpha_2',n-1}$, respectively. To obtain the ARL from Equation (7) and ASS from Equation (8) for the ARL-unbiased S^2-chart under the RS, the probabilities p_a, p_b, p_c, p_d, p_e and p_f can be obtained from the set of Equations (5), respectively by replacing B_1, B_2, R_1, R_2 with B_1', B_2', R_1', R_2', respectively.

To construct the control limits for an ARL-unbiased S^2-chart under RS, we set (i) IC ARL equals to the nominal ARL_0, (Equation 11a) (ii) IC ASS equals to desired ASS_0 (Equation 11b), and (iii) the first derivative of the ARL function with respect to δ at $\delta = 1$ equal to zero so that the ARL function is maximized at the IC state of the process, i.e., δ=1, (Equation 11c). Thus, the design constants γ, α_1' and α_2' can be obtained by solving the following set of Equations.

$$\mathrm{ARL}_{\mathrm{RS}}(1) = \mathrm{ARL}_0 \tag{11a}$$

$$\mathrm{ASS}_{\mathrm{RS}}(1) = \mathrm{ASS}_0 \tag{11b}$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{ARL}_{\mathrm{RS}}(\delta)|_{\delta=1} = 0 \tag{11c}$$

where, $\mathrm{ARL}_{\mathrm{RS}}(\delta)|_{\delta=1} = \frac{\left(p_a(\delta)+p_f(\delta)\right)\left[-p_b'(\delta)-p_e'(\delta)\right]-\left(1-p_b(\delta)-p_e(\delta)\right)\left[p_a'(\delta)-p_f'(\delta)\right]}{(p_a(\delta)+p_f(\delta))^2}$ with $p_a'(\delta)|_{\delta=1} = -\left(-\frac{B_1}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{B_1}{\delta}\right)$, $p_b'(\delta)|_{\delta=1} = \left(-\frac{B_1}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{B_1}{\delta}\right) - \left(-\frac{R_1}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{R_1}{\delta}\right)$, $p_e'(\delta)|_{\delta=1} = \left(-\frac{R_2}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{R_2}{\delta}\right) - \left(-\frac{B_2}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{B_2}{\delta}\right)$ and $p_f'(\delta)|_{\delta=1} = \left(-\frac{B_2}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{B_2}{\delta}\right)$. Since the values of the design constants cannot be obtained analytically, we used the numerical iterative procedure in 'R' software to solve them.

Note that for $\alpha_1' = \alpha_2'$, the ARL-unbiased S^2-chart under the RS reduces to the ARL-unbiased S^2-chart under the SS and consequently, $\mathrm{UCL}'_{\mathrm{RS}} = \mathrm{URL}'_{\mathrm{RS}}$ and $\mathrm{LCL}'_{\mathrm{RS}} = \mathrm{LRL}'_{\mathrm{RS}}$. Therefore, the control limits of the ARL-unbiased S^2-chart under SS can be obtained as follows.

$$\mathrm{LCL}'_{\mathrm{RS}} = \frac{\sigma_0^2}{n-1}B_1' = \mathrm{LCL}' \quad \text{and} \quad \mathrm{UCL}'_{\mathrm{RS}} = \frac{\sigma_0^2}{n-1}B_2' = \mathrm{UCL}' \tag{12}$$

where the constants α_1' $(0 < \alpha_1' < 1)$ and $\gamma \ge 1$ can be obtained to solve the following conditions. Please see also, Knoth and Morais (2015).

$$\mathrm{ARL}(1) = \mathrm{ARL}_0 \tag{13a}$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{ARL}(\delta)|_{\delta=1} = 0 \tag{13b}$$

where $\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{ARL}(\delta)|_{\delta=1} = \left[\left(-\frac{B_2'}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{B_2'}{\delta}\right) - \left(-\frac{B_1'}{\delta^2}\right)f_{\chi^2_{n-1}}\left(\frac{B_1'}{\delta}\right)\right] / \left[1 + F_{\chi^2_{n-1}}\left(\frac{B_1'}{\delta}\right) - F_{\chi^2_{n-1}}\left(\frac{B_2'}{\delta}\right)\right]^2$. Once we obtain the design constants, we can calculate the control limits of the ARL-unbiased S^2-chart using Equation (12).

5. Performance comparison of the proposed ARL-unbiased S^2-chart under repetitive sampling and standard ARL-unbiased S^2-chart

To evaluate the proposed ARL-unbiased S^2-chart under the RS and compare with the standard ARL unbiased S^2-chart, we obtain the control limits for $ARL_0 = 370$ and $ASS_0 = 4.4, 4.8$ (for $n = 4$) and 7.7, 8.4 (for $n = 7$). The corresponding ARL and ASS values for both charts are reported in Table 2 for different δ values. It can be observed that the OOC ARL values are smaller than the IC ARL for all shift sizes which guarantees a higher chance of detecting shifts in the process variance than giving a false alarm. The charts under the RS outperform the standard ARL-unbiased S^2-chart in terms of smaller OOC ARL values. For example, when $n = 4, \delta = 0.7$, the standard ARL-unbiased S^2-chart has ARL value 254.70 whereas it is 245.72 for the proposed chart with $ASS_0 = 4.4$ and 238.35 with

$ASS_0 = 4.8$. It is worth to note that unlike the ARL-biased chart, the ARL-unbiased chart with the RS performs better than the corresponding ARL unbiased S^2-chart for all shifts. Moreover, the larger n and ASS_0 result in an improved performance. The design of the proposed chart is flexible, and a user can design at his/her own choice how much he/she is willing to pay the price in terms of large ASS_0 to get the OOC signal quickly. It is worth mentioning that the chart becomes inferior to the ARL-biased S^2-chart in the deterioration case. It may seem unreasonable because one wish to have a chart which is able to detect the situations that are responsible for bad quality items. Final choice, however, is with the management. The present study provides choices to the user/management that he/she can choose an appropriate chart in his/her favorable conditions.

Table 2: *Un-biased performance of the equal-tailed S^2-chart under SS and RS using the metric ARL and ASS for nominal $ARL_0 = 370$, $ASS_0 = 4.4, 4.8$ at $n = 4$ and $ASS_0 = 7.7, 8.4$ at n=7, respectively.*

	$n = 4$						$n = 7$					
	ARL-unbiased S^2		Proposed ARL-unbiased chart				ARL-unbiased S^2		Proposed ARL-unbiased chart			
			ASS = 4.4		ASS = 4.8				ASS = 7.7		ASS = 8.4	
δ	ARL	ASS	ARL	ASS	ARL	ASS	ARL	ASS	ARL	ASS	ARL	ASS
0.1	15.36	4.00	3.99	16.87	1.96	37.27	2.22	7.00	1.01	16.19	1.00	17.12
0.3	73.38	4.00	53.57	6.04	41.87	8.43	19.51	7.00	9.11	16.32	5.96	26.96
0.5	155.25	4.00	138.82	4.93	126.79	5.90	71.11	7.00	55.60	9.82	47.33	12.57
0.7	254.70	4.00	245.72	4.57	238.35	5.15	175.64	7.00	163.42	8.29	155.56	9.51
0.9	351.05	4.00	349.52	4.43	347.91	4.86	330.21	7.00	327.60	7.78	325.74	8.56
1	370.00	4.00	369.85	4.40	369.92	4.80	370.00	7.00	369.98	7.70	369.91	8.40
1.1	348.38	4.00	346.90	4.39	346.79	4.77	325.17	7.00	322.80	7.69	321.15	8.37
1.3	224.97	4.00	218.74	4.41	216.02	4.79	149.61	7.00	142.72	7.84	138.79	8.55
1.5	122.57	4.00	115.69	4.48	112.55	4.87	64.36	7.00	58.41	8.13	55.59	8.95
1.7	69.44	4.00	63.64	4.57	61.12	4.99	32.50	7.00	27.97	8.50	26.13	9.46
3	9.21	4.00	7.29	5.17	6.71	5.75	4.07	7.00	2.87	10.15	2.58	11.46
4	4.81	4.00	3.65	5.37	3.35	5.95	2.32	7.00	1.67	9.84	1.55	10.75
k	5.821054		5.674593		5.614001		3.556330		3.495460		3.435879	
α_1	0.000396		0.000368		0.000341		0.000593		0.000547		0.000508	
α_2			0.013988		0.025540				0.020769		0.038080	

6. Example

To illustrate an application of the proposed ARL-biased and -unbiased S^2-chart using the RS, we use data considered in a case study of syringes with a self-contained, single dose of an injectable drug produced by a pharmaceutical company. The conclusions of the case study and associated data have published in Franklin and Mukherjee [10]. The critical quality characteristic was the length at which the cap of the syringe was tacked. For further details, we refer the interested readers to follow Franklin and Mukherjee [10]. Based on the capability analysis with $C_{pk} = 1.04$, the management determined that the process was minimally capable and susceptible to producing the length desired. Therefore, it was decided to monitor the process via control charts to find any reasonable cause(s).

To monitor the process under consideration, the case study considered total 47 samples each of size 5 which are reported in Table 2 of Frankline and Mukherjee [10]. In their study, the R-chart was used to monitor the process variability. The quality practitioner used the first 15 points to construct the control limits of the $\bar{X}$- and R- charts and plotted these 15 points on the charts. The charts show that the process is OOC in both centre and variation as well. However, these OOC signals were not noticed. We use the S^2- chart instead of R-chart and find that the S^2-chart does not give any OOC signal implying that the variability is IC. The next 15 samples were collected for phase II monitoring, and it was noticed that the caps were not perfectly tacked. Thus, the technician was called for the adjustments. Taking the first attempt of the machine adjustment into consideration, 31st sample was collected and observed that still the machine was not at its targeted position and to the next, second attempt were made to adjust the machine. Sample number 32nd were collected and plotted on the R chart to observe the adjustment effect. Quality practitioner observed that the second adjustment was also no better and the technician called for the third time. The third try was successful in the sense that the next 15 samples collected and plotted, and no point was beyond the control limits, and no action was taken.

However, a careful examination of both charts indicates that the process is not IC. To examine the sensitivity of variation, here, S^2-chart under RS (proposed chart) has been used instead of R-chart. For this purpose, we implement the ARL-biased and -unbiased charts design under the RS with nominal $(\mathrm{ARL}_0,\ \mathrm{ASS}_0) = (370.4, 5.5)$ for $n = 5$. The estimates from the first 15 samples were used to construct the control limits of the proposed S^2-chart under the RS. We get variance estimate $\sigma^2 = 0.000134$. The set of control limits $(\mathrm{LCL}_{RS}, \mathrm{UCL}_{RS})$ and the repetitive limit $(\mathrm{LRL}_{RS}, \mathrm{URL}_{RS})$ of the ARL-biased and -unbiased S^2-charts under the RS are found to be $3.37 \times 10^{-6}, 6.03 \times 10^{-4}$ and $(2.29 \times 10^{-5},\ 3.23 \times 10^{-4})$ and $(4.34 \times 10^{-6},\ 6.79 \times 10^{-4})$ and $(3.04 \times 10^{-5},\ 4.04 \times 10^{-4})$, respectively, which are depicted in Figure 1 with the $\mathrm{CL} = 1.34 \times 10^{-4}$. For the better visualization, we have taken the y-axis on log scale. Red and dashed lines represent the control limits of the ARL-unbiased S^2-chart under RS, whereas black and dashed lines represent the control limits of the ARL-biased S^2-chart under the RS. It can be seen from Figure 1 (Panels 2 and 3) that unlike the R-chart, both ARL-biased and -unbiased S^2-charts under the RS produce OOC signals. Thus, the process for the first 15 samples may be considered IC. The next 17 sample points clearly indicate the signal at the points 23 and 28. After the adjustment, the proposed chart also gives a signal at the 41st point. In the case study, the quality practitioner observed that the process was OOC signal based on the supplementary runs rules. However, a less qualified operator overlooked these non-random patterns. The proposed chart gives more objective assessment to decide the OOC signal because the points are lying below the control limits.

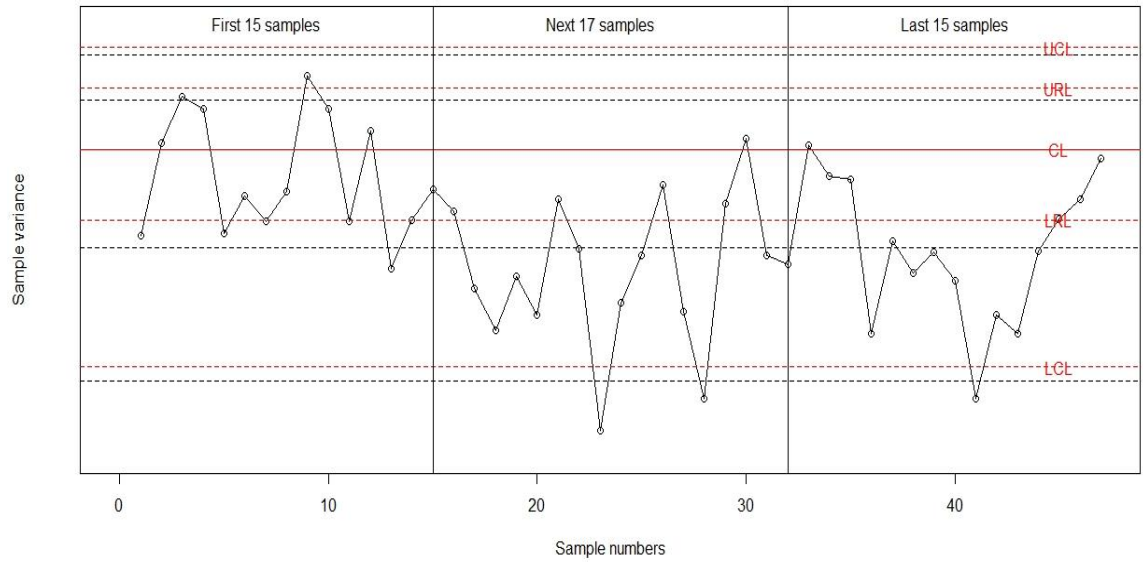


Figure 1: *The ARL-unbiased S^2-chart under the RS (red dashed lines) and ARL-biased S^2-chart under the RS (black dashed lines).*

7. Conclusion

In this paper, we consider the ARL-biased (equal-tailed) and -unbiased S^2-charts to monitor the changes (both upward and downward) in the process variance. The study shows that the ability of both charts to early detect changes can be enhanced by using the RS with maintaining simplicity of design and operation and features of the Shewhart's type charts. The ARL-biased and -unbiased S^2-charts using the RS perform better than their counterparts ARL-biased and -unbiased S^2-charts under the SS. The ARL-biased chart using the RS outperforms the ARL-unbiased chart using the RS in detecting an increase in the variance, though the latter ensures an early detection of both downward and upward shifts than it raises the false alarm. The study provides alternatives of choosing their designs according to their need. Indeed, if the management is aspiring for a continual improvement of the process, he may opt the ARL-unbiased S^2-chart under the RS.

Further, many research topics maybe of interest as a follow-up. The present study considers the case when the process variance is known, which may not be suitable in some real applications. It is well accepted that the parameter estimation affects the chart's performance in a negative way which also needs to be investigated. Following the recent literature, the conditional and unconditional performance of the proposed charts can be investigated given a Phase I sample. Finally, all the calculations are performed using the R statistical software and the programs are available from the authors on request.

References

[1] Ahmad, L., Aslam, M., and Jun, C. H. (2014). Designing of X-bar control charts based on process capability index using repetitive sampling. *Transactions of the Institute of Measurement and Control*, 36: 367-374.

[2] Alevizakos, V., Chatterjee, K., Koukouvinos, C., and Lappa, A. (2023). The effect of parameters estimation on the performance of variables control charts under repetitive sampling. *Communications in Statistics-Theory and Methods*, 52:2379-2401.

[3] Aslam, M. (2019). Control chart for variance using repetitive sampling under neutrosophic statistical interval system. *IEEE Access*, 7:25253-25262.

[4] Aslam, M., Azam, M., and Jun, C. H. (2014). New attributes and variables control charts under repetitive sampling. *Industrial Engineering and Management Systems*, 13:101-106.

[5] Aslam, M., Khan, N., and Jun, C. H. (2015). A new S^2 control chart using repetitive sampling. *Journal of Applied Statistics*, 42:2485-2496.

[6] Aslam, M., Khan, N., Azam, M., and Jun, C. H. (2014). Designing of a new monitoring t-chart using repetitive sampling. *Information sciences*, 269:210-216.

[7] Castagliola P, Celano G, and Fichera S (2009). A new CUSUM-S^2 control chart for monitoring the process variance. *Journal of Quality in Maintenance Engineering*, 0:0-0.

[8] Chang, T. C., and Gan, F. F. (1994). Optimal designs of one-sided EWMA charts for monitoring a process variance. *Journal of statistical Computation and Simulation*, 49:33-48.

[9] Chang, T. C., and Gan, F. F. (1995). A cumulative sum control chart for monitoring process variance. *Journal of Quality Technology*, 27:109-119.

[10] Franklin, L. A., and Mukherjee, S. N. (1999). An SPC case study on stabilizing syringe lengths. *Quality Engineering*, 12:65-71.

[11] He, D., and Grigoryan, A. (2002). Construction of double sampling s-control charts for agile manufacturing. *Quality and reliability engineering international*, 18:343-355.

[12] Khoo, M. B. (2004). S^2 control chart based on double sampling. *International Journal of Pure and Applied Mathematics*, 13:249-258.

[13] Knoth S and Morais M C (2015). On ARL-unbiased control charts. In *Frontiers in Statistical Quality Control*, 11:95-117.

[14] Lawson, J. (2019). Phase II monitoring of variability using Cusum and EWMA charts with individual observations. *Quality Engineering*, 31:417-429.

[15] Montgomery, D. C. (2020). *Introduction to statistical quality control*. John Wiley & Sons.

[16] Pignatiello Jr, J. J., Acosta-Mejia, C. A., and Rao, B. V. (1995). The performance of control charts for monitoring process dispersion. In *Proceedings of the 4th Industrial Engineering Research Conference*, 0:320-328

[17] Sarmiento, M. G., Chakraborti, S., and Epprecht, E. K. (2018). Exact two-sided statistical tolerance limits for sample variances. *Quality and Reliability Engineering International*, 34:1238-1253.

[18] Sherman, R. E. (1965). Design and evaluation of a repetitive group sampling plan. *Technometrics*, 7:11-21.

[19] Singh, T., and Kumar, N. (2022). Improved Phase II exponential chart under repetitive sampling. *International Journal of Agricultural and Statistical Sciences*, 0:0-0.

[20] Zhang, L., Bebbington, M. S., Lai, C. D., and Govindaraju, K. (2005). On statistical design of the S^2 control chart. *Communications in Statistics-theory and Methods*, 34:229-244.

EVALUATION OF PERFORMANCE MEASURES FOR RELIABLE AND SECURE PHISHING DETECTION SYSTEM

Pratikkumar A Barot[1*], Sunil A Patel[2], H B Jethva[3]

•

Department of Computer Engineering, Gujarat Technological University, Ahmedabad, India

[1]pratikabarot@gmail.com

[2]dp.sapatel@gmail.com

[3]hbjethva@gmail.com

Abstract

Phishing is an illegal act and security breach which acquires a user's confidential information without consent. Anti-phishing techniques used to detect and prevent such malicious acts to provide data safety to the end user. Researchers proposed an anti-phishing solution with the help of techniques like the blacklist record, heuristic function, visual similarity, and machine learning algorithm. In recent times many researchers proposed machine learning techniques for phishing detection and achieve more than 90% accuracy. However, there is reliability issue in the accuracy measures used by the researchers. In real life, the phishing dataset is unbalanced. Most of the researchers ignore this data quality during their research work design. In the case of unbalanced data, traditional accuracy measure does not give proper performance evaluation. It shows biased performance evaluations. In this paper, we experimented with an unbalanced dataset of phishing detection and did detailed result analysis to highlight the reliability issue of traditional performance evaluation measures for unbalanced data classification. We experiment with four classification algorithms and found that more than 90% of accuracy does not entitle any classifier as secure and safe if the dataset is unbalanced. Our work highlights the data factors and algorithmic limitations that compromise the system security and data safety.

Keywords: Reliable phishing detection, Cyber security, Unbalanced data, Performance evaluation, Cyber safety.

1. Introduction

The use of internet-based services increases day by day. However, the internet brings some hidden security threats to users. One of the threats is Phishing. Phishing is a cybercrime that acquires the user's confidential information without the consents of the user. Phishing websites looks like its legitimate counterpart and spoof user activity to obtain user's personal information [1]. Other phishing techniques which compromise the data safety are email phishing, SMS phishing, voice phishing, and website phishing [2].

The Anti-Phishing working group (APWG) [3] is a non-profit organization. APWG examines and publishes reports of phishing attacks. According to the APWG report, in the 1st quarter, 0f 2018 total 2,63,538 phishing attempts gets detected, and in the 2nd quarter of 2018, total 2,33,040 phishing attempts gets detected. The online banking and payment sector experiences the most number of phishing attacks.

Many researchers performed studies to proposed reliable and secure phishing detection tools [4–6]. However, the unbalanced nature of the phishing dataset does not much explored by researchers. The phishing dataset is unbalanced dataset with two class labels: Legitimate and Phishing. It has more phishing instances as compared to legitimate instances. In most of the

existing research work, the legitimate class label is in a minority. Unbalanced data classification needs more focus when the misclassification cost of one class is higher than another class [7]. In the phishing dataset, the misclassification cost of phishing URLs is much higher than legitimate URLs. It makes phishing data classification more complex. Further the unbalanced data makes the traditional evaluation parameters to present misleading performance evaluation [8].

In this paper, we performed a detailed study and in-depth result analysis to highlight the shortcomings of performance measurement in existing phishing detection researches. To propose an accurate phishing detection system, identification and validation of the performance evaluation parameter is primary requirement. The biased performance measures present biased evaluation of the security system. Especially, for the unbalanced data, we need a balanced and unbiased performance evaluation technique [7].

This paper starts with the introduction section. The second section discuss the related work of phishing detection system. Third section gives dataset information, and fourth section give detail of implementation methodology for our experiment. Fifth section present the unbiased and reliable performance measures, and compares the results with traditional performance measures to highlight the shortcoming of widely used metrics for reliable evaluation of machine learning based security systems.

2. Related Work

Machine learning is widely use in designing of cyber security system. It is increasingly used for system and data safety through effective fault detection and bug management [9]. Various anti-phishing techniques have been use for the cyber security. Researchers uses techniques such as user training, black-list approach, heuristic-based approach, visual similarity-based approach, and machine learning approach for the anti-phishing. The subsequent section summarises the widely used anti-phishing techniques.

a) User education or simulated training

Many researchers worked on the development of mobile games which helps to educate people and to raise awareness about phishing among the uses [4]. The approach is limited to users who love gaming and due to this; this approach alone is not sufficient for phishing detection [10].

b) Blacklist approach

In this approach, a dataset is created and maintained for a list of phishing URLs. Target URL is checked in this dataset. If found in a phishing dataset then it is detected as phishing otherwise it is called legitimate [5,11].

c) Heuristic-based approach

In this approach, the heuristic function is applied to extracted features or to extract features for phishing detection. However this approach is not able to detect all new attacks and it is easier to bypass once the attacker knows the heuristic algorithm or features used [12].

d) Visual similarity-based approach

The limitation of this is approach is more time is needed for image comparison and more space is needed to store all images [6].

e) Machine learning-based approach

Machine Learning is the most promising technique approach than other anti-phishing solutions [13]. Machine Learning-based classifiers are efficient classifiers and achieved an accuracy of more than 90% [14,15]. The machine learning-based anti-phishing technique is robust and more accurate as compared to other techniques [13,14]. Classification is one of the machine learning techniques for label prediction. Classification algorithms like support vector machine (SVM), Naïve Bayesian, KNN, decision tree, and ensemble classification algorithm like random forest are

widely used. Random forest is a collection of more than one decision tree algorithm and used an ensemble technique for final prediction. Many authors claim that the random forest gives better accuracy in phishing website detection [2,15–19].

Jain et al. [1] proposed a URL based machine learning technique for phishing detection. The authors use SVM and achieved 91.28% accuracy. However, in the case of redundant or irrelevant features performance gets decreased. Authors use accuracy for performance evaluation which present majority biased performance evaluation [20].

Alejandro et al. [18] proposed a neural network algorithm for the detection of phishing URLs. They proposed a recurrent neural network (RNN) based technique. RNN requires more time to train the model. Sudhanshu et al. [21] proposed a rule-based classification technique for phishing website detection. They found that the association classification algorithm performs well as compared to other algorithms. Bayu et al. [17] presented a comparison of ensemble approaches like random forest, rotation forest, gradient boosted machine, and extreme gradient boosting against decision tree. The result shows the superior performance of random forest as compared to other classifiers.

Anand Desai et al. [22] uses SVM, KNN, and random forest classifier for malicious web content detection from balance data. Jain et al. [2] proposed an anti-phishing approach that extracts features from the client-side. The proposed approach is fast but it extracts features only from URL and source code. This approach has a limitation as it can detect webpages written in HTML only. Ahmad et al. [23] proposed three new features to improve the accuracy rate for phishing website detection. Liu et al. [19] proposed an approach that focuses on character frequency features. In this, they have combined statistical analysis of URL with a machine learning technique. However, the authors use majority class centric performance evaluation parameters in their experiments. In 2016, Tahir et al. [24] proposed a hybrid model for the classification of the phishing website.

In 2015, Bhagyashree et al. [25] proposed a feature-based approach. Features like WHOIS, Page Rank, Alexa rank, and Phish Tank-based features are used for disguising phishing and non-phishing website.

Most of the researchers claim more than 90% accuracy in phishing detection. However, the phishing dataset is unbalanced. An unbalanced dataset has unequal class distribution. Accurate detection of both the class is important in the case of an unbalanced phishing dataset. As per Barot et al. [20], traditional accuracy measure gives biased result for unbalanced data. So validation and verification of performance evaluation for the phishing detection technique is an important task. In existing researches the 90% accuracy is mostly due to the accurate classification of majority class.

Traditional classification algorithms are designed for balanced data and thus do not perform well for unbalanced data [8]. Unbalanced data has skewed class distribution. In binary-class unbalanced data, one class is in majority and another class is in the minority [26]. Imbalance ratio (IR) is used to measure the level of imbalance and it is derived as the ratio of majority class instances and minority class instances [26]. Unbalanced data need special care to improve classification accuracy [27].

As per Barot et al. [20], an accurate and unbiased evaluation of unbalanced data classification is an important task. Traditional metrics are not suitable for performance evaluation of unbalanced data classification. Especially, when the misclassification cost of minority class is huge as compared to the misclassification of the majority class, we need unbiased metrics for performance evaluation. In the case of the phishing dataset, data is unbalanced and misclassification cost is different for both the minority and majority class. So we need unbiased performance evaluation measures. Barot et al. [20] proposed a balanced and unbiased performance evaluation for unbalanced data classification.

Table 1: *The original Data source of phishing and legitimate URL.*

Dataset	No of instances	Category
PhishTank [28]	30251	Phishing
DMOZ [29]	3494	Legitimate

3. Dataset

In a real-time scenario, legitimate URLs are more than the phishing URLs. It requires a robust phishing detection system that accurately identifies phishing URLs among the large numbers of legitimate URLs. This empirical study is designe to verify the performance evaluation of phishing detection from unbalanced phishing data. We use PhishTank [28] for phishing and DMOZ [29] as a non-phishing dataset. There are a total of 30251 instances of phishing and 3494 instances of legitimate class. We combine these two datasets into one dataset. The resultant dataset is unbalanced in which the phishing class is in majority, and the legitimate is in minority.

Table 1 shows detail of the original data source of phishing and legitimate URLs. For our experiment, we create a pre-labelled unbalanced dataset that comprises extracted features of URLs (label 1 for phishing and 0 for legitimate). Total 11 features namely, "IP Address", "presence of subdomain", "'@' Present in URL or not", "Presence of dash(-) in URL or not", "Length of URL", "Suspicious words in URL", "Embedded Domain", "presence of HTTPS Protocol", "HTTP_Count", "DNS lookup", and "Inconsistent URL" are extracted from the URLs. URL is treated as inconsistent if the domain name does not match in the WHOIS database.

We created a dataset by taking 23000 phishing URLs from PhishTank [28] and 2000 legitimate URLs from DMOZ [29]. The final dataset contains a total of 25000 URL instances. The phishing dataset is unbalanced and legitimate instances are in minority.

4. Implementation Methodology

For our experiment, we have used the Naïve Bayesian, Sequential minimal optimization (SMO), decision tree, and random forest algorithms. The main aim of this experiment is to check algorithms reliability by analyzing the impact of uneven distributions of classes on classification accuracy. We used the WEKA library for the implementation. We had kept the phishing dataset unbalanced to highlight the main reasons of the high accuracy value even if minority class is totally misclassified. The poor performance of minority class detection hidden under the accurate detection of the majority instance makes the security system unreliable.

Initially, phishing and legitimate URLs given to the features extractor that extracts the features. The extracted features then used for the detection of phishing URLs. The extracted features used to train the naïve Bayesian, SMO, random forest, and decision tree classifiers.

4.1. Main Steps of Proposed Approach

The major steps of our experimental study are:

Step 1) Dataset preparation: Dataset is prepared by collecting phishing URLs from PhishTank and legitimate URLs from the DMOZ directory.

Step 2) Data pre-processing: the dataset is pre-processed to remove noise and to fill missing values. The missing value is handled by taking attribute mean.

Step 3) Feature extraction: extraction of features from the website URL by using a feature extraction algorithm.

Step 4) Train model: Traditional algorithms are implemented in JAVA using the WEKA library.

Using training dataset classifiers are trained.

Step 5) Test Model: Trained model is tested with test split and performance is evaluated.

Step 6) Result and evaluation: analysis of result and checking reliability of performance matrices.

We have used 10-fold cross-validation for training and testing. Figure 1 shows the flow of our empirical work.

As shown in Figure 1, first we collect legitimate and phishing URLs. We applied feature extraction and feature selection methods to create dataset for classification algorithms. Then we train the traditional classification model from the training set and test the model using the test set. Finally, we present the results and discuss the importance of the unbiased performance measure for unbalanced dataset classification.

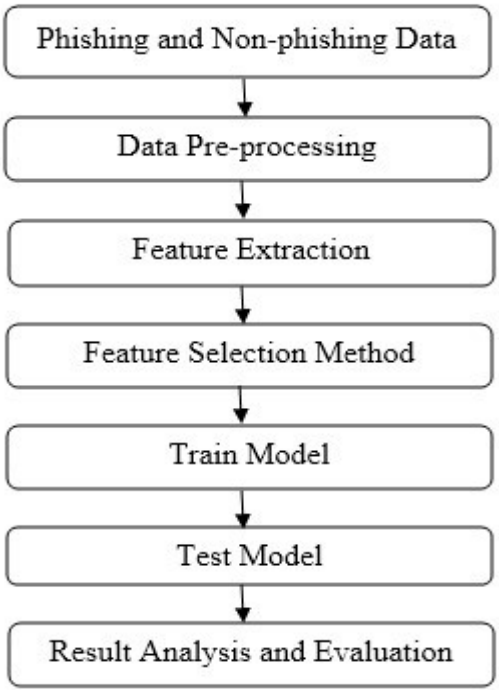


Figure 1: *Implementation flow*

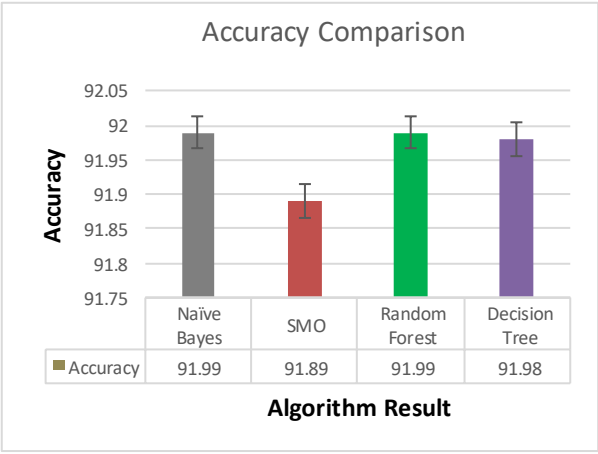


Figure 2: *Performance (Accuracy) Comparison of Classifiers*

5. Result and Discussion

Most of the existing researches show high accuracy of phishing detection. Our experiment also get similar performance in terms of high accuracy. Figure 2 shows the results comparison of naïve Bayesian, SMO, decision tree, and random forest classifiers.

Despite the high accuracy, detailed result analysis reveals that the algorithms are not an optimal choice for phishing detection because of unbalanced nature of the phishing dataset. The detailed results (TP_rate, precision, and F-measure) of naïve Bayesian, random forest and decision tree given in Table 2. The naïve Bayesian, random forest and decision tree shows identical results for phishing detection.

In Table 2, the accuracy, TP rate, precision and f-measure value for the individual class is shown to highlight the shortcoming of the traditional measures. As shown in Table 2, the accuracy, TP rate, precision, and f-measure for the minority class is 0, and for majority class, it is more than 90%, that produce the weighted average close to 90%. If classifiers are selected just by looking into the weighted average value, then it increase misclassification cost due to the misclassification of all legitimate instances. Legitimate instances are small in number (minority class) and thus ignored by all three classifiers.

The confusion matrix in Table 3 shows that one instance of the majority class gets misclassified as minority class, and all legitimate (minority) class instances gets misclassified as phishing class (majority) instances. Algorithms that show a similar pattern of the results should not considered reliable algorithm for phishing detection due to their inability of correct predictions of minority class.

From the result, we observed that the high accuracy of classification algorithms is due to the high predictive performance of the majority class. Classification accuracy of minority class is close to zero due to the high imbalance ratio of the phishing dataset and the biasing of traditional classifiers towards the majority class. Many researchers consider overall accuracy as a performance measure and show more than 90% accuracy to claim good performance of their algorithms. However, the result analysis suggest that majority class biased accuracy is not an appropriate performance measure for the unbalanced data classification [8]. The security system designed using such machine learning algorithms are unreliable and do not gives proper safe environment.

Table 2: *Detailed Result*

	Accuracy	TP Rate	Precision	F-measure
Phishing	0.999	0.999	0.920	0.958
Legitimate	0.000	0.000	0.000	0.000
Weighted Avg.	0.919	0.920	0.846	0.882

Table 3: *Confusion Matrix of Naïve Bayesian.*

	Phishing	Legitimate
Phishing	22999	1
Legitimate	**2000**	**0**

In our experiment, all the traditional classification algorithms claims around 90% for the accuracy, TP rate, precision, and f-measure. However, none of the classifier is optimal because of the tendency of the classifiers to consider all the instances as phishing instances. As the phishing

class is in majority and covers 92% of the dataset, if the classifier considers all instances as phishing, then it automatically achieves almost 90% accuracy. However, it hides the poor prediction of legitimate URL.

In the case of class imbalance, such high value of performance measures is misleading. The actual low accuracy of minority class detection outnumbered by the high accuracy of the majority class detection. Such classifiers are unreliable and not provide optimal solution for secure system design. Such machine learning algorithms should not employ even though they gives more than 90% accuracy. Selection of classifier just because it shows more than 90% accuracy ends up by blocking of all the URLs – both legitimate and phishing URLs get blocked as each URL detected as a phishing URL. Such a bias algorithm presents bias results and shows high accuracy because of the dominance of majority class instances.

Barot et al. [20] proposed a new unbiased evaluation parameters for unbalanced data. They proposed two measures called B-mean and IR based weighted mean named IRWMean. The B-mean gives a more balance performance evaluation while IRWMean gives minority class-biased performance evaluation. The mathematical equations for the B-mean and IRWMean are given in Eq. (1) & Eq. (2).

$$IRWMean = (IR \times TN_rate) + (1/IR \times TP_rate) \div (IR + 1/IR) \quad (1)$$

$$B\text{-}mean = ((IR \times TN_rate) + (1/IR \times TP_rate)) \div (((IR + 1/IR) + Acc) \div 2)) \quad (2)$$

Table 4, shows performance evaluation of the naïve Bayesian algorithm in terms of the B-mean and IRWMean for the phishing dataset. As shown in the table, B-mean is 0.46. The value of B-mean is very low as compared to the accuracy, precision and f-measure values given in Table 2. This is because of poor prediction of the minority class. The B-mean consider the imbalanced ratio to determine the misclassification cost of the minority class and majority class. Although the all majority class instances are correctly classified, the B-mean is 0.46 to indicate the misclassification of costly minority class. The B-mean show more balanced performance evaluation while the accuracy, f-measure and precision are majority class biased measures.

The value of IRWMean is 0.016 which indicate total misclassification of costly minority class. The IRWMean considers negligible benefit of correct prediction of majority class. Proper tuning of the weight according the domain specific misclassification cost gives more balanced performance evaluation.

Figure 3 shows a comparison of traditional performance parameters with IRWMean and B-mean. As none of the minority instances is correctly classified the TN_rate (for the legitimate class) is zero. But still, the precision, f-measure, and accuracy are too high and they give biased and misleading results. From the accuracy, precision, or f-measure, we are not able to discover that the TN_rate is zero and none of the minority instances is correctly predicted.

Table 4: *IRWMean and B-mean Calculation for Phishing Detection*

Imbalance Ratio (IR)	*1/IR*	*IRWMean*	*Acc*	*B-mean*
7.33	*0.136*	*0.016*	*0.91*	*0.46*

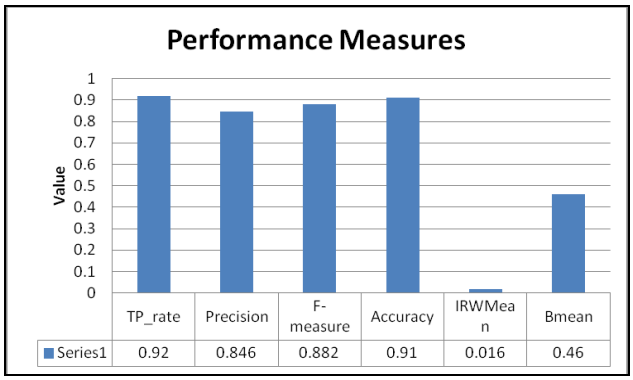


Figure 3: *Comparison of Performance Measures*

IRWMean is 0.016 which is close to the poor accuracy of minority class. IRWMean is an important parameter for performance evaluation when the misclassification cost of minority class is huge as compared to the majority class. The IRWMean gives true performance evaluation when there is a huge difference between the misclassification cost of minority and majority class. The B-mean gives more balanced performance evaluation based on the predictive accuracy of minority and majority classes and their misclassification cost. Considering the unbalanced nature and importance of both the classes of phishing datasets, the B-mean is more appropriate measure for performance evaluation.

Similar pattern of class unbalance is observed in medical datasets. In the dataset of cancer patients, if the classifier categorizes all patients as free from cancer, it automatically claims more than 90% of accuracy because in the cancer dataset cancer-free patients are in majority and cover more than 90% of the dataset. However, such classifiers cost life loss due to the misleading performance measures. In this type of applications of unbalanced data, IRWMean can give a more valid performance evaluation. In the medical domain, reliable, accurate and unbiased performance evaluation is important as misclassification of minority samples cost a loss of life.

6. Conclusion

Many researchers proposed works for the phishing detection. Most of the study claims secure and safe environment with more than 90% accuracy. However, in the case of class imbalance, high accuracy is misleading. The actual accuracy of minority class detection gets outnumbered by the high accuracy of the majority class that makes the security system unreliable. The accurate detection of the majority class, which covers large portion of the target dataset, influence the accuracy measure. In the case of the unbalanced dataset, where the misclassification cost of minority class is huge, the majority-biased misleading accuracy result into increased misclassification cost. Our experiment suggests that the performance measures should be carefully selected when the misclassification cost is uneven in unbalanced data classification. More than 90% of accuracy donot entitle any system to be reliable and secure. Comprehensive and generalized performance evaluation with unbiased measures is necessary for the reliable cyber security system. The B-mean and IRWMean performance measure consider the misclassification cost and gives un-biased and reliable performance evaluation. In the future, we will experiment with a medical dataset that generally observe large variations in the misclassification cost of minority and majority classes. In the medical domain reliability of machine learning based system is very important to save human lives.

References

[1] Jain AK, Gupta BB. PHISH-SAFE: URL features-based phishing detection system using machine learning. Adv Intell Syst Comput 2018;729:467–74. https://doi.org/10.1007/978-981-10-8536-9_44.

[2] Ankit Kumar Jain BBG. Towards detection of phishing websites on client-side using machine learning based approach. Springer Sci Bus Media 2017.

[3] Aaron G. Phishing Activity Trends Report 2nd Quarter. Anti-Phishing Work Gr 2019:1–12.

[4] Arachchilage NAG, Love S, Beznosov K. Phishing threat avoidance behaviour: An empirical investigation. Comput Human Behav 2016;60:185–97. https://doi.org/10.1016/j.chb.2016.02.065.

[5] Sheng S, Wardman B, Warner G, Cranor LF, Hong J, Zhang C. An empirical analysis of phishing blacklists. 6th Conf. Email Anti-Spam, CEAS 2009, 2009.

[6] Jain AK, Gupta BB. Phishing detection: Analysis of visual similarity based approaches. Secur Commun Networks 2017;2017. https://doi.org/10.1155/2017/5421046.

[7] Barot PA, Jethva HB. ImbTree: Minority Class Sensitive Weighted Decision Tree for Classification of Unbalanced Data. Int J Intell Syst Appl Eng 2021;9:152–8. https://doi.org/10.18201/IJISAE.2021473633.

[8] Barot P, Jethva H. MiNB: Minority Sensitive Naïve Bayesian Algorithm for Multi-Class Classification of Unbalanced Data. Int Arab J Inf Technol 2022;19:609–16. https://doi.org/10.34028/iajit/19/4/5.

[9] Chmielowski L, Konstantynov P, Luczak R, Kucharzak M, Burduk R. a Novel Method for Software Bug Report Assignment. Reliab Theory Appl 2023;18:579–88. https://doi.org/10.24412/1932-2321-2023-273-579-588.

[10] Sheng S, Magnien B, Kumaraguru P, Acquisti A, Cranor LF, Hong J, et al. Anti-Phishing Phil: The design and evaluation of a game that teaches people not to fall for phish. ACM Int. Conf. Proceeding Ser., vol. 229, 2007, p. 88–99. https://doi.org/10.1145/1280680.1280692.

[11] Jain AK, Gupta BB. A novel approach to protect against phishing attacks at client side using auto-updated white-list. Eurasip J Inf Secur 2016;2016:1–11. https://doi.org/10.1186/s13635-016-0034-3.

[12] Nguyen LAT, To BL, Nguyen HK, Nguyen MH. Detecting phishing web sites: A heuristic URL-based approach. Int. Conf. Adv. Technol. Commun., 2013, p. 597–602. https://doi.org/10.1109/ATC.2013.6698185.

[13] Qabajeh I, Thabtah F, Chiclana F. A recent review of conventional vs. automated cybersecurity anti-phishing techniques. Comput Sci Rev 2018;29:44–55. https://doi.org/10.1016/j.cosrev.2018.05.003.

[14] Khonji M, Iraqi Y JA. Phishing Detection: A Literature Survey. IEEE Commun Surv Tutor 2013:2091–121.

[15] Abbas AR, Singh S, Kau M. Detection of phishing websites using machine learning. Lect Notes Networks Syst 2020;89:1307–14. https://doi.org/10.1007/978-981-15-0146-3_128.

[16] Subasi A, Molah E, Almkallawi F, Chaudhery TJ. Intelligent phishing website detection using random forest classifier. 2017 Int Conf Electr Comput Technol Appl ICECTA 2017 2017;2018-Janua:1–5. https://doi.org/10.1109/ICECTA.2017.8252051.

[17] Bayu Adhi Tama K-HR. A Comparative Study of Phishing Websites Classification Based on Classifier Ensembles. J Korea Multimed Soc 2018;21.

[18] Alejandro Correa Bahnsen ECB. Classifying Phishing URLs Using Recurrent Neural Networks. IEEE 2017.

[19] Chunlin Liu, Bo Lang, L Wang YZ. Finding effective classifier for malicious URL detection. ACM 2018.

[20] Pratik A Barot HBJ. Empirical Study of Evaluation Metric to Proposed New Balanced Metric (B-Mean) for Unbiased Performance Evaluation of Imbalanced Data Classification. IJRAR 2018.

[21] Gautam S, Rani K, Joshi B. Detecting phishing websites using rule-based classification

algorithm: a comparison. Lect Notes Networks Syst 2018;9:21–33. https://doi.org/10.1007/978-981-10-3932-4_3.

[22] Desai A, Jatakia J, Naik R, Raul N. Malicious web content detection using machine leaning. RTEICT 2017 - 2nd IEEE Int Conf Recent Trends Electron Inf Commun Technol Proc 2017;2018-Janua:1432–6. https://doi.org/10.1109/RTEICT.2017.8256834.

[23] Ahmad Abunadi, Anazida Zainal OA. Feature Extraction Process: A Phishing Detection Approach. IEEE 2013.

[24] M. Amaad Ul Haq Tahir, Sohail Asghar, Ayesha Zafar SG. A Hybrid Model to Detect Phishing-Sites using Supervised Learning Algorithms. Int. Conf. Comput. Sci. Comput. Intell. IEEE, 2016.

[25] E. B, K. T. Phishing URL Detection: A Machine Learning and Web Mining-based Approach. Int J Comput Appl 2015;123:46–50. https://doi.org/10.5120/ijca2015905665.

[26] Barot PA, Jethva HB. Statistical Study to Prove Importance of Causal Relationship Extraction in Rare Class Classi fi cation 2018;1. https://doi.org/10.1007/978-3-319-63673-3.

[27] BAROT PA, JETHVA HB. Mgini-improved decision tree using minority class sensitive splitting criterion for imbalanced data of covid-19. J Inf Sci Eng 2021;37:1097–108. https://doi.org/10.6688/JISE.202109_37(5).0008.

[28] Phishing dataset PhishTank n.d. https://www.phishtank.com/developer_info.php.

[29] Legitimate Url Dataset DMOZ n.d. https://www.dmoz.org/.

DEVELOPING A NEW HUNTSBERGER TYPE SHRINKAGE ESTIMATOR FOR THE ENTROPY OF EXPONENTIAL DISTRIBUTION UNDER DIFFERENT LOSS FUNCTIONS

Priyanka Sahni, Rajeev Kumar
•
Department of Mathematics, Maharishi Dayanand University, Rohtak.
sahnipriyanka2@gmail.com, profrajeevmdu@gmail.com

Abstract

The aim of the paper is to develop a new Huntsberger type shrinkage estimator for the entropy function of the exponential distribution. The present paper proposes a Huntsberger type shrinkage estimator for the entropy function of the exponential distribution. This Huntsberger type shrinkage entropy estimator is based on test statistic, which eliminates arbitrariness of choice of shrinkage factor. For the developed estimator risk expressions under LINEX loss function and squared error loss function have been calculated. To assess the efficacy of the proposed estimator, numerical computations are performed, and graphical analysis is carried out for risk and relative risks for the proposed estimator. It is also compared with the existing best estimator for distinct degrees of asymmetry and different levels of significance. Based on the criteria of relative risk, it is found that the proposed Huntsberger type shrinkage estimator is better than the existing estimator for the entropy function of the exponential distribution for smaller values of level of significance and degrees of freedom.

Keywords: Exponential distribution, entropy function, shrinkage estimation, progressive censoring type sample, LINEX loss function, squared error loss function.

I. Introduction

The Exponential distribution is widely used models in reliability and life-testing research. It has been extensively examined by researchers in terms of inferential issues and its application in these fields. Many academics have studied how to estimate exponential distribution's parameters using both classical and Bayesian techniques. For example Bain [5], Chandrasekar et al. [8], Jaheen [12] and Ahmadi et al. [3], along with other references, have contributed to this body of knowledge.

If f and F be the probability density function and the distribution function of the random variable X respectively, then by Shannon [18], entropy function is given as

$$H(f) = E[-\log(f(X))] \quad (1)$$

For sharply peaked distribution entropy is very low and is much higher when the probability is spread out. Many authors worked on the estimation entropy for different life distributions. Noteworthy work in this direction may be refereed from Lazo and Rathee [15], Misra et al. [16], Jeevanand and Abdul- Sathar [13] and Kayal and Kumar [14] etc.

Suppose the random variable X has the probability distribution $f(x,\theta)$ where interest is to estimate entropy function as a function of θ. According to Thomson [20], shrinkage estimation can be accomplished by altering the usual estimator of the unknown parameter θ by bringing it closer to θ_0. Researchers have addressed several shrinkage estimators for various parameters or parametric functions under various sorts of distributions in statistical literature. Huntsberger [11] introduced weighted shrinkage estimator of the form

$\tilde{\theta}_\phi = \phi(\hat{\theta})\hat{\theta} + (1-\phi(\hat{\theta}))\theta_0$,

where $\phi(.),\ 0 \le \phi(.) \le 1,$ represents a weighted function specifying the degree of belief in θ_0.

In this paper, we shall concentrate on obtaining a new Huntsberger type shrinkage estimation of entropy function under symmetric/asymmetric loss functions for progressive type-II censored sample, when the underlying distribution is assumed to follow an exponential distribution.
The form of density we consider is

$$f(x,\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \qquad x \ge 0,\ \theta > 0 \tag{2}$$

Progressive censoring is indeed a useful scheme in the area of reliability and life time research. The number of authors including Cohen [9], Gibbons and Vance [10], Viveros and Balakrishan [22], Balakrishan and Aggarwala [6], Aggarwala [4], Adubisi and Adubisi [2], have contributed to the literature on inference problems related to progressive censoring for various probability distributions.

I. Shrinkage Estimators of H (f)

For exponential distribution with mean θ, the entropy function can be calculated as

$$H(f) = 1 + \ln(\theta) \tag{3}$$

Since $H(f)$ is linear function of $\ln(\theta)$, estimating $H(f)$ is correspondent to estimating $\ln(\theta)$. We shall write $I(\theta) = \ln(\theta)$ so that $H(f) = 1 + I(\theta)$. We will now talk about estimation of $I(\theta)$.

From the exponential distribution given in (2), Let $X_{1:m:n}, X_{2:m:n}, \ldots\ldots, X_{m:m:n}$ be Type II progressive censored sample. The progressive censored sample's joint density is then calculated (see balakrishan and Aggarwala [6])

$$f(x_{1:m:n}, x_{2:m:n}, \ldots\ldots, x_{m:m:n}) = C\prod_{i=1}^{m} f(x_{i:m:n})(1-F(x_{i:m:n}))^{R_i},\ 0 \le x_{1:m:n} \le x_{2:m:n} \le \ldots\ldots \le x_{m:m:n}, \tag{4}$$

where

$$C = n(n-R_1-1)(n-R_1-R_2-2)\ldots\ldots(n-R_1-R_2-\ldots-R_{m-1}-m+1)$$

Now substituting f and F in (4), we get

$$f(x_{1:m:n}, x_{2:m:n}, \ldots\ldots, x_{m:m:n}) = C\left(\frac{1}{\theta}\right)^m \exp\left(-\frac{\sum_{i=1}^{m}(R_i+1)x_{i:m:n}}{\theta}\right),\ 0 \le x_{1:m:n} \le x_{2:m:n} \le \ldots\ldots \le x_{m:m:n} \tag{5}$$

Then MLE of θ can easily be obtained as

$$\hat{\theta} = \frac{\sum_{i=1}^{m}(R_i+1)x_{i:m:n}}{m} \tag{6}$$

Since $I(\theta)$ is continous function of θ, the MLE of $I(\theta)$ is obtained by replacing θ by its MLE $\hat{\theta}$ in $I(\theta)$. The MLE of entropy function for the exponential distribution is then

$$\hat{H}(f) = 1 + \ln(\hat{\theta}) \tag{7}$$

We can demonstrate that the distribution of $\hat{\theta}$ has

$$f(\hat{\theta};\theta) = \left(\frac{m}{\theta}\right)^m \frac{\hat{\theta}^{m-1}\exp(-\frac{m\hat{\theta}}{\theta})}{\Gamma(m)},\ \hat{\theta} > 0 \tag{8}$$

A symmetric loss function treats underestimation and overestimation equally, penalizing both types of errors in the same manner. However, in certain situations, the consequences of underestimation and overestimation may not be the same. To address this issue, many authors have used and promoted the usage of 'asymmetric' loss functions, particularly when discussing claim settlements and other related topics. Prominent researchers such as Varian [19], Zellner [23], Basu and Ebrahimi [7], Adegoke et al. [1] have highlighted the convenience and superiority of using asymmetric loss functions in various scenarios.

The loss function proposed by Varian [21] is defined as:

$$L(\Delta) = b(e^{a\Delta} - a\Delta - 1) \tag{9}$$

where $\Delta = \hat{\theta} - \theta$, **b** denotes scale parameter and **a** denotes shape parameter. When overestimation is more critical than underestimation then the positive value of **a** is used and for other cases, its negative value is used.

In section 2, the shrinkage estimator is defined. The expressions for the risk(s) are provided in section 3. In Section 4, the Relative Risk(s) are determined. Finally, in section 5, the proposed estimator is compared with the best estimator available.

II. Proposed Estimator

Srivastava and Shah [19] have proposed a shrinkage estimator of scale parameter in exponential distribution. The key contribution of their estimator is the removal of arbitrariness in the choice of shrinkage factor 'k' by making it dependent on the test statistics. Sahni and Kumar [17] proposed a Huntsberger type shrinkage estimator for the entropy of the exponential distribution by taking 'k' dependent on the test statistic. There could be several other choice of 'k'. Thus taking idea of the various choices of shrinkage factors and also with the help of sample and prior guess information a new Huntsberger type shrinkage entropy estimator for mean of exponential distribution can be proposed as follows:

$$\tilde{I}_1(\theta) = \begin{cases} \left(\frac{2m\hat{\theta}}{\theta_0\chi^2}\right)^2 \ln(\hat{\theta}) + \left(1 - \left(\frac{2m\hat{\theta}}{\theta_0\chi^2}\right)^2\right)\ln(\theta_0); & \text{if } \chi_1^2 \le \frac{2m\hat{\theta}}{\theta_0\chi^2} \le \chi_2^2 \\ \ln(\hat{\theta}) \quad ; & \text{Otherwise} \end{cases} \tag{10}$$

where k depends on the test statistic and is given as $k = \left(\frac{2m\hat{\theta}}{\theta_0\chi^2}\right)^2$ and $\chi^2 = (\chi_2^2 - \chi_1^2)$.

III. Derivation of Risk(s)

1. Risk of MLE, $\hat{I}(\theta)$

Risk of the estimator $\hat{I}(\theta)$ with respect to LLF is defined as follows:

$$R_{LLF}(\hat{I}(\theta)) = E(\hat{I}(\theta)/LLF)$$

$$= \int_0^\infty \left(\exp\left(a\left(\ln(\hat{\theta}) - \ln(\theta) \right) \right) - a\left(\ln(\hat{\theta}) - \ln(\theta) \right) - 1 \right) f(\hat{\theta};\, \theta) d(\hat{\theta})$$

$$= \int_0^\infty \left(\exp\left(a\left(\ln(\frac{\hat{\theta}}{\theta}) \right) \right) - a\left(\ln(\frac{\hat{\theta}}{\theta}) \right) - 1 \right) f(\hat{\theta};\, \theta) d(\hat{\theta})$$

Now, by using the transformation $x = \frac{m\hat{\theta}}{\theta}$ and substituting in the integral above, we get

$$R_{LLF}(\hat{I}(\theta)) = \frac{\Gamma(m+a)}{m^a \Gamma(m)} - a(\Psi(m) - \ln(m) - 1) \tag{11}$$

where

$$\Psi(n) = \frac{\frac{d}{dn}\Gamma(n)}{\Gamma(n)}$$

Also, under SELF, the risk of estimator $\hat{I}(\theta)$ is obtained as

$$R_{SELF}(\hat{I}(\theta)) = E\left(\ln(\hat{\theta}) - \ln(\theta) \right)^2$$

$$= \int_0^\infty \left(\ln(\hat{\theta}) - \ln(\theta) \right)^2 f(\hat{\theta};\, \theta) d(\hat{\theta}) = \int_0^\infty \left(\ln(\frac{\hat{\theta}}{\theta}) \right)^2 f(\hat{\theta};\, \theta) d(\hat{\theta})$$

$$= G(0, \infty, (\log(x))^2) - 2\ln(m)\psi(m) + (\ln(m))^2, \tag{12}$$

where

$G(t_1, t_2, W) = \frac{\int_{t_1}^{t_2} W x^{n-1} e^{-x} dx}{\Gamma(n)}$ and W is a function of x.

2. Risk of Shrinkage Estimator $\tilde{I}_1(\theta)$

Risk of the estimator $\tilde{I}_1(\theta)$ with respect to LLF is defined as follows:

$$R_{LLF}(\tilde{I}_1(\theta)) = E(\tilde{I}_1(\theta) / LLF)$$

$$= \int_{r_1}^{r_2} \begin{pmatrix} \exp(a((\frac{2m\hat{\theta}}{\theta_0 \chi^2})^2 \ln(\hat{\theta}) + (1 - (\frac{2m\hat{\theta}}{\theta_0 \chi^2})^2) \ln(\theta_0) - \ln(\theta))) \\ -a((\frac{2m\hat{\theta}}{\theta_0 \chi^2})^2 \ln(\hat{\theta}) + (1 - (\frac{2m\hat{\theta}}{\theta_0 \chi^2})^2) \ln(\theta_0) - \ln(\theta)) - 1 \end{pmatrix} f(\hat{\theta};\, \theta) d\hat{\theta}$$

$$+ \int_0^\infty \left(\exp(a(\ln(\hat{\theta}) - \ln(\theta))) - a(\ln(\hat{\theta}) - \ln(\theta)) - 1 \right) f(\hat{\theta};\, \theta) d\hat{\theta}$$

$$- \int_{r_2}^{r_1} \left(\exp(a(\ln(\hat{\theta}) - \ln(\theta))) - a(\ln(\hat{\theta}) - \ln(\theta)) - 1 \right) f(\hat{\theta};\, \theta) d\hat{\theta}$$

where r_1 and r_2 are the boundaries of the acceptance region of a test of the hypothesis $H_0 : \theta = \theta_0$ against the alternative $H_1 : \theta \neq \theta_0$. Define $r_1 = \frac{\theta_0 \chi_1^2}{2m}, r_2 = \frac{\theta_0 \chi_2^2}{2m}$, where χ_1^2 and χ_2^2 are respectively lower and upper α^{th} percentile values of the chi-square distribution with 2m degrees of freedom. Again, letting $x = \frac{m\hat{\theta}}{\theta}$ and solving the integrals in the expression for LLF we get

$$R_{LLF}(\hat{I}_1(\theta)) = I_1 + I_2 + I_3 + \frac{4am(m+1)\log(m\phi)}{(\phi\chi^2)^2}\left[I(r_2', m+2) - I(r_1', m+2)\right] - aG(0,\infty,\ln(x))$$
$$-\frac{\Gamma(a+m)}{\Gamma(m)\, m^a}\left[I(r_2', a+m) - I(r_1', a+m)\right] + \frac{\Gamma(a+m)}{\Gamma(m)\, m^a} - a\ln(m\phi)\left[I(r_2', m) - I(r_1', m)\right] + a\ln m - 1, \quad (13)$$

where

$$I_1 = \int_{r_1'}^{r_2'} \phi^a \left(\frac{t}{m\phi}\right)^{\frac{4at^2}{(\phi\chi^2)^2}} \frac{e^{-t}t^{m-1}}{\Gamma(m)}dt,\ I_2 = \frac{-4am(m+1)}{(\phi\chi^2)^2}\int_{r_1'}^{r_2'} (\log t)\frac{e^{-t}t^{m+1}}{\Gamma(m+2)}dt \text{ and } I_3 = \int_{r_1'}^{r_2'} a(\log t)\frac{e^{-t}t^{m-1}}{\Gamma(m)}dt$$

where $r_1' = \frac{2\chi_1^2}{\lambda}$, $r_2' = \frac{2\chi_2^2}{\lambda}$, $\phi = \frac{\theta_0}{\theta}$ and I(x, n) is the cumulative distribution function of gamma distribution given as

$$I(x,n) = \frac{\int_0^{\infty} t^{n-1}e^{-t}dt}{\Gamma(n)}$$

Under SELF, risk of the estimator $\tilde{I}_1(\theta)$ is defined as follows:

$$R_{SELF}(\tilde{I}_1(\theta)) = E\left(\tilde{I}_1(\theta) - \ln(\theta)\right)^2$$
$$= \int_{r_1}^{r_2}\left((\frac{2m\hat{\theta}}{\theta_0\chi^2})^2 \ln(\hat{\theta}) + (1-(\frac{2m\hat{\theta}}{\theta_0\chi^2})^2)\ln(\theta_0) - \ln(\theta)\right)^2 f(\hat{\theta};\theta)d\hat{\theta}$$
$$+ \int_0^{\infty}\left(\ln(\hat{\theta}) - \ln(\theta)\right)^2 f(\hat{\theta};\theta)d\hat{\theta} - \int_{r_1}^{r_2}\left(\ln(\hat{\theta}) - \ln(\theta)\right)^2 f(\hat{\theta};\theta)d\hat{\theta}$$

Again by using the transformation $x = \frac{m\hat{\theta}}{\theta}$ and substituting in the integrals above, we get

$$R_{SELF}(\tilde{I}_1(\theta)) = I_4 + I_5 + I_6 + \left((\ln\phi)^2 - (\ln m)^2\right)\left[I(r_2', m) - I(r_1', m)\right] + 2(\ln m)G(r_1', r_2', (\ln x))$$
$$\left(-2\ln m\right)G(0,\infty,(\ln x)) - G\left(r_1', r_2', (\ln x)^2\right) - \frac{8m(m+1)(\ln m\phi)(\ln\phi)}{\phi^2\chi^4}\left[I(r_2', (m+2)) - I(r_1', (m+2))\right]$$
$$+ \frac{16m(m+1)(m+2)(m+3)(\ln m\phi)^2}{\phi^4\chi^8}\left[I(r_2', (m+4)) - I(r_1', (m+4))\right] + G(0,\infty,(\ln x)^2) + (\ln m)^2, \quad (14)$$

where

$$I_4 = \frac{16\Gamma(m+4)}{\phi^4\chi^8\Gamma(m)}\int_{r_1'}^{r_2'}(\ln t)^2\frac{e^{-t}t^{m+3}}{\Gamma(m+4)}dt,\ I_5 = -\frac{32\ln(m\phi)\Gamma(m+4)}{\phi^4\chi^8\Gamma(m)}\int_{r_1'}^{r_2'}(\ln t)\frac{e^{-t}t^{m+3}}{\Gamma(m+4)}dt \quad \text{and}$$

$$I_6 = \frac{8\ln(\phi)\Gamma(m+2)}{\phi^2\chi^4\Gamma(m)}\int_{r_1'}^{r_2'}(\ln t)\frac{e^{-t}t^{m+1}}{\Gamma(m+2)}dt$$

IV. Relative Risk(s)

To investigate the properties of the proposed estimator under LLF and SELF, we can compare the relative risks of the estimator with the MLE $\hat{I}(\theta)$.

The relative risk of $\tilde{I}_1(\theta)$ under LLF compared to $\hat{I}(\theta)$ is

$$RR_{LLF}(\tilde{I}_1(\theta)) = \frac{R_{LLF}(\hat{I}(\theta))}{R_{LLF}(\tilde{I}_1(\theta))}$$

Additionally, under SELF, the relative risk of $\tilde{I}_1(\theta)$ w.r.t. $\hat{I}(\theta)$

$$RR_{SELF}(\tilde{I}_1(\theta)) = \frac{R_{SELF}(\hat{I}(\theta))}{R_{SELF}(\tilde{I}_1(\theta))}$$

V. Numerical Computations And Graphical Analysis

We observe that the expressions $RR_{LLF}(\tilde{I}_1(\theta))$, $RR_{SELF}(\tilde{I}_1(\theta))$ depend on m, a, ϕ and α. To show the performance of this considered estimator under LLF and SELF, we have taken some values of these constants as given in Sahni and Kumar [17], i.e. a= -1,-2,-3, 1, 2, 3, m= 5, 8, α = 0.01, 0.05, ϕ = 0.2(0.2)1.6.

Tables I and Table II and Figures 1 to 9 present the behaviour of relative risks of the estimators w.r.t α for varing values of m and a.

i. For $m = 5$, $\alpha = 1\%$ and for all values of 'a' (+ve as well as -ve), $\tilde{I}_1(\theta)$ yield better results than the conventional estimator for the whole scale of ϕ.

ii. Further if we switch α to 5%, the same type of behaviour comes under notice for RR. However, the magnitude of relative risk values was smaller as compared to $\alpha = 1\%$ values.

iii. We have also taken $\alpha = 10\%$ in order to explore the pattern at a higher level of significance and it is found that $\tilde{I}_1(\theta)$ still gives the better results as compared to the conventional estimator but the magnitude of RR values become lower but even then it remains mostly above unity.

iv. After comparing these relative risk values, a lower value i.e. $\alpha = 1\%$ is preferred. Similarly, when varing the value of 'm', higher relative risk values were obtained for $m = 5$ compared to other values of m as 8, 10 and 12. Thus, a smaller 'm' is advised. Higher RR shows better control over risk. Therefore, we can conclude that selecting appropriate values of 'a' and ' α ' will result in a higher gain in terms of performance of $\tilde{I}_1(\theta)$.

Table 1: *Relative risk of estimator* $\tilde{I}_1(\theta)$ *under LLF*

$\alpha = 0.01$		ϕ							
m	a	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6
5	-1	0.8691	0.8752	1.4756	3.3650	5.0900	3.1115	1.7132	1.0904
	-2	0.8377	0.9944	1.8405	3.6257	4.0222	2.6767	1.7404	1.2428
	-3	0.8736	1.2787	2.1754	2.7952	2.4724	1.9211	1.5151	1.2499
	1	0.9372	0.8379	1.1346	2.2949	4.3252	2.9016	1.3766	0.7639
	2	0.9577	0.8565	1.0623	1.9232	3.5375	2.6712	1.2620	0.6679
	3	0.9719	0.8839	1.0193	1.6401	2.8481	2.4762	1.2146	0.6180
8	-1	1.0069	0.8980	1.0572	2.1134	3.9030	2.4284	1.2104	0.7311
	-2	0.9973	0.9040	1.2027	2.5358	4.0023	2.4178	1.3193	0.8545
	-3	0.9767	0.9473	1.4398	2.9114	3.5163	2.2076	1.3708	0.9713
	1	1.0087	0.9249	0.9117	1.5052	2.9003	2.1365	0.9868	0.5372
	2	1.0071	0.9420	0.8762	1.2987	2.3910	1.9649	0.9106	0.4742
	3	1.0054	0.9573	0.8546	1.1334	1.9524	1.8058	0.8639	0.4319
11	-1	1.0187	1.0611	0.9111	1.4573	2.8868	2.0297	0.9805	0.5833
	-2	1.0267	1.0523	0.9918	1.7352	3.2296	2.0971	1.0709	0.6729
	-3	1.0372	1.0475	1.1122	2.0790	3.3339	2.0673	1.1487	0.7707
	1	1.0093	1.0687	0.8137	1.0657	2.0586	1.7619	0.8195	0.4447
	2	1.0066	1.0655	0.7829	0.9228	1.6948	1.6147	0.7613	0.3962
	3	1.0048	1.0588	0.76	0.8022	1.3789	1.4688	0.7199	0.36

$\alpha = 0.05$									
5	-1	1.0199	1.1053	1.4350	2.0083	2.1672	1.6757	1.1896	0.8884
	-2	1.0424	1.2268	1.6473	2.0502	1.9381	1.5303	1.1891	0.9676
	-3	1.1248	1.4121	1.6966	1.7260	1.5322	1.3070	1.1320	1.0110
	1	1.0110	1.0230	1.1380	1.5109	1.8990	1.6317	1.0834	0.7190
	2	1.0084	1.0128	1.0600	1.2972	1.6186	1.5025	1.0252	0.6601
	3	1.0060	1.0084	1.0108	1.1343	1.3559	1.3457	0.9790	0.6262
8	-1	1.0342	1.1634	1.1458	1.4046	1.7189	1.4285	0.9832	0.7179
	-2	1.0519	1.1999	1.2771	1.6149	1.7994	1.4232	1.0246	0.7930
	-3	1.0808	1.2642	1.4479	1.7602	1.7307	1.3597	1.0488	0.8670
	1	1.0158	1.1184	0.9860	1.0348	1.3173	1.2794	0.8806	0.5906
	2	1.0111	1.0987	0.9389	0.8979	1.0998	1.1502	0.8300	0.5442
	3	1.0078	1.0794	0.9064	0.7884	0.9040	1.0001	0.7804	0.5101
11	-1	1.0134	1.2545	1.0347	1.0059	1.3095	1.2578	0.8958	0.6649
	-2	1.0219	1.2928	1.1210	1.1756	1.4639	1.2951	0.9356	0.7257
	-3	1.0371	1.3347	1.2319	1.36	1.5541	1.2935	0.9693	0.7898
	1	1.0056	1.1827	0.9107	0.7387	0.9443	1.0796	0.8089	0.5627
	2	1.0039	1.1499	0.8650	0.6358	0.7728	0.9503	0.7631	0.5232
	3	1.0028	1.12	0.8277	0.5483	0.6191	0.8031	0.7127	0.4911

Table 2: *Relative risk of estimator* $\tilde{I}_1(\theta)$ *under SELF*

α=0.01				ϕ				
m	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6
5	0.9076	0.8377	1.2584	2.7861	5.0209	3.1094	1.5457	0.9090
8	1.0094	0.9085	0.9679	1.7697	3.4465	2.3068	1.0904	0.6229
11	1.0131	1.0673	0.8548	1.2395	2.4663	1.9064	0.8938	0.5068
α=0.05								
5	1.0140	1.0478	1.2579	1.7682	2.1238	1.7027	1.1452	0.7987
8	1.0231	1.1390	1.0523	1.2039	1.5375	1.3767	0.9329	0.6491
11	1.0085	1.2178	0.9664	0.8608	1.1278	1.1841	0,8531	0.6102
α=0.1								
5	1.0384	1.1420	1.2480	1.4216	1.4917	1.2946	1.0070	0.7862
8	1.0206	1.2009	1.0874	1.0031	1.0871	1.0675	0.8857	0.7147
11	1.0059	1.2144	1.0072	0.7265	0.7954	0.9163	0.8559	0.7344

5.1. ***Graphs of Relative Risk for*** $\tilde{I}_1(\theta)$ ***under LLF***

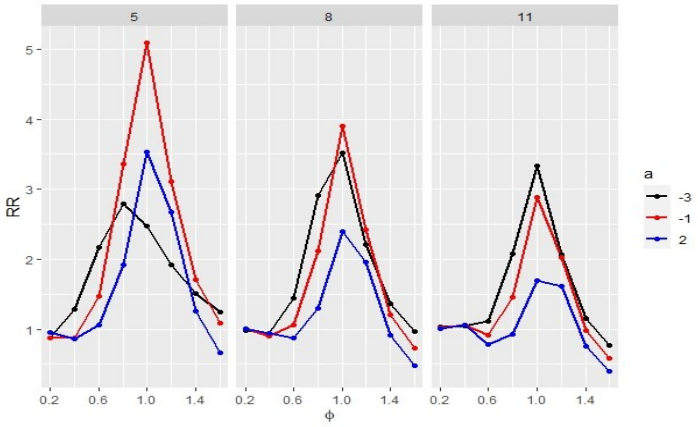


Figure 1: *For α=0.01*

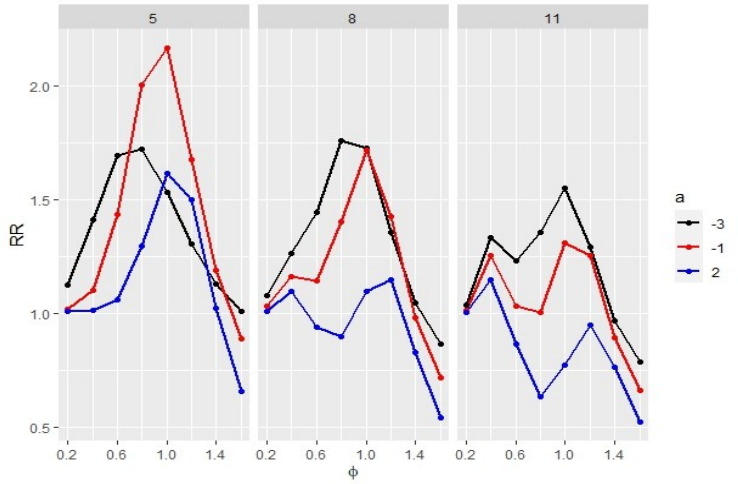


Figure 2: *For α=0.05*

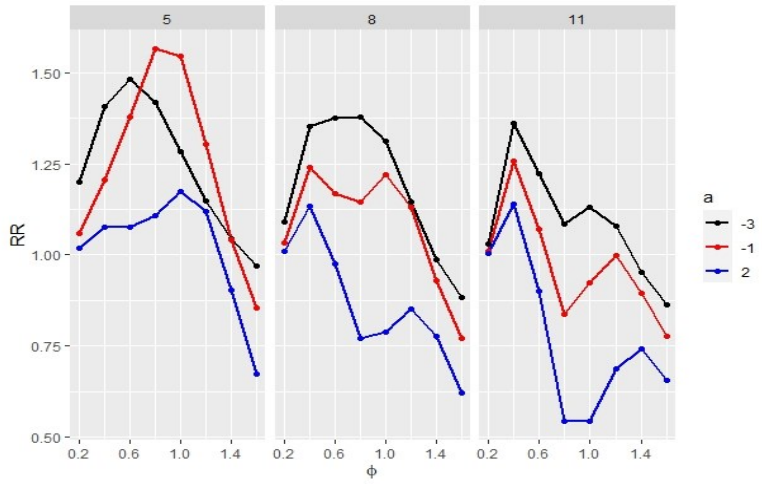


Figure 3: *For α=0.1*

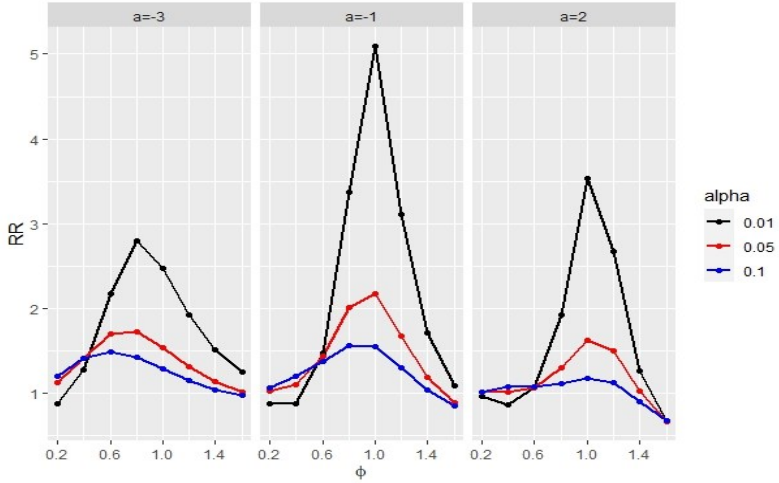


Figure 4: *For m=5*

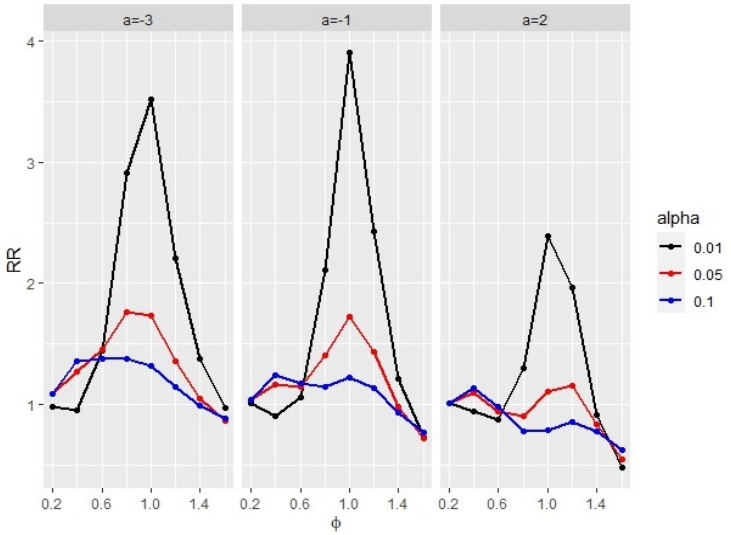


Figure 5: *For m=8*

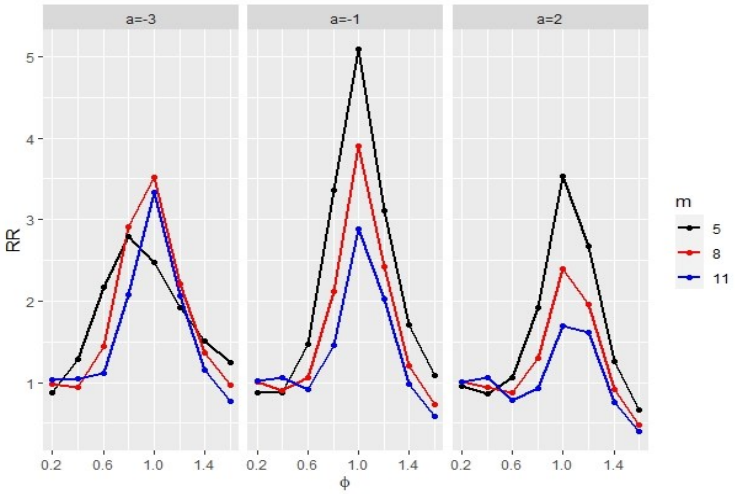


Figure 6: *For α=0.01*

5.2. ***Graphs of Relative Risk for*** $\tilde{I}_1(\theta)$ ***under SELF***

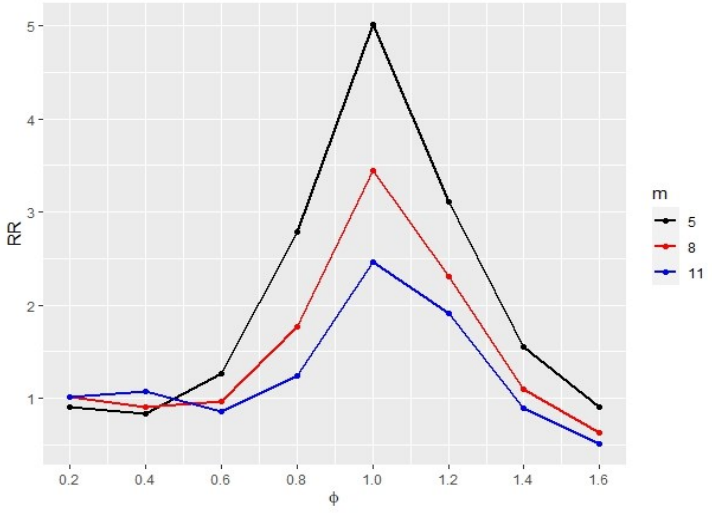


Figure 7: *For α=0.01*

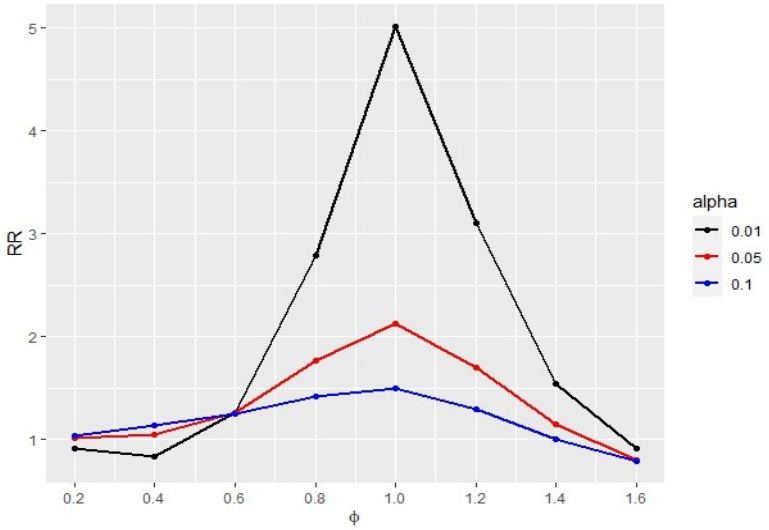


Figure 8: *For m=5*

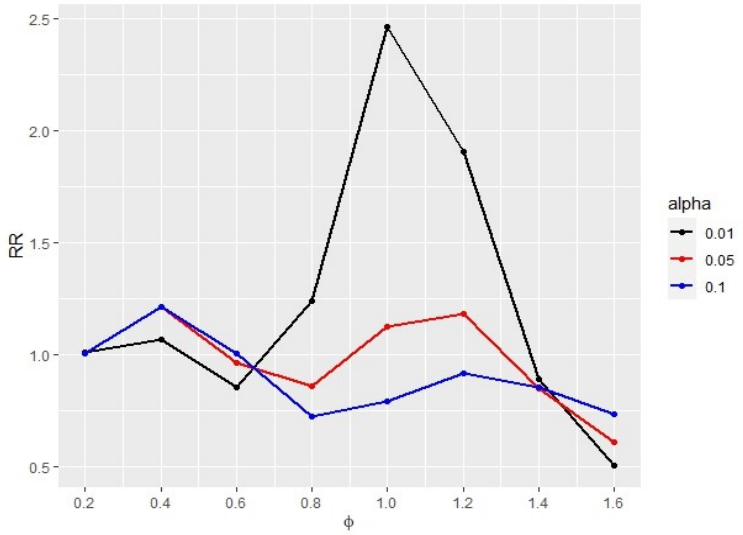


Figure 9: *For m=11*

VI. Conclusion

In this paper, a new Huntsberger type shrinkage entropy estimator for Exponential distribution have been proposed and its properties have been examined under different loss functions. On the basis of relative risk, it is found that the proposed estimator gives better results for smaller values of degrees of freedom and level of significance. And it is also concluded that the proposed estimator gives better results than the estimator proposed by Sahni and Kumar [17], when the estimated value is close to the actual value.

References

[1] Adegoke, T.M., Obisesan, K.O., Oladoja, O. M. and Adegoke, G.K. (2023). Bayesian and classical estimations of transmuted inverse Gompertz distribution. *Reliability: Theory & Applications*, *2(73):* 24-38. https://doi.org/10.24412/1932-2321-2023-273-207-222.

[2] Adubisi, O. D., Adubisi, C. E. (2023). Different estimation methods for Topp-Leone-(A) model: interference and applicationsto complete and censored datasets. *Reliability: Theory & Applications, 2(73):* 24-38. https://doi.org/10.24412/1932-2321-2023-273-100-117.

[3] Ahmadi, J., Doostparast, M. and Parsian, A. (2005). Estimation and prediction in a two

parameter exponential distribution based on k-record values under LINEX loss function. *Communications in Statistics: Theory and Methods*, 34:795-805.

[4] Aggarwala, R. (2001). Progressive Censoring: A review in Hand Book of Statistics. Advances in Reliability, 20:373-429.

[5] Bain, L.J. (1978). Statistical Analysis of Reliability and Life Testing Models. New York: Marcer Dekker.

[6] Balakrishan, N. and Aggarwala, R. (2000). Progressive Censoring: Theory and Applications. Birkhauser and Publisher. https://doi.org/10.1007/978-1-4612-1334-5.

[7] Basu, A.P. and Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of Statistical Planning and Inference*, 29:21-31.

[8] Chandrasekar, B., Leo Alexander, T. and Balakrishan, N. (2002). Equivalent Estimation for Parameters of Exponential Distributions Based Type II progressively Censored Samples. Communications in Statistics: Theory and Methods, 31(10):1675-1686.

[9] Cohen, A.C. (1963). Progressively Censored Samples in Life Testing. Technometrics, 5:327-329.

[10] Gibbons, D.I. and Vance, L.C. (1983). Estimators for the 2-Parameter Weibull Distribution with Progressively Censored Samples. IEEE Transactions on Reliability, R-32:95-99.

[11] Huntsberger, D.V. (1955). A generalization of a preliminary testing procedure for pooling data. *The Annals of Mathematical Statistics*, 26:734-743.

[12] Jaheen, Z.F. (2004). Empirical Bayes Analysis of Record Statistics Based on LINEX and Quadratic Loss Functions. Computers and Mathematics with Applications, 47:947-954.

[13] Jeevanand, E.S. and Abdul-Sathar, E.I. (2009). Estimation of residual entropy function for exponential distribution from censored samples. *ProbStat Forum.*, 2:68-77.

[14] Kayal, S. and Kumar, S. (2011). Estimating the entropy of an exponential population under the LINEX loss function. *Journal of Indian Statistical Association*, 49:91-112.

[15] Lazo, A.C.G.V. and Rathie, P.N. (1978). On the entropy of continous distribution. *IEEE Transactions on Information Theory*, 24:120-122. https://doi.org/10.1109/tit.1978.1055832.

[16] Misra, N., Singh, H. and Demchuk, E. (2005). Estimation of the entropy of multivariate normal distribution. *Journal of Multivariate Analysis*, 92:324-342. https://doi.org/10.1016/j.jmva.2003.10.003. [17] Sahni, P. and Kumar, R. (2022). Huntsberger Type Shrinkage Entropy Estimator for Mean of Exponential Distribution under Different Loss Functions. International Journal of Agricultural and Statistical Sciences, 18:913-920. https://connectjournals.com/03899.2022.18.913.

[18] Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379-423. .https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

[19] Srivastava, R. and Shah, T. (2010). Shrinkage Testimators of Scale Parameter for Exponential Model under Asymmetric Loss function. Journal of Reliability and Statistical Studies, 3(2):11-25.

[20] Thompson, J.R. (1968). Some Shrinkage Techniques for Estimating the Mean. JASA, 63: 113-122. https://doi.org/10.1080/01621459.1968.11009226.

[21] Varian, H.R. (1975). A bayesian approach to real estate assessment studies in bayesian econometrics and statistics in honour of L.J. Savage Eds. S. E. Fienberg and A. Zeliner. Amsterdam. North Holland, 195-208.

[22] Viveros, R. and Balakrishan, N. (1994). Interval Estimation of Life Characteristics from Progressively Censored Data. Technometrics, 36:84-91.

[23] Zellner, A. (1986). Bayesian Estimation and Prediction using Asymmetric Loss Functions. Journal of the American Statistical Association, 81:446-451. https://doi.org/10.1080/01621459.1986.10478289.

THE SABUR DISTRIBUTION: PROPERTIES AND APPLICATION RELATED TO ENGINEERING DATA

•

Aijaz Ahmad[1*], Afaq Ahmad[2] and Aafaq A. Rather[3]

[1]Department of Mathematics, Bhagwant University, Ajmer, India.
[2]Department of mathematical science, IUST, Awantipora, Kashmir, India.
[3]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune, India.

ahmadaijaz4488@gmail.com[1*] , aderaafaq@gmail.com[2] , aafaq7741@gmail.com[3]

Abstract

This paper introduces a novel probability distribution called the Sabur distribution (SD), characterized by two parameters. It offers a comprehensive analysis of this distribution, encompassing various properties such as moments, moment-generating functions, deviations from the mean and median, mode and median, Bonferroni and Lorenz curves, Renyi entropy, order statistics, hazard rate functions, and mean residual functions. Furthermore, the paper delves into the graphical representation of the probability density function, cumulative distribution function and hazard rate function to provide a visual understanding of their behavior. The distribution's parameters are estimated using the well-known method of maximum likelihood estimation. The paper also showcases the practical applicability of the Sabur distribution through real-world examples, underscoring its performance and relevance in various scenarios.

Keywords: Moments, moment generating function, reliability measures, mean deviations, maximum likelihood function.

Subject classification: 60E05, 62E15.

1. Introduction

Statistical distributions hold great importance in fields such as biomedicine, engineering, economics, and various scientific domains. Two widely recognized distributions, namely the exponential distribution and the gamma distribution, are often used as lifetime distributions for analyzing statistical data. Among these, the exponential distribution stands out due to its singular parameter and several intriguing statistical properties, notably its memory less property and constant hazard rate characteristic. In the realm of statistics, numerous extensions of these distributions have been developed to enhance their flexibility and applicability. One notable contribution to this literature is attributed to Lindley in [10]. He introduced a one-parameter lifetime distribution characterized by the following probability density function:

$$f(y,\beta) = \frac{\beta^2}{(1+\beta)}(1+y)e^{-\beta y} \quad ; y > 0\,, \beta > 0$$

In recent years, researchers have made significant advancements in the study of the Lindley distribution and have proposed various one- and two-parameter distributions to model complex datasets effectively. A notable contribution was made by Ghitney et al. [8], who conducted an extensive study on the Lindley distribution. They demonstrated that the Lindley distribution

outperforms the exponential distribution when applied to modelling waiting times before bank customer service. Additionally, they highlighted that the contours of the hazard rate function for the Lindley distribution show an increasing trend, while the mean residual life function is a decreasing function of the random variable. Zakerzadeh and Dolati [16] and Nadarajah et al. [12] extended the Lindley distribution by introducing new parameters and evaluating the performance of these extended distributions using various datasets. Over the years, several authors have made contributions to modify the Lindley distribution. For instance, Merovci [11] introduced the transmuted Lindley distribution and discussed its various properties. Sharma et al. [14, 15] introduced the inverse of the Lindley distribution and examined its unique characteristics. Shanker et al. [13] developed a novel lifetime distribution called the Akash distribution, which demonstrated superior performance compared to both the exponential and Lindley distributions. Ahmad et al. [1] introduced the transmuted inverse Lindley distributions and conducted analyses of their properties, Ahmad et al [2, 3], Bhaumik, D. K. et al. [5], Flaih, A et al. [6]. Each of these distributions comes with its own set of advantages and limitations when applied to analyzing complex data.

In this paper, the authors aim to introduce a new two-parameter distribution that offers greater flexibility and improved results compared to existing distributions. The probability density function of this newly established two-parameter distribution is as follows

$$f(y,\alpha,\beta)=\frac{\beta^2}{\alpha\beta+\beta^2+1}\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}\quad ;y>0,\alpha,\beta>0 \tag{2.1}$$

The proposed distribution is named as Sabur distribution which is a combination of two distributions, Exponential distribution having scale parameter β and gamma distribution having shape parameter 3 with scale parameter β.With combining proportion as $\frac{\beta(\alpha+\beta)}{\alpha\beta+\beta^2+1}$

$$f(y,\alpha,\beta)=\pi\phi_1(y,\beta)+(1-\pi)\phi_2(y,\beta)$$

Where $\pi=\frac{\beta(\alpha+\beta)}{\alpha\beta+\beta^2+1}$

$$\phi_1(y,\beta)=\beta e^{-\beta y}\,,\qquad \phi_2(y,3,\beta)=\frac{\beta^3}{2}y^2e^{-\beta y}$$

The cumulative distribution function of (1.1) is given as

$$F(y,\alpha,\beta)=1-\left[1+\frac{\beta^2y^2+2\beta y}{2(\alpha\beta+\beta^2+1)}\right]e^{-\beta y}\,;y>0,\alpha,\beta>0 \tag{2.1}$$

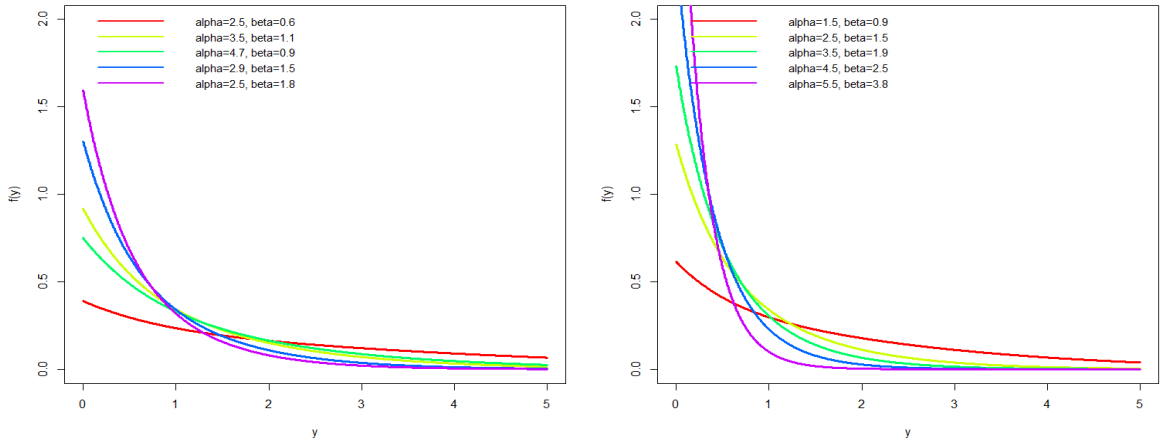


Fig 1: *The graph of p.d.f of SD for different values of parameters.*

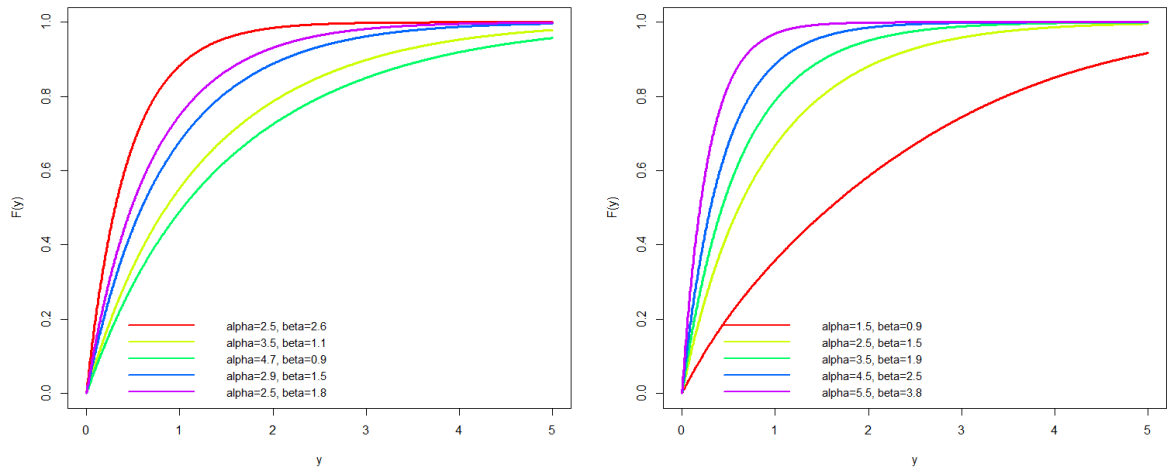


Fig 2: *The graph of c.d.f of SD for different values of parameters.*

2. Statistical Properties

In this section different properties of the Sabur distribution has been discussed such as moments, moment generating function, mode and median.

2.1 Moments of Sabur Distribution

Let us consider Y be a random variable follows the Sabur distribution then the r^{th} moment of the distribution denoted by μ_r' is given as

$$\mu_r' = E(Y^r) = \int_0^\infty y^r f(y,\alpha,\beta)dy$$
$$= \int_0^\infty y^r \frac{\beta^2}{\alpha\beta+\beta^2+1}\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}\ dy$$
$$= \frac{\beta^2}{\alpha\beta+\beta^2+1}\int_0^\infty \left\{(\alpha+\beta)y^r+\frac{\beta}{2}\ y^{r+2}\right\}e^{-\beta y}\ dy$$
$$= \frac{\beta^2}{\alpha\beta+\beta^2+1}\left[\frac{(\alpha+\beta)\Gamma(r+1)}{\beta^{r+1}}+\frac{\Gamma(r+3)}{2\beta^{r+2}}\right]$$

Substituting $r = 1,2,3,4$, we obtain first four moments of the distribution about origin.

$$\mu_1' = \frac{\alpha\beta+\beta^2+3}{\beta(\alpha\beta+\beta^2+1)}\ ,\mu_2' = \frac{2(\alpha\beta+\beta^2+6)}{\beta^2(\alpha\beta+\beta^2+1)}$$
$$\mu_3' = \frac{6(\alpha\beta+\beta^2+10)}{\beta^3(\alpha\beta+\beta^2+1)}\ ,\mu_4' = \frac{24(\alpha\beta+\beta^2+15)}{\beta^4(\alpha\beta+\beta^2+1)}$$

Therefore, the mean and variance of Sabur distribution is given as

$$\mu = E(Y) = \frac{\alpha\beta+\beta^2+3}{\beta(\alpha\beta+\beta^2+1)}$$

The central moments of Sabur distribution can be obtained by using above raw moments

$$\mu_2 = \sigma^2 = \mu_2' - (\mu_1')^2 = \frac{\alpha^2\beta^2+2\alpha\beta^3+8\alpha\beta+\beta^4+8\beta^2+3}{\beta^2(\alpha\beta+\beta^2+1)^2}$$

$$\mu_3 = \frac{2\beta^6+60\beta^4+6\alpha\beta^5+4\alpha^2\beta^4+2\alpha(\alpha^2+\alpha+59)\beta^3+2(30\alpha^2+39)\beta^2+78\alpha\beta+36}{\beta^3(\alpha\beta+\beta^2+1)^3}$$

$$\mu_4 = \frac{\left\{\begin{array}{c} 9\beta^8 + 222\beta^6 - 24\alpha\beta^8 - 12\alpha\beta^7(4\alpha+3) - 6\alpha^2\beta^6(4\alpha-13) \\ -6\alpha\beta^5(2\alpha^2 - 33\alpha - 95) + \alpha^2\beta^4(2\alpha^2 + 24\alpha + 894) + \alpha\beta^3(534\alpha^2 - 144\alpha + 852) \\ +\beta^2(702\alpha + 690) + 812\alpha\beta + 495 \end{array}\right\}}{\beta^4(\alpha\beta + \beta^2 + 1)^4}$$

Coefficient of variation (C.V)= $\frac{\sigma}{\mu_1'} = \frac{\sqrt{\alpha^2\beta^2+2\alpha\beta^3+8\alpha\beta+\beta^4+8\beta^2+3}}{\beta(\alpha\beta+\beta^2+1)(\alpha\beta+\beta^2+3)}$

Coefficient of skewness ($\sqrt{\beta_1}$)= $\frac{\mu_3}{(\mu_2)^{3/2}} = \frac{2\beta^6+60\beta^4+6\alpha\beta^5+4\alpha^2\beta^4+2\alpha(\alpha^2+\alpha+59)\beta^3+2(30\alpha^2+39)\beta^2+78\alpha\beta+36}{(\alpha^2\beta^2+2\alpha\beta^3+8\alpha\beta+\beta^4+8\beta^2+3)^{3/2}}$

Coefficient of kurtosis (β_2) = $\frac{\mu_4}{(\mu_2)^2} = \frac{\left\{\begin{array}{c} 9\beta^8+222\beta^6-24\alpha\beta^8-12\alpha\beta^7(4\alpha+3)-6\alpha^2\beta^6(4\alpha-13) \\ -6\alpha\beta^5(2\alpha^2-33\alpha-95)+\alpha^2\beta^4(2\alpha^2+24\alpha+894)+\alpha\beta^3(534\alpha^2-144\alpha+852) \\ +\beta^2(702\alpha+690)+812\alpha\beta+495 \end{array}\right\}}{(\alpha^2\beta^2+2\alpha\beta^3+8\alpha\beta+\beta^4+8\beta^2+3)}$

Index of dispersion(γ) = $\frac{\sigma^2}{\mu_1'} = \frac{\alpha^2\beta^2+2\alpha\beta^3+8\alpha\beta+\beta^4+8\beta^2+3}{\beta(\alpha\beta+\beta^2+1)(\alpha\beta+\beta^2+3)}$

2.2. Moment Generating Function of Sabur Distribution

Let us consider Y be a random variable follows the Sabur distribution then moment generating function of the distribution denoted by $M_Y(t)$ is given as

$$\begin{aligned} M_Y(t) = E(e^{ty}) &= \int_0^\infty e^{ty} f(y,\alpha,\beta)dy \\ &= \frac{\beta^2}{\alpha\beta+\beta^2+1}\int_0^\infty \left(\alpha+\beta+\frac{\beta}{2}y^2\right) e^{-(\beta-t)y}dy \\ &= \frac{\beta^2}{\alpha\beta+\beta^2+1}\left\{\frac{(\alpha+\beta)}{(\beta-t)}+\frac{\beta}{(\beta-t)^3}\right\} \\ &= \frac{1}{\alpha\beta+\beta^2+1}\left\{(\alpha\beta+\beta^2)\sum_{k=0}^{\infty}\left(\frac{t}{\beta}\right)^k + \sum_{k=0}^{\infty}\binom{k+2}{k}\left(\frac{t}{\beta}\right)^k\right\} \\ &= \sum_{k=0}^{\infty}\frac{2(\alpha\beta+\beta^2)+(k+1)(k+2)}{2(\alpha\beta+\beta^2+1)}\left(\frac{t}{\beta}\right)^k \end{aligned}$$

2.3. Mode and Median of Sabur Distribution

The value or number in a data set, which are occurring repeatedly may be termed as mode while median is the middle value or number in a data set arranged in ascending order.

Taking logarithm to the pdf of Sabur distribution,we get

$$\log f(y,\alpha,\beta) = 2\log\beta - \log(\alpha\beta+\beta^2+1) + \log\left(\alpha+\beta+\frac{\beta}{2}y^2\right) - \beta y$$

Differentiate w.r.t y ,we get

$$\frac{\partial \log f(y,\alpha,\beta)}{\partial y} = \frac{\beta y}{(\alpha\beta+\beta^2+1)} - \beta$$

Equating $\frac{\partial \log f(y,\alpha,\beta)}{\partial y} = 0$, we get

$$\frac{\beta y}{(\alpha\beta+\beta^2+1)} - \beta = 0 \Rightarrow y = \frac{1 \pm \sqrt{1-2\beta(\alpha+\beta)}}{\beta}$$

$$M_0 = y_0 = \frac{1 \pm \sqrt{1-2\beta(\alpha+\beta)}}{\beta}$$

Using the empirical formula, we obtain median as

$$\begin{aligned} Median &= \frac{1}{3}M_0 + \frac{2}{3}\mu \\ &= \frac{1 \pm \sqrt{1-2\beta(\alpha+\beta)}}{3\beta} + \frac{2(\alpha\beta+\beta^2+3)}{3\beta(\alpha\beta+\beta^2+1)} \end{aligned}$$

3. Reliability Measures

Suppose Y be a continuous random variable with cdf (y) , $y \geq 0$ .then its reliability function which is also called survival function is defined as

$$S(y) = p_r(Y > y) = \int_0^{\infty} f(y)\, dy = 1 - F(y)$$

Therefore, the survival function is given

$$S(y, \alpha, \beta) = 1 - F(y, \alpha, \beta) = \left[1 + \frac{\beta^2 y^2 + 2\beta y}{2(\alpha\beta + \beta^2 + 1)}\right] e^{-\beta y} \quad (3.1)$$

The hazard function of a random variable y is given as

$$H(y, \alpha, \beta) = \frac{f(y, \alpha, \beta)}{S(y, \alpha, \beta)} \quad (3.2)$$

Using equation (1.1) and equation (3.1) in (3.2), we get

$$H(y, \alpha, \beta) = \frac{2\beta^2 \left(\alpha + \beta + \frac{\beta}{2} y^2\right)}{[\beta^2 y^2 + 2\beta y + 2(\alpha\beta + \beta^2 + 1)]}$$

Also, the reverse hazard function denoted as $h_r(y, \alpha, \beta)$ can be obtained as

$$h_r(y, \alpha, \beta) = \frac{f(y, \alpha, \beta)}{F(y, \alpha, \beta)} \quad (3.3)$$

Using (1.1) and (1.2) in equation (3.3), we get

$$h_r(y, \alpha) = \frac{2\beta^2 \left(\alpha + \beta + \frac{\beta}{2} y^2\right) e^{-\beta y}}{2(\alpha\beta + \beta^2 + 1) - [\beta^2 y^2 + 2\beta y + 2(\alpha\beta + \beta^2 + 1)] e^{-\beta y}}$$

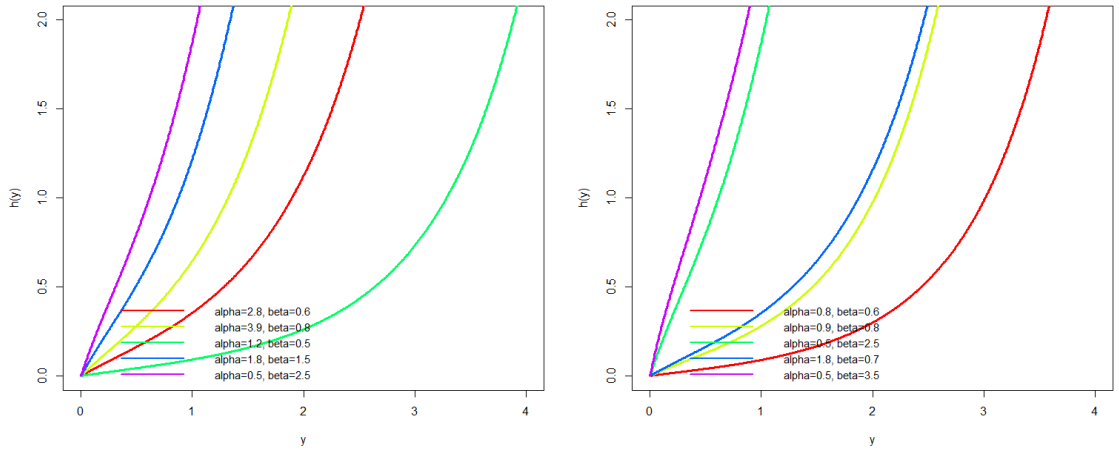


Fig 3: *The graph of hazard rate function of SD for different values of parameters.*

The mean residual function denoted by$m(y)$, and is defined as

$$m(y) = \frac{1}{1 - F(y)} \int_y^{\infty} 1 - F(z)\, dz$$

Therefore, the mean residual function of Sabur distribution is given by

$$m(y) = \frac{\beta^2 y^2 + 4y\beta + 2(\alpha\beta + \beta^2 + 3)}{\beta[\beta^2 y^2 + 2y\beta + 2(\alpha\beta + \beta^2 + 1)]}$$

We observe that $H(0) = f(0) = \frac{\alpha\beta^2}{\alpha\beta + \beta^2 + 1}$ and $m(0) = \mu = \frac{\alpha\beta + \beta^2 + 3}{\beta(\alpha\beta + \beta^2 + 1)}$

4. Order Statistics of Sabur Distribution

Let us consider $Y_1, Y_2 \ldots Y_n$ be random sample of sample size n from sabur distribution with pdf (1.1) and cdf (1.2).Then the pdf of k^{th} order statistics is given by

$$f_{Y_{(k)}}(y) = \frac{n!}{(k-1)!\,(n-k)!}\,[F(y)]^{k-1}\,[1-F(y)]^{n-k}\,f(y) \quad , k = 1,2,3,\ldots,n \quad (4.1)$$

Now substituting the equation (1.1) and (1.2) in equation (4.1), we obtain the k^{th} order statistics as

$$f_{Y_{(k)}}(y) = \frac{n!\,\beta^2\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}}{(k-1)!\,(n-k)!\;(\alpha\beta+\beta^2+1)}\left\{1-\left[1+\frac{\beta^2y^2+2\beta y}{2(\alpha\beta+\beta^2+1)}\right]e^{-\beta y}\right\}^{k-1} \left\{\left[1+\frac{\beta^2y^2+2\beta y}{2(\alpha\beta+\beta^2+1)}\right]e^{-\beta y}\right\}^{n-k} \quad (4.2)$$

The pdf of first order statistics Y_1 is given as

$$f_{Y_{(1)}}(y) = \frac{n\beta^2\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}}{(\alpha\beta+\beta^2+1)}\left\{\left[1+\frac{\beta^2y^2+2\beta y}{2(\alpha\beta+\beta^2+1)}\right]e^{-\beta y}\right\}^{n-1}$$

And the pdf of n^{th} order statistics Y_n is given as

$$f_{Y_{(n)}}(y) = \frac{n\beta^2\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}}{(\alpha\beta+\beta^2+1)}\left\{1-\left[1+\frac{\beta^2y^2+2\beta y}{2(\alpha\beta+\beta^2+1)}\right]e^{-\beta y}\right\}^{n-1}$$

5. Renyi Entropy

If Y is a continuous random variable having probability density function $f(y,\alpha,\beta)$, then Renyi entropy is defined as

$$T_R(\delta) = \frac{1}{1-\delta}\,log\left\{\int_0^\infty f^\delta(y)dy\right\}$$

where $\delta > 0$ and $\delta \neq 1$

Thus, the Renyi entropy for Sabur distribution (1.1) , is given as

$$T_R(\delta) = \frac{1}{1-\delta}\,log\left\{\int_0^\infty\left[\frac{\beta^2}{\alpha\beta+\beta^2+1}\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}\right]^\delta dy\right\}$$

$$= \frac{1}{1-\delta}\,log\left\{\frac{\beta^{2\delta}(\alpha+\beta)^\delta}{(\alpha\beta+\beta^2+1)^\delta}\int_0^\infty\left(1+\frac{\beta}{2(\alpha+\beta)}y^2\right)^\delta e^{-\beta\delta y}dy\right\}$$

$$= \frac{1}{1-\delta}\,log\left\{\frac{\beta^{2\delta}(\alpha+\beta)^\delta}{(\alpha\beta+\beta^2+1)^\delta}\int_0^\infty\sum_{r=0}^\infty\binom{\delta}{r}\left(\frac{\beta}{2(\alpha+\beta)}y^2\right)^r e^{-\beta\delta y}dy\right\}$$

$$= \frac{1}{1-\delta}\,log\left\{\sum_{r=0}^\infty\binom{\delta}{r}\frac{\beta^{2\delta+r}(\alpha+\beta)^{\delta-r}}{2^r(\alpha\beta+\beta^2+1)^\delta}\int_0^\infty y^{2r}e^{-\beta\delta y}dy\right\}$$

$$= \frac{1}{1-\delta}\,log\left\{\sum_{r=0}^\infty\binom{\delta}{r}\frac{\beta^{2\delta+r}\,(\alpha+\beta)^{\delta-r}}{2^r(\alpha\beta+\beta^2+1)^\delta}\,\frac{\Gamma(2r+1)}{(\beta\delta)^{2r+1}}\right\}$$

$$= \frac{1}{1-\delta}\,log\left\{\sum_{r=0}^\infty\binom{\delta}{r}\frac{\beta^{2\delta-(r+1)}\,(\alpha+\beta)^{\delta-r}}{2^{r-1}(\alpha\beta+\beta^2+1)^\delta}\,\frac{\mathrm{r}\,\Gamma(2r)}{\delta^{2r+1}}\right\}$$

6. Mean Deviation from Mean of Sabur Distribution

The quantity of scattering in a population is evidently measured to some extent by the totality of the deviations. Let Y be a random variable from Sabur distribution with mean μ then the mean deviation from mean is defined as.

$$D(\mu) = E(|Y-\mu|) = \int_0^\infty |Y-\mu|\,f(y)\,dy$$

$$= \int_0^{\mu} (\mu - y)\, f(y)\, dy + \int_{\mu}^{\infty} (y - \mu)\, f(y) dy$$
$$= \mu \int_0^{\mu} f(y)\, dy - \int_0^{\mu} y\, f(y)\, dy + \int_{\mu}^{\infty} y f(y)\, dy - \int_{\mu}^{\infty} \mu f(y)\, dy$$
$$= \mu F(\mu) - \int_0^{\mu} y\, f(y)\, dy - \mu[1 - F(\mu)] + \int_{\mu}^{\infty} y f(y)\, dx$$
$$= 2\mu F(\mu) - 2\int_0^{\mu} y f(y)\, dy \tag{6.1}$$

Now

$$\int_0^{\mu} y f(y)\, dy = \frac{\beta^2}{\alpha\beta + \beta^2 + 1} \int_0^{\mu} y \left(\alpha + \beta + \frac{\beta}{2} y^2\right) e^{-\beta y}\, dy$$

After solving the integral, we get

$$\int_0^{\mu} y f(y)\, dy = \mu - \left\{\frac{\mu^3\beta^3 + 3\mu^2\beta^2 + 6\mu\beta + \beta(\alpha + \beta)(\mu\beta + 1) + 6}{2\beta(\alpha\beta + \beta^2 + 1)}\right\} e^{-\beta\mu} \tag{6.2}$$

Now substituting equation (6.2) in equation (6.1), we get

$$D(\mu) = \frac{\{\mu^2\beta^2 + 6\mu\beta + \beta^2(1 - \mu\alpha - \mu\beta) + \beta(\alpha - 2\mu) + 6\}}{\beta(\alpha\beta + \beta^2 + 1)}\, e^{-\beta\mu}$$

7. Mean Deviation from Median of Sabur Distribution

Let Y be a random variable from Sabur distribution with median M then the mean deviation from median is defined as.

$$D(M) = E(|Y - M|) = \int_0^{M} (M - y)\, dy + \int_{M}^{\infty} (y - M)\, dy$$
$$= MF(M) - \int_0^{M} y f(x)\, dy - M[1 - F(M)] + \int_{M}^{\infty} y f(y) dy$$
$$= \mu - 2\int_0^{M} y f(y)\, dy \tag{7.1}$$

Now

$$\int_0^{M} y f(y)\, dy = \frac{\beta^2}{\alpha\beta + \beta^2 + 1} \int_0^{M} y \left(\alpha + \beta + \frac{\beta}{2} y^2\right) e^{-\beta y}\, dy$$

After solving the integral, we get

$$\int_0^{M} y f(y)\, dy = \mu - \left\{\frac{M^3\beta^3 + 3M^2\beta^2 + 6M\beta + \beta(M + \beta)(M\beta + 1) + 6}{2\beta(\alpha\beta + \beta^2 + 1)}\right\} e^{-\beta M} \tag{7.2}$$

Now substituting equation (7.2) in equation (7.1), we get

$$D(M) = \left\{\frac{M^3\beta^3 + 3M^2\beta^2 + 6M\beta + \beta(M + \beta)(M\beta + 1) + 6}{\beta(\alpha\beta + \beta^2 + 1)}\right\} e^{-\beta M} - \mu$$

8. Bonferroni and Lorenz Curves

In economics the relation between poverty and economy is well studied by using Bonferroni and Lorenz curves. Besides that these curves have been used in different fields such as reliability, insurance and biomedicine.

The Bonferroni curve, $B(s)$is given as.

$$B(s) = \frac{1}{s\mu} \int_0^{t} y\, f(y)\, dy \tag{8.1}$$

Or

$$B(s) = \frac{1}{s\mu} \int_0^{s} F^{-1}(y) dy$$

And Lorenz curve, $L(s)$ is given as.

$$L(s) = \frac{1}{\mu}\int_0^t yf(y)\,dy \tag{8.2}$$

Or

$$L(s) = \frac{1}{\mu}\int_0^s F^{-1}(y)\,dy$$

Where $E(X) = \mu$ and $t = F^{-1}(s)$.

Now

$$\int_0^t yf(y)\,dy = \mu - \left\{\frac{t^3\beta^3 + 3t^2\beta^2 + 6t\beta + \beta(\alpha+\beta)(t\beta+1) + 6}{2\beta(\alpha\beta+\beta^2+1)}\right\}e^{-\beta t} \tag{8.3}$$

Substituting equation (8.3) in equations (8.1) and (8.2), we get

$$B(s) = \frac{1}{s}\left[1 - \left\{\frac{t^3\beta^3 + 3t^2\beta^2 + 6t\beta + \beta(\alpha+\beta)(t\beta+1) + 6}{2\beta(\alpha\beta+\beta^2+1)\mu}\right\}e^{-\beta t}\right]$$

And

$$L(s) = \left[1 - \left\{\frac{t^3\beta^3 + 3t^2\beta^2 + 6t\beta + \beta(\alpha+\beta)(t\beta+1) + 6}{2\beta(\alpha\beta+\beta^2+1)\mu}\right\}e^{-\beta t}\right]$$

9. Estimation of Parameters of Sabur Distribution

Suppose $Y_1, Y_2, Y_3, \ldots Y_n$ be random samples of size n from Sabur distribution. Then the likelihood function of Sabur distribution is given as.

$$\begin{aligned} l &= \prod_{i=1}^{n} f(y_i, \alpha.\beta) \\ &= \prod_{i=1}^{n}\left\{\frac{\beta^2}{\alpha\beta+\beta^2+1}\left(\alpha+\beta+\frac{\beta}{2}y^2\right)e^{-\beta y}\right\} \\ &= \left(\frac{\beta^2}{\alpha\beta+\beta^2+1}\right)^n \prod_{i=1}^{n}\left(\alpha+\beta+\frac{\beta}{2}y_i^2\right)e^{-\beta\sum_{i=1}^{n}y_i} \end{aligned}$$

The log likelihood function is given by

$$\log l = 2n\log\beta - n\log(\alpha\beta+\beta^2+1) + \sum_{i=1}^{n}\log\left(\alpha+\beta+\frac{\beta}{2}y_i^2\right) - \beta\sum_{i=1}^{n}y_i$$

Now, differentiating partially w. r. t parameters α and β respectively we get

$$\frac{\partial \log l}{\partial \alpha} = \frac{-n\beta}{\alpha\beta+\beta^2+1} + \sum_{i=1}^{n}\frac{1}{\alpha+\beta+\frac{\beta}{2}y_i^2}$$

$$\frac{\partial \log l}{\partial \beta} = \frac{2n}{\beta} - \frac{n(\alpha+2\beta)}{\alpha\beta+\beta^2+1} + \sum_{i=1}^{n}\frac{\left(1+\frac{y_i^2}{2}\right)}{\alpha+\beta+\frac{\beta}{2}y_i^2} - \sum_{i=1}^{n}y_i$$

Now solving $\frac{\partial \log l}{\partial \alpha} = 0$, $\frac{\partial \log l}{\partial \beta} = 0$, we get

$$\frac{-n\beta}{\alpha\beta+\beta^2+1} + \sum_{i=1}^{n}\frac{1}{\alpha+\beta+\frac{\beta}{2}y_i^2} = 0 \tag{9.1}$$

Also

$$\frac{2n}{\beta} - \frac{n(\alpha+2\beta)}{\alpha\beta+\beta^2+1} + \sum_{i=1}^{n}\frac{\left(1+\frac{y_i^2}{2}\right)}{\alpha+\beta+\frac{\beta}{2}y_i^2} - n\bar{y} = 0 \tag{9.2}$$

It is obvious that equations (9.1) and (9.2), are not in closed form, hence cannot be solved analytically for α and . In order to find the value of α and β it is imperative to apply iterative methods. The MLE of the parameters denoted as $\hat{\theta}(\hat{\alpha},\ \hat{\beta})$ of $\theta(\alpha,\beta)$ can be obtained by using Newton-Raphson method, bisection method, secant method etc.

Since the MLE of $\hat{\theta}$ follows asymptotically normal distribution which is given as

$$\sqrt{n}\left(\hat{\theta}-\theta\right) \rightarrow N\left(0, I^{-1}(\theta)\right) \tag{9.3}$$

Where $I^{-1}(\theta)$ is the limiting variance – covariance matrix of $\hat{\theta}$ and $I(\theta)$ is a 2×2 Fisher information matrix I,e

$$I(\theta) = -\frac{1}{n}\begin{bmatrix} E\left(\frac{\partial^2 \log l}{\partial^2 \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha\, \partial \beta}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \beta\, \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial^2 \beta}\right) \end{bmatrix}$$

Where

$$\frac{\partial^2 \log l}{\partial^2 \alpha} = \frac{n\beta^2}{(\alpha\beta+\beta^2+1)^2} - \sum_{i=1}^{n} \frac{1}{\left(\alpha+\beta+\frac{\beta}{2}y_i^2\right)^2}$$

$$\frac{\partial^2 \log l}{\partial^2 \beta} = \frac{-2n}{\beta^2} + \frac{n(\alpha^2+2\beta^2+2\alpha\beta)}{\alpha\beta+\beta^2+1} - \sum_{i=1}^{n} \frac{\left(1+\frac{y_i^2}{2}\right)^2}{\left(\alpha+\beta+\frac{\beta}{2}y_i^2\right)^2}$$

$$\frac{\partial^2 \log l}{\partial\alpha\partial\beta} = \frac{\partial^2 \log l}{\partial\beta\partial\alpha} = \frac{n(\beta^2-1)}{\alpha\beta+\beta^2+1} - \sum_{i=1}^{n} \frac{\left(1+\frac{y_i^2}{2}\right)}{\alpha+\beta+\frac{\beta}{2}y_i^2}$$

Hence the approximate $100(1-\psi)\%$ confidence interval for α and β are respectively given by

$$\hat{\alpha} \pm z_{\frac{\psi}{2}}\sqrt{I_{\alpha\alpha}^{-1}(\hat{\theta})} \;\; , \hat{\beta} \pm z_{\frac{\psi}{2}}\sqrt{I_{\beta\beta}^{-1}(\hat{\theta})}$$

Where $z_{\frac{\psi}{2}}$ is the ψ^{th} denotes percentile the standard distribution

10. Application

In this section, the importance and flexibility of the formulated distribution is illustrated by using a real life data set. And the distribution is compared with Lindley distribution (LD), Shanker Distribution (SHD), Exponential distribution (ED), inverse Lindley distribution (ILD) and Nadarajah-Haghighi distribution (HD). In order to compare the two distribution models, we consider the criteria like AIC (Akaike information criterion), CAIC (corrected Akaike information criterion) and BIC (Bayesian information criterion. The better distribution corresponds to lesser AIC, CAIC and BIC values.

Data: The data set is the strength data of glass of the aircraft window reported by Fuller *et al* [7]. The data are

18.83, 20.80, 21.657, 23.03, 23.23, 24.05, 24.321, 25.50, 25.52, 25.80, 26.69, 26.77, 26.78, 27.05, 27.67, 29.90, 31.11, 33.20, 33.73, 33.76, 33.89, 34.76, 35.75, 35.91, 36.98, 37.08, 37.09, 39.58, 44.045, 45.29, 45.381.

From above Table 1, it has been observed that the Sabur distribution have the lesser AIC, CAIC, -logL and BIC values. Hence we can conclude that Sabur distribution leads to a better fit as compared to Lindley distribution (LD), Shanker Distribution (SHD), Exponential distribution (ED), inverse Lindley distribution (ILD) and Nadarajah-Haghighi distribution (HD)

Table 1: *MLE's, - ln L, AIC, CAIC and BIC of the fitted distributions of data sets*

Model	Parameter Estimates	S.E	-log L	AIC	CAIC	BIC
Sabur Distribution	$\boldsymbol{\alpha = 0.3501}$ $\boldsymbol{\beta = 0.0967}$	$\boldsymbol{\alpha = 0.0124}$ $\boldsymbol{\beta = 0.0080}$	**120.44**	**244.89**	**245.32**	**247.76**
LD	$\alpha = 0.0630$	$\alpha = 0.0412$	127.0	256.0	256.1	257.4
SD	$\alpha = 0.0647$	$\alpha = 0.0475$	126.15	254.3	254.5	255.8
ED	$\alpha = 0.0355$	$\alpha = 0.1507$	137.25	276.7	276.8	277.9
ILD	$\alpha = 30.153$	$\alpha = 5.2523$	137.24	276.49	276.63	277.92
NHD	$\alpha = 0.0026$ $\beta = 0.0008$	$\alpha = 9.8991$ $\beta = 3.2255$	128.59	261.19	264.06	261.62

.

11. Concluding Remarks

In this paper, we introduce a new two-parameter lifetime distribution called the "Sabur distribution. We explore various mathematical properties of this distribution, including its shape, moments, hazard rate, mean residual life functions, mean deviations, and order statistics. Additionally, we derive expressions for the Bonferroni and Lorenz curves as well as the Renyi entropy measure for the proposed distribution. Furthermore, we discuss the method of maximum likelihood estimation for estimating the distribution's parameter. To demonstrate the practical utility and superiority of the Sabur distribution over existing alternatives such as the Shanker, Nadarajah-Haghighi, exponential, Lindley, and inverse Lindley distributions, we perform goodness-of-fit tests using criteria like the Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC) on real lifetime datasets.

References

1. Aijaz, A., Ahmad, A, Tripathi, R. (2020), Transmuted inverse Lindley distribution: Statistical properties Applications, Science, technology and development, 9(7), 1-10.
2. Aijaz. A, Jallal. M, Qurat Ul Ain.S, Tripathi .R (2020). The Hamza distribution with Statistical properties and Applications. Asian Journal of Probability and Statistics, 8(1):28-42.
3. Aijaz, A., Ahmad, A, Tripathi, R.(2020). Sauleh distribution with Statistical Properties and Applications. Annal Biostat and Biomed Appli. 4(1), 1-5
4. Bonferroni. C.E, (1930). Elementi di Statisca gnerarale, Seeber, Firenze.
5. Bhaumik, D. K., Kapur, K. and Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples, Technometrics, 51(3), 326-334.
6. Flaih, A., Elsalloukh, H., Mendi, E., Milanova, M., 2012. The exponentiated inverted Weibull distribution. Appl. Math. Inf. Sci 6 (2), 167–171.
7. Fuller, E.J., Frieman, S., Quinn, J., Quinn, G., and Carter, W. (1994): Fracture mechanics approach to the design of glass aircraft windows: A case study, SPIE Proc 2286, 419-430.
8. Ghitany, M.E., Atieh, B., Nadarajah, S., (2008). Lindley distribution and its applications, Mathematics Computing and Simulation, 78, 493-506.

9. Gross, A. J., Clark, V. A., (1975). Survival Distributions: Reliability Applications in the Biometrical Sciences, John Wiley, New York
10. Lindley, D, V. (1958). Fiducial distributions and Bayes' theorem, Journal of Royal Statistical Society, Series B, 20, 102-107.
11. Merovci, F., (2013) Transmuted Lindley distribution, International Journal of Open problems in Computer Science and Mathematics, 6, 63-72.
12. Nadarajah, S., BAKOUCH, H. S., TAHMASBI, R., (2011). A Generalized Lindley distribution, Sankhya Series B, 73, 331-359.
13. Shanker, R., (2015). Akash distribution and Its Application, International Journal of Probability and Statistics, 4 (3), 65-75.
14. Sharma, V., SINGH, S., SINGH, U., AGIWAL, V., (2015). The inverse Lindley distribution – a stress-strength reliability model with applications to head and neck cancer data, Journal of Industrial & Production Engineering, 32 (3), 1623-173.
15. Sharma, S, K, Singh, U., Sharma, V.K., (2014). The Truncated Lindley distribution-interference and Application, Journal of Statistics Applications & Probability, 3 (2), 219-228.
16. Zakerzadeh, H., Dolati, A., (2009). Generalized Lindley distribution, Journal of Mathematical extension, 3 (2), pp. 13-25.

COST & PROFIT ANALYSIS OF TWO-DIMENSIONAL STATE M/M/2 QUEUING MODEL WITH CORRELATED SERVERS, MULTIPLE VACATION, BALKING AND CATASTROPHES

Sharvan Kumar[1], Indra[2]

[1*, 2] Department of Statistics and O. R, Kurukshetra University, Kurukshetra, Haryana, India-136119

[1]kumarsharvan067@gmail.com , [2]indra@kuk.ac.in

Abstract

The present study obtains the time-dependent solution of a two-dimensional state Markovian queuing model with infinite capacity, correlated servers, multiple vacation, balking and catastrophes. Inter arrival times follow an exponential distribution with parameters $\boldsymbol{\lambda}$ *and service times follow Bivariate exponential distribution BVE (*$\boldsymbol{\mu}$, $\boldsymbol{\mu}$, $\boldsymbol{\nu}$*) where* $\boldsymbol{\mu}$ *is the service time parameter and* $\boldsymbol{\nu}$ *is the correlation parameter. Both the servers go on vacation with probability one when there are no units in the system and the servers keeps on taking a sequence of vacations of random length each time the system becomes empty, till it finds at least one unit in the system to start each busy period referred as multiple vacation. The unit finds a long queue and decides not to join it; may be considered as balking. All the units are ejected from the system when catastrophes occur and the system becomes temporarily unavailable. The system reactivates when new units arrive. Occurrence of catastrophes follow Poisson distribution with rate* ξ*. Laplace transform approach has been used to find the time-dependent solution. By using differential-difference equations, the recursive expressions for probabilities of exactly* $\boldsymbol{i}$ *arrivals and* $\boldsymbol{j}$ *departures by time* $\boldsymbol{t}$ *are obtained. The probabilities of this model are consistent to the results of "Pegden & Rosenshine". The model estimates the total expected cost, total expected profit and obtained the optimal values by varying time* $\boldsymbol{t}$ *for cost and profit. These important key measures give a greater understanding of the model behaviour. Numerical analysis and graphical representations have been done by using Maple software.*

Keywords: Correlated servers, Multiple vacation, Balking, Catastrophes

1. Introduction

A two-dimensional state model has been used to deal with complicated transient analysis of some queuing problem. This model is used to examine the queuing system for exact number of arrivals and departures by given time t. In case of a one-dimensional state model, it is difficult to determine how many units have entered, left or waiting units in the system, while the two-dimensional state model exactly identifies the numbers of units that have entered, left, or waiting in the system. The idea of two-dimensional state model for the M/M/1 queue was first given by Pegden & Rosenshine [4]. After that, the two-dimensional state model has attracted the attention of a lot of researchers.

A system of queues in series or in parallel should ordinarily be studied taking into account the interdependence of servers, but this leads to very complicated mathematics even in very simple case of systems. So to reduce such complications of analysis the servers are considered to be independent. But this independence of servers cause impact in time bound operations such as vehicle inspection counters, toll booths, large bars and cafeterias *etc*. where for efficient system functioning the correlation between the servers contributes significantly. Nishida *et al.* [3] investigated a two-server Markovian queue assuming the correlation between the servers and obtained steady-state results for a limited waiting space capacity of two units. Sharma [6] investigated the transient solution to this problem again using only two units waiting spaces capacity. Sharma and Maheswar [8] developed a computable matrix approach to study a correlated two-server Markovian queue with finite waiting space. They also derived waiting time distribution for steady-state and obtained the transient probabilities through steady-state by using a matrix approach and Laplace transform approach.

Various studies have been conducted to evaluate different performance measures to verify the robustness of the system in which a server takes a break for a random period of time *i.e.* vacation. When the server returns from a vacation and finds the empty queue, it immediately goes on another vacation *i.e.* multiple vacation and if it finds at least one waiting unit, then it will commence service according to the prevailing service policy. Different queuing systems with multiple vacation have been extensively investigated and effectively used in several fields including industries, computer & communication systems, telecommunication systems *etc.* Different types of vacation policies are available in literature such as single vacation, multiple vacation and working vacations. Researches on multiple vacation systems have grown tremendously in the last several years. Cooper [2] was the first to study the vacation model and determined the mean waiting time for a unit arrive at a queue served in cyclic order. Doshi [5] and Wu & Zhang [15] have done outstanding researches on queuing system with vacations and released some excellent surveys. Xu and Zhang [12] considered the Markovian multi-server queue with a single vacation (e, d)-policy. They also formulated the system as a quasi-birth-and-death process and computed the various stationary performance measures. Altman and Yechiali [13] studied the customer's impatience in queues with server vacations. Kalidaas *et al.* [18] obtained the time-dependent solution of a single server queue with multiple vacation. Ammar [19] analysed M/M/1 queue with impatient units and multiple vacation. Sharma and Indra [24] investigated the dynamic aspects of a two-dimensional state single server Markovian queuing system with multiple vacation and reneging. Gahlawat *et al*. [25] studied the time-dependent first in first out queuing model with a single intermittently available server and variable-sized bulk arrivals and bulk departures by using the Laplace transform and inverse transform approaches.

Queues with balking have numerous applications in everyday life. Balking occurs if units avoid joining the queue, when they perceive the queue to be too long. Long queues at cash counters, ticket booths, banks, barber shops, grocery stores, toll plaza *etc*. Kumar *et al.* [7] obtained the time-dependent solution of an M/M/1 queue with balking. Chauhan and Sharma [10] derived an expression of the probability distribution for the number of customers in the service station for the M/M/r queuing model with balking and reneging. Zhang and Yue [11] analysed the M/M/1/N queuing system with balking, reneging and server vacation. Sharma and Kumar [17] studied a single-server Markovian feedback queuing system with balking. Bouchentouf and Medjahari [23] presented a single-server feedback queuing system under two differentiated multiple vacation with balking and obtained steady-state probabilities for the model. They also derived some important performance measures, including the average number of customers in the system, the average number of customers in the queue, the average balking rate *etc*.

Queuing systems with catastrophes are also getting a lot of attention nowadays and may be used to solve a wide range of real-world problems. Catastrophes may occur at any time, resulting in the loss of units and the deactivation of the service centre, because they are totally unpredictable in

nature. Such type of queues with catastrophes plays an important role in computer programs, telecommunication and ticket counters *etc.* For example, virus or hacker attacking a computer system or program causing the system fail or become idle. Chao [9] obtained steady-state probability of the queue size and a product form solution of a queuing network system with catastrophes. Kumar *et al.* [14] obtained time-dependent solution for M/M/1 queuing system with catastrophes. Kalidass *et al.* [16] derived explicit closed form analytical expressions for the time-dependent probabilities of the system size. Dharamraja and Kumar [20] studied Markovian queuing system with heterogeneous servers and catastrophes. Chakravarthy [21] studied delayed catastrophic model in steady state using the matrix analytic method. Sampath and Liu [22] studied an M/M/1 queue with reneging, catastrophes, server failures and repairs using modified Bessel function, Laplace transform and probability generating function approach. Souza and Rodriguez [26] worked on fractional M/M/1 queue model with catastrophes.

With above concepts in mind, we analyse a two-dimensional state M/M/2 queuing model with correlated servers, multiple vacation, balking and catastrophes.

Consider a situation in a company, where two colleagues work independently on the same project *i.e.* they are not able to share the information of project with each other and not helping each other. Then it will take a long time to complete the project, some information will be lost due to communication gap and the results obtained are not much reliable. But if both of them work together (interdependent servers) *i.e.* they will share all information of project and help each other, then there are more chances that it will take less time and results obtained will be more reliable. Hence interdependent servers are more reliable than the independent servers. When servers work interdependently then they termed as correlated servers. After project completion, the colleagues may take a break, when they find there is no further work, considered as vacation. During the project, if someone wants to work with these colleagues on different project but due to their busy schedule decides not to join them; it may be considered as balking. If due to disease or any other reason the colleagues are not working this may be considered as catastrophes.

The present paper has been structured as follows. In section 1 introduction and in section 2 the model assumptions, notations and description are given. In section 3 the differential-difference equations to find out the time-dependent solution are given and section 4 describes important performance measures. Section 5 investigates the total expected cost function and total expected profit function for the given queuing system. In section 6, we present the numerical results in the form of tables and section 7 contains the tables and graphs to illustrate the impact of various factors on performance measures. The last section contains discussion on the findings and suggestions for further work.

2. Model Assumptions, Notations and Description

- Arrivals follow Poisson distribution with parameter $\boldsymbol{\lambda}$.
- There are two servers and the service times follow Bivariate exponential distribution BVE* ($\boldsymbol{\mu}$, $\boldsymbol{\mu}$, $\boldsymbol{\nu}$) where $\boldsymbol{\mu}$ is the service time parameter and $\boldsymbol{\nu}$ is the correlation parameter.
- The vacation time of the server follows an exponential distribution with parameter $\boldsymbol{w}$.
- On arrival a unit either decides to join the queue with probability $\boldsymbol{\beta}$ or not to join the queue with probability $\boldsymbol{1-\beta}$.
- Occurrence of catastrophes follows Poisson distribution with parameter ξ.
- Various stochastic processes involved in the system are statistically independent of each other.

  *introduced by Marshall and Olkin [1]

Initially, the system starts with zero units and the server is on vacation, *i.e.*

$$P_{0,0,V}(0)=1 \quad ; \quad P_{0,0,B}(0)=0 \tag{1}$$

$$\delta_{i,j}=\begin{cases}1 & ;for \quad i=j\\ 0 & ;for \quad i\neq j\end{cases} \quad ; \quad \sum_{i}^{j}=0 \quad for \quad j<i \tag{2}$$

The Two-Dimensional State Model

$P_{i,j,V}(t)$ = The probability that there are exactly ***i*** arrivals and ***j*** departures by time ***t*** and the server is on vacation.

$P_{i,j,B}(t)$ = The probability that there are exactly ***i*** arrivals and ***j*** departures by time ***t*** and the server is busy in relation to the queue.

$P_{i,j}(t)$ = The probability that there are exactly ***i*** arrivals and ***j*** departures by time ***t***.

3. The Differential-Difference Equations for the Queuing Model under Study

$$\frac{d}{dt}P_{i,i,V}(t)=-\lambda\beta P_{i,i,V}(t)+(\mu+\nu)P_{i,i\text{-}1,B}(t)(1-\delta_{i,0})+\nu P_{i,i\text{-}2,B}(t)(1-\delta_{i,0}-\delta_{i,1})+\xi(1-P_{i,i,V}(t)) \quad i\geq 0 \tag{3}$$

$$\frac{d}{dt}P_{i+1,i,B}(t)=-(\lambda\beta+\mu+\nu+\xi)P_{i+1,i,B}(t)+2\mu P_{i+1,i\text{-}1,B}(t)(1-\delta_{i,0})+\nu P_{i+1,i\text{-}2,B}(t)(1-\delta_{i,0}-\delta_{i,1})+wP_{i+1,i,V}(t) \quad i\geq 0 \tag{4}$$

$$\frac{d}{dt}P_{i,j,V}(t)=-(\lambda\beta+w+\xi)P_{i,j,V}(t)+\lambda\beta P_{i\text{-}1,j,V}(t) \quad i>j\geq 0 \tag{5}$$

$$\frac{d}{dt}P_{i,j,B}(t)=-(\lambda\beta+2\mu+\nu+\xi)P_{i,j,B}(t)+\lambda\beta P_{i\text{-}1,j,B}(t)(1-\delta_{i\text{-}1,j})+2\mu P_{i,j\text{-}1,B}(t)(1-\delta_{j,0}) \quad i>j+1 \tag{6}$$
$$+\nu P_{i,j\text{-}2,B}(t)+wP_{i,j,V}(t)$$

The preceding equations (3) to (6) are solved by taking the Laplace transforms together with initial condition

$$\overline{P}_{0,0,V}(s)=\frac{(\xi+s)}{s(s+\lambda\beta+\xi)} \tag{7}$$

$$\overline{P}_{i,0,V}(s)=\frac{(\lambda\beta)^{i}(\xi+s)}{s(s+\lambda\beta+\xi)(s+\lambda\beta+w+\xi)^{i}} \quad i>0 \tag{8}$$

$$\overline{P}_{i,i,V}(s)=\left(\frac{\mu+\nu}{s+\lambda\beta+\xi}\right)P_{j,j\text{-}1,B}(s)+\left(\frac{\nu}{s+\lambda\beta+\xi}\right)P_{j,j\text{-}2,B}(s) \quad i>0 \tag{9}$$

$$\overline{P}_{i,0,B}(s)=\frac{w(\lambda\beta)^{i}(\xi+s)}{s(s+\lambda\beta+\xi)(s+\lambda\beta+w+\xi)(s+\lambda\beta+\mu+\nu+\xi)(s+\lambda\beta+2\mu+\nu+\xi)^{i-1}} + w(\lambda\beta)^{i}\sum_{m=1}^{i-1}\frac{(\xi+s)}{s(s+\lambda\beta+\xi)(s+\lambda\beta+w+\xi)^{m+1}(s+\lambda\beta+2\mu+\nu+\xi)^{i-m}} \quad i\geq 1 \quad (10)$$

$$\overline{P}_{i+1,i,B}(s)=\left(\frac{2\mu}{s+\lambda\beta+\mu+\nu+\xi}\right)P_{i+1,i-1,B}(s)+\left(\frac{\nu}{s+\lambda\beta+\mu+\nu+\xi}\right)P_{i+1,i-2,B}(s)+\left(\frac{(\mu+\nu)w\lambda\beta}{(s+\lambda\beta+\xi)(s+\lambda\beta+w+\xi)(s+\lambda\beta+\mu+\nu+\xi)}\right)P_{i-1,i-1,B}(s) \quad i>0 \quad (11)$$

$$\overline{P}_{i,j,V}(s)=\left(\frac{\mu+\nu}{s+\lambda\beta+\xi}\right)\left(\frac{\lambda\beta}{s+\lambda\beta+w+\xi}\right)^{i-j}P_{j,j-1,B}(s)+\left(\frac{\lambda\beta}{s+\lambda\beta+w+\xi}\right)^{i-j}\left(\frac{\nu}{s+\lambda\beta+\xi}\right)P_{j,j-2,B}(s) \quad i>j\geq 0 \quad (12)$$

$$\overline{P}_{i,j,B}(s)=\left(\frac{\lambda\beta}{s+\lambda\beta+2\mu+\nu+\xi}\right)P_{i-1,j,B}(s)+\left(\frac{2\mu}{s+\lambda\beta+2\mu+\nu+\xi}\right)P_{i,j-1,B}(s)+\left(\frac{\nu}{s+\lambda\beta+2\mu+\nu+\xi}\right)P_{i,j-2,B}(s) + \left(\frac{\mu+\nu}{s+\lambda\beta+\xi}\right)\left(\frac{w}{s+\lambda\beta+2\mu+\nu+\xi}\right)\left(\frac{\lambda\beta}{s+\lambda\beta+w+\xi}\right)^{i-j}\left[P_{j,j-1,B}(s)+P_{j,j-2,B}(s)\right] \quad i>j+1,\ j>0 \quad (13)$$

It is seen that

$$\sum_{i=0}^{\infty}\sum_{j=0}^{i}\left[\overline{P}_{i,j,V}(s)+\overline{P}_{i,j,B}(s)(1-\delta_{i,j})\right]=\frac{1}{s} \quad (14)$$

and hence

$$\sum_{i=0}^{\infty}\sum_{j=0}^{i}\left[P_{i,j,V}(t)+P_{i,j,B}(t)(1-\delta_{i,j})\right]=1 \quad (15)$$

a verification.

4. Performance Measures

(i) The Laplace transform of $P_{i.}(t)$ of the probability that exactly ***i*** units arrive by time ***t***; when initially there are no unit in the system is given by

$$\overline{P}_{i.}(s)=\sum_{j=0}^{i}\left[\overline{P}_{i,j,V}(s)+\overline{P}_{i,j,B}(s)(1-\delta_{i,j})\right]=\sum_{j=0}^{i}\overline{P}_{i,j}(s)=\frac{(\lambda\beta)^{i}}{(s+\lambda\beta)^{i+1}}$$

And its inverse Laplace transform is: $P_{i.}(t)=\frac{e^{-\lambda\beta t}(\lambda\beta t)^{i}}{i!}$ (16)

The arrivals follow a Poisson distribution as the probability of the total number of arrivals is not affected by vacation time of the server.

(ii) $P_{.j}(t)$ is the probability that exactly ***j*** units have been served by time ***t.*** In terms of $P_{i,j}(t)$ we have

$$P_{.j}(t)=\sum_{i=j}^{\infty}P_{i,j}(t) \quad (17)$$

(iii) The Laplace transform of mean number of arrival is: $\sum_{i=0}^{\infty} i\overline{P}_{i.}(s) = \frac{\lambda\beta}{(s^2)}$ (18)

And its inverse Laplace transform is: $\sum_{i=0}^{\infty} iP_{i.}(t) = \lambda\beta t$ (19)

(iv)The mean number of units in the queue is calculated as follows

$$Q_L(t) = \sum_{N=0}^{\infty} NP_V(t) + \sum_{N=1}^{\infty} (N-1)P_B(t) \quad (20)$$

Where $N = i–j$.

5. Cost Function and Profit Function

For the given queuing system, the following notations have been used to represent various costs to find out the total expected cost and total expected profit per unit time:

Let

C_H: Cost per unit time for unit in the queue.

C_B: Cost per unit time for a busy server.

C_μ: Cost of service per unit time.

C_V: Cost per unit time when the server is on vacation.

If I is the total expected amount of income generated by delivering a service per unit time then

(i) Total expected cost per unit at time t is given by

$$TC(t) = C_H * Q_L(t) + C_B * P_B(t) + C_V * P_V(t) + \mu * C_\mu \quad (21)$$

(ii) Total expected income per unit at time t is given by

$$TE_I(t) = I * \mu * (1 - P_V(t)) = I * \mu * P_B(t) \quad (22)$$

(iii) Total expected profit per unit at time t is given by

$$TE_P(t) = TE_I(t) - TC(t) \quad (23)$$

6. Numerical Results

6.1. Numerical Validity Check

(i) For the state when the server is on vacation

$$P_V(t) = \sum_{j=0}^{i} P_{i,j,V}(t) \quad (24)$$

(ii) For the state when the server is busy in relation to the queue

$$P_B(t) = \sum_{j=0}^{i-1} P_{i,j,B}(t) \quad (25)$$

(iii) The probability $P_{i.}(t)$ that exactly i units arrive by time t is

$$\sum_{j=0}^{i} P_{i,j}(t) = \sum_{j=0}^{i} P_{i,j,V}(t) + \sum_{j=0}^{i-1} P_{i,j,B}(t) \quad (26)$$

(iv) A numerical validity check of inversion of $\overline{P}_{i,j}(s)$ is based on the relationship

$$\Pr\{\boldsymbol{i} \text{ arrivals in } (0, \boldsymbol{t})\} = \frac{e^{-\lambda\beta t}(\lambda\beta t)^i}{i!} = \sum_{j=0}^{\infty} P_{i,j}(t) = P_{i.}(t) \tag{27}$$

The probabilities of this model shown in last column of table 1 given below are consistent to the last column of "Pegden & Rosenshine" [4] by keeping constant values of w*=1,* ξ*=0,* β*=1 and* v*=0.25 shown in table*

Table-1: *Numerical validity check of inversion of* $\overline{P}_{i,j}(s)$

λ	μ	t	i	$\frac{e^{-\lambda t} * (\lambda t)^i}{i!}$	$\sum_{j=0}^{i} P_{i,j,V}(t)$	$\sum_{j=0}^{i-1} P_{i,j,B}(t)$	$\sum_{j=0}^{i} P_{i,j}(t)$
1	2	3	1	0.149361	0.129196	0.020165	0.149361
1	2	3	3	0.224041	0.158076	0.065965	0.224041
1	2	3	5	0.100818	0.057803	0.043016	0.100818
2	2	3	1	0.014873	0.012865	0.002008	0.014873
2	2	3	3	0.089235	0.062961	0.026274	0.089235
2	2	3	5	0.160623	0.092090	0.068533	0.160623
1	2	4	1	0.073263	0.065390	0.007873	0.073263
1	2	4	3	0.195367	0.148001	0.047366	0.195367
1	2	4	5	0.156294	0.100998	0.055296	0.156294
2	2	4	1	0.002683	0.002395	0.000288	0.002683
2	2	4	3	0.028626	0.021686	0.006940	0.028626
2	2	4	5	0.091604	0.059195	0.032409	0.091604
2	4	4	5	0.091604	0.073396	0.018208	0.091604
1	2	4	4	0.195367	0.136810	0.058557	0.195367
1	2	3	6	0.050409	0.025824	0.024585	0.050409

7. Sensitivity Analysis

This part focuses on the impact of the arrival rate (λ), service rate (μ), vacation rate (w), correlation parameter (v), balking probability (1-β) and catastrophes rate (ξ) on the probability when the server is on vacation ($P_V(t)$), probability when the server is busy ($P_B(t)$), expected queue length ($Q_L(t)$), total expected cost ($TC(t)$), total expected income ($TE_I(t)$) and total expected profit ($TE_P(t)$) at time t. To determine the numerical results for the sensitivity of the queuing system one parameter varied while keeping all the other parameters fixed and taking cost per unit time for unit in the queue=10, cost per unit time for a busy server=8, cost per unit time when the server is on vacation=5, cost of service per unit time=4, total expected amount of income=100 and number of units in the system=8.

7.1. Impact of Arrival Rate (λ)

We examine the behaviour of the queuing system using measures of effectiveness along with cost and profit analysis by varying arrival parameter with time, while keeping all other parameters fixed; μ=5, w=3, ν=0.25, ξ=0.0001 and β=1. In table 2, we observe that as the value of λ increases with time t, $P_B(t)$, $Q_L(t)$, $TC(t)$, $TE_I(t)$ and $TE_P(t)$ increases but $P_V(t)$ decreases.

Table-2: *Measures of Effectiveness versus* λ

t	λ	$P_V(t)$	$P_B(t)$	$Q_L(t)$	$TC(t)$	$TE_I(t)$	$TE_P(t)$
1	1.00	0.8550827	0.1449161	0.4618549	30.0532913	72.45805	42.4047587
2		0.8451239	0.1546386	0.4824289	30.2870173	77.31930	47.0322827
3		0.8427343	0.1534636	0.4779723	30.2211033	76.73180	46.5106967
4		0.8414541	0.1511897	0.4597185	30.0139731	75.59485	45.5808769
5		0.8403975	0.1485374	0.4210111	29.6003977	74.26870	44.6683023
1	1.10	0.8427120	0.1572855	0.5062882	30.5347260	78.64275	48.1080240
2		0.8321542	0.1673763	0.5277639	30.7774204	83.68815	52.9107296
3		0.8279867	0.1661032	0.5192708	30.6614671	83.05160	52.3901329
4		0.8278007	0.1634085	0.4891622	30.3378935	81.70425	51.3663565
5		0.8264318	0.1599687	0.4312130	29.7240386	79.98435	50.2603114
1	1.20	0.8306335	0.1693615	0.5505522	31.0135815	84.68075	53.6671685
2		0.8194239	0.1797142	0.5725572	31.2604051	89.85710	58.5966949
3		0.8165607	0.1777708	0.5579035	31.0840049	88.88540	57.8013951
4		0.8147421	0.1744598	0.5117878	30.5872669	87.22990	56.6426331
5		0.8130834	0.1702170	0.4412499	29.8396520	85.10850	55.2688480

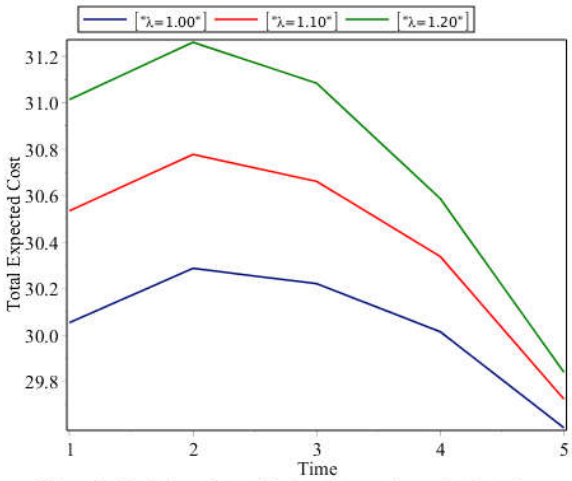


Figure 1: *Variation of cost with time t by varying arrival rate λ*

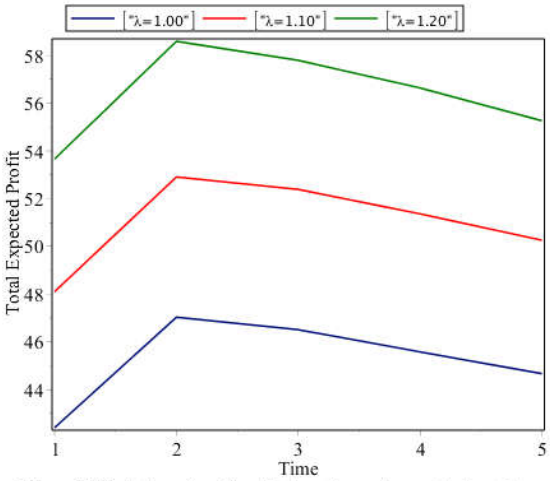


Figure 2: *Variation of profit with time t by varying arrival rate λ*

Figure 1 shows the variation of cost with time by varying arrival rate while keeping the other parameters fixed.

Figure 2 shows the variation of profit with time by varying arrival rate while keeping the other parameters fixed.

7.2. Impact of Service Rate (μ)

The behaviour of the queuing system measures of effectiveness along with cost and profit analysis by varying μ with time t, while keeping all other parameters fixed; λ=1, w=3, ν=0.25, ξ=0.0001 and β=1. In table 3, we observe that as the value of μ increases with time t, $P_B(t)$, $Q_L(t)$, $TC(t)$, $TE_I(t)$ and $TE_P(t)$ increases but $P_V(t)$ decreases.

Table-3: *Measures of Effectiveness versus* μ

t	μ	$P_V(t)$	$P_B(t)$	$Q_L(t)$	$TC(t)$	$TE_I(t)$	$TE_P(t)$
1	2.00	0.7369325	0.2630663	0.6177736	19.9669289	52.613260	32.6463311
2		0.6762899	0.3234726	0.7090209	21.0594393	64.694520	43.6350807
3		0.6699756	0.3262223	0.7084844	21.0445004	65.244460	44.1999596
4		0.6667285	0.3239153	0.6797388	20.7273529	64.783060	44.0557071
5		0.6649823	0.3189526	0.6203355	20.0798873	63.790520	43.7106327
1	2.50	0.7655438	0.2344550	0.5762758	21.4661170	58.613750	37.1476330
2		0.7243328	0.2754297	0.6356142	22.1812436	68.857425	46.6761814
3		0.7211500	0.2750479	0.6303259	22.1093922	68.761975	46.6525828
4		0.7177601	0.2728837	0.6050702	21.8275721	68.220925	46.3933529
5		0.7134529	0.2684820	0.5531662	21.2567825	67.120500	45.8637175
1	3.00	0.7897314	0.2102673	0.5431578	23.0623734	63.080190	40.0178166
2		0.7609462	0.2388163	0.5847948	23.5632094	71.644890	48.0816806
3		0.7587053	0.2374926	0.5788371	23.4818383	71.247780	47.7659417
4		0.7561697	0.2354741	0.5561210	23.2258513	70.642230	47.4163787
5		0.7501976	0.2317373	0.5088841	22.6937274	69.521190	46.8274626

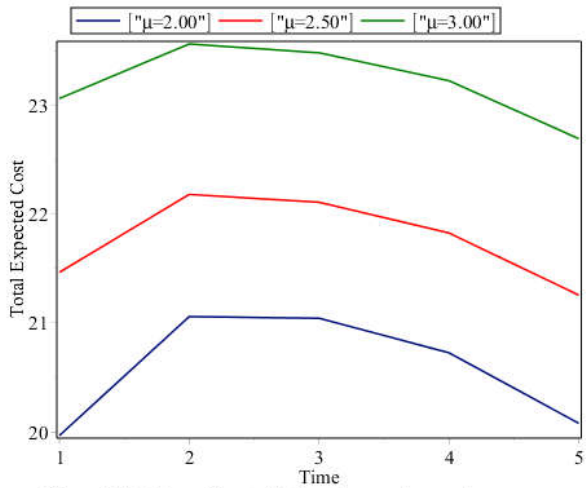


Figure 3:*Variation of cost with time t by varying service rate* μ

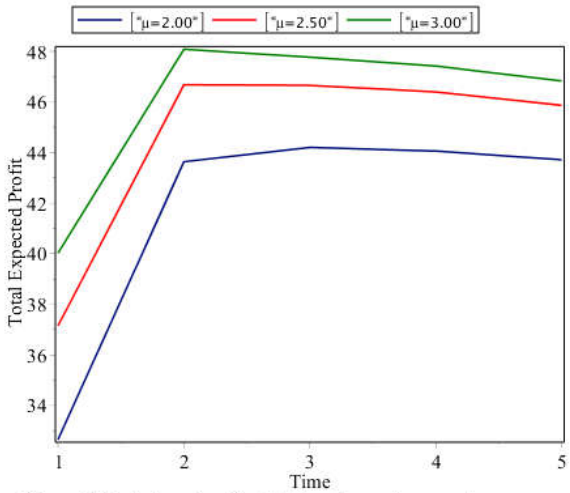


Figure 4: *Variation of profit with time t by varying service rate* μ

Figure 3 shows the variation of cost with time by varying service rate while keeping the other parameters fixed**.**

Figure 4 shows the variation of profit with time by varying service rate while keeping the other parameters fixed.

7.3. Impact of Vacation Rate (w)

We observe that the behaviour of the queuing system measures of effectiveness along with cost and profit by varying w with time t, while keeping all other parameters fixed; λ=1, μ=5, ν=0.25, ξ=0.0001 and β=1. In table 4, we observe that as the value of w increases with time t, $P_B(t)$, $Q_L(t)$, $TC(t)$, $TE_I(t)$ and $TE_P(t)$ increases but $P_V(t)$ decreases.

Table-4: *Measures of Effectiveness versus w*

t	w	$P_V(t)$	$P_B(t)$	$Q_L(t)$	$TC(t)$	$TE_I(t)$	$TE_P(t)$
1	3.00	0.8550827	0.1449161	0.4618549	30.0532913	72.45805	42.4047587
2		0.8451239	0.1546386	0.4824289	30.2870173	77.31930	47.0322827
3		0.8427343	0.1534636	0.4779723	30.2211033	76.73180	46.5106967
4		0.8414541	0.1511897	0.4597185	30.0139731	75.59485	45.5808769
5		0.8403975	0.1495374	0.4210111	29.6083977	74.76870	45.1603023
1	4.00	0.8462857	0.1537131	0.4015245	29.4763783	76.85655	47.3801717
2		0.8411568	0.1586057	0.4083021	29.5576506	79.30285	49.7451994
3		0.8389988	0.1571991	0.4040524	29.4931108	78.59955	49.1064392
4		0.8378927	0.1547511	0.3886032	29.3135043	77.37555	48.0620457
5		0.8321931	0.1517418	0.3558749	28.9336489	75.87090	46.9372511
1	5.00	0.8415912	0.1584076	0.3613974	29.0891908	79.20380	50.1146092
2		0.8386494	0.1611131	0.3640154	29.1223058	80.55655	51.4342442
3		0.8365297	0.1596682	0.3602046	29.0620401	79.83410	50.7720599
4		0.8345421	0.1575017	0.3464072	28.8967961	78.75085	49.8540539
5		0.8300883	0.1548467	0.3171915	28.5611301	77.42335	48.8622199

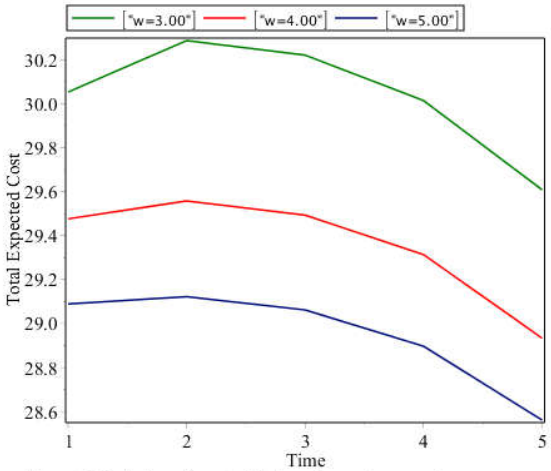


Figure 5: *Variation of cost with time t by varying vacation rate w*

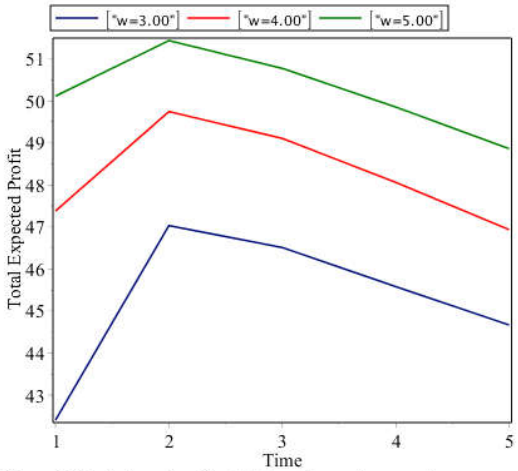


Figure 6: *Variation of profit with time t by varying vacation rate w*

Figure 5 shows the variation of cost with time by varying vacation rate while keeping the other parameters fixed.

Figure 6 shows the variation of profit with time by varying vacation rate while keeping the other parameters fixed.

7.4. Impact of Correlation Parameter (ν)

We see that the behaviour of the queuing system using measures of effectiveness along with cost and profit analysis by varying ν with time t, while keeping all other parameters fixed; λ=1, μ=5, w=3, ξ=0.0001 and β=1. In table 5, we observe that as the value of ν increases with time t, $P_B(t)$, $Q_L(t)$, $TC(t)$, $TE_I(t)$ and $TE_P(t)$ increases but $P_V(t)$ decreases.

Table-5: *Measures of Effectiveness versus* ν

t	ν	$P_V(t)$	$P_B(t)$	$Q_L(t)$	$TC(t)$	$TE_I(t)$	$TE_P(t)$
1	0.25	0.8550827	0.1449161	0.4618549	30.0532913	72.45805	42.4047587
2		0.8451239	0.1546386	0.4824289	30.2870173	77.31930	47.0322827
3		0.8427343	0.1534636	0.4779723	30.2211033	76.73180	46.5106967
4		0.8414541	0.1511897	0.4597185	30.0139731	75.59485	45.5808769
5		0.8403975	0.1475374	0.4210111	29.5923977	73.76870	44.1763023
1	0.50	0.8612902	0.1387085	0.4555353	29.9714720	69.35425	39.3827780
2		0.8522410	0.1475215	0.4756205	30.1975820	73.76075	43.5631680
3		0.8497855	0.1464124	0.4712878	30.1331047	73.20620	43.0730953
4		0.8472639	0.1433799	0.4532707	29.9160657	71.68995	41.7738843
5		0.8456738	0.1392611	0.4150683	29.4931408	69.63055	40.1374092
1	0.75	0.8670184	0.1329803	0.4497140	29.8960744	66.49015	36.5940756
2		0.8587324	0.1410301	0.4694068	30.1159708	70.51505	40.3990792
3		0.8562189	0.1399790	0.4651801	30.0527275	69.98950	39.9367725
4		0.8544773	0.1361665	0.4473789	29.8355075	68.08325	38.2477425
5		0.8524003	0.1315346	0.4096382	29.4106603	65.76730	36.3566397

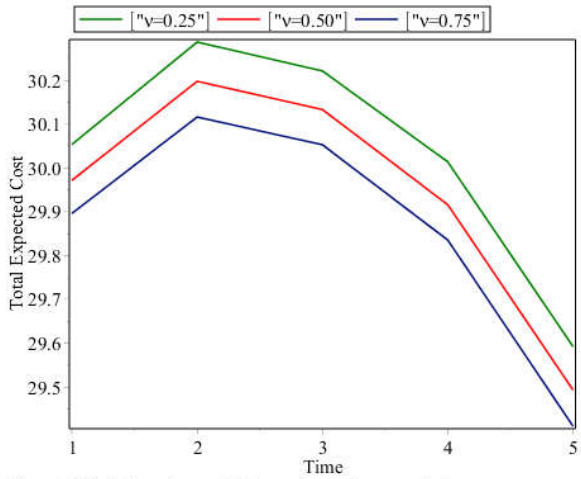


Figure 7:*Variation of cost with time t by varying correlation parameter v*

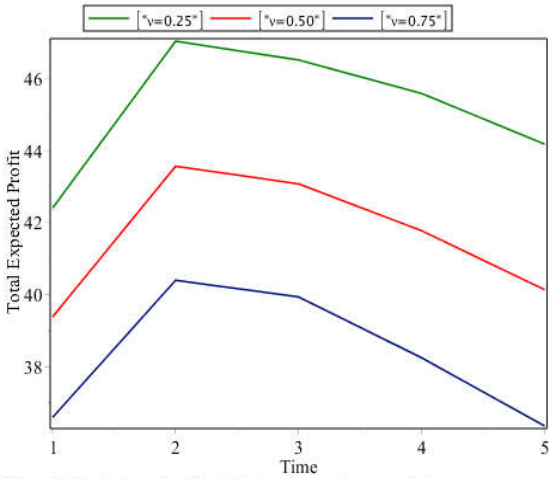


Figure 8: *Variation of profit with time t by varying correlation parameter v*

Figure 7 shows the variation of cost with time by varying correlation parameter while keeping the other parameters fixed**.**

Figure 8 shows the variation of profit with time by varying correlation parameter while keeping the other parameters fixed.

7.5. Impact of Joining Probability (β)

We see that the behaviour of the queuing system measures of effectiveness along with cost and profit analysis by varying β with time t. While keeping all other parameters fixed; λ=1, μ=5, w=3, ν=0.25 and ξ=0.0001. In table 6, we observe that as the value of β increases with time t, $P_B(t)$, $Q_L(t)$, $TC(t)$, $TE_I(t)$ and $TE_P(t)$ increases but $P_V(t)$ decreases.

Table-6: *Measures of Effectiveness versus β*

t	β	$P_V(t)$	$P_B(t)$	$Q_L(t)$	$TC(t)$	$TE_I(t)$	$TE_P(t)$
1	0.65	0.8876633	0.0990264	0.2868039	28.0985667	49.51320	21.4146333
2		0.8840590	0.1068084	0.3068008	28.3427702	53.40420	25.0614298
3		0.8880948	0.1069225	0.3132331	28.4281850	53.46125	25.0330650
4		0.8895493	0.1065419	0.3154762	28.4548437	53.27095	24.8161063
5		0.8871414	0.1053500	0.3133391	28.4118980	52.67500	24.2631020
1	0.75	0.8767085	0.1125818	0.3349279	28.6334759	56.29090	27.6574241
2		0.8720983	0.1210185	0.3563601	28.8922405	60.50925	31.6170095
3		0.8751685	0.1210047	0.3615710	28.9595901	60.50235	31.5427599
4		0.8746624	0.1200356	0.3610913	28.9445098	60.01780	31.0732902
5		0.8732121	0.1183004	0.3533446	28.8459097	59.15020	30.3042903
1	0.85	0.8671069	0.1257705	0.3844829	29.1865275	62.88525	33.6987225
2		0.8608738	0.1347799	0.4064078	29.4466862	67.38995	37.9432638
3		0.8623335	0.1344946	0.4090881	29.4785053	67.24730	37.7687947
4		0.8615314	0.1334039	0.4037567	29.4124552	66.70195	37.2894948
5		0.8598611	0.1310590	0.3868133	29.2159105	65.52950	36.3135895

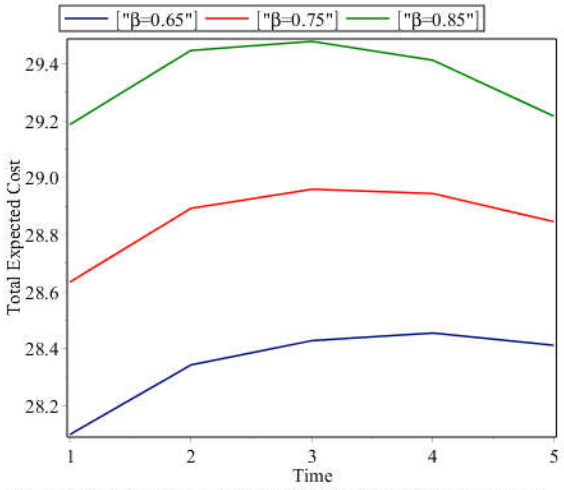


Figure 9:*Variation of cost with time t by varying joining probability β*

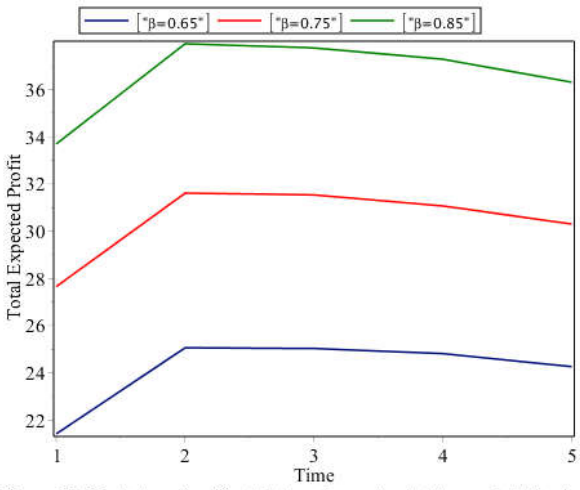


Figure 10: *Variation of profit with time t by varying joining probability β*

Figure 9 shows the variation of cost with time by varying joining probability while keeping the other parameters fixed.

Figure 10 shows the variation of profit with time by varying joining probability while keeping the other parameters fixed.

7.6. Impact of Catastrophes Rate (ξ)

We see that the behaviour of the queuing system measures of effectiveness along with cost and profit analysis by varying ξ with time t, while keeping all other parameters fixed; λ=1, μ=5, w=3, ν=0.25 and β=1. In table 7, we observe that as the value of ξ increases with time t, $P_B(t)$, $Q_L(t)$, $TC(t)$, $TE_I(t)$ and $TE_P(t)$ increases but $P_V(t)$ decreases.

Table-7: *Measures of Effectiveness versus* ξ

t	ξ	$P_V(t)$	$P_B(t)$	$Q_L(t)$	$TC(t)$	$TE_I(t)$	$TE_P(t)$
1	0.0001	0.8550827	0.1449161	0.4618549	30.0532913	72.45805	42.4047587
2		0.8451239	0.1546386	0.4824289	30.2870173	77.31930	47.0322827
3		0.8427343	0.1534636	0.4779723	30.2211033	76.73180	46.5106967
4		0.8414541	0.1491897	0.4597185	29.9979731	74.59485	44.5968769
5		0.8403975	0.1445374	0.4210111	29.5683977	72.26870	42.7003023
1	0.0002	0.8550885	0.1449103	0.4618404	30.0531289	72.45515	42.4020211
2		0.8451309	0.1546316	0.4824119	30.2868263	77.31580	47.0289737
3		0.8427420	0.1534569	0.4779565	30.2209302	76.72845	46.5075198
4		0.8414661	0.1491850	0.4597091	29.9979015	74.59250	44.5945985
5		0.8404256	0.1445380	0.4210193	29.5686250	72.26900	42.7003750
1	0.0003	0.8550943	0.1449045	0.4618259	30.0529665	72.45225	42.3992835
2		0.8451380	0.1546246	0.4823948	30.2866348	77.31230	47.0256652
3		0.8427496	0.1534503	0.4779406	30.2207564	76.72515	46.5043936
4		0.8414782	0.1491802	0.4596998	29.9978306	74.59010	44.5922694
5		0.8404537	0.1445385	0.4210276	29.5688525	72.26925	42.7003975

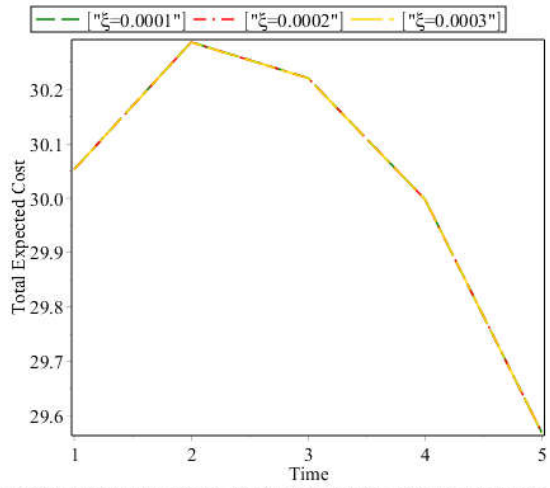


Figure 11:*Variation of cost with time t by varying catastrophes rate* ξ

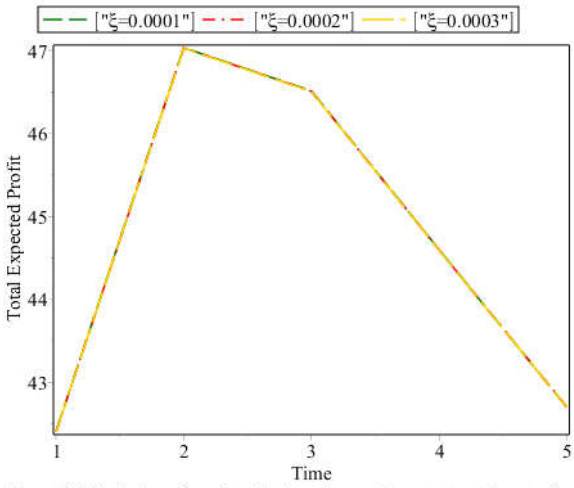


Figure 12: *Variation of profit with time t by varying catastrophes rate* ξ

Figure 11 shows the variation of cost with time by varying catastrophes rate while keeping the other parameters fixed.

Figure 12 shows the variation of profit with time by varying catastrophes rate while keeping the other parameters fixed.

8. Discussion

Figure 1 & figure 2 show the variation of cost & profit respectively with time t by varying λ(=1.00, 1.10, 1.20). The value of both cost & profit increases with increase in t upto t(=2.00) then decreases slightly. Hence we get the optimal value at t=5 when λ=1.00 and t=2 when λ=1.20 for minimum cost and maximum profit respectively.

Figure 3 show the variation of cost with time t by varying μ(=2.00, 2.50, 3.00). The value of cost increases with increase in t upto t(=2.00) then decreases slightly. The variation in profit with time t represented in figure 4 by varying μ(=2.00, 2.50, 3.00). The profit increases with increase in time up to (i) t=3 when μ=2.00 (ii) t=2 when μ=2.50, 3.00 respectively then decreases slightly. Hence we get the optimal value at t=1 when μ=2.00 and t=2 when μ=3.00 for minimum cost and maximum profit respectively.

Figure 5 & figure 6 show the variation of cost & profit respectively with time t by varying w(=3.00, 4.00, 5.00). The value of both cost & profit increases with increase in t upto t(=2.00) then decreases slightly. Hence we get the optimal value at t=5 when w=5.00 and t=2 when w=5.00 for minimum cost and maximum profit respectively.

Figure 7 & figure 8 show the variation of cost & profit respectively with time t by varying v(=0.25, 0.50, 0.75). The value of both cost & profit increases with increase in t upto t(=2.00) then decreases slightly. Hence we get the optimal value at t=5 when v=0.75 and t=2 when v=0.25 for minimum cost and maximum profit respectively.

Figure 9 show the variation of cost with time t by varying β(=0.65, 0.75, 0.85). The value of cost increases with increase in t upto (i) t=4 when β=0.65 (ii) t=3.00 when β=0.75, 0.85 respectively then decreases slightly. The variation in profit with time t represented in figure 10 by varying β(=0.65, 0.75, 0.85). The profit increases with increase in time up to t(=2.00) then decreases slightly. Hence we get the optimal value at t=1 when β=0.65 and t=2 when β=0.85 for minimum cost and maximum profit respectively.

Figure 11 & figure 12 show the variation of cost & profit respectively with time t by varying ξ(=0.0001, 0.0002, 0.0003). The value of both cost & profit increases with increase in t upto t(=2.00) then decreases slightly. Hence we get the optimal value at t=5 when ξ=0.0001 and t=2 when ξ=0.0001 for minimum cost and maximum profit respectively. Finally, the variation in rate of catastrophes shows the minor effect on cost and profit.

9. Conclusions and Future Work

The time-dependent solution, for the two-dimensional state M/M/2 queuing model with correlated servers, multiple vacation, balking and catastrophes, has been obtained. The model estimates the total expected cost and total expected profit, the best optimal value is at t=1 when service rate(=2.00) and t=2 when arrival rate(=1.20) for minimum cost and maximum profit respectively. These key measures give a greater understanding of model behaviour. Finally, the numerical analysis clearly demonstrates the meaningful impact of the correlated servers and multiple vacation on the system performances. This model finds its applications in communication networks, computer networks, supermarkets, hospital administrations, financial sector and many others.

As part of future study, this model may be examined further for Non-Markovian queues, bulk queues, tandem queues *etc*.

10. Acknowledgement

First and foremost, I am deeply indebted to my supervisor, Prof. (Dr.) Indra Rani for her unwavering guidance and expertise. Her insightful feedback and constant encouragement played a pivotal role in shaping this research. I am grateful to the Department of Statistics and Operational Research, Kurukshetra University, Kurukshetra for providing the resources and research facilities necessary for this study. Additionally, I would like to acknowledge the Kurukshetra University for providing the University research scholarship which made this research possible.

References

[1] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal American Statistical Association*, 62(317):30-40.

[2] Cooper, R. B. (1970). Queues served in cyclic order waiting times. *Bell System Technical Journal*, 49(3):399-413.

[3] Nishida, T., Watanabe, R. and Tahara, A. (1974). Poisson queue with correlated two servers. *Technology Reports of Osaka University*, 24:1193.

[4] Pegden, C. D. and Rosenshine, M. (1982). Some new results for the M/M/1 queue. *Management Science*, 28(7):821-828.

[5] Doshi, B.T. (1986). Queuing systems with vacations- a survey. *Queuing Systems*, 1(1):29-66.

[6] Sharma, O.P. (1990). Markovian queues. *Elis Horwood*, ltd, England.

[7] Kumar, B. K., Parthasarathy, P. R. and Sharafali, M. (1993). Transient solution of an M/M/1 queue with balking. *Queuing Systems*, 13(4):441-448.

[8] Sharma, O. P., and Maheswar, M. V. R. (1994). Analysis of M| M| 2| N queue with correlated servers. *Optimization*, 30(4):379-385.

[9] Chao, X. (1995). A queuing network system with catastrophes and product form solution. *Operations Research Letters*, 18(2):75-79.

[10] Chauhan, M. S., and Sharma, G. C. (1996). Profit analysis of M/M/R queueing model with balking and reneging. *Monte Carlo Methods and Applications*, 2(2):139-144.

[11] Zhang, Y., Yue, D. and Yue, W. (2005). Analysis of an M/M/1/N queue with balking, reneging and server vacations. *In Proceedings of the 5th International Symposium on OR and its Applications*, 63(8):37-47.

[12] Xu, X. and Zhang, Z.G. (2006). Analysis of multi-server queue with a single vacation (e, d)-policy. *Performance Evaluation*, 63(8):825-838.

[13] Altman, E. and Yechiali, U. (2006). Analysis of customers' impatience in queues with server vacations. *Queueing systems*, 52(4):261-279.

[14] Kumar, B., Krishnamoorthy, A., Madheswari, S. and Basha, S. (2007). Transient analysis of a single server queue with catastrophes, failures and repairs. *Queueing Systems*, 56(3):133-141.

[15] Wu, J.C. and Zhang, Z. G. (2010). Recent developments in vacation queuing models: a short survey. *International Journal of Operations Research*, 7(4):3-8.

[16] Kalidass, K., Gopinath, S., Gnanaraj, J. and Ramanath, K. (2012). Time dependent analysis of an M/M/1/N queue with catastrophes and a repairable server. *Opsearch*, 49(1):39-61.

[17] Sharma, S. K. and Kumar, R. (2012). A markovian feedback queue with retention of reneged customers and balking. *Advanced Modeling and Optimization*, 14(3):681-688.

[18] Kalidaas, K., Gnanaraj, J., Gopinath, S. and Kasturi, R. (2014). Transient analysis of an M/M/1 queue with a repairable server and multiple vacations. *International Journal of Mathematics in Operational Research*, 6(2):193-216.

[19] Ammar, S. I. (2015). Transient analysis of an M/M/1 queue with impatient behaviour and multiple vacations. *Applied Mathematics and Computation*, 260:97-105.

[20] Dharamraja, S. and Kumar, R. (2015). Transient solution of a markovian queuing system with heterogeneous servers and catastrophes. *Opsearch*, 52(4):810-826.

[21] Chakravarthy, S. R. (2017). A catastrophic queuing system with delayed action. *Applied Mathematical System modelling*, 46:631-649.

[22] Suranga Sampath, M. I. G. and Liu, J. (2018). Transient analysis of an M/M/1 queue with reneging, catastrophes, server failures and repairs. *Bulletin of the Iranian Mathematical Society*, 44(3):585-603.

[23] Bouchentouf, A. A. and Medjahri, L. (2019). Performance and economic evaluation of differentiated multiple vacation queueing system with feedback and balked customers. *Applications and Applied Mathematics: An International Journal*, 14(1):46-62.

[24] Sharma, R. and Indra (2020). Dynamic aspect of two-dimensional single server markovian queueing model with multiple vacations and reneging. *In Journal of Physics: Conference Series*, 1531(1):012060.

[25] Gahlawat, V. K., Rahul, K., Mor, R. S., & Malik, M. (2021) A two state time-dependent bulk queue model with intermittently available server. *Reliability: Theory & Applications*, 16(2(64)):114-122.

[26] De Oliveira Souza, M. and Rodriguez, P. M. (2021). On a fractional queueing model with catastrophes. *Applied Mathematics and Computation*, 410:126468.

INNOVATIVE METHODS OF ENSURING THE FUNCTIONAL SAFETY OF TRAIN CONTROL SYSTEMS

I.B. Shubinsky[1], E.N. Rozenberg[2], H. Schäbe[3]

[1]DsC, prof., NIIAS, Moscow, Russia, Igor-shubinsky@yandex.ru
[2]DsC, prof., NIIAS, Moscow, Russia, s,lavruhina@vniias.ru
[3]Doctor of Natural Sciences, TÜV Rheinland InterTraffic, Cologne, Germany, dr.hendrik.schaebe@gmail.com

Abstract

The paper examines the specificity of the modern intelligent control systems. The Big Data technology and Data Science algorithms open up great potential in train traffic management based on hazard prevention. An example is given of high reliability and acceptable accuracy of hazardous railway infrastructure failure prediction using methods based on artificial intelligence. A great deal of attention is given to economical methods of ensuring the required levels of functional safety of train control systems. For that purpose, the efficiency of the digital twin-based method was evaluated. It is shown that, under certain conditions, this method allows significantly reducing the cost of a control system while achieving an acceptably high level of functional safety. The method of virtual second channels is based on the same principle of using information redundancy rather than hardware redundancy. The paper presents and analyses the method of virtual second channels in respect to an axle counter-based train control system. It is established that it is possible to ensure a safety integrity level of an entire control system with a virtual second channel at least as high as SIL3. The above methods ensure, on the one hand, a reduction of the amount of equipment and significantly lower cost of the systems and, on the other hand, requires the creation of additional software and substantiation of the acceptability of the achieved level of functional safety. This matter is within the competence of the developer of the control system.

Keywords: Control system, functional safety, artificial intelligence, safety case, wrong-side failure prediction, digital twin, virtual channels.

1. Introduction

Amid the ongoing digital transformation, the development of modern computer-based control and management systems in railway transportation involves an accelerated deployment of a whole number of innovative solutions and a wide use of off-the-shelf products, which eventually makes systems more complex and opens up a number of opportunities in terms of automation, including automatic train operation, and can affect the functional safety indicators. In order to maintain these parameters at a specified level and to minimize the impact of human factor, the railway community is increasingly using formal methods and automated means of engineering, diagnostics and monitoring at all stages of a system's life cycle. A major factor of safety and dependability is the standardisation of a system's architecture, interfaces, open-source design and testing software, including the standardisation of approaches to remote lab testing of products by different manufacturers to prove the reliability of operation at the boundaries of systems of various

manufacturers [1]. In the railway industry, all of the above caused the development of an intelligent system for adaptive automated train traffic management (IMS). Figure 1 shows the basic structure diagram of the IMS that is designed for the purpose of managing train traffic.

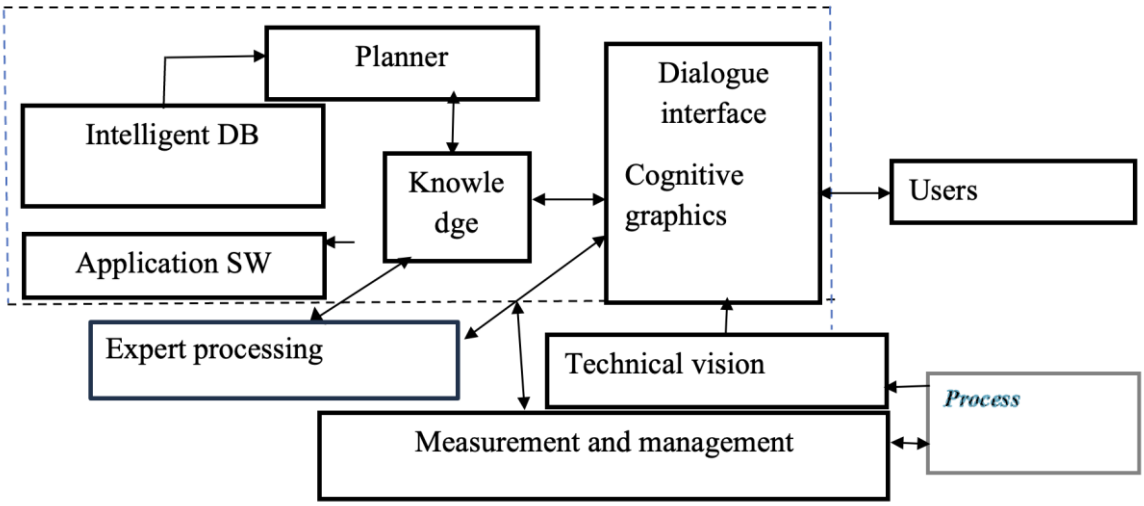


Fig. 1. Structure diagram of the intelligent automated management system

The generic part of the intelligent automated management system (IMS) (in Fig. 1, shown with a dashed line) contains knowledge and data bases, a planner, application software and a dialogue interface. The structure of this part of the system is static, is practically not exposed to external disturbances, operates with algorithms with time-invariant parameters, etc. Automatic train operation is implemented using the IMS that, along with the generic part, contains elaborate machine vision, adaptive measurement, control and cognitive graphics facilities. The machine vision and cognitive graphics are implemented partially through machine learning of neural networks. Artificial intelligence methods play an important role in recognising complex objects or under difficult circumstances. All of that as a whole defines the following specific features of the IMS [2]:

1. Branching system architecture.
2. Availability of machine vision and effect of weather conditions.
3. Close information interaction between the system and the environment via the information communication channels.
4. Presence of a large and not always definite number of vulnerabilities within a system closely connected to the environment.
5. A high probability of evolving environmental effects and resulting changed system behaviour.
6. The random nature of the control algorithm parameters as the result of neural network training using the incoming information flows and accumulated databases.
7. Branching software of both the generic part of the system, and, especially, rolling stock detection and control facilities.

It should be noted that one of the key features of IMS is that, along the branching architecture, the connections within the system change significantly. The latter noticeably reduces the capability to prove the safety of the intelligent system.

The use of accumulated knowledge and generation on new knowledge originated as far back as in

the 1980's. The main goal was to generate expert evaluations for the purpose of solving a number of applied problems, primarily those associated with control/management. The development of computer technology allowed using Big Data in railway transportation for the purpose of collecting and processing massive amounts of raw data [3-5], etc. For the purpose of working with such data, mathematical statistics and Data Science-based machine learning methods are used, which allows generating prediction models and eventually optimally solving the tasks at hand.

Activities aimed at increasing the intellectual level of railway transportation facilities condition analytics using Big Data-based methods [6 – 9, 11 – 16], etc. cover a number of tasks, such as predictive maintenance of the Lastochka EMUs, route optimisation, passenger flow management, locomotive operation and repair, prediction of infrastructure facility condition. The accumulated experience of automation of data collection and management, application of Data Science algorithms allows not only predicting objective events, but defining the development strategies of data collection systems. It is the initial data that ultimately define the correctness of the prediction results. Moreover, as of late more often than not the nature, scope and discreteness of initial data define the exact problem that can be solved. For such objects as the Lastochka EMUs and Sapsan high-speed trains that feature a large number of condition monitoring sensors, the uncertainty of an object's condition between the data acquisition periods is minimised. That allows employing a number of data analysis methods, solving various problems from the evaluation of the probability of an event to a train's up time.

The above capabilities allow solving totally new problems as part of functional safety analysis and preparation of safety cases of such innovative systems. This paper provides a number of examples of successful application of modern methods.

2. Predicting wrong-side failures using artificial intelligence

Information on infrastructure facilities of the Russian Railways lacks temporal continuity as the majority of data collection processes is not automated and is performed manually. As regards track facilities, automated systems for diagnostics and remote monitoring of track condition are widely used, which enables predictive analysis with large volumes of initial data. For each of the three Russian railways selected for the research, predictive analysis models were constructed based on the following classification algorithms: XGboost, RForest, SVM, kNN, AdaBoost and Logit. Those algorithms differ in terms of their conceptual approach and mathematical content, which allows matching an optimal algorithm to each data sample. For each model, target hyperparameters were selected for subsequent optimization as part of cross-validation.

Accuracy analysis of each of the above machine learning algorithms, as well as comparison of the accuracy of various binary classification algorithms were conducted based on the following measure:

TP, number of correctly predicted category objects with mark "1".

Note: Mark "1" denotes the onset of a wrong-side failure.

FN, number of objects with true category "1", yet predicted "0".

Note: Mark "0" denotes the operable state (absence of hazardous failure).

FP, number of objects with true category "0", yet predicted "1".

TN, number of correctly predicted categories with mark "0".

The primary quality measures of the binary classification models are:

– Overall accuracy of the algorithm that defines the classifier's general efficiency in terms of its ability to provide correct answers:

$$AC = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

– False alarm that shows the efficiency of a classifier for the purpose of predicting deviations from the normal state:

$$FPR = \frac{FP}{FP + TN} \tag{2}$$

– Accuracy of the algorithm that shows the correctly predicted proportion of objects recognized as objects with mark "1":

$$PR = \frac{TP}{TP + FP} \tag{3}$$

– Completeness of the algorithm that shows the correctly predicted proportion of objects that are effectively marked "1":

$$RE = \frac{TP}{TP + FN} \tag{4}$$

– The *F* measure of the algorithm is the harmonic mean of accuracy and completeness:

$$F = \frac{2PR \cdot RE}{PR + RE} \tag{5}$$

– Area under the ROC curve (AUC) is a global quality characteristic whose values lays between 0 and 1. Value 0.5 corresponds to random guessing, while value 1 corresponds to faultless recognition. *AUC is the area under the ROC curve. The ROC curve characterises the ratio between the proportion of false positive classifications (FPR) and the proportion of correct positive classifications (RE). The ROC curve is a sufficiently complex measure of an algorithm's accuracy*.

Table 1 shows the results of the comparison of classification methods per key accuracy indicators *PR* (the more the better) and *FPR* (the less the better). For the purpose of objective model comparison, the probability *threshold* for all methods was set to 0.1. The comparison was done for all the railways over the test sample.

Table 1. Precision indicators for various models of predictive analysis on the test sample for all railways

Parameter	XGBoost	RForest	SVM	kNN	AdaBoost	Logit
PR	0.7790	0.7772	0.6510	0.7112	0.7511	0.7037
FPR	0.0800	0.0802	0.0839	0.0909	0.0799	0.0853

As it can be seen from Table 1, the XGBoost, RandomForest and AdaBoost methods are superior to SVM, kNN and Logit in terms of accuracy. The XGBoost decision tree-based gradient boosting model can be considered the most superior in terms of quality. This method ensures the highest probability of correct answer with the lowest probability of false alarm.

In [3] and [14] show the results of a numerical experiment of railway line categorization based on failure prediction.

Table 2. Results of the experiments using the XGBoost model for the selected three Russian railways with the *threshold* 2 (equal probabilities of error with mark "0" and mark "1") for the control period

Description	Gorky Railway	Northern Railway	Kuybyshev Railway
Prediction-to-fact comparison matrix | | `1` | `0` | | `1` | *TP* | *FN* | | `0` | *FP* | *TN* | ***TN,*** number of correctly predicted `0` ***FN,*** number of incorrectly predicted `0` ***FP,*** number of incorrectly predicted `1` ***TP,*** number of correctly predicted `1`	| | `1` | `0` | | `1` | 413 | 1031 | | `0` | 116 | 3498 | 1`, failure `0`, operable state	| | `1` | `0` | | `1` ` | 76 | 1808 | | `0` ` | 23 | 6174 | `1`, failure `0`, operable state	| | `1` | `0` | | `1` ` | 168 | 1154 | | `0` ` | 57 | 3321 | `1`, failure `0`, operable state
AC	0.7732	0.7734	0.7423
TPR	0.2860	0.0403	0.1271
PR	0.7807	0.7677	0.7467
FPR	0.0321	0.0037	0.0169
F measure	0.4187	0.0767	0.2172

The developed models of predictive analysis of wrong-side failures of track facilities were tested over 8 months of 2020 and showed good results. The accuracy of prediction was evaluated over the control period. The June prediction used the May data from the test period, while the July prediction used the June data.

The classification problem was solved using a number of machine learning algorithms. Learning samples were generated. The quality of the models was analysed in the context of the conducted research.

Table 2 shows an example of the quality of evaluation of wrong-side failure prediction with *threshold 2* for the track services of the three railways over the control period. The control period consists of two months of 2020: June and July.

An analysis of Table 2 shows that, for instance, in the Gorky Railway, the number of correctly predicted category objects with mark "1" is *TP* = 413, while the number of correctly predicted category objects with mark "0" is *TN* = 3498. The same table quotes actual data for incorrectly predicted marks "0" and "1". Calculations using formulas (1) – (5) have established that the classifier's efficiency in terms of its ability to give correct answers is *AC* = 0.77. The probability of false alarm is *FPR* = 0.03. Similar results were obtained for the other two railways.

Similar research of feasibility of predictive analysis was conducted as regards signalling and power supply assets.

Further improvement of the accuracy of hazardous event prediction is done using data *sampling* (retrieval of a subset of observations of interest in terms of analysis out of the massive set of initial data). The use of Big Data technology and Data Science algorithms, including data sampling methods, allowed raising the accuracy of prediction of hazardous states within a control/management system to 93.4% [10].

3. Functional safety of systems with digital twins

In railway transportation, economically-efficient innovative methods of ensuring functional safety are actively developing alongside predictive analytics. That is due to the fact that the key solutions for ensuring functional safety of railway systems consist in the application of multi-channel hardware and multi-version software, which significantly increases the system's cost and often limits its mass deployment. We must note certain difficulties in terms of upgrading and modifying such systems due to the requirement to redesign their redundant components subject to new needs. Those are due to the systems' adaptation to another type of rolling stock, variation of track tonnage, change of line class, etc.

The current problem consists in designing affordable mass-producible train control and protection systems that comply with the stricter requirements for functional safety. In order to achieve the price/safety balance, a number of methods are used, including virtual information processing channels, digital twins, and other methods and their combinations.

This section considers the feasibility of a train control and/or protection system consisting of a SIL1 or SIL2 initial object and an external circuit of digital twins intended for enabling the desired level of functional safety.

A digital twin is understood as an entity containing:

- a mathematical model of the initial object;
- a software implementation of the model that performs all the operating functions of the initial object;
- the results of model verification and proof of its adequacy to the initial system, as well as a list of hazardous and potentially hazardous states defining the allowed duration of wrong-side failures of the initial object;
- operational documentation.

A digital twin is generated as a computer model consisting of three interconnected levels:

- objective containing the computer model of the control/management system hardware components involved in the implementation of the system's operation algorithm, with associated models of executive and measurement devices;
- logical containing the simulation model of the operating algorithm of the train control and/or protection system;
- visual, where data is visualized, user control commands are generated.

Ensuring the adequacy of the virtual model to the real railway facility is the key element of the design of a train protection system [20].

As an object, let us examine a railway signalling system. An automatic block and power (computer-based) interlocking system in a station contains sensors that acquire information on the performance of track circuits (voltage level at the inputs of the receivers). Describing the operation of such sensors is in itself a complicated task, as the aim is to select an optimum out of track occupancy observation modes, broken rail detection, cab signalling strength.

However, such processes are well-studied and reduced to standard requirements that ensure traffic

safety. Accordingly, for the purpose of virtual modelling, their mathematical description may be used as part of predictive diagnostics of the events that constitute the continuous process of their operation. For the next level of the virtual model that represents the discrete-event operation of the simulated object, it suffices to only have the value of the process parameters exceeding the norms of operability and safety.

The discrete-event operation in the virtual model is well-represented with a discrete automaton, for which safety criteria have also been developed that are based on the monotonicity of the control functions. Similarly, the virtual model of the operation of individual rolling stock systems can be constructed. Thus, the operation of the brake line of a train in the braking mode comes down to valve opening, which causes the loss of pressure in the line and brake operation. The process itself is described by complex differential equations of airflow propagation throughout the length of the train. Estimating the consequences in terms of safety only requires to have the criteria of pressure drop in the tail car within the specified time. Given that the braking mode itself may be classified as service, full service and emergency, in the virtual model, the respective variants must be generated. It should be noted that the mechanical action of brake blocks against the wheels due to the discharging of the brake line can be described by limiting temporal characteristics that affect the length of braking.

For the next level of the virtual model, i.e., the level of train protection system description, it suffices to have the time marks of the beginning of brake valve opening and end of the train deceleration or its stopping. Thus, the virtual model of a digital twin combines simplified continuous mathematical models of continuous processes in transformation of information and the associated discrete-event models. The external circuit of the train control and/or protection system under consideration is formed by two same-type digital twins in a dual-channel configuration with independent channel inputs and outputs and a secure comparator [19]. The methods of designing the dual-channel configuration for the purpose of ensuring functional safety are described in the above standards.

The external circuit is connected to the initial single-channel object of the interference-immune and intrusion-protected communication channel (Fig. 2). The system whose diagram is shown in Fig. 2 falls into the category of vital technical systems (e.g., train control and/or protection system) that are to meet stricter requirements in terms of functional safety. This diagram applies to trackside systems. In mobile systems, digital twins can be connected via a radio channel.

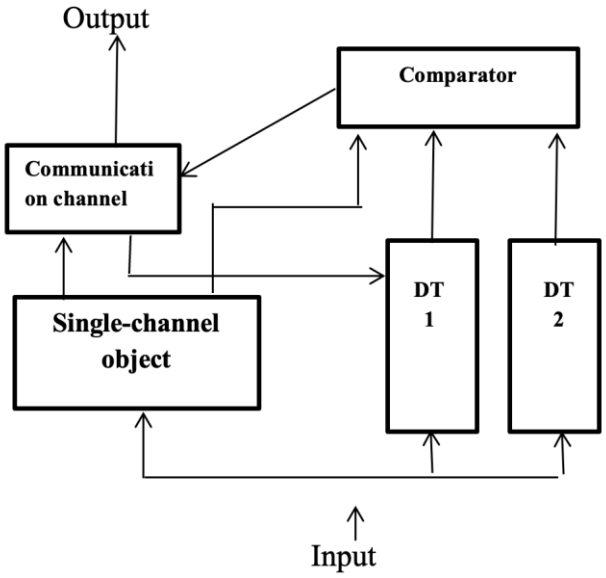


Fig. 2. Summarized structure diagram of a technical system with a digital twin

The introduction of digital twins into a vital technical system raises concerns and requires a more substantial safety case of such a system. The same is true in case of artificial intelligence elements [18, 20].

Criteria of right-side failure of a system with digital twins:

1. Non-matching performance of the initial object and digital twins caused by an undetected failure of the initial object or one of the digital twins. System restoration
2. Failure of one of the digital twins. System restoration

Criteria of wrong-side failure of a system with digital twins:

object failure and error in the delivery of the control command to involve the digital twins or failure of the initial object and digital twins

The adopted assumptions, safety model of the system with digital twins, and findings of the analytical research are reported in [19]. This paper established that, with an error not exceeding the first order of smallness, the wrong-side failure rate of a system with digital twins is defined by the following expression:

$$\lambda F=1/TWF=2\lambda 2\,\bar{\nu}$$

where $\lambda 2$ is the failure rate of the digital twin, $\bar{\nu}$ is the probability of erroneous transmission of a control command by digital twins.

It must be taken into account that in order to ensure an important assumption, i.e., "the failure rate of a digital twin is 2 or 3 orders of magnitude lower than that of the initial object", it is important that the software was designed using methods that comply with higher safety integrity levels, for instance, 2 of 3 safety integrity levels higher. Alternatively, the failure rate is to be proven statistically (see, for instance, Braband et al. [20]).

Ensuring the compliance with the EN 50159 requirements for the communication channel safety, implies that the probability of timely and faultless communication of the command to involve the digital twins tends to one. Therefore, a probability $\bar{\nu}$ of incorrect delivery of digital twin control command close to 0 can be achieved. Subsequently, by using digital twins, the safety of the initial object in terms of wrong-side failure rate may be improved by several orders of magnitude. Indeed, let us examine the relation of the wrong-side failure rates of the initial object ($\lambda_O = \lambda_1$) to the wrong-side failure rate of the system: $Э=\lambda F / \lambda WF = \lambda 1 / 2\lambda 2\,\bar{\nu}$. As $\lambda_1 >> \lambda_2$ and $\bar{\nu} \to 0$, our assertion is correct.

Thus, the transformation of the initial object into a system with digital twins allows significantly reducing the wrong-side failure rate. The introduction of digital twins into the system is a new, not yet tested way of ensuring system safety. Naturally, it requires a substantial safety case. That is associated with significant additional expenditures. The decision on the benefits of additional costs is taken by the customer and system developer together. At the same time, it must be taken into consideration that in case of mass production of technical systems, the effect of additional costs is reduced and the effect of significantly improved safety is maintained.

4. Ensuring the functional safety of locomotive control system using a virtual information processing channel

The structure of the computer-based train separation system is based on the principle of multi-channel information processing. Subsequently, if we examine information communication from the source to the receiver, the device that generates the information and the one that receives it must have at least two channels.

The experience in the design of such systems accumulated in the railway industry is sufficient for confirming the requirement for a train protection system with a dual-channel architecture. By way of example let us point out the experience of operation of a dual-channel train protection system in Russia's railways. They operate over 30 000 onboard train protection systems (KLUB-U), more than 300 stations with computer-based interlocking systems, more than 1000 km of computer-based automatic block systems. All of those systems follow the dual-channel principle. Over more than 30 years of those systems' operation not a single case of wrong-side failure has occurred.

Meanwhile, the matter of information redundancy is as obvious as that of the hardware redundancy.

[21] examines the principle of software duality in the context of vital information processing, whereas a specific task involves the development of at least two different algorithms and design of different software codes. The best-known examples of the technology's application include onboard and trackside devices by Bombardier, while, in case of more vital systems, the company uses a combination of dual-channel hardware and and dual-version programming. The implementation of this safety principle causes additional expenditures that grow exponentially along with the number of computer-based interlocking systems.

Let us examine the applicability of the new approach to ensuring safe train control using single-channel axle counters that involves comparing current axle counter data with the information embedded in the receiver. The information comparison procedures perform the function of the second axle counter channel thus enabling the creation of a system with dual-version programming rather than one with a dual-channel onboard train control system, but, to an extent, with dual-channel axle counters.

Let us imagine that within an onboard system that complies with all functional safety requirements there is certain unique information. Such information, for instance, may represent the number of axles in the train that travels along a line. The safety condition consists in the fact that, ahead of the train, there are k axle-counting stations that, having counted the number of axles that crossed the station, transmit this information to the locomotive for comparison with the information recorded in the vital onboard device. In case the measured number of the train's axles matches the predefined information with respect to the coordinate of the *i-th* ($i = 1,2,\ldots,k$) axle-counting station, a decision with a given level of safety regarding the vacancy of the line behind the train's tail is taken. Each *i*-th axle-counting station out of a series of vacant line sections has its own geographical coordinate and, consecutively, the summation of the distances between such coordinates defines the distance between the head of the next train and the last axle-counting station behind the tail of the preceding train. If the information does not match, then the section occupied by the first train is extended to the next axle-counting station.

Hereinafter the common matter of information communication between trains over a digital channel for vital information communication is not considered, as compliance with the safety level is assumed (coding, encryption etc.).

That poses the question of the necessity of the conventional solution that is the redundant source of information in the form of an axle counter. Indeed, out of the variety of axle counter failures there is only one hazardous situation when an error in the count of a train's axles is concealed by such an

axle counter failure that will eventually cause matching information onboard the train. Obviously, a whole number of conditions are at play, including the different number of axles in trains, the complete variety of axle counter failures, out of which the one concealing event is to manifest itself.

While calculating the failure probabilities it may additionally be taken into consideration that the time interval, within which an axle counter is to fail, is defined only by the moment in time when a train crosses it, while the rest of the time the axle counter may be disconnected from the power source and, accordingly, not susceptible to internal failures.

Therefore, a legitimate question is whether single-channel sources of information (in this case, axle counters) can be used instead of dual-channel ones, thus significantly reducing the cost of hardware, while having a dual-channel receiver (in this case, an information receiver onboard the train) as the reference device. In order to answer that question, [22] examines two safety models. The *first model* considers a SIL4 onboard control system, a dual-channel, dual-version system. All k axle counters have a dual-channel architecture, which enables SIL 3 safety. The *second model* involves the onboard control system also being SIL4, a dual-channel, dual-version system. However, all k axle counters have a single-channel architecture, which allows significantly reducing their cost. At the same time, there are concerns regarding the sufficiency of the safety integrity of the axle counters and the entire system. In order to ensure the safety integrity level of the axle counters of at least SIL 3, as it is the case with the dual-channel architecture, it is required to ensure the corresponding high level of correct detection of axle counter failure. This problem is solved by comparing the current railway coordinates of a locomotive and axle counter, the moment in time when the train crosses the respective axle counter, by calculating the current axle counter number and the number of axles of the passing train. The results are constantly compared. This action aims to enable coordinate identification by a device with an appropriate safety integrity level. Such devices are part of today's train equipment.

The above actions allow endowing axle counters with virtual second information processing channels.

Mathematical descriptions of the models, the adopted assumptions and obtained analytic expressions for the functional safety indicators of the locomotive control system represented by the first and the second models are given in [22]. The second model is different from the first one in that the axle counters have a second, virtual channel. Therefore, for a comparative evaluation of the safety of a control system with an actual or a virtual second channel in the axle counters it suffices evaluating the dependence of the system safety indicator on the failure detection parameters in the first and second cases. An obvious safety indicator could be mean time to wrong-side system failure (T_{ws}). In [22], it is shown that the probability of correct axle counter failure detection α is the most significant parameter of failure detection in a system with an actual second channel in the axle counter. In a control system with a virtual second channel in the axle counters, the probability of the absence of faulse coinsidence of calculation results of the locomotive onboard system and the axle counters ν. Fig. 3 and 4 show the dependences of the safety indicator on the parameters of detection of axle counter failures α and ν respectively.

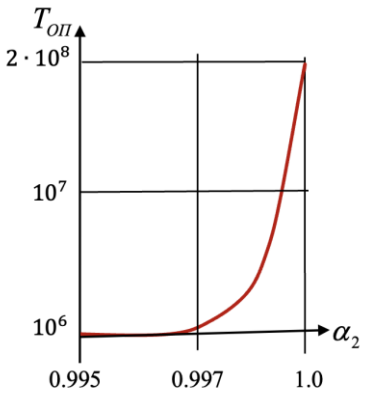


Fig. 3. Graph of the system's mean time to wrong-side failure T_{WS} against the probability of correct detection of axle counter failures α if the axle counter failure rate λ_{ac}= 10^{-6} 1/h

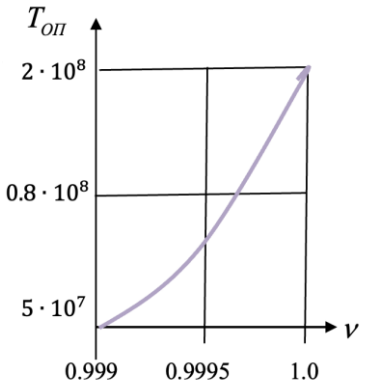


Fig. 4. Graph of the system's mean time to wrong-side failure T_{WS} against the probability of absence of false coincidence between the locomotive's onboard system data and axle counter data ν if the axle counter failure rate λ_{ac}= 10^{-6} 1/h

The comparison of the graphs in Fig. 3 and 4 shows that the maximum value of mean time to wrong-side system failure T_{WS} (model 1) equal to $1.7*10^8$ hours is also achievable for a system represented by model 2 if $\nu = 0.9999$. Such probability of the absence of false coincidence between the locomotive's onboard system data and axle counter data can be made possible [22] by increasing the depth of background testing, improving the algorithms for calculating dynamic data for a better efficiency of comparison of the output data of the locomotive onboard system and the axle counters, as well as through extended safety codes.

5. Conclusion

A high level of functional safety of railway systems can be achieved even as part of a single-channel architecture as long as such economically efficient innovative methods as digital twins or virtual information processing channels are used. The above methods ensure, on the one hand, a reduction of the amount of equipment and significantly lower cost of the systems and, on the other hand, requires the creation of additional software and substantiation of the acceptability of the achieved level of functional safety. There is still a concern that the cost of additional activities is comparable to the obtained economic effect of migration towards single-channel architectures. This matter requires further research.

The innovative methods of ensuring functional safety are efficient both at the design stage, and in the course of system operation. Artificial intelligence-based methods might well allow achieving high accuracy and reliability of hazard prediction, thus enabling their anticipation and prevention. A combination of innovative methods will cause an accumulated effect in ensuring functional safety of railway facilities.

References:

1. Ozerov A.V. Towards safer rail control, command and signalling in the context of digitization. Dependability 2020;2:54-64.
2. Shubinsky I.B., Rozenberg E.N. General provisions of the substantiation of functional safety of intelligent systems in railway transportation. Dependability 2023;23(3):38-45.
3. Thaduri A., Galar D., Kumar U. Railway assets: a potential domain for big data analytics. Proc. Comput. Sci. 2015;53:457-467.

4. Shubinsky I.B., Zamyshliaev A.M. [Managing technical assets in railway transportation]. Moscow: VNIITI RAN; 2021. (in Russ.)
5. Shubinsky I.B., Zamychlaev A.M. Technical Asset Management for Railway Transport (Using the URRAN Approach). Springer: International Series in Operations Research & Management Science, 10.1007/978-3-030-90029;2022.
6. Lasisi A., Attoh-Okine N. Principal components analysis and track quality index: a machine learning approach. Transp. Res. Part C Emerg. Technol. 2018;91:230-248.
7. Santur Y., Karakose M., Akin E. Random forest based diagnosis approach for rail fault inspection in railways. National Conference on Electrical, Electronics and Biomedical Engineering 2016:714-719.
8. Nakhaee M.C., Hiemstra D., Stoelinga M., van Noort M. The Recent Applications of Machine Learning in Rail Track Maintenance: A Survey. Lecture Notes in Computer Science 2019:91-105.
9. Famurewa S.M., Zhang L., Asplund M. Maintenance analytics for railway infrastructure decision support. Journal Qual. Maint. Eng. 2017;23:310-325.
10. Pronevich O.B., Zaitsev M.V. [Intelligent methods of improving the prediction accuracy of rare, hazardous events in railway transportation]. Zheleznodorozhny transport 2012;3:54-65.
11. Ozerov A.V., Olshansky A.M., Kuropteva A.P. Predictive analytics using data science in railway transport. Nauka i teknologii zheleznykh dorog 2020;4:63-76. (in Russ.)
12. [Big Data helps RZD prevent rolling stock malfunctions]. (accessed 15.06.2023). Available at: https://ru-bezh.ru. (in Russ.)
13. Baranov L.A., Balakina E.P., Zhang Y. Prediction error analysis for intelligent management and predictive diagnostics systems. Dependability 2023;23(2):12-18.
14. [RZD to use Big Data to analyse passenger behaviour]. (accessed 15.06.2023). Available at: https://news.myseldon.com. (in Russ.)
15. Zabolotskaya E.A., Izimov D.V., Drivolskaya N.A. [Development of an algorithm for automated Big Data processing as part of railway station parameter prediction]. In: [Development of economics in transportation: new vectors in the post-pandemic period. Proceedings of the international science and practise conference]. Saint Petersburg; 2020. P. 112-120. (in Russ.)
16. Rozenberg E.N., Olshansky A.M., Ozerov A.V., Safronov R.A. Big Data-based methods for functional safety case preparation. Dependability 2022;22(2):38-46.
17. Shubinsky I.B., Zamyshliaev A.M., Pronevich O.B., Ignatov A.N., Platonov E.N. Application of machine learning methods for predicting hazardous failures of railway track assets. Dependability 2020;2(73):43-53.
18. Braband J. A practical guide to safety analysis methods. Railway Signalling + Telecommunication 2001;9:41-45.
19. Shubinsky I.B., Schäbe H., Rozenberg E.N. On the functional safety of a complex technical control system with digital twins. Dependability 2021;1:38-44.
20. Braband J., Gall H., Schäbe H. Proven in use for software: assigning an SIL based on statistics. In: Mahboob Q., Zio E., editors. Handbook of RAMS in Railway systems – Theory and Practice. Boca Raton, Taylor and Francis; 2018. P. 337-350.
21. Iyudu K.A. [Dependability, supervision and diagnostics of computer systems]. Moscow: Vysshaya Shkola; 1989. (in Russ.)
22. Shubinsky I.B., Rozenberg E.N., Korovin A.S., Penkova N.G. On a method for ensuring functional safety of a system with single-channel information processing. Dependability 2022;22(3):44-52.
23. Braband, J.; Schäbe, H. Propagation of uncertainty in railway signaling risk analysis. In: Podofillini et al., editors. Safety and Reliability of Complex Engineered Systems. Proc. ESREL 2015, Taylor & Francis Group (London); 2015. P. 2623-2626.

A FUZZY INNOVATIVE ORDERING PLAN USING STOCK DEPENDENT HOLDING COST OF INSPECTION WITH SHORTAGES IN TIME RELIABILITY DEMAND USING TFN

Sivan V[1*], Thirugnanasambandam K[2], Sivasankar N[3], Sanidari[4]

•

[1]Department of Mathematics, Saveetha Engineering College Chennai- 602105, India.
[*]shivan.ve@gmail.com
[2]Department of Mathematics,Muthurangam Govt Arts College Vellore- 632002,India.
kthirugnanasambandam@gmail.com
[3]Department of Economics, Pondicherry University, Puducherry 605014- India.
sivasankarjh@gmail.com
[4]Department of Mathematical Sciences, Kaduna State University, Kaduna, Nigeria.
sanisdari@yahoo.com

Abstract

The considerations in this paper are, the demand is consistent with time deterioration, the holding cost is dependent based on the quantity of stock available in the system, and the ordering cost is linear and time-dependent. This system should be considered in terms of fuzziness. It is assumed that the shortages are permitted partially, the order is inspected, defective items are identified, by using penalty cost, the defective items should be minimized. Under the classical model and fuzzy environment, the mathematical equation is arrived at to find the optimal solution of total relevant cost with optimal order quantity and time using triangular fuzzy numbers. Defuzzification has been accomplished through the use of the signed distance method of integration. The solutions have been arrived and the model numerical problem of three levels of values (lower, medium and upper) in parametric changes has been verified. Using Sensitivity analysis, the solution is used to validate the changes in different parameter values of the system. To demonstrate the convexity of the TRC function over time, it has used a three-dimensional mesh graph.

Keywords: Ordering plan, Triangular Fuzzy Numbers, Stock depending holding cost, Varying order Cost.

1. Introduction

In any type of business, maintenance of the stock plays crucial role. The stock should be with effective quality (with freshness); it means deterioration should be very less. In this paper, the demand is estimated so that the deterioration is reliable with time period. If the demand increases automatically the deterioration is to be minimized with time. The order placement depends on the demand of the system. Suppose the quantity supplied as part of order quantity is less than the demand, it will lead to shortage. The ordered stock should be in good condition if there is any defective product or service supplied, the firm will incur loss and also the goodwill of the customer. In order to meet the shortages, the lost sale cost is added, the items should be inspected. Sometimes when items are supplied by delay, then the penalty cost will also be added in the

process. Here the ordering cost is not fixed as it is linearly time-dependent. Also, the holding cost is not fixed for the entire time period, whenever the stock reduces, the holding cost also reduces.

Abhishek et al., [1] developed a paper in a fuzzy economic production quantity model deteriorating production depends proportional to population, selling price and advertisement, in this paper he explain clearly how demand is work with population and selling price with advertisement. Dutta .D. et al., [2] presented a article in Optimal Inventory Shortages Fuzziness in Demand,. the model is developed in crisp environment. After that it is convert in to fuzzy environment. All the functions convert in to (TFN) and the fuzzy trapezoidal number. In order to Defuzzification, the SDM used. The EOQ, optimal total cost derived and in both environment. Magfura Pervin et al., [10] explained the combined vendor buyer of quadratic demand inspection ,preservation technology applied the vendor applied the (PT) to reduce deterioration cost using this technique to reduce the total cost and fount the optimal total cost. Pavan Kumar [3] deals with Optimal inventory model with shortages applied fuzzy environment. The shortages were allowed partially backlogged method were applied for manage the stock for genuine customer is was very useful for manufacturer. Sankar Kumar Roy et al., [4] established a model of Imperfection and inspection and varying demand in trade credit using inspection policy easily found the defect of the items and supply to the customer and get the goodwill of the customer. This model is very useful for improve the relationship of customer and supplier. Sivan.V et al., [5] formulated a model of retailer supplier of price dependent demand; To demand will improve automatically when we reduce the cost of the item. Srabani Shee et al.,[6] proposed a model in Fuzzy Supply Chain Varying Holding Cost of supplier and retailer , supplier will more benefit then the retailer since retailer will spend more amount for holding the items so the holding cost of retailer is more than the supplier all the calculation doing by fuzzy and crisp environment. Thirugnanasambandam. et al.,[7] developed model of estimation of EOQ Model negative exponential Demand of linear term. The drugs are maintained stock and problem formed using negative exponential demand so day by day the demand is diminishing. Two types of demand functions formulated and calculated the optimal total cost and more quantity. Tripathi [8] investigated the innovative stock sensitive demand of EOQ for deterioration by means of inconsistent here the newly found stock dependent holding cost using the model holding cost to minimized and linear and nonlinear holding cost considered and verified with parametric changes. Sudip Adak et al., [9] established inventory model reliability dependent partial backordering in fuzziness. Here the deterioration is minimized using the demand if demand is increases automatically the deterioration is reduced and partial shortages are balanced with backlogging here also the results were found and compared with crisp and fuzzy environment.

The present paper has eleven sections. Basic definitions and fuzzy preliminaries are followed by introduction provided. In Section 3, notations and assumptions are introduced. The problems are described and formulated in the fourth section. In Section 5, comes out with numerical solutions and sample problems. The Sensitivity Analysis, Graphical representation and the impact of parametric changes are portrayed in the sixth section. In the section seven represents detailed observation. In Section 8, the Inventory model in fuzzy environment is formulated. Some numerical problems are using Triangular Fuzzy Numbers with different data sets are solved. Illustrative examples are given in the Section nine. In Section 10, comparative studies of crisp and fuzzy optimal values are explored. Conclusions and further developments are distinguished in the final section.

2. Definitions and Fuzzy Preliminaries

Definition 1. Membership value : A fuzzy set $\tilde{\mho}$ is a universe of discourse. The following set of pairs is defined as X. $\tilde{\mho} = \{(x, \mu_{\tilde{\mho}}(x))/x \in R\}$, where $\mu_{\tilde{\mho}}(x) : X \longrightarrow [0,1]$ is a mapping called **membership value** or degree of membership of $x \in R$ in the fuzzy set $\tilde{\mho}$.

Definition 2. Convex : A fuzzy set $\tilde{\mho}$ of the universe if and only if the discourse X is Convex,

$\forall x_1, x_2, \in R$ The following set of pairs is defined as X.

$$\mu_{\tilde{\mho}}(\rho x_1 + (1-\rho)x_2)) \geq \min\left[\mu_{\tilde{\mho}}(x_1), \mu_{\tilde{\mho}}(x_2)\right], \text{ when } 0 \leq \rho \leq 1.$$

Definition 3. Normal Fuzzy Set : A fuzzy set $\tilde{\mho}$ of the universe X is referred to as a Normal Fuzzy Set, meaning that at least one exists $x \in R$ such that $\mu_{\tilde{\mho}}(x) = 1$.

Definition 4. Triangular Fuzzy Number (TFN) : The Triangular Fuzzy Number $\tilde{\mho} = [a_{F1}, a_{F2}, a_{F3}]$ and is formed its continuous membership function $\mu_{\tilde{\mho}}(x) : X \longrightarrow [0,1]$ is,

$$\mu_{\tilde{\mho}}(x) = f(x) = \begin{cases} \dfrac{x - a_{F1}}{a_{F2} - a_{F1}}, & \text{for } a_{F1} \leq x \leq a_{F2}; \\ \dfrac{a_{F3} - x}{a_{F3} - a_{F2}}, & \text{for } a_{F2} \leq x \leq a_{F3}; \\ 0, & \text{Otherwise;} \end{cases}$$

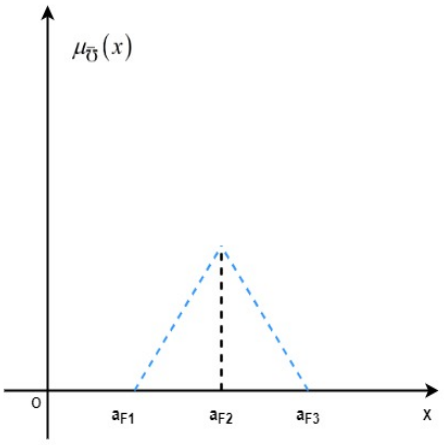


Figure 1: *Triangular Fuzzy Number*

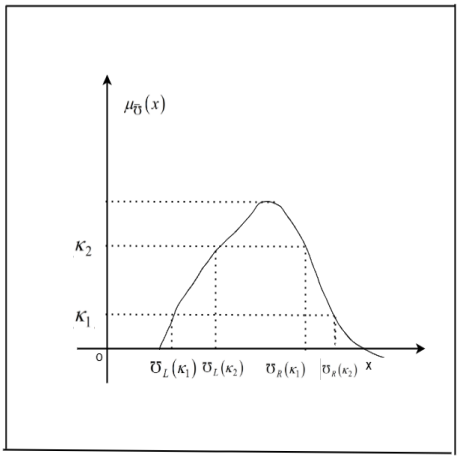


Figure 2: *Fuzzy number with cuts*

Definition 5. Signed Distance Method : Signed Distance Method: Defuzzification of $\tilde{\mho}$ can be discovered using the Signed Distance Method. If $\tilde{\mho}$ is a TFN then Sign distance from $\tilde{\mho}$ to 0 is described as:

$$d(\tilde{\mho}, 0) = \frac{1}{2}\int_0^1 [\tilde{\mho}_L(k), \tilde{\mho}_R(k), 0]\, dk.$$

3. Notations and Assumption

3.1. Notations

1. $I_1(t)$ The inventory level in the time period $0 \leq t \leq t_1$
2. $I_2(t)$ The inventory level in the time period $t_1 \leq t \leq T$
3. t_1 The time stock reached to zero
4. T The total cycle time
5. c_{FO} Ordering cost depend of time dependent
6. $\tilde{c_{\mathrm{FO}}}$ Fuzzy Ordering cost depends of time dependent
7. $\theta(t) = \frac{t}{a^{\lambda}}$ Deterioration period $a, \lambda \geq 1$
8. c_{F2} : Deterioration cost per unit time

9. $\tilde{c_{F2}}$: Fuzzy deterioration cost per unit time
10. c_{F3} : Holding cost per unit time
11. $\tilde{c_{F3}}$: Fuzzy holding cost per unit time
12. c_{F4} : Shortage cost per unit time
13. $\tilde{c_{F4}}$: Fuzzy shortage cost per unit time
14. c_{F5} : Inspection cost per unit time
15. $\tilde{c_{F5}}$: Fuzzy inspection cost per unit time
16. c_{F6} : Penalty cost per unit time
17. $\tilde{c_{F6}}$: Fuzzy penalty cost per unit time
18. Q_F : The maximum order level in the time period $(0 \leq t \leq t_1)$
19. $TRC(t_1, T)$: The total relevant cost per cycle
20. $\tilde{TRC}(t_1, T)$: The Fuzzified total relevant cost per cycle
21. β_F : Unit of shortage cost $(0 \leq \beta_F \leq 1)$
22. μ_F : Unit of lost sale cost $(\mu_F \geq 0)$
23. δ_F : Unit of penalty cost $\delta_F \geq 0$
24. λ : Shape parameter $\lambda \geq 0$
25. D_I : The total items deteriorated

3.2. Assumption

1. The demand function is written $D(t) = p_1 t^a + p_2$ for $0 \leq t \leq t_1$, $t_1 \leq t \leq T$ for $p_1, p_2 \geq 0$
2. Deterioration per cycle $DC_F = \frac{t}{a^\lambda}$, $a \geq 1, \lambda \geq 1$
3. Ordering cost $c_{\text{FO}} = r_1 t_1 + r_2$, $r_1 \geq 0$, $r_2 > 0$
4. T is the complete cycle periods time horizon
5. t_1 is the period of time when inventory level reduces to finish
6. Lead time is negligible.
7. Shortages partially allowed is β_F, $0 \leq \beta_F < 1$ is backordered.
8. In some situation the demand may be considered high, in that period to maintain good relationship with customer and in to consider the inspection policy.
9. In the given cycle time, in the beginning of the process full inventory level is considered.
10. During the shortage period same demand is to be considered.

3.3. Decision Variables

- t_1 : The time Period first level $0 \leq t \leq t_1$
- T : The time Period second level $t_1 \leq t \leq T$

3.4. Objective Functions

1. TRC(t_1, T) : Total relevant cost per cycle
2. β_F : Cost of preservation technology investment per unit time
3. Q_F : The maximum order in the time period $0 \leq t \leq T$
4. t_1, T : Optimal time periods in $0 \leq t \leq t_1$, $t_1 \leq t \leq T$

4. Problem Description and Mathematical Equation

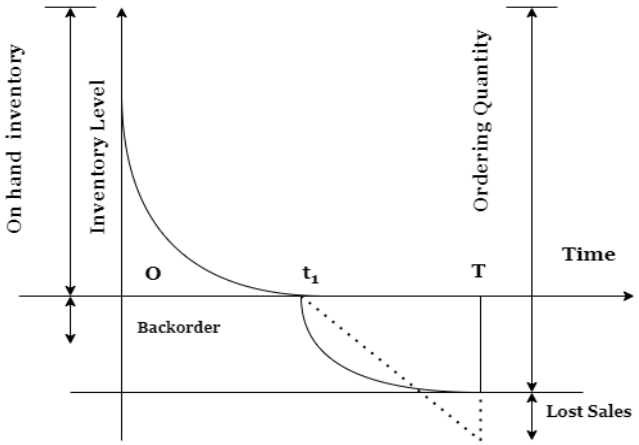


Figure 3: *Inventory level Vs Time.*

4.1. Problem description.

Initially the inventory level is Q_F, because of demand and deterioration the level of inventory is gradually reduced in time period $t = t_1$, so the shortages in this inventory is taken partially.

4.2. Mathematical Equation.

The following formula were used to create the Mathematical model for this paper using differential equations. In shortage period also, the same demand function is maintained.

$$\frac{dI_1(t)}{\mathrm{dt}} + \frac{t}{a^{\lambda}} I_1(t) \quad = \quad -\ (p_1 t^a + p_2)\ ,\ for\ \ 0 \leq t \leq t_1, \tag{1}$$

$$\frac{dI_1(t)}{\mathrm{dt}} \quad = \quad -\ \beta\,(p_1 t^a + p_2)\,,\ for\ \ t_1 \leq t \leq T, \tag{2}$$

Using the initial and the boundary conditions, let us find $I_1(t)$, and $I_2(t)$. In $I_1(t)$ and $I_2(t)$ put $t = t_1$, $t = T$ and $I_1(t_1) = 0$, $I_2(T) = 0$ and get the solution. In $I_1(t)$ put $t = 0$ get order quantity Q_F therefore $I_1(0) = Q_F$.

Solution of 1 and 2

$$\begin{aligned} I_1(t) &= \{p_2(t_1-t)\} + \frac{p_2}{6a^\lambda}(t_1^3-t^3) + \frac{p_1}{a+1}(t_1^{a+1}-t^{a+1}) + \frac{p_1}{2a^{\lambda+1}+6a^\lambda} \\ &\quad \times (t_1^{a+3}-t^{a+3}) + \frac{p_2}{2a^\lambda}(t^3-t_1t^2) + \frac{p_2}{12a^{2\lambda}}(t^5-t^2t_1^3) \\ &\quad + \frac{p_1}{(a+1)2a^\lambda}(t^{a+3}-t_1^{a+1}t^2) + \frac{p_1}{(2a^{\lambda+1}+6a^\lambda)(2a^\lambda)} \\ &\quad \times (t^{a+5}-t_1^{a+3}t^2) \end{aligned} \tag{3}$$

$$\begin{aligned} I_1(0) &= Q_F \\ &= \{p_2(t_1)\} + \frac{p_2}{6a^\lambda}(t_1^3) + \frac{p_1}{a+1}(t_1^{a+1}) + \frac{p_1}{2a^{\lambda+1}+6a^\lambda}(t_1^{a+3}) \end{aligned} \tag{4}$$

$$I_2(t) = \beta\left\{\frac{p_1}{a+1}[(T^{a+1}-t^{a+1}) + p_2(T-t)]\right\} \tag{5}$$

The ordering cost is given by,

$$DC = c_{\mathrm{FO}} = r_1t_1 + r_2, \qquad r_1, r_2 > 0 \tag{6}$$

The total number of pieces that becomes deteriorated throughout that period of interval $0 \le t \le t_1$ is formed by,

$$\begin{aligned} D_I &= Q - \int_0^{t_1} D(t)\,dt \\ &= p_2(t_1) + \frac{p_2}{6a^\lambda}(t_1^3) + \frac{p_1}{a+1}(t_1^{a+1}) + \frac{p_1}{2a^{\lambda+1}+6a^\lambda}(t_1^{a+3}) - \int_0^{t_1}(p_1t^a + p_2)\,dt \\ &= \frac{p_2}{6a^\lambda}(t_1^3) + \frac{p_1}{2a^{\lambda+1}+6a^\lambda}t_1^{a+3} \end{aligned}$$

Therefore the deteriorating cost is formed by,

$$c_{F2}\left\{\frac{p_2}{6a^\lambda}(t_1^3) + \frac{p_1}{2a^{\lambda+1}+6a^\lambda}t_1^{a+3}\right\} \tag{7}$$

The Holding cost (HC) during the interval $[0, t_1]$ is formed by,

$$\begin{aligned} HC &= c_3\int_0^{t_1} e + f[I_1(t)]dt \\ &= c_{F3}\left(et_1 + f\left[\left(\frac{p_2}{2}\right)t_1^2 + \left(\frac{p_2}{8a^\lambda}\right)\frac{t_1^4}{2} + \frac{p_1}{a+2}t_1^{a+2} + \left(\frac{p_1}{2a^{\lambda+1}+6a^\lambda}\right)\right.\right. \\ &\quad \times \left(\frac{a+3}{a+4}\right)t_1^{a+4} - \left(\frac{p_2}{72a^{2\lambda}}\right)t_1^6 - \left(\frac{p_1}{3(2a^{\lambda+1}+2a^\lambda}\right)\left(\frac{a+1}{a+4}\right)t_1^{a+4} \\ &\quad \left.\left. - \left(\frac{p_1}{6a^\lambda(2a^{\lambda+1}+6a^\lambda}\right)\left(\frac{a+3}{a+6}\right)t_1^{a+6}\right]\right) \end{aligned} \tag{8}$$

Shortage Cost is formed by,

$$\begin{aligned} Sh.C &= c_{F4}\beta_F\int_{t_1}^{T}(p_1t^a + p_2)\,\mathrm{dt} \\ &= c_{F4}\beta_F\left\{\frac{p_1}{(a+1)}[T^{a+1} - {t_1}^{a+1}] + p_2(T-t_1)\right\} \end{aligned} \tag{9}$$

The lost sale cost is formed by,

$$\begin{aligned} Lo.SC &= \mu_F(1-\beta_F)\int_{t_1}^{T}(p_1t^a+p_2)\,\mathrm{dt} \\ &= \mu_F(1-\beta_F)\left\{\frac{p_1}{(a+1)}[T^{a+1}-{t_1}^{a+1}]+p_2(T-t_1)\right\} \end{aligned} \tag{10}$$

The Inspection cost in the interval $[0,t_1]$ is given by,

$$\begin{aligned} In.C &= c_{F5}Q_F \\ &= c_{F5}\left\{p_2t_1+\frac{p_2}{6a^\lambda}t_1^3+\frac{p_1}{(a+1)}t_1^{a+1}+\frac{p_1}{2a^{\lambda+1}+6a^\lambda}t_1^{a+3}\right\} \end{aligned} \tag{11}$$

The penalty cost (PC) during the interval $[0,t_1]$ is formed by,

$$\begin{aligned} Pn.C &= c_{F6}\delta_F\int_0^{t_1}(p_1t^a+p_2)\,\mathrm{dt} \\ &= c_{F6}\delta_F\left(\frac{p_1}{(a+1)}t_1^{a+1}+p_2t_1\right) \end{aligned} \tag{12}$$

The Total Average Relevant Cost,

$$\begin{aligned} \mathrm{TRC}(t_1,T) &= \frac{1}{T}\Big\{\text{ Ordering Cost + Deteriorating Cost + Holding Cost +} \\ &\qquad \text{Shortage Cost + Lost Cost + Inspection cost + Penalty cost}\Big\} \\ &= \frac{1}{T}\left[r_1t_1+r_2+c_{F2}\left\{\frac{p_2}{6a^\lambda}(t_1^3)+\frac{p_1}{2a^{\lambda+1}+6a^\lambda}t_1^{a+3}\right\}\right. \\ &\quad +c_{F3}\left(et_1+f\left[\left(\frac{p_2}{2}\right)t_1^2+\left(\frac{p_2}{8a^\lambda}\right)\frac{t_1^4}{2}+\frac{p_1}{a+2}t_1^{a+2}\right]\right) \\ &\quad +\left(\frac{p_1}{2a^{\lambda+1}+6a^\lambda}\right)\left(\frac{a+3}{a+4}\right)t_1^{a+4}-\left(\frac{p_2}{72a^{2\lambda}}\right)t_1^6 \\ &\quad -\left(\frac{p_1}{3(2a^{\lambda+1}+2a^\lambda)}\right)\left(\frac{a+1}{a+4}\right)t_1^{a+4} \\ &\quad -\left(\frac{p_1}{6a^\lambda(2a^{\lambda+1}+6a^\lambda)}\right)\left(\frac{a+3}{a+6}\right)t_1^{a+6} \\ &\quad +c_{F4}\beta_F\left\{\frac{p_1}{(a+1)}[T^{a+1}-{t_1}^{a+1}]+p_2(T-t_1)\right\} \\ &\quad +\mu_F(1-\beta_F)\left\{\frac{p_1}{(a+1)}[T^{a+1}-{t_1}^{a+1}]+p_2(T-t_1)\right\} \\ &\quad +c_{F5}\left\{p_2t_1+\frac{p_2}{6a^\lambda}t_1^3+\frac{p_1}{(a+1)}t_1^{a+1}+\frac{p_1}{2a^{\lambda+1}+6a^\lambda}t_1^{a+3}\right\} \\ &\quad \left.+c_{F6}\delta_F\left(\frac{p_1}{(a+1)}t_1^{a+1}+p_2t_1\right)\right] \end{aligned} \tag{13}$$

For the convenience let us do this substitution,

$$\varphi_1 = \frac{p_2}{6a^{\lambda}} \qquad \varphi_2 = \frac{p_1}{2a^{\lambda+1}+6a^{\lambda}} \quad \varphi_3 = \frac{p_2}{2}$$

$$\varphi_4 = \frac{p_2}{8a^{\lambda}} \qquad \varphi_5 = \frac{p_1}{a+2} \qquad \varphi_6 = \frac{p_1}{2a^{\lambda+1}+6a^{\lambda}}\left(\frac{a+3}{a+4}\right)$$

$$\varphi_7 = \frac{p_2}{24a^{\lambda}} \qquad \varphi_8 = \frac{p_2}{72a^{2\lambda}} \qquad \varphi_9 = \frac{p_1}{3\left(2a^{\lambda+1}+6a^{\lambda}\right)}\left(\frac{a+1}{a+4}\right)$$

$$\varphi_{10} = \frac{p_1}{6a^{\lambda}\left(2a^{\lambda+1}+6a^{\lambda}\right)}\left(\frac{a+3}{a+6}\right) \quad \varphi_{11} = \frac{p_1}{(a+1)}$$

Substitute the above φ_1, φ_2,φ_4,φ_5,φ_6,φ_7,φ_8,φ_9,φ_{10},φ_{11} in equation (13)
The Total Relevant Cost will become

$$\begin{aligned}
\mathrm{TRC}(t_1,T) \;=\; & \frac{1}{T}\Big[r_1t_1 + r_2 + c_{F2}(\varphi_1 t_1^3 + \varphi_2 t_1^{a+3}) \\
& + c_{F3}\left(et_1 + f\left[\varphi_3 t_1^2 + \varphi_1\frac{t_1^4}{2} + \varphi_5 t_1^{a+2} + \varphi_6 t_1^{a+4} - \varphi_8 t_1^6 - \varphi_9 t_1^{a+4}\right.\right. \\
& \left.\left. - \varphi_{10} t_1^{a+6}\right]\right) + c_{F4}\beta_F\left\{\varphi_{11}[T^{a+1} - {t_1}^{a+1}] + p_2(T-t_1)\right\} \\
& + \mu_F(1-\beta_F)[\varphi_{11}[T^{a+1} - {t_1}^{a+1}] + p_2(T-t_1) \\
& + c_{F5}\left\{p_2t_1 + \varphi_1 t_1^3 + \varphi_{11} t_1^{a+1} + \varphi_2 t_1^{a+3}\right\} \\
& + c_{F6}\delta_F\left(\varphi_{11} t_1^{a+1} + p_2t_1\right)\Big] \qquad (14)
\end{aligned}$$

$$\begin{aligned}
\mathrm{TRC}(t_1,T) \;=\; & \frac{1}{T}\Big[r_1t_1 + r_2 + c_{F2}(\varphi_1 t_1^3 + \varphi_2 t_1^{a+3}) \\
& + c_{F3}\left(et_1 + f\left[\varphi_3 t_1^2 + \varphi_1\frac{t_1^4}{2} + \varphi_5 t_1^{a+2} + \varphi_6 t_1^{a+4} - \varphi_8 t_1^6 - \varphi_9 t_1^{a+4}\right.\right. \\
& \left.\left. - \varphi_{10} t_1^{a+6}\right]\right) + [c_{F4}\beta_F + \mu(1-\beta)]\left\{\varphi_{11}[T^{a+1} - {t_1}^{a+1}]\right. \\
& \left. + p_2(T-t_1)\right\} + c_{F5}\left\{p_2t_1 + \varphi_1 t_1^3 + \varphi_{11} t_1^{a+1} + \varphi_2 t_1^{a+3}\right\} \\
& + c_{F6}\delta_F\left(\varphi_{11} t_1^{a+1} + p_2t_1\right)\Big] \qquad (15)
\end{aligned}$$

5. Numerical Solutions and Sample Problems

5.1. Numerical Solutions of Fuzzy Innovative Ordering Plan

For the solution purpose, MATLAB R2018b and Excel solver are used to find all the optimal solutions, all the graphs and convex mesh using MATLAB R2018b software.

To find the solution of the equation (15) using the below necessary and sufficient condition. The necessary condition for the least value of TRC (t_1, T) are,

$$\frac{\partial\left(\mathrm{TRC}\left(t_1,T\right)\right)}{\partial\, t_1} = 0$$

and

$$\frac{\partial\left(\mathrm{TRC}\left(t_1,T\right)\right)}{\partial\ T}=0$$

The Sufficient condition for optimal TRC (t_1,T), $t_1>0,\ T>0$.

$$\frac{\partial^2(\mathrm{TRC})}{\partial t_1^2}>0$$

and

$$\frac{\partial^2(\mathrm{TRC})}{\partial T^2}>0$$

$$\begin{vmatrix} \frac{\partial^2(\mathrm{TRC})}{\partial t_1^2} & \frac{\partial^2(\mathrm{TRC})}{\partial t_1 \partial T} \\ \frac{\partial^2(\mathrm{TRC})}{\partial T \partial t_1} & \frac{\partial^2(\mathrm{TRC})}{\partial T^2} \end{vmatrix} > 0$$

Therefore the optimal solutions of t_1^*, T^*, Q_F^* and TRC^* are found and given in the table.

Table 1: *The optimal solution using crisp*

t_1^*	T^*	Q_F^*	$\mathrm{TRC}^*(t_1,T)$(Rs.)
0.83244	0.86903	10.1671	148.603

5.2. Sample Problems

Example 1. Let's take the input: p_1= 4.2, p_2 = 9.5 , a =1.2, e = 1.8,f= 0.2, β_F= 0.004, δ_F= 3.85, μ_F=11, λ = 3.99, c_{F2} =Rs.15, c_{F3} = Rs.10, c_{F4} =Rs.5 , c_{F5} =Rs.0.5, c_{F6} =Rs.1.03, r_{F1} = Rs.9.8, r_{F2} = Rs.40.

The Optimal Solutions are Q_F^* = 9.58728, t_1^* = 0.82247,T^* =0.90941 & TRC^* = Rs.145.406.

Example 2. Let's take the input: p_1= 4.2, p_2 = 9.99, a =1.2 , e = 1.8, f = 0.2, β_F= 0.004, δ_F= 3.85, μ_F=11,λ = 3.99, c_{F2} =Rs.15, c_{F3} = Rs.10, c_{F4} =Rs.5, c_{F5} =Rs.0.5, c_{F6} =Rs.1.03, r_{F1} =Rs. 9.8, r_{F2} = Rs.40.

The Optimal Solutions are Q_F^* = 10.1671, t_1^* = 0.83244, T^* = 0.86903 & TRC^* = Rs.148.603

Example 3. Let's take the input: p_1 = 4.2, p_2 = 10.25, a = 1.2, e = 1.8, f = 0.2, β_F = 0.004, δ_F = 3.85, μ_F = 11, λ = 3.99, c_{F2} =Rs.15, c_{F3} = Rs.10, c_{F4} =Rs.5, c_{F5} =Rs. 0.5, c_{F6} =Rs.1.03, r_{F1} = Rs.9.8, r_{F2} = Rs.40.

The Optimal Solutions are Q_F^* = 10.4737, t_1^* = 0.83734, T^* = 0.84657 & TRC^* = Rs.150.252.

5.3. Convexity of the optimal function

Convexity of Optimal total cost TRC(t_1,T) versus t_1 and T using Matlab R2018b are shown graphically Figure (4) and Figure (6).

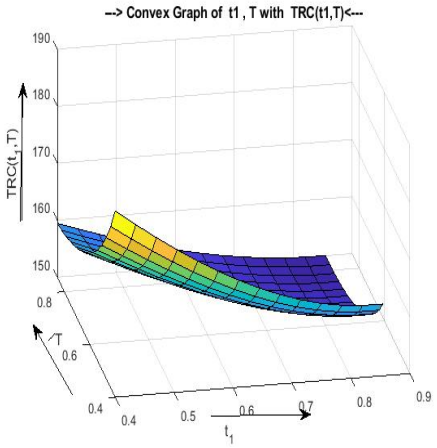


Figure 4: *Convexity Graph of t_1, T with $TRC(t_1, T)$*

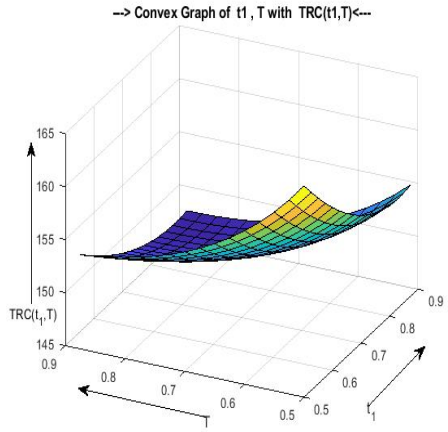


Figure 5: *Convexity Graph of t_1, T with $TRC(t_1, T)$*

6. Sensitivity Analysis and Graphical representation

6.1. Sensitivity analysis of Fuzzy Innovative Economic Order Quantity

Table 2: *Sensitivity analysis of time reliability demand (Parameters P_1, P_2, a, e and f)*

Parameter	Changed values	t_1^*	T^*	Q_F^*	TRC^* (Rs.)
P_1	4.1580	0.8318	0.8698	10.1431	148.2532
	4.2000	0.8324	0.8647	10.1671	148.3721
	4.2420	0.8331	0.8598	10.1912	148.4893
	4.2840	0.8337	0.8548	10.2152	148.6049
P_2	10.0879	0.8343	0.8563	10.2826	148.9947
	10.1898	0.8362	0.8474	10.4028	149.6376
	10.2917	0.8381	0.8384	10.5228	150.2749
	10.4975	0.8418	0.8199	10.765	151.5447
a	1.2120	0.8459	0.8566	10.3447	147.8646
	1.2240	0.8595	0.8485	10.5236	147.3534
	1.2362	0.8733	0.8403	10.7076	146.8281
	1.2485	0.8872	0.8321	10.893	146.2988
e	1.6940	0.8422	0.8454	10.32	147.3337
	1.7464	0.8324	0.8647	10.1671	148.3721
	1.7820	0.8324	0.8647	10.1671	147.8557
	1.8000	0.8341	0.8615	10.1932	148.1983
f	0.18822	0.846	0.8532	10.3799	147.7543
	0.19404	0.8393	0.859	10.2739	148.0635
	0.19800	0.8347	0.8628	10.2028	148.2694
	0.20000	0.8324	0.8647	10.1671	148.3721

Table 3: *Sensitivity analysis of time reliability demand (Parameters β_F, δ_F, μ_F, λ and c_{F2})*

Parameter	Changed values	t_1^*	T^*	Q_F^*	TRC^* (Rs.)
β_F	0.00202	0.8339	0.8619	10.1902	148.3776
	0.00288	0.8333	0.8631	10.1801	148.3753
	0.00360	0.8327	0.8642	10.1718	148.3732
	0.00400	0.8324	0.8647	10.1671	148.3721
β_F	3.7357	0.8469	0.8397	10.3946	147.0331
	3.7734	0.8422	0.8481	10.3197	147.4823
	3.8115	0.8373	0.8565	10.2438	147.9286
	3.8500	0.8324	0.8647	10.1671	148.3721
μ_F	11.1100	0.8459	0.8377	10.3787	148.3928
	11.3322	0.8592	0.8098	10.5892	148.3414
	11.3333	0.8724	0.7807	10.7987	148.2116
	11.4433	0.8853	0.7504	11.0072	147.9958
λ	4.02990	0.8349	0.8633	10.2017	148.2932
	4.07020	0.8374	0.8618	10.2368	148.2132
	4.11090	0.84	0.8603	10.2723	148.1320
	4.15201	0.8426	0.8587	10.3082	148.0497
c_{F2}	14.5545	0.8419	0.8591	10.3146	148.0689
	14.7015	0.8387	0.861	10.2652	148.1704
	14.8500	0.8356	0.8629	10.2161	148.2714
	15.0000	0.8324	0.8647	10.1671	148.3721

Table 4: *Sensitivity analysis of time reliability demand (Parameters c_{F3}, c_{F4} c_{F5}, c_{F6}, r_{F1} and r_{F2})*

Parameter	Changed values	t_1^*	T^*	Q_F^*	TRC^* (Rs.)
c_{F3}	9.7030	0.8442	0.8492	10.3512	147.5376
	9.8010	0.8403	0.8544	10.2902	147.8166
	9.9000	0.8364	0.8596	10.2289	148.0947
	10.0000	0.8324	0.8647	10.1671	148.3721
c_{F4}	5.0500	0.8325	0.8647	10.1675	148.3722
	5.1005	0.8325	0.8646	10.1679	148.3723
	5.1515	0.8325	0.8646	10.1683	148.3724
	5.2030	0.8325	0.8645	10.1687	148.3725
c_{F5}	0.4851	0.8346	0.8615	10.2007	148.1968
	0.4901	0.8339	0.8626	10.1896	148.2548
	0.4950	0.8332	0.8636	10.1784	148.3132
	0.5000	0.8324	0.8647	10.1671	148.3721
c_{F6}	0.9994	0.8498	0.8352	10.4391	146.7880
	1.0095	0.845	0.8436	10.3642	147.2414
	1.0197	0.8373	0.8565	10.2438	147.9286
	1.0300	0.8353	0.8604	10.2117	148.1393
r_{F1}	9.5089	0.8351	0.8595	10.2093	148.0905
	9.6050	0.8342	0.8612	10.1954	148.1837
	9.7020	0.8333	0.863	10.1813	148.2776
	9.8000	0.8324	0.8647	10.1671	148.3721
r_{F2}	39.4020	0.8324	0.8561	10.1671	147.9095
	39.8000	0.8324	0.8647	10.1671	148.3721
	40.1980	0.8324	0.8733	10.1671	148.8301
	40.6000	0.8324	0.8818	10.1671	149.2881

6.2. The graphical representation using Matlab 2018b

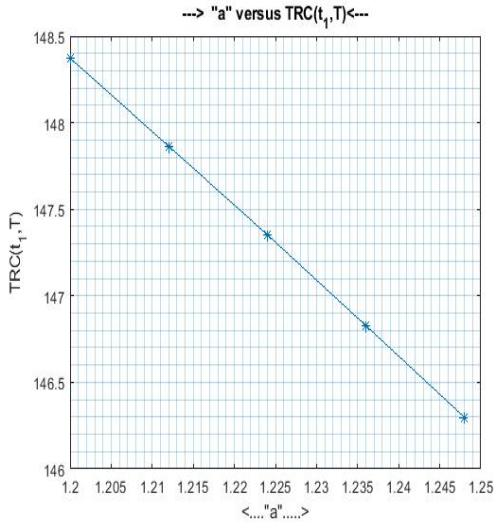


Figure 6: *The impact of a is compared with $TRC(t_1, T)$*

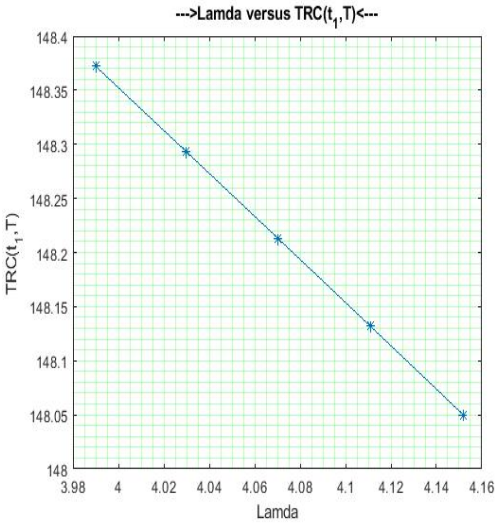


Figure 7: *The impact of λ is compared with $TRC(t_1, T)$*

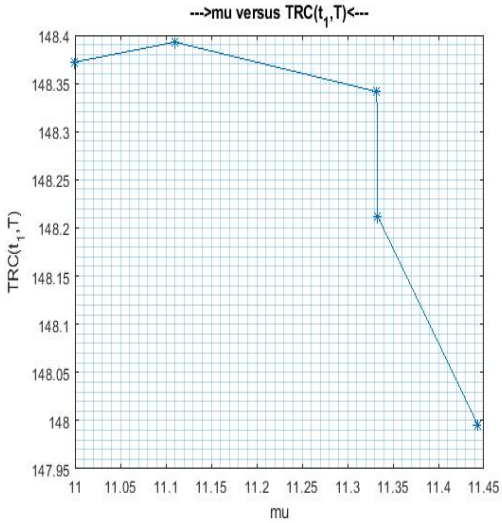


Figure 8: *The impact of μ_F (LSC) is compared with $TRC(t_1, T)$*

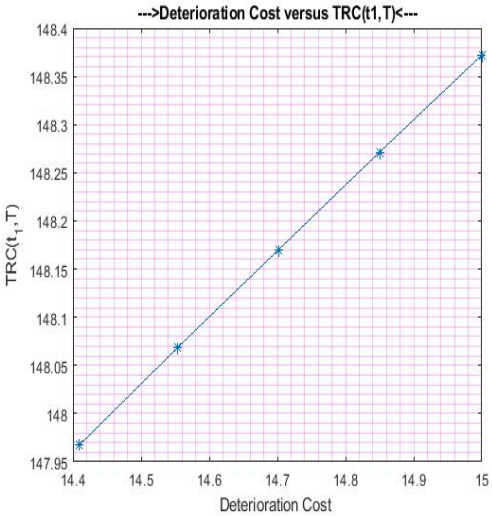


Figure 9: *The impact of Deteriorating cost is compared with $TRC(t_1, T)$*

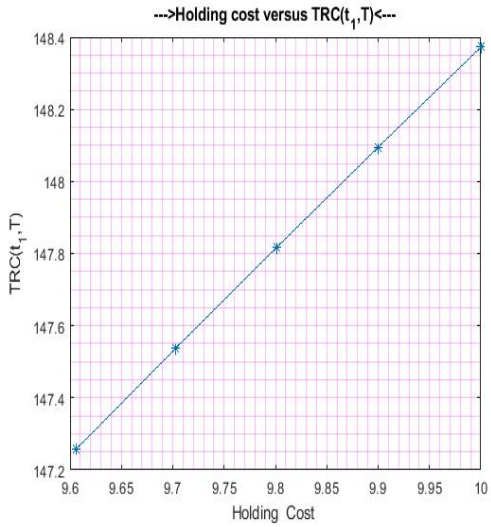


Figure 10: *The impact of Holding cost is compared with* $TRC(t_1, T)$

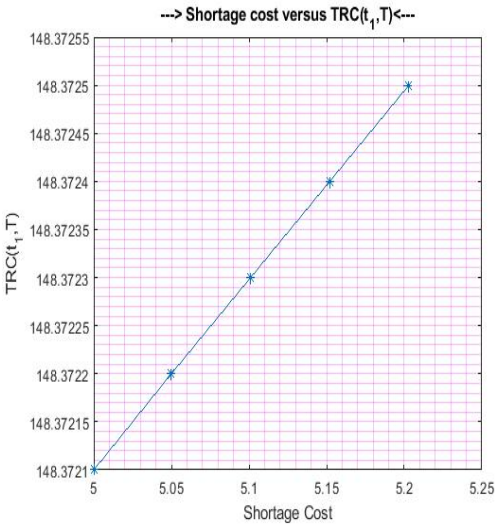


Figure 11: *The impact of Shortage cost is compared with* $TRC(t_1, T)$

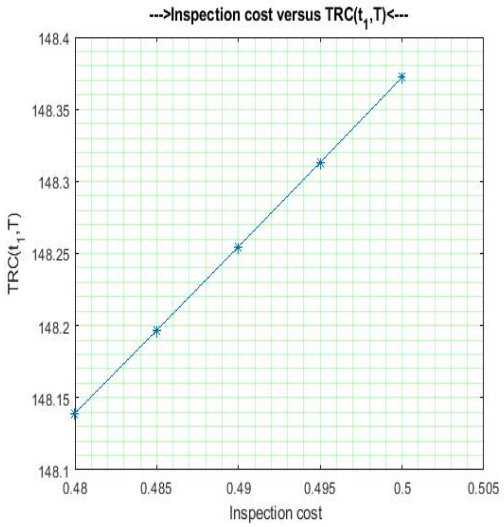


Figure 12: *The impact of Inspection cost is compared with* $TRC(t_1, T)$

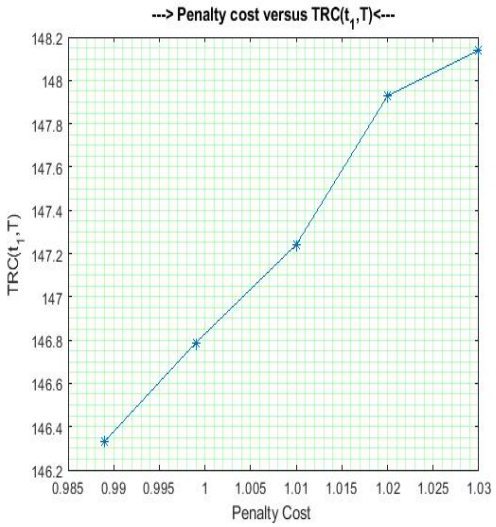


Figure 13: *The impact of Panelty cost is compared with* $TRC(t_1, T)$

7. Observations using table values

Here the investigations are done by using tabular values, let us observe the following progress.

1. While p_2 is raising, the following values t_1, T, Q_F and TRC are oscillating.
2. While the values of a, μ_F & λ are raising, the value of t_1 is raising, T is reducing, Q_F is mounting and TRC is gradually turning down.
3. During the augmentation of the following values, e, f, δ_F, c_{F2}, c_{F3}, c_{F5}, c_{F6} and r_{F1}, t_1 is diminishing, T is growing, Q_F is turning down and TRC is gradually raising.
4. During the mounting of the c_{F4}, the following value of $t_1$is raising , T is growing, Q_F turns up and TRC is gradually leading.
5. While the value of $p_1$is raising, the following values t_1 is raising, T is reducing, Q_F is mounting and TRC is gradually raising
6. While the value of β_F is raising, t_1 is reducing, T is raising, Q_F and TRC are turning down
7. While the value of $p_1$is raising, t_1 is raising, T is reducing, Q_F is mounting and TRC is gradually raising
8. While the value of r_{F2} is raising, the same values of t_1 are repeated, T is raising, the same values of Q_F are repeated and TRC is gradually raising.

8. The Proposed Inventory Model Produced in a Fuzzy Environment

Due to the decision making problem, sometimes the output will be uncertaint and vague, and so some new ideas can be applied to meet the difficulties in characterizing the vagueness and uncertainty. Let us apply the fuzzy environment using Triangular Fuzzy Numbers,

$$\begin{aligned} TRC(t_1,T) &= \frac{1}{T}\Bigg[r_1t_1 + r_2 + c_{F2}(\varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3}) \\ &+ c_{F3}\Bigg\{e\,t_1 + f\Bigg[\varphi_3\, t_1^2 + \varphi_1\,\frac{t_1^4}{2} + \varphi_5\, t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 \\ &- \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\Bigg]\Bigg\} + c_{F4}\beta_F\left\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T-t_1)\right\} \\ &+ \mu_F(1-\beta_F)\,\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T-t_1) + c_{F5}\Bigg\{p_2\,t_1 + \varphi_1 t_1^3 \\ &+ \varphi_{11}\,t_1^{a+1} + \varphi_2\,t_1^{a+3}\Bigg\} + c_{F6}\delta_F(\varphi_{11}\,t_1^{a+1} + p_2t_1\Bigg] \end{aligned} \tag{16}$$

For the convenience, let us do the following suitable substitution.

$$A_1 = \varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3}$$

$$A_2 = e\,t_1 + f\left[\varphi_3\, t_1^2 + \varphi_1\,\frac{t_1^4}{2} + \varphi_5\, t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 - \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\right]$$

$$A_3 = \beta_F\left\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T-t_1)\right\}$$

$$A_4 = \mu(1-\beta_F)\,[\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T-t_1)]$$

$$A_5 = p_2\,t_1 + \varphi_1 t_1^3 + \varphi_{11}\,t_1^{a+1} + \varphi_2\,t_1^{a+3}$$

$$A_6 = \delta_F(\varphi_{11}\,t_1^{a+1} + p_2t_1)$$

In equation (16), substitute the above A_1, A_2, A_3, A_4, A_5, A_6

$$TRC\ (t_1, T) = \frac{1}{T}\Big[r_1 t_1 + r_2 + c_{F2}A_1 + c_{F3}A_2 + c_{F4}A_3 + A_4 + c_{F5}A_5 + c_{F6}A_6\Big] \tag{17}$$

The parameters and costs should be fuzzified using Triangular Fuzzy Number (TFN).

$\widetilde{r_{F1}} = (r_{11},\ r_{12},\ r_{13})$, $\widetilde{r_{F2}} = (r_{21},\ r_{22},\ r_{23})$,

$\widetilde{c_{F2}} = (c_{F21},\ c_{F22}, c_{F23},)$ $\widetilde{c_{F3}} = (c_{F31},\ c_{F32}, c_{F33})$, $\widetilde{c_{F4}} = (c_{F41},\ c_{F42}, c_{F43})$,

$\widetilde{c_{F5}} = (c_{F51},\ c_{F52}, c_{F53})$, $\widetilde{c_{F6}} = (c_{F61},\ c_{F62}, c_{F63})$

$$\begin{aligned} \widetilde{TRC}(t_1, T) &= \frac{1}{T}\Big[(\widetilde{r}_{F1}\, t_1 + \widetilde{r}_{F2}) + \Big\{\tilde{c}_{F2}\,(\varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3})\Big\} \\ &+ \tilde{c}_{F3}\,(e\, t_1 + f\Big[\varphi_3\, t_1^2 + \varphi_1\, \frac{t_1^4}{2} + \varphi_5\, t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 \\ &- \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\Big]\Big\} + \tilde{c}_{F4}\,\beta_F\,\Big\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1)\Big\} \\ &+ \mu_F(1-\beta_F)\,\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1) + \tilde{c}_{F5}\Big\{p_2\, t_1 + \varphi_1 t_1^3 \\ &+ \varphi_{11}\, t_1^{a+1} + \varphi_2\, t_1^{a+3}\Big\} + \tilde{c}_{F6}\delta_F(\varphi_{11}\, t_1^{a+1} + p_2 t_1\Big] \end{aligned} \tag{18}$$

$$\begin{aligned} \widetilde{TRC}(t_1, T) &= \frac{1}{T}\Big[((r_{11},\ r_{12},\ r_{13})t_1 + (r_{21},\ r_{22},\ r_{23})) + \Big\{(c_{F21},\ c_{F22}, c_{F23}) \\ &\times(\varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3})\Big\} + (c_{F31},\ c_{F32}, c_{F33})(e\, t_1 + f\Big[\varphi_3\, t_1^2 + \varphi_1\, \frac{t_1^4}{2} \\ &+\varphi_5\, t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 - \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\Big]\Big\} \\ &+(c_{F41},\ c_{F42}, c_{F43})\beta_F\,\Big\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1)\Big\} \\ &+\mu(1-\beta)\,\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1) + (c_{F51},\ c_{F52}, c_{F53}) \\ &\times\Big\{p_2\, t_1 + \varphi_1 t_1^3 + \varphi_{11}\, t_1^{a+1} + \varphi_2\, t_1^{a+3}\Big\} + (c_{F61},\ c_{F62}, c_{F63})\,\delta_F \\ &\times(\varphi_{11}\, t_1^{a+1} + p_2 t_1\Big] \\ &= (U_F, V_F, W_F) \end{aligned} \tag{19}$$

where,

$$\begin{aligned} U_F &= \frac{1}{T}\Big[(r_{11})t_1 + r_{21} + c_{F21}\Big\{\varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3}\Big\} + c_{F31}\, e\, t_1 + f\Big[\varphi_3\, t_1^2 + \varphi_1\, \frac{t_1^4}{2} \\ &+\varphi_5\, t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 - \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\Big] + \Big[c_{F41}\,\beta_F + \mu_F(1-\beta_F)\Big] \\ &\times\Big\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1)\Big\} + c_{F51}\Big\{p_2\, t_1 + \varphi_1 t_1^3 + \varphi_{11}\, t_1^{a+1} \end{aligned}$$

$$+\varphi_2\, t_1^{a+3}\Big\} + c_{F61}\,\delta_F(\varphi_{11}\, t_1^{a+1} + p_2 t_1)\Big] \tag{20}$$

$$\begin{aligned} V_F &= \frac{1}{T}\Big[(r_{12})t_1 + r_{22} + c_{F22}\Big\{\varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3}\Big\} + c_{F32}\, e\, t_1 + f\Big[\varphi_3\, t_1^2 + \varphi_1\, \frac{t_1^4}{2} \\ &+\varphi_5 t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 - \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\Big] + \Big[c_{F42}\,\beta_F + \mu_F(1-\beta_F)\Big] \\ &\times\Big\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1)\Big\} + c_{F52}\Big\{p_2\, t_1 + \varphi_1 t_1^3 + \varphi_{11}\, t_1^{a+1} \\ &+\varphi_2\, t_1^{a+3}\Big\} + c_{F62}\delta_F(\varphi_{11}\, t_1^{a+1} + p_2 t_1)\Big] \end{aligned} \tag{21}$$

$$\begin{aligned} W_F &= \frac{1}{T}\Big[(r_{13})t_1 + r_{23} + \Big\{c_{F23}(\varphi_1\, t_1^3 + \varphi_2\, t_1^{a+3})\Big\} + c_{F33}\, e\, t_1 + f\Big[\varphi_3\, t_1^2 + \varphi_1\, \frac{t_1^4}{2} \\ &+\varphi_5\, t_1^{a+2} + \varphi_6\, t_1^{a+4} - \varphi_8\, t_1^6 - \varphi_9\, t_1^{a+4} - \varphi_{10}\, t_1^{a+6}\Big] + \Big[c_{F43}\,\beta_F + \mu_F(1-\beta_F)\Big] \\ &\times\Big\{\varphi_{11}\,[T^{a+1} - {t_1}^{a+1}] + p_2\,(T - t_1)\Big\} + c_{F53}\Big\{p_2\, t_1 + \varphi_1 t_1^3 + \varphi_{11}\, t_1^{a+1} \\ &+\varphi_2\, t_1^{a+3}\Big\} + c_{F63}\,\delta_F(\varphi_{11}\, t_1^{a+1} + p_2 t_1)\Big] \end{aligned} \tag{22}$$

The κ-cuts $\mho_L(\kappa)$ & $\mho_R(\kappa)$ of Triangular Fuzzy Numbers. $TRC\tilde{(}t_1, T)$ are given by

$$\begin{aligned} \mho_L(\kappa) &= U_F + (V_F - U_F)\kappa \\ &= \frac{1}{T}\Big[(r_{11})t_1 + r_{21} + c_{F21}\, A_1 + c_{F31}\, A_2 + c_{F41}\, A_3 + A_4 \\ &+c_{F51}\, A_5 + c_{F61}\, A_6 + \Big\{(r_{21} - r_{11})t_1 + (r_{22} - r_{21}) + (c_{F22} - c_{F21})A_1 \\ &+(c_{F32} - c_{F31})A_2 + (c_{F42} - c_{F41})A_3 + A_4 + (c_{F52} - c_{F51})A_5 \\ &+(c_{F62} - c_{F61})A_6\Big\}\kappa\Big] \end{aligned} \tag{23}$$

$$\begin{aligned} \mho_R(\kappa) &= W_F - (W_F - V_F)\kappa \\ &= \frac{1}{T}\Big[(r_{13})t_1 + r_{23} + c_{F23}\, A_1 + c_{F33}\, A_2 + c_{F43}\, A_3 + A_4 \\ &+c_{F53}\, A_5 + c_{F63}\, A_6 + \Big\{(r_{13} - r_{12})t_1 + (r_{23} - r_{23}) + (c_{F23} - c_{F23})A_1 \\ &+(c_{F33} - c_{F33})A_2 + (c_{F43} - c_{F43})A_3 + A_4 + (c_{F53} - c_{F53})A_5 \\ &+(c_{F63} - c_{F62})A_6\Big\}\kappa\Big] \end{aligned} \tag{24}$$

By apply the Signed Distance Method, the defuzzified value of average TRC, using the fuzzy number

$$\begin{aligned} TRC(t_1,T) &= \frac{1}{2}\left[\int_0^1 \{\upsilon_L\kappa+\upsilon_R\kappa\}\,d\kappa\right] \\ &= \frac{1}{4T}\Big[(r_{11}+2r_{12}+r_{13})t_1+(r_{21}+2r_{22}+r_{23})+(c_{F21}+2c_{F22} \\ &\quad +c_{F23})A_1+(c_{F31}+2c_{F32}+c_{F33})A_2+[c_{F41}\ +2c_{F42}+c_{F43}]\,A_3 \\ &\quad +4A_4+(c_{F51}\ +2c_{F52}+c_{F53})A_5+(c_{F61}+2c_{F62} \\ &\quad +c_{F63}\,)A_6\Big] \end{aligned} \tag{25}$$

9. Solutions and numerical problems using triangular fuzzy numbers of different data

9.1. Solutions using triangular fuzzy numbers

For the solution purpose of equation (19), MATLAB R2018b and Excel 2010 solver are used to find all the optimal solutions.
For optimization let us do the following:

The necessary condition for the least value of $TRC\,(t_1,T)$ are,

$$\frac{\partial(\widetilde{TRC}\,(t_1,T)}{\partial t_1}=0 \text{ and } \frac{\partial(\widetilde{TRC}\,(t_1,T)}{\partial T}=0$$

The sufficient condition for optimal $TRC\,(t_1,T)$, $t_1>0, T>0$.

$$\begin{vmatrix} \dfrac{\partial^2(\widetilde{TRC})}{\partial t_1^2} & \dfrac{\partial^2(\widetilde{TRC})}{\partial t_1\,\partial T} \\ \dfrac{\partial^2(\widetilde{TRC})}{\partial T\,\partial t_1} & \dfrac{\partial^2(\widetilde{TRC})}{\partial T^2} \end{vmatrix} > 0$$

Therefore the optimal fuzzy solutions of t_1^*, T^*, Q_F^* and TRC^* are found and given in the table.

Table 5: *Optimal solution using fuzzy Numbers*

${t_1}^*$	T^*	Q_F^*	$TRC(t_1,T)^*(Rs.)$
0.54526521	0.6455852	20.4035902	135.998429

9.2. Sample problems using triangular fuzzy numbers

Example 4. Let's take the input: p_1 = 8.25, p_2 = 31.75, a = 1.25, e = 1.825, f = 0.385, β_F = 0.0044, δ_F = 4.2, μ_F = 13.75, λ = 4.25, $\widetilde{c_{F2}}=(8,11.5,15)$, $\widetilde{c_{F3}}=(2,2.5,3)$, $\widetilde{c_{F4}}=(1.25,2.125,3)$, $\widetilde{c_{F5}}=(0.2,0.3,0.4)$, $\widetilde{c_{F6}}=(0.22,0.33,0.44)$, $\widetilde{r_{F1}}$ = (8.55,10.525,12.5), $\widetilde{r_{F2}}=(20.5,21.5,22.5)$.
The optimal solutions are Q_F^* = 19.6661, t_1^* =0.57313, T^* =0.70023, & TRC^* = Rs.127.092.

Example 5. Let's take the input: p_1 = 8.25, p_2 = 34.75, a = 1.2, e = 1.825, f = 0.385, β_F = 0.0044, δ_F = 4.2, μ_F = 13.75, λ = 4.25, $\widetilde{c_{F2}}=(8,11.5,15)$, $\widetilde{c_{F3}}=(2,2.5,3)$, $\widetilde{c_{F4}}=(1.25,2.125,3)$,

$\widetilde{c_{F5}} = (0.2, 0.3, 0.4)$, $\widetilde{c_{F6}} = (0.22, 0.33, 0.44)$, $\widetilde{r_{F1}}$ =(8.55,10.525,12.5),$\widetilde{r_{F2}} = (20.5, 21.5, 22.5)$.
The optimal solutions are Q_F^* = 20.4035902, t_1^* =0.54526521, T^* = 0.6455852, & TRC^* = Rs.135.998429.

Example 6. Let's take the input: p_1 = 8.25, p_2 = 36, a = 1.2, e = 1.825 , f = 0.385, β_F = 0.0044, δ_F = 4.2, μ_F = 13.75, λ = 4.25, $\widetilde{c_{F2}} = (8, 11.5, 15)$, $\widetilde{c_{F3}} = (2, 2.5, 3)$, $\widetilde{c_{F4}} = (1.25, 2.125, 3)$, $\widetilde{c_{F5}} = (0.2, 0.3, 0.4)$, $\widetilde{c_{F6}} = (0.22, 0.33, 0.44)$, $\widetilde{r_{F1}}$ = (8.55,10.525,12.5), $\widetilde{r_{F2}} = (20.5, 21.5, 22.5)$.
The optimal solutions are Q_F^* = 21.2316, t_1^* =0.54833, T^* =0.60289, & TRC^*= Rs.138.968.

10. Comparison of Crisp and Fuzzy Optimal Solutions

Table 6: *Comparison of crisp and fuzzy solutions*

	$t_1{}^*$	T^*	Q_F^*	$TRC(t_1, T)^*(Rs.)$
Crisp	0.83244	0.86903	10.1671	148.603
Fuzzy	0.54526521	0.6455852	20.4035902	135.998429

11. Conclusion & Extending investigation scope

In this study, an attempt is made to formulate an inventory model of innovative economic order with the quantity of items. The considerations in this paper are (i) the demand is consistent with time deterioration,(ii) the holding cost has been used as dependent on the amount of stock available in the system, and (iii) the ordering cost is linear and time-dependent. This system should be considered in terms of crisp and fuzziness. It is assumed that the shortages are permitted partially and the quantity ordered is inspected to reduce defective items. To use the penalty cost delay of supplying items should be minimized. Under the classical model and fuzzy environment, a mathematical equation is arrived. The optimal solution of total relevant cost with optimal order quantity and time using triangular fuzzy numbers has been found. Defuzzification has been accomplished through the use of the signed distance method of integration. The solutions have been arrived at and verified by using model with a few numerical problems of three levels of values (lower, medium, and upper) in parametric changes. Sensitivity analysis is used to validate the changes in different values of the system's parameters. To demonstrate the convexity of the total relevant cost function over time, a three-dimensional mesh graph has been used. This model can be modified and developed further by changing the demand into probabilistic, price, advertisement dependent etc.

References

[1] Abhishek Kanti Biswas and Sahidul Islam, 2019, "A fuzzy epq model for non-instantaneous deteriorating items where production depends on demand which is proportional to population, selling price as well as advertisement". Independent journal of management and production, Volume (10), issue (5), pp. 1679-1703.

[2] Dutta D and Kumar P, 2013, "Optimal Policy for an Inventory Model without Shortages considering Fuzziness in Demand Holding Cost and Ordering Cost". International Journal of Advanced and Innovative Research, Volume (2), issue (3), pp. 320-325.

[3] Pavan Kumar, 2021, "Optimal policies for inventory model with shortages time-varying holding and ordering costs in trapezoidal fuzzy environment". Independent journal of management and production, Volume (12), issue (2), pp. 557-574.

[4] Sankar Kumar Roy and Magfura Pervin, 2019, " Imperfection with inspection policy and variable demand under trade-credit a deteriorating inventory model". numerical Algebra Control and optimization, Volume (10), issue (05), pp. 45-74.

[5] Sivan V and Thirugnanasambandam K, 2020, "A Joint activity of single retailer and single supplier with price dependent demand in changeable decision of lead time with shortages". Advances in Mathematics Scientific Journal, Volume (9), issue (7), pp. 4641-4658.
[6] Srabani Shee and Tripti Chakrabarti, 2020, "A Fuzzy Two-Echelon Supply Chain Model for Deteriorating Items with Time Varying Holding Cost Involving Lead Time as a Decision Variable". Optimization and Inventory Management, Asset Analytics, Springer book series (ASAN), chap (21) pp. 391-406.
[7] Thirugnanasambandam K and Sivan V, 2020, "Estimation of EOQ Model for Drug Inventory in Wellness Industries Using Time Dependent Negative Exponential Demand and Holding Cost of Linear Equation Without Constant Term". Journal of advanced research and dynamical control system, Volume (12), issue (5), pp. 91-106.
[8] Tripathi R.P., 2020, "Innovative investigation of stock sensitive demand induced economic order quantity (EOQ) model for deterioration by means of inconsistent holding cost functions". Springer Nature Switzerland AGSN Applied Sciences, Volume (2), issue (4), pp. 1-8.
[9] Sudip Adak and Mahapatra G.S., 2018, "An inventory model of time and reliability Dependent demand with deterioration and partial backordering under fuzziness". Recent Advances in Intelligent Information Systems and Applied Mathematics, Book series (SCI, volume 863), pp. 610-627.
[10] Magfura Pervin, Sankar Kumar Roy and Gerhard Wilhelm Weber, 2019, "An integrated vendor-buyer model with quadratic demand under inspection policy and preservation technology". Hacettepe Journal of Mathematics and Statistics Hacet, Volume (20), issue(10) pp. 1-22.

RT&A, No 4 (76)
Volume 18, December 2023

STOCHASTIC ANALYSIS OF JUICE PLANT SUBJECT TO REPAIR FACILITY

Amit Kumar[1] and Pinki Kumari[2]

•

[1]Department of Mathematics, Govt. College, Satnali, Haryana
[2]Department of Physics, Lord University, Chikani, Alwar, Rajasthan
prof.amitmalik@gmail.com, prof.pinkimalik@gmail.com

Abstract

The performance of a juice plant is analyzed by using the base state and the regenerative point graphical technique. The juice plant under consideration consists of three distinct units. It is considered that units A and B may be in a complete failed state through partial failure mode but unit C is in only partially failed state. If one of the units A or B or C partially fails then the system works in a reduced state. When any unit is completely failed then the system is in failed state and no unit can fail further when the system is in a failed state. A technician is always available to repair the failed unit. In this paper, the failure time and repair time follow general distributions. Tables are used to describe the reliability measures such as mean time to system failure, availability and profit values of juice plant.

Keywords: Reliability, juice plant, repair time, mean sojourn time and profit.

I. Introduction

Nowadays, manufacturers have to produce their products continuously to meet the increasing demands of their products which are possible by making their productions as efficient as possible. This paper discusses the MTSF, availability and profit of a juice plant with priority in repair using the regenerative point graphical technique under specified conditions. A large amount of research work has been done on repairable systems such that Bao and Mays [2] analyzed the hydraulic reliability of water distribution systems under demand, pipe roughness and pressure head. Gnedenko and Igor [7] explored reliability and probability studies for engineering purposes. Jack and Murthy [9] discovered the role of limited warranty and extended warranty for the product. Wang and Zhang [19] examined the repairable system of two non identical components under repair facility using geometric distributions. Diaz et al. [5] threw light on the warranty cost management system. Kumar and Goel [15] explored the idea of an imperfect switch on redundant systems in banking industry. Goyal [8] described the availability and behavior of single unit system under preventive maintenance and degradation after complete failure using RPGT. Kumar and Goel [14] analyzed the preventive maintenance in two unit cold standby system under general distributions. Malik and Rathee [17] threw light on the two parallel units system under preventive maintenance and maximum operation time. Kashid and Kumar [11] examined the availability of two unit system under degradation and subject to the repair facility. Kumar et al. [12] evaluate the effects of washing unit in the paper industry by using the regenerative point graphical technique. Levitin et al. [16] explored the results of optimal preventive replacement of failed units in a cold standby system by using the poisson process. Agarwal et al. [1] analyzed the performance and reliability of water treatment plant under repair facility. Barak et al. [3] threw light on the

availability and profit values of milk plant under repair facility. Kumar et al. [13] described the cold standby redundant system under repair and refreshment facilities subject to inspection. Chaudhary and Sharma [4] explored the parallel non identical units system that gives priority to repair over preventive maintenance. Garg and Garg [6] analyzed the reliability and profit values of briquette machine under neglected faults like sound and overheating. Jia et al. [10] explored the two unit system under demand and energy storage techniques. Sengar and Mangey [18] examined the performance of complicated systems under inspection using copula methodology.

II. System Assumptions

There are following system assumptions:

- The juice plant consists of three distinct units.
- Unit A consists of a washing and storage tank.
- Unit B has grinding, blending, evaporation and pasteurization.
- Unit C has bottling, labeling and packing units.
- It is considered that units A and B may be in a complete failed state through partial failure mode but unit C is in only partially failed state.
- Failure and repair times follow general distributions.
- The failed unit works like a new unit after repair.

III. System Notations

There are following system notations:

$i \xrightarrow{Sr} j$	r^{th} directed simple path from state 'i' to state 'j' where 'r' takes the positive integral values for different directions from state 'i' to state 'j'.
$\xi \xrightarrow{s\,f\,f} i$	A directed simple failure free path from state ξ to state 'i'.
$m-cycle$	A circuit (may be formed through regenerative or non regenerative / failed state) whose terminals are at the regenerative state 'm'.
$m-\overline{cycle}$	A circuit (may be formed through the unfailed regenerative or non regenerative state) whose terminals are at the regenerative 'm' state.
$U_{k,k}$	Probability factor of the state 'k' reachable from the terminal state 'k' of 'k' cycle.
$U_{\overline{k,k}}$	The probability factor of state 'k' reachable from the terminal state 'k' of $k\ \overline{cycle}$.
μ_i	Mean sojourn time spent in the state 'i' before visiting any other states.
μ_i'	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state 'i' at t=0.
η_i	Expected waiting time spent while doing a job given that the system entered to the regenerative state 'i' at t=0.
$A/\overline{A}/a$	System's first unit is in the operative state/reduced state/failed state.
$B/\overline{B}/b$	System's second unit is in the operative state/reduced state/failed state.
$C/\overline{C}/c$	System's third unit is in the operative state/reduced state/failed state.
$\lambda_1, \lambda_2, \lambda_3$	The constant partial failure rate of the unit A/B/C respectively.
λ_4, λ_5	The constant complete failure rate of the unit A/B respectively.
w_1, w_2, w_3	Fixed repair rate of the unit A/B/C after partial failure respectively.
w_4, w_5	Fixed repair rate of unit A/B after the complete failure respectively.
○ ⬭ ▭	Upstate/ reduced state/ failed state

IV. Circuits Descriptions

The individual circuit description is given by the table 1:

Table 1: *Circuit Descriptions*

Primary, Secondary and Tertiary Circuit at the vertex (i)			
i	(C1)	(C2)	(C3)
0	(0,1,0), (0,2,0), (0,3,0) (0,1,4,0), (0,2,5,0)	Nil	Nil
1	(1,0,1)	(0,2,0), (0,3,0)	Nil
2	(2,0,2)	(0,1,0), (0,3,0)	Nil
3	(3,0,3)	(0,1,0), (0,2,0)	Nil
4	(4,0,1,4)	(0,1,0), (0,2,0) (0,3,0), (1,0,1)	(2,0,2), (3,0,3)
5	(5,0,2,5)	(0,1,0), (0,2,0) (0,3,0), (2,0,2)	(1,0,1), (3,0,3)

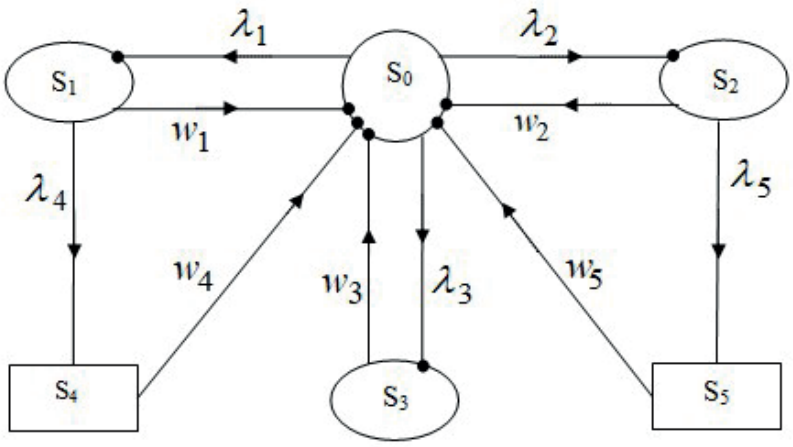


Figure 1: *State Transition Diagram*

where, $S_0 = ABC$, $S_1 = \bar{A}BC$, $S_2 = A\bar{B}C$, $S_3 = AB\bar{C}$, $S_4 = aBC$, $S_5 = AbC$

V. Transition Probabilities

The transition probabilities are following

$p_{0,1} = \lambda_1 /(\lambda_1 + \lambda_2 + \lambda_3)$, $p_{0,2} = \lambda_2 /(\lambda_1 + \lambda_2 + \lambda_3)$, $p_{0,3} = \lambda_3 /(\lambda_1 + \lambda_2 + \lambda_3)$

$p_{1,0} = w_1 /(w_1 + \lambda_4)$, $p_{1,4} = \lambda_4 /(w_1 + \lambda_4)$, $p_{2,0} = w_2 /(w_2 + \lambda_5)$

$$p_{2,5} = \lambda_5 /(w_2 + \lambda_5),\ p_{3,0} = p_{4,0} = p_{5,0} = 1 \tag{1}$$

It has been conclusively established that

$$p_{01} + p_{03} = 1,\ p_{10} + p_{12} + p_{14} = 1,\ p_{21} + p_{27} = 1,\ p_{31} + p_{38} = 1$$
$$p_{41} + p_{45} = 1,\ p_{56} = p_{76} = p_{86} = 1,\ p_{31} + p_{31.8(65)^n} = 1$$
$$p_{10} + p_{12} + p_{11.4} + p_{11.4(56)^n} = 1,\ p_{21} + p_{21.7(65)^n} = 1 \tag{2}$$

VI. Mean Sojourn Time

Let μ_i represents the mean sojourn time. Mathematically, the time taken by a system in a particular state becomes

$$\mu_i = \sum_j m_{i,j} = \int_0^\infty P(T > t)\,dt\ .$$

and $\mu_0 = 1/(\lambda_1 + \lambda_2 + \lambda_3), \mu_1 = 1/(w_1 + \lambda_4)\ ,\ \mu_2 = 1/(w_2 + \lambda_5)$

$$\mu_3(t) = 1/(w_3),\ \mu_4 = 1/(w_4),\ \mu_5 = 1/(w_5) \tag{3}$$

VII. Evaluation of Parameters

All reliability parameters (such as mean time to system failure, availability, busy period of the server and expected number of visits) are determined by using the regenerative point graphical technique. The probability factors of all the reachable states from the base state '0' are given below

$$U_{0,0} = (0,1,0) + (0,2,0) + (0,3,0) = 1,\ U_{0,1} = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3},\ U_{0,2} = \frac{\lambda_2}{\lambda_1 + \lambda_2 + +\lambda_3}$$

$$U_{0,3} = \frac{\lambda_3}{(\lambda_1 + \lambda_2 + \lambda_3)},\ U_{0,4} = \frac{\lambda_1\lambda_4}{(\lambda_1 + \lambda_2 + \lambda_3)(w_1 + \lambda_4)},\ U_{0,5} = \frac{\lambda_2\lambda_5}{(\lambda_1 + \lambda_2 + \lambda_3)(w_2 + \lambda_5)}$$

I. Mean Time to System Failure (MTSF)

The regenerative un-failed states (i=0, 1, 2, 3) to which the system can transit (with initial state 0) before entering to any failed state (using base state ξ=0) then MTSF becomes

$$T_0 = \left[\sum_{i=0}^{3} Sr\left\{\frac{\left\{pr(0 \xrightarrow{Sr(sff)} i)\right\}.\mu_i}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}}\right\}\right] \div \left[1 - \sum Sr\left\{\frac{\left\{pr(0 \xrightarrow{Sr(sff)} 0)\right\}}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}}\right\}\right]$$

$$T_0 = \frac{(w_1 + \lambda_4)(w_2 + \lambda_5)(w_3 + \lambda_3) + w_3[\lambda_1(w_2 + \lambda_5) + \lambda_2(w_1 + \lambda_4)]}{w_3[(\lambda_1 + \lambda_2 + \lambda_3)(w_1 + \lambda_4)(w_2 + \lambda_5) - \lambda_1 w_1(w_2 + \lambda_5) - \lambda_2 w_2(w_1 + \lambda_4)]} \tag{4}$$

II. Availability of the system

The system is available for use at regenerative states j=0, 1, 2, 3 with ξ=0 then the availability of system is defined as

$$A_0 = \left[\sum_{j=0}^{3} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\}.f_j.\mu_j}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}} \right\} \right] \div \left[\sum_{i=0}^{5} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}} \right\} \right]$$

$$A_0 = \frac{w_4 w_5[(w_1+\lambda_4)(w_2+\lambda_5)(w_3+\lambda_3) + w_3\{\lambda_1(w_2+\lambda_5)+\lambda_2(w_1+\lambda_4)\}]}{\left[\begin{array}{r} w_4 w_5(w_1+\lambda_4)(w_2+\lambda_5)(w_3+\lambda_3) + \lambda_1 w_3 w_5(w_2+\lambda_5)(w_4+\lambda_4) \\ + \lambda_2 w_3 w_4(w_1+\lambda_4)(w_5+\lambda_5) \end{array}\right]} \quad (5)$$

III. Busy Period of the Server

The server is busy due to repair of the failed unit at regenerative states *j*= 1, 2, 3, 4, 5 with ξ = 0 then the fraction of time for which the server remains busy is defined as

$$B_0 = \left[\sum_{j=1}^{5} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\}.\eta_j}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}} \right\} \right] \div \left[\sum_{i=0}^{5} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}} \right\} \right]$$

$$B_0 = \frac{\left[\begin{array}{r} w_4 w_5 \lambda_3(w_1+\lambda_4)(w_2+\lambda_5) + \lambda_1 w_3 w_5(w_2+\lambda_5)(w_4+\lambda_4) \\ + \lambda_2 w_3 w_4(w_1+\lambda_4)(w_5+\lambda_5) \end{array}\right]}{\left[\begin{array}{r} w_4 w_5(w_1+\lambda_4)(w_2+\lambda_5)(w_3+\lambda_3) + \lambda_1 w_3 w_5(w_2+\lambda_5)(w_4+\lambda_4) \\ + \lambda_2 w_3 w_4(w_1+\lambda_4)(w_5+\lambda_5) \end{array}\right]} \quad (6)$$

IV. Estimated number of visits made by the server

The repairman visits at regenerative states *j*= 1, 2, 3 with ξ=0 then the number of visits by the repairman is defined as

$$V_0 = \left[\sum_{j=1}^{3} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\}}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}} \right\} \right] \div \left[\sum_{i=0}^{5} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}} \right\} \right]$$

$$V_0 = \frac{\left[\begin{array}{r} w_4 w_5 \lambda_3(w_1+\lambda_4)(w_2+\lambda_5)) + \lambda_1 w_3 w_4 w_5(w_2+\lambda_5) \\ + \lambda_2 w_3 w_4 w_5(w_1+\lambda_4) \end{array}\right]}{\left[\begin{array}{r} w_4 w_5(w_1+\lambda_4)(w_2+\lambda_5)(w_3+\lambda_3) + \lambda_1 w_3 w_5(w_2+\lambda_5)(w_4+\lambda_4) \\ + \lambda_2 w_3 w_4(w_1+\lambda_4)(w_5+\lambda_5) \end{array}\right]} \quad (7)$$

V. Profit Analysis

The profit function may be used to do a profit analysis of the system and it is given by

$$P = E_0 A_0 - E_1 B_0 - E_2 V_0 \quad (8)$$

where, $E_0 = 5000$ (Pay per unit uptime of the system)

$E_1 = 1000$ (Charge per unit time for which server is busy due to repair)

$E_2 = 500$ (Charge per visit of the server)

VIII. Discussion

Table 2 describes the nature of mean time to system failure of the juice plant having an increasing trend corresponding to increment in repair rate (w_2). In this table, the values of parameters λ_1=0.2, λ_2=0.3, λ_3=0.15, λ_4=0.25, λ_5=0.4, w_1=0.35, w_3=0.4, w_4=0.45, w_5=0.5 respectively taking as constant for the simplicity. When λ_1=0.2 changing into λ_1=0.25; λ_2=0.3 changing into λ_2=0.35 and λ_3=0.15 changing into λ_3=0.2 then MTSF values have decreasing trends.

Table 2: *MTSF vs. Repair Rate*

$w_2 \downarrow$	λ_1=0.2, λ_2=0.3 λ_3=0.15, λ_4=0.25 λ_5=0.4, w_1=0.35 w_3=0.4, w_4=0.45 w_5=0.5	λ_1=0.25	λ_2=0.35	λ_3=0.2
0.1	4.0647311	3.8613371	3.218144	3.775294
0.2	4.1010786	3.8894957	3.235483	3.806505
0.3	4.1312067	3.9129156	3.250148	3.833119
0.4	4.1565858	3.9327001	3.262713	3.856081
0.5	4.1782569	3.9496349	3.273599	3.876094
0.6	4.1969772	3.9642941	3.283122	3.893693
0.7	4.2133109	3.9771076	3.291522	3.90929
0.8	4.2276872	3.9884034	3.298987	3.923207
0.9	4.240438	3.998436	3.305664	3.935703
1	4.2518242	4.007406	3.311673	3.946984

Table 3 explores the increasing trends of availability corresponding to increments in repair rate (w_2) where the system's other parameters possess constant values. When the failure rate of unit changes (λ_1=0.2 to 0.25), (λ_2=0.3 to 0.35) and (λ_3=0.15 to 0.2) then the availability of system declines.

Table 3: *Availability vs. Repair Rate*

$w_2 \downarrow$	λ_1=0.2, λ_2=0.3 λ_3=0.15, λ_4=0.25 λ_5=0.4, w_1=0.35 w_3=0.4, w_4=0.45 w_5=0.5	λ_1=0.25	λ_2=0.35	λ_3=0.2
0.1	0.25028	0.214076	0.185918	0.190417
0.2	0.364639	0.319498	0.281347	0.293845
0.3	0.427923	0.381465	0.338725	0.358019
0.4	0.467056	0.421888	0.376681	0.401314
0.5	0.493101	0.450148	0.403461	0.432266
0.6	0.511375	0.470912	0.423258	0.455361
0.7	0.524719	0.48675	0.438421	0.473168
0.8	0.534774	0.499191	0.450366	0.487264
0.9	0.542545	0.509198	0.45999	0.498663
1	0.548679	0.517404	0.46789	0.508046

Table 4 explores the trend of profit values with respect to repair rate (w_2) and its value increase corresponding to increments in repair rate (w_2) where the system's other parameters possess constant values. When the failure rate of unit changes (λ_1=0.2 to 0.25), (λ_2=0.3 to 0.35) and (λ_3=0.15 to 0.2) then the profit of system declines.

Table 4: *Profit vs. Repair Rate*

w_2 ↓	λ_1=0.2, λ_2=0.3 λ_3=0.15, λ_4=0.25 λ_5=0.4, w_1=0.35 w_3=0.4, w_4=0.45 w_5=0.5	λ_1=0.25	λ_2=0.35	λ_3=0.2
0.1	2692.9	2046.417	1588.805	1690.491
0.2	4428.411	3640.771	3027.839	3256.171
0.3	5389.1	4576.626	3891.418	4226.372
0.4	5983.344	5186.276	4461.578	4880.064
0.5	6378.977	5611.935	4863.085	5346.797
0.6	6656.653	5924.286	5159.353	5694.6
0.7	6859.482	6162.259	5385.877	5962.46
0.8	7012.365	6348.978	5563.998	6174.235
0.9	7130.568	6498.993	5707.276	6345.293
1	7223.9	6621.895	5824.711	6485.948

IX. Conclusion

The performance of the juice plant is discussed using the regenerative point graphical technique (RPGT). The above tables explore that when the repair rate increases then the MTSF, system's availability and profit values also increase but when the failure rate increases then the MTSF, availability and profit values decrease. It is clear that RPGT is helpful for industries to analyze the behaviour of the products and components of system.

References

[1] Agrawal, A., Garg, D., Kumar, A. and Kumar, R. (2021). Performance analysis of the water treatment reverses osmosis plant. *Reliability Theory and Applications*, *16*(3): 16-25.
[2] Bao, Y. and Mays, L. W. (1990). Model for water distribution system reliability. *Journal of Hydraulic Engineering*, 116(9): 1119-1137.
[3] Barak, M. S., Garg, R. and Kumar, A. (2021). Reliability measures analysis of a milk plant using RPGT. *Life Cycle Reliability and Safety Engineering*, 10(3): 295-302.
[4] Chaudhary, P. and Sharma, A. (2022): A two non-identical unit parallel system with priority in repair and correlated life times. *Reliability Theory and Application,* 17(1): 113-122.
[5] Diaz, V. G., Gomez, J. F., Lopez, M., Crespo, A. and Leon, P. M. (2009). Warranty cost models State-of-Art: A practical review to the framework of warranty cost management. *ESREL:* 2051–2059.
[6] Garg, D. and Garg, R. (2022). Reliability analysis and profit optimization of briquette machine by considering neglected faults. *Reliability Theory and Application,* 17(67): 210-222.
[7] Gnedenko B. and Igor A.U. (1995). Probabilistic reliability engineering. *John Wiley and Sons.*

[8] Goyal, V. (2015). Availability modeling and behavioral analysis of a single unit system under preventive maintenance and degradation after complete failure using RPGT. *Arya Bhatta Journal of Mathematics and Informatics*, 7(2): 381-390.

[9] Jack, N. and Murthy, D. P. (2007). A flexible extended warranty and related optimal strategies. *Journal of the Operational Research Society*, 58(12): 1612–1620.

[10] Jia, H., Peng, R., Yang, L., Wu, T., Liu, D. and Li, Y. (2022). Reliability evaluation of demand based warm standby systems with capacity storage. *Reliability Engineering and System Safety*, 218: (108132)

[11] Kashid, D. U. and Kumar, R. (2017). Availability modeling of two units system subject to degradation when intermediate repair is feasible using RPGT. *International Multidisciplinary Conference on Commerce, Management, Technology, Engineering and Environmental Sciences & International Conference on Humanities and Social Sciences, (Andheri) Mumbai*.

[12] Kumar, A., Garg, D. and Goel, P. (2019). Mathematical modeling and behavioral analysis of a washing unit in paper mill. *International Journal of System Assurance Engineering and Management*, 10: 1639-1645.

[13] Kumar, A., Garg, R. and Barak, M. S. (2021). Reliability measures analysis of a cold standby system subject to refreshment. *International Journal of System Assurance Engineering and Management*, 14: 147-155.

[14] Kumar, J. and Goel, M. (2016). Availability and profit analysis of a two-unit cold standby system for general distribution. *Cogent Mathematics*, *3*(1): 1262937.

[15] Kumar, S. and Goel, P. (2014). Availability analysis of two different units system with a standby having imperfect switch over device in banking industry. *Arya Bhatta Journal of Mathematics and Informatics*, *6*(2): 299-304.

[16] Levitin, G., Finkelstein, M. and Xiang, Y. (2020). Optimal preventive replacement for cold standby systems with reusable elements. *Reliability Engineering and System Safety*, 204: (107135).

[17] Malik, S. C. and Rathee, R. (2016). Reliability modelling of a parallel system with maximum operation and repair times. *International Journal of Operational Research*, 25(1): 131-142.

[18] Sengar S. and Mangey R. (2022). Reliability and performance analysis of a complex manufacturing system with inspection facility using copula methodology. *Reliability Theory & Applications*, 17(71): 494-508.

[19] Wang, G. J. and Zhang, Y. L. (2007). An optimal replacement policy for repairable cold standby system with priority in use. *International Journal of Systems Science*, *38*(12): 1021-1027.

STOCHASTIC MODEL ON EARLY-STAGE BREAST CANCER WITH TWO TYPES OF TREATMENTS

SUMAN [1], RAJEEV KUMAR [2]

•

[1, 2] Department of Mathematics, Maharshi Dayanand University, Rohtak
[1] suman.uiet@mdurohtak.ac.in, [2] profrajeevmdu@gmail.com

Abstract

The aim of the paper is to study effectiveness of the different treatments of early-stage breast cancer through analysis of a stochastic model. The early-stage breast cancer is a term used to describe breast cancer that is detected at an early stage of development, typically before it has spread to other parts of the body. Early detection of breast cancer is critical as it greatly increases the chances of successful treatment and saves lives. At early-stage breast cancer of the patient, two types of treatment namely, tamoxifen and tamoxifen combined with radiation therapy are commonly used. As it is essential to consider innovative and cost-effective strategies for early detection and treatment. Investigations through analysis of the stochastic model on early-stage breast cancer with these two types of treatments may help the stakeholders. Keeping this in view, in the present paper, a stochastic model is developed for the early-stage breast cancer considering two treatment types, namely tamoxifen and tamoxifen combined with radiation therapy. The model is analyzed by Markov process and regenerative point technique. Mean sojourn time refers to the average amount of time spent by a patient in a particular state before transitioning to another state and mean survival time refers to the average time a patient survives after diagnosis of breast cancer. Mean sojourn time and mean survival time have been calculated. Sensitivity analysis is a technique to understand how changes in input variables or parameters affect the output or outcome of a model and it helps assess the robustness, reliability, and stability of a model by quantifying the impact of variations in input factors. The paper also includes sensitivity and relative sensitivity analyses of the model which explore the impact of different parameters on the survivability of the patient. The MATLAB software has been used for numerical computing and plotting various graphs. The investigation through our analysis of the stochastic model shows that the mean survival time lessens with the rise in the rates of transition and mean survival time from the treatment, tamoxifen plus radiation is higher than the treatment, tamoxifen only. It is concluded that tamoxifen plus radiation is more effective and useful than only tamoxifen for treatment of early-stage breast cancer.

Keywords: Breast cancer, stochastic model, Markov process, regenerative point technique, mean sojourn time, mean survival time and sensitivity analysis

1. Introduction

Cancer is a leading global cause of adult deaths. According to IARC (International Agency for Research on Cancer), India reported approximately 635,000 cancer-related deaths in 2008. Breast cancer ranks as the most frequently diagnosed malignancy among women globally, with 2.3 million cases and 685,000 deaths in 2020. As of the end of 2020, 7.8 million women who had battled were still alive, making it the most common cancer globally.

A lot of work has been done on Markov modelling on breast cancer. Duffy et al. [6] proposed a Markov model indicating that progression to an advanced state is notably faster among individuals aged 40-49. Johnstone et al. [9] analyzed the historical data of untreated breast cancer patients Subsequently, Anthony et al. [1] and Cong et al. [4] found that the combined treatment of tamoxifen and radiation proves more effective in preventing breast cancer recurrence compared to tamoxifen alone. According to Schairer et al. [13], the probability of death, whether from breast cancer or other causes, was frequently higher in black patients compared to white patients. Understanding the parameters of cancer progression is crucial for evaluating screening policies' effectiveness, as highlighted by Harvey et al. [14]. These parameters include transition rates, preclinical sojourn time, sensitivity and the influence of various risk factors on the progression of cancer. Several studies have previously examined breast cancer and its facets. In 2014, Ventura et al. [15] demonstrated the use of a multi-state Markov model to analyze the progression of breast cancer, employing various methods for parametric estimation. Anthony et al. [3] evaluated that radiotherapy reduces the risk of breast and axillary recurrence in early-stage breast cancer when combined with breast-conserving surgery (BCS) and Tamoxifen, but it does not appear to significantly impact distant recurrence or overall survival. Dey et al. [5] provided insights into breast cancer history, risk factors, symptoms, and global and Indian mortality rates, along with available treatments. Grover et al. [7] developed a three-state Markov model based on CA15-3 marker ranges to track disease progression in breast cancer patients. Furthermore, Huang et al. [8] established a breast cancer transition model based on the Chinese population's natural history validating its applicability. Ruiz-Castro et al. [13] developed a discrete-time piecewise Markov model to study the behavior of a multi-state illness. Bayer et al. [2] developed and conducted an analysis of a Markov model designed to simulate and evaluate the treatment strategies for cancer. In 2022, Mubarik et al. [10] made estimations regarding the future trends of breast cancer-related mortality in East and South Asian countries. Moreover Newman et al. [11] discussed the breast cancer burden in low and middle-income countries.

The aim of the paper is to study effectiveness of the different treatments of early-stage breast cancer through analysis of a stochastic model. The early-stage breast cancer is a term used to describe breast cancer that is detected at an early stage of development, typically before it has spread to other parts of the body. Early detection of breast cancer is critical as it greatly increases the chances of successful treatment and saves lives. At early-stage breast cancer of the patient, two types of treatment namely, tamoxifen and tamoxifen combined with radiation therapy are commonly used. As it is essential to consider innovative and cost-effective strategies for early detection and treatment. Investigations through analysis of the stochastic model on early-stage breast cancer with these two types of treatments may help the stakeholders.

In the present paper, we develop a stochastic model for the early-stage breast cancer considering two treatment types, namely tamoxifen and tamoxifen combined with radiation therapy. Mean sojourn time and mean survival time have been calculated. In section 2, we described the stochastic model and its assumptions. In section 3, we define the states of the model and various notations used in the model. In section 4 and section 5, we find the steady-state probabilities and mean sojourn times. Section 6 and section 7 deal with computation of unconditional mean time and mean survival time (MST). A particular case has been considered in section 8. Numerical calculation and graphical analysis are done in section 9. Section 10 includes sensitivity analysis and relative sensitivity analysis. Finally, conclusion is presented in section 11.

2. Model Description and Assumptions

The present paper introduces a comprehensive six-state markov model for breast cancer. Any normal person may be infected with breast cancer symptoms. Whenever an early-stage breast cancer is diagnosed in a person then without taking the possibility of breast surgery, two types of treatments, namely tamoxifen (say, treatment-1) and tamoxifen plus radiation (say, treatment-2) have been considered. Here only these two treatments are assumed to be available to the patient that may perfectly cure the cancer. When the patient recovers using any of the treatments, then he/she will go to normal state, otherwise he/she will be in death state as the facility to carry out breast surgery of the patient is not available. Various assumptions for the model are as under:

- Initially the person is normal.
- Transition rates follow exponential distribution and other rates follow general distribution.
- All random variables are independent.

3. Model Development

Various states of the model and notations of different parameters are described in table 1 and table 2 respectively.

Table 1: *States specification*

States	Description
S_0	Normal State
S_1	The state when breast cancer is diagnosed.
S_2	The state in which treatment-1 is given to breast cancer patient.
S_3	The state in which treatment-2 is given to the breast cancer patient
S_4	The state in which no treatment is given to the breast cancer patient
S_5	Death state

Table 2: *Notations*

Notations	Description
λ_1	Transition rate from normal to diagnosed state
λ_2	Transition rate from normal to death state
λ_3	Transition rate from treatment-1 to death state
λ_4	Transition rate from treatment-2 to death state
λ_5	Transition rate from no treatment to death state
$h_1(t) / H_1(t)$	p.d.f. /c.d.f. of time of recovery from treatment-1.
$h_2(t) / H_2(t)$	p.d.f. /c.d.f. of time of recovery from treatment-2.
p / q	Probability that treatment is given/not given to the patient.
p_1 / q_1	Probability that treatment-1/treatment-2 is given to the patient.

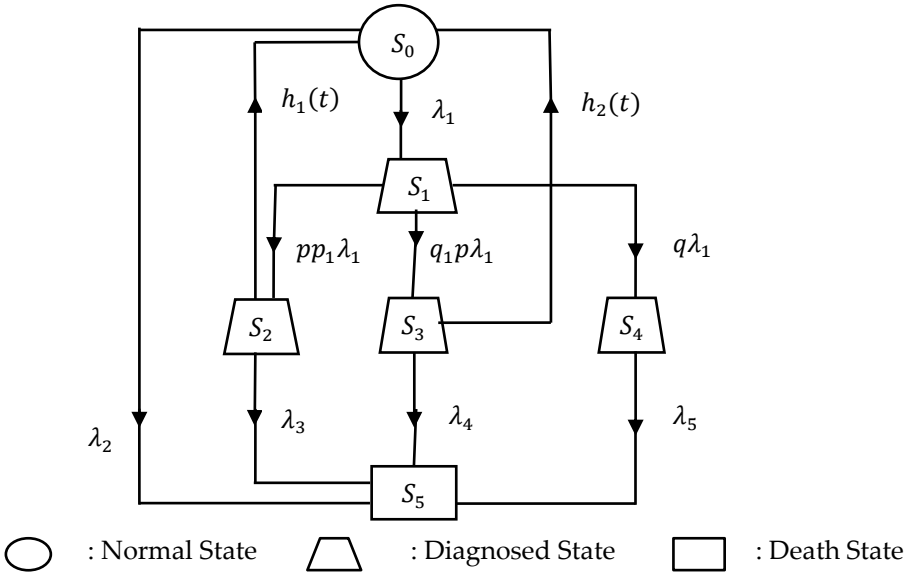


Figure 1. *State -Transition Diagram*

4. Transition Probabilities

The transition probabilities are

$q_{01}(t) = \lambda_1 e^{-(\lambda_1+\lambda_2)t}$; $q_{05}(t) = \lambda_2 e^{-(\lambda_1+\lambda_2)t}$; $q_{12}(t) = p_1 p\lambda_1 e^{-\lambda_1 t}$

$q_{13}(t) = q_1 p\lambda_1 e^{-\lambda_1 t}$; $q_{14}(t) = q\lambda_1 e^{-\lambda_1 t}$; $q_{20}(t) = h_1(t)\, e^{-\lambda_3 t}$

$q_{25}(t) = \lambda_3 e^{-\lambda_3 t}\overline{H_1(t)}$; $q_{30}(t) = h_2(t) e^{-\lambda_4 t}$; $q_{35}(t) = \lambda_4 e^{-\lambda_4 t}\overline{H_2(t)}$

$q_{45}(t) = \lambda_5 e^{-\lambda_5 t}$.

The steady-state probabilities $p_{ij} = \lim_{s\to 0} L\{q_{ij}(t)\}$ are obtained as

$p_{01} = \frac{\lambda_1}{\lambda_1+\lambda_2}$; $p_{05} = \frac{\lambda_2}{\lambda_1+\lambda_2}$; $p_{12} = pp_1$; $p_{13} = pq_1$

$p_{14} = q$; $p_{20} = h_1^*(\lambda_3)$; $p_{30} = h_2^*(\lambda_4)$; $p_{25} = 1 - h_1^*(\lambda_3)$

$p_{35} = 1 - h_2^*(\lambda_4)$; $p_{45} = 1$.

Clearly,

$p_{01} + p_{05} = 1$; $p_{12} + p_{13} + p_{14} = 1$; $p_{20} + p_{25} = 1$

$p_{30} + p_{35} = 1$; $p_{45} = 1$

5. Mean Sojourn Time

Expected time taken by the patient in state i before transiting to any other state is called mean sojourn time in that state.

It is denoted by μ_i and is given by

$$\mu_i = \int_0^\infty P(T_i > t)\,dt = \int_0^\infty R(t)\,dt$$

Then, we have

$$\mu_0 = \frac{1}{\lambda_1 + \lambda_2} \quad ; \quad \mu_1 = \frac{1}{\lambda_1} \quad ; \quad \mu_2 = \frac{1 - h_1^*(\lambda_3)}{\lambda_3}$$

$$\mu_3 = \frac{1 - h_2^*(\lambda_4)}{\lambda_4} \quad ; \quad \mu_4 = \frac{1}{\lambda_5}$$

6. Unconditional Mean Time

Unconditional mean time m_{ij} is mathematically stated as $m_{ij} = \int_0^\infty t\, q_{ij}(t)\,dt = -q_{ij}^{*'}(0)$.

Then, we have

$$m_{01} = \frac{\lambda_1}{(\lambda_1 + \lambda_2)^2} \quad ; \quad m_{05} = \frac{\lambda_2}{(\lambda_1 + \lambda_2)^2} \quad ; \quad m_{12} = \frac{p p_1}{\lambda_1}$$

$$m_{13} = \frac{p q_1}{\lambda_1} \quad ; \quad m_{14} = \frac{q}{\lambda_1} \quad ; \quad m_{20} = -h_1^{*'}(\lambda_3)$$

$$m_{25} = h_1^{*'}(\lambda_3) + \frac{1}{\lambda_3} - \frac{h_1^*(\lambda_3)}{\lambda_3}; \quad m_{30} = -h_2^{*'}(\lambda_4) \quad ; \quad m_{35} = h_2^{*'}(\lambda_4) + \frac{1}{\lambda_4} - \frac{h_2^*(\lambda_4)}{\lambda_4}$$

$$m_{45} = \frac{1}{\lambda_5}$$

Thus,

$$m_{01} + m_{05} = \mu_0 \quad ; \quad m_{12} + m_{13} + m_{14} = \mu_1 \quad ; \quad m_{20} + m_{25} = \mu_2$$

$$m_{30} + m_{35} = \mu_3 \quad ; \quad m_{45} = \mu_4$$

7. Mean Survival Time

Let $\phi_i(t)$ denotes the cumulative distribution function of first passage time from S_i to death state.

The following recursive relations are obtained for $\phi_i(t)$:

$$\phi_0(t) = Q_{01}(t) Ⓢ \phi_1(t) + Q_{05}(t)$$
$$\phi_1(t) = Q_{12}(t) Ⓢ \phi_2(t) + Q_{13}(t) Ⓢ \phi_3(t) + Q_{14}(t) Ⓢ \phi_4(t)$$
$$\phi_2(t) = Q_{20}(t) Ⓢ \phi_0(t) + Q_{25}(t)$$
$$\phi_3(t) = Q_{30}(t) Ⓢ \phi_0(t) + Q_{35}(t)$$
$$\phi_4(t) = Q_{45}(t)$$

Taking Laplace Stielje's Transform (L.S.T.) on both sides of above equations and solve for $\phi_0^{**}(s)$, we have

$$\phi_0^{**}(s) = \frac{N(s)}{D(s)} \qquad (1)$$

where $\phi_0^{**}(s)$ *is* Laplace Stielje's Transform of $\phi_0(t)$,

$$N(s) = Q_{05}^{**}(s) + Q_{01}^{**}(s)\left(Q_{12}^{**}(s)Q_{25}^{**}(s) + Q_{13}^{**}(s)Q_{35}^{**}(s) + Q_{14}^{**}(s)Q_{45}^{**}(s)\right)$$

and $D(s) = 1 - Q_{01}^{**}(s)Q_{20}^{**}(s)Q_{12}^{**}(s) - Q_{01}^{**}(s)Q_{30}^{**}(s)Q_{13}^{**}(s).$

Now, the mean survival time is given by

$$T_0 = \lim_{s \to 0} \frac{1 - \phi_0^{**}(s)}{s}$$

Using L'Hospital's rule and putting the value of $\phi_0^{**}(s)$ from equation (1), we get

$$T_0 = \frac{N}{D},$$

where

$$N = \mu_0 + \mu_1 p_{01} + \mu_2 p_{01} p_{12} + \mu_3 p_{01} p_{13} + \mu_4 p_{01} p_{14}$$

and $D = 1 - p_{01} p_{12} p_{20} - p_{01} p_{13} p_{30}$.

8. Particular Case

The following particular case is considered for analysis purpose:

$h_1(t) = \beta_1 e^{-\beta_1 (t)}$ and $h_2(t) = \beta_2 e^{-\beta_2 (t)}$, where β_1 and β_2 are recovery rate from treatment-1 and treatment-2, respectively. The transition probabilities are given by

$$p_{01} = \frac{\lambda_1}{\lambda_1 + \lambda_2} \quad ; \quad p_{05} = \frac{\lambda_2}{\lambda_1 + \lambda_2} \quad ; \quad p_{12} = pp_1 \quad ; \quad p_{13} = pq_1$$

$$p_{14} = q \quad ; \quad p_{20} = h_1^*(\lambda_3) \quad ; \quad p_{30} = h_2^*(\lambda_4) .$$

Mean sojourn time is given by

$$\mu_0 = \frac{1}{\lambda_1 + \lambda_2} \quad ; \quad \mu_1 = \frac{1}{\lambda_1} \quad ; \quad \mu_2 = \frac{1 - h_1^*(\lambda_3)}{\lambda_3} = \frac{1 - p_{20}}{\lambda_3}$$

$$\mu_3 = \frac{1 - h_2^*(\lambda_4)}{\lambda_4} = \frac{1 - p_{30}}{\lambda_4} \quad ; \quad \mu_4 = \frac{1}{\lambda_5},$$

where $h_1^*(\lambda_3) = p_{20} = \frac{\beta_1}{\lambda_3 + \beta_1}$ and $h_2^*(\lambda_4) = p_{30} = \frac{\beta_2}{\lambda_4 + \beta_2}$.

9. Numerical Computation and Graphical Analysis

For the numerical computation and graphical analysis, the above particular case is considered. The transition rates (λ_1, λ_2) are taken as given in Harvey et al. (2013) whereas other parameters $(\lambda_3, \lambda_4, \lambda_5, p_1, q_1, p, q)$ are assumed here. Various graphs have been plotted for mean survival time taking varying values to the parameters involved in its expression.

In the figures 2, 4, 6, 8 and 10, graphs exhibit the nature of mean survival time (T_0) versus transition rates $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 for varying recovery rate β_1 *whereas* in the figures 3, 5, 7, 9 and 11, graphs exhibit the nature of mean survival time (T_0) versus transition rates $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 for varying recovery rate β_2. In figure 12, graph presents the nature of mean survival time versus transition rate λ_1 for varying values of β_1 and β_2 .

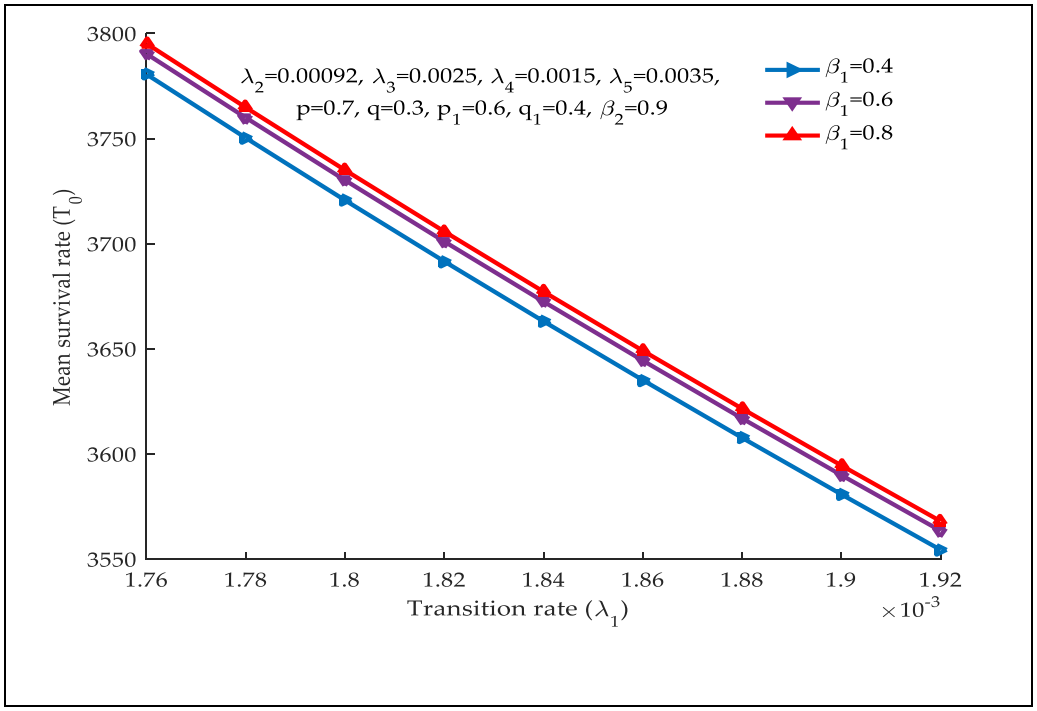


Figure 2. *Mean Survival Time* (T_0) *versus transition rate* (λ_1) *for varying recovery rate* (β_1)

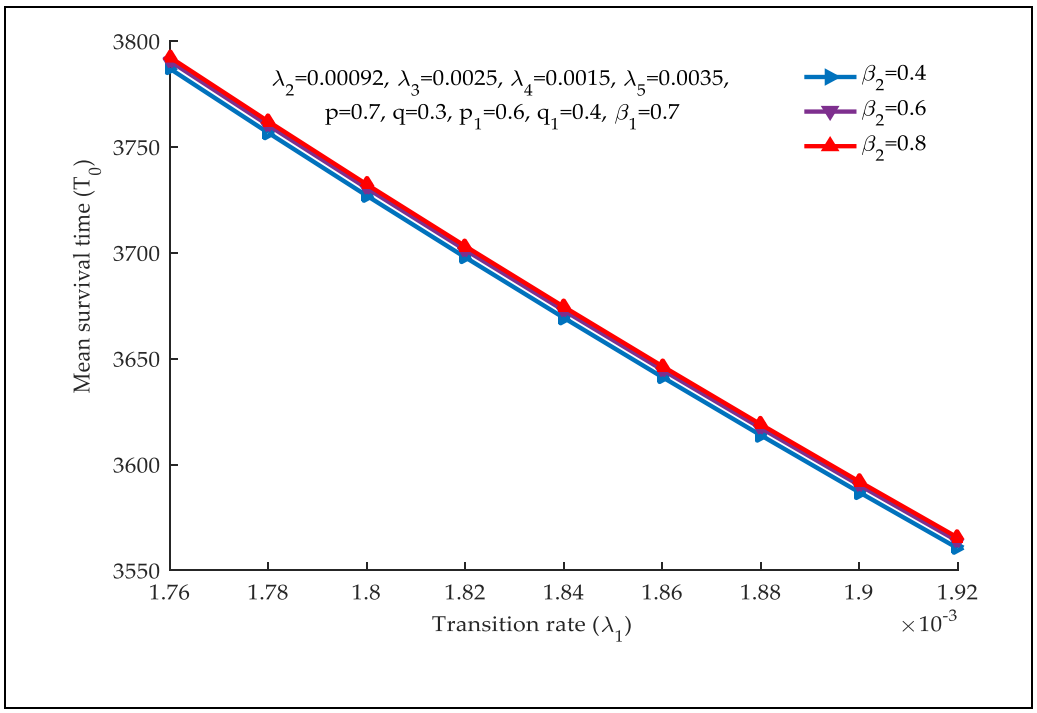


Figure 3. *Mean Survival Time* (T_0) *versus transition rate* (λ_1) *for varying recovery rate* (β_2)

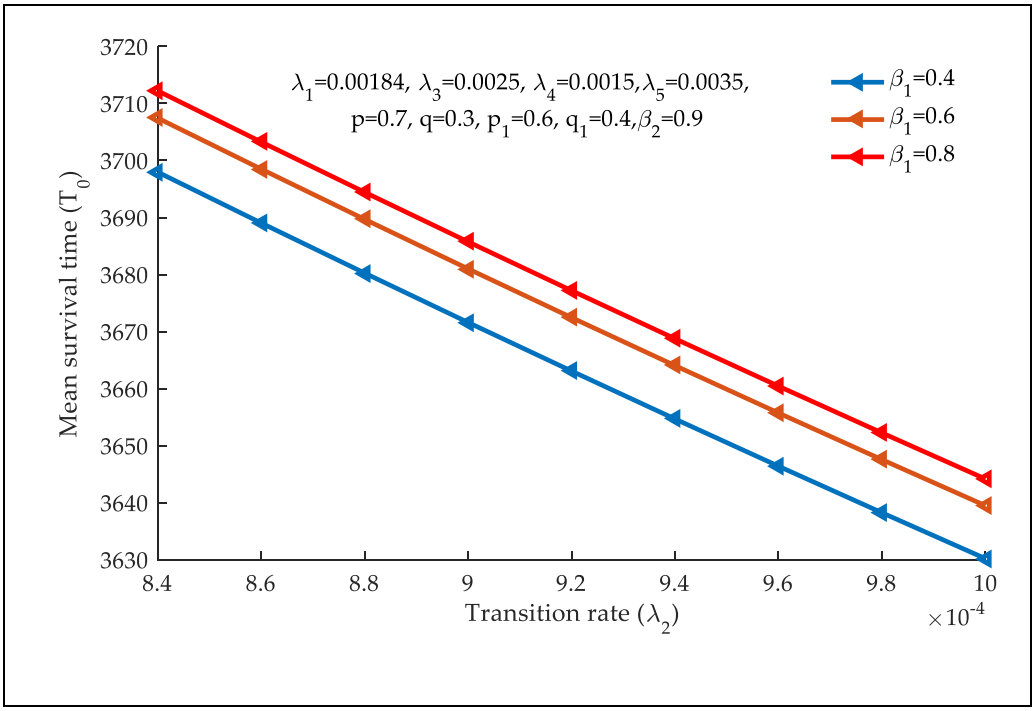


Figure 4. *Mean Survival Time* (T_0) *versus transition rate* (λ_2) *for varying recovery rate* (β_1)

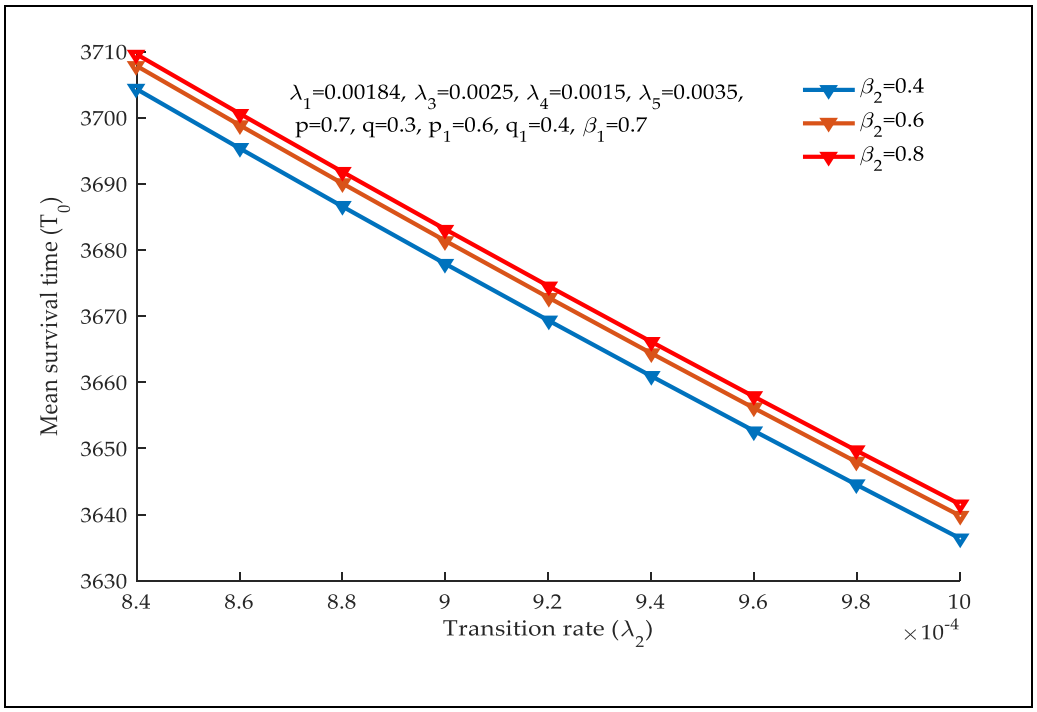


Figure 5. *Mean Survival Time* (T_0) *versus transition rate* (λ_2) *for varying recovery rate* (β_2)

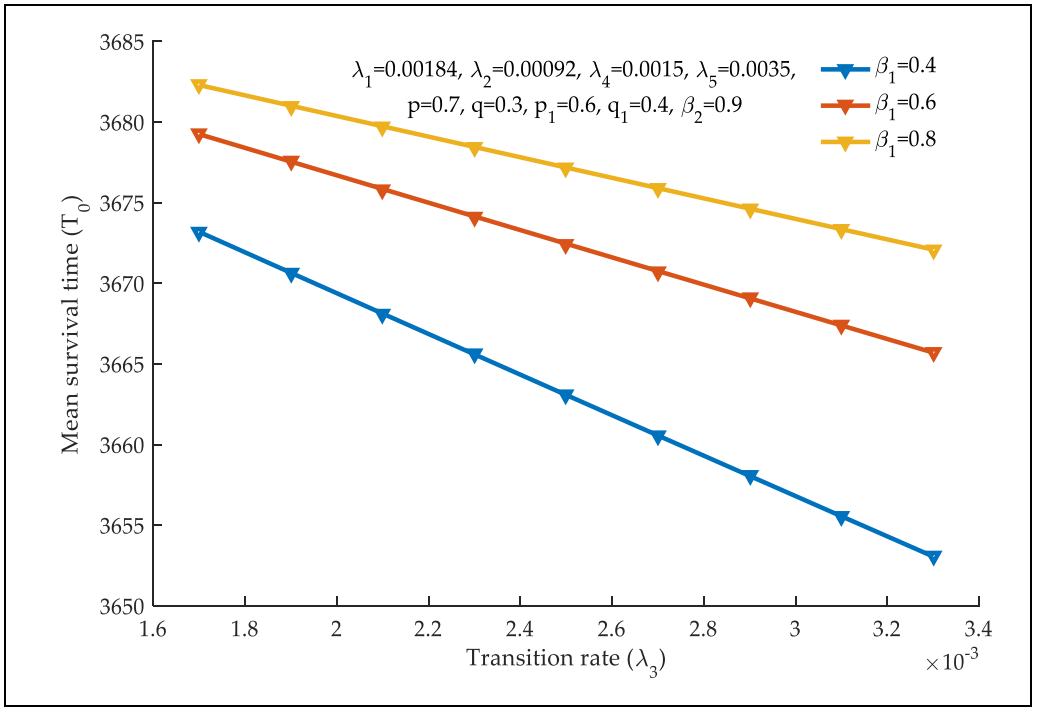


Figure 6. *Mean Survival Time* (T_0) *versus transition rate* (λ_3) *for varying recovery rate* (β_1)

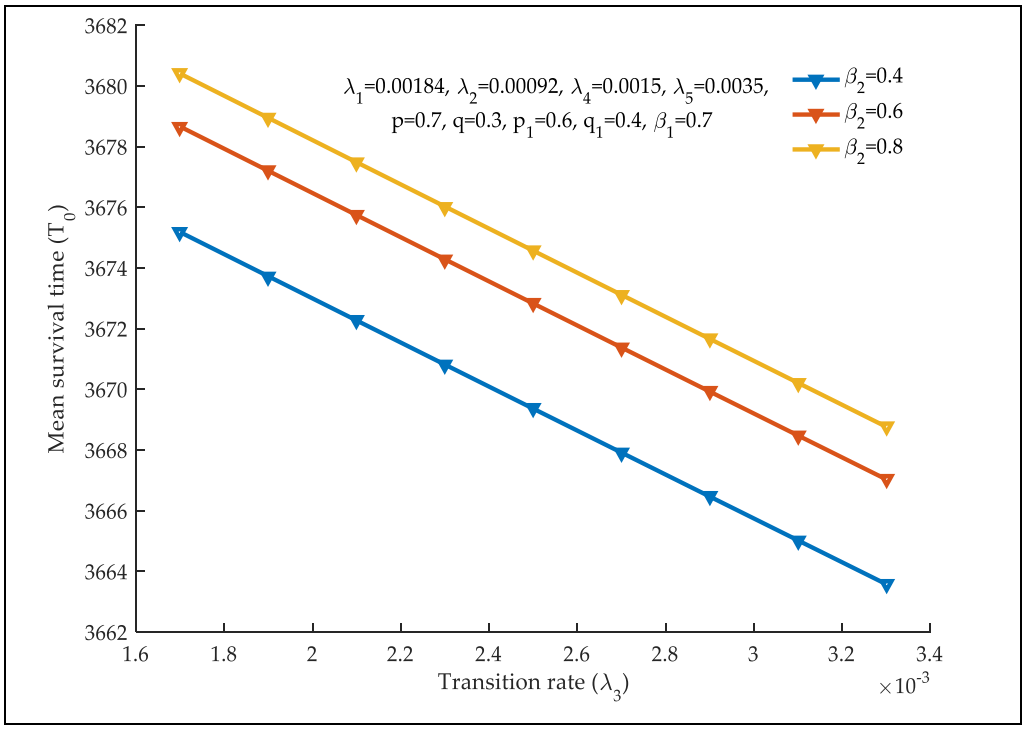


Figure 7. *Mean Survival Time* (T_0) *versus transition rate* (λ_3) *for varying recovery rate* (β_2)

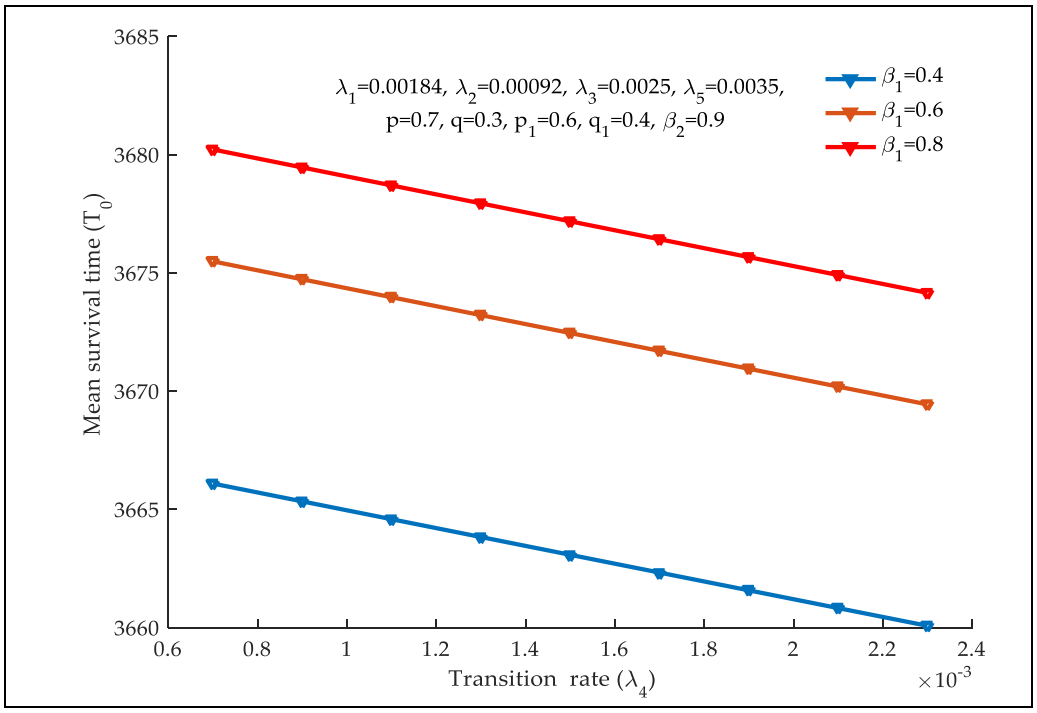


Figure 8. *Mean Survival Time* (T_0) *versus transition rate* (λ_4) *for varying recovery rate* (β_1)

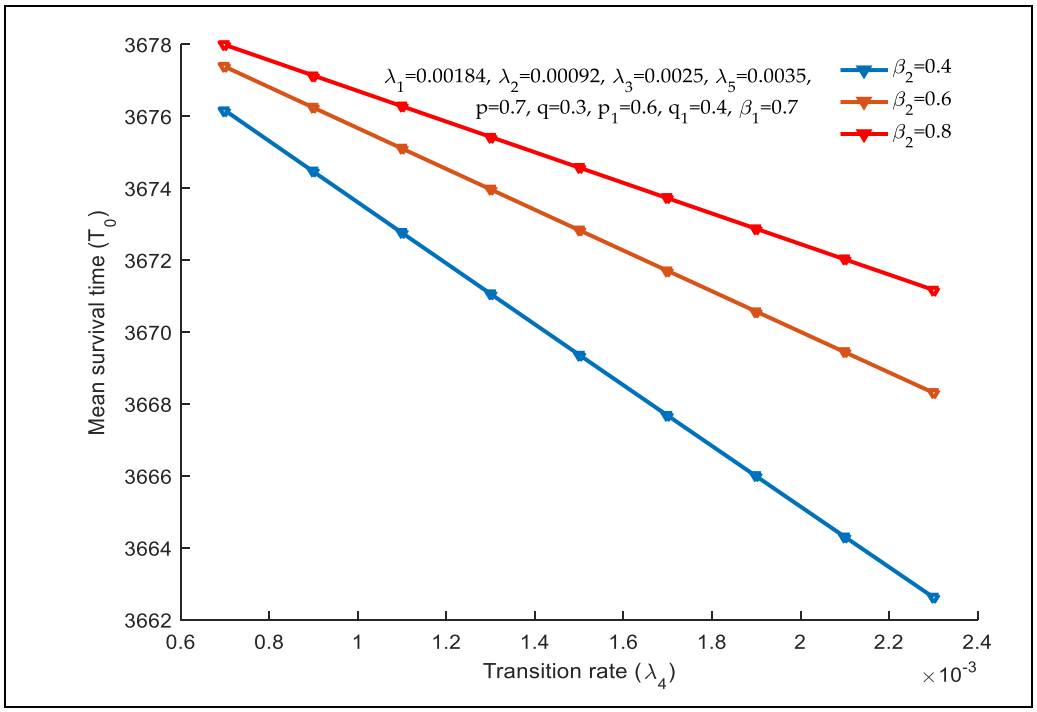


Figure 9. *Mean Survival Time* (T_0) *versus transition rate* (λ_4) *for varying recovery rate* (β_2)

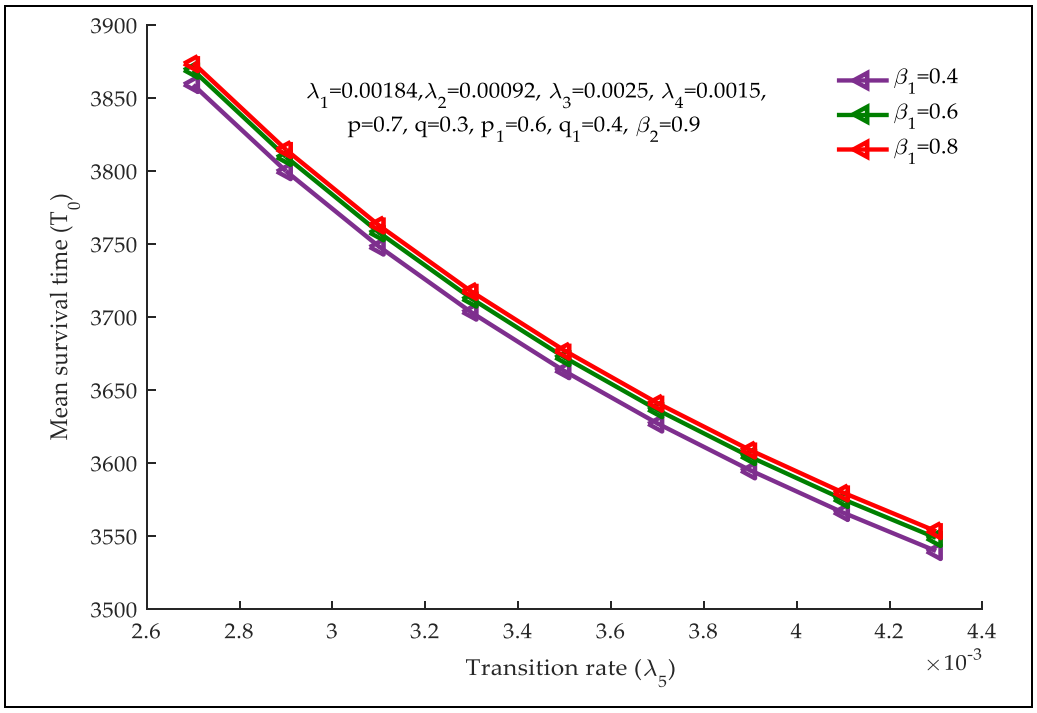


Figure 10. *Mean Survival Time* (T_0) *versus transition rate* (λ_5) *for varying recovery rate* (β_1)

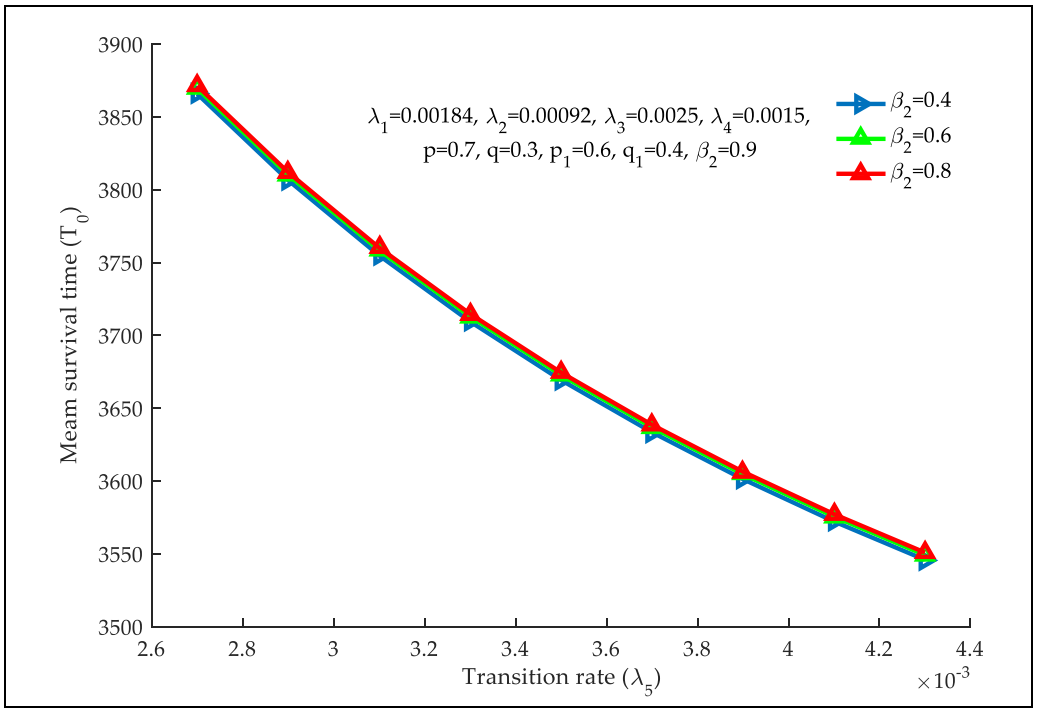


Figure 11. *Mean Survival Time* (T_0) *versus transition rate* (λ_5) *for varying recovery rate* (β_2)

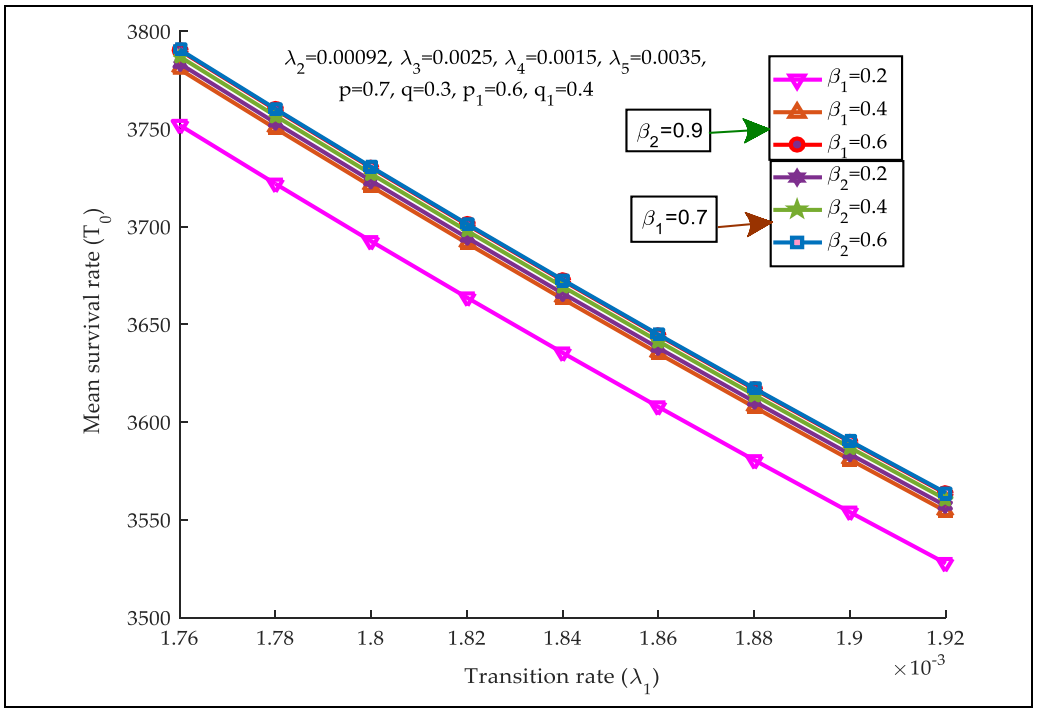


Figure 12. *Mean Survival Time* (T_0) *versus transition rate* (λ_1) *for varying recovery rates* β_1 *and* β_2

The following interpretations have been drawn from the plotted graphs from the figure 2 to figure12. It can be observed that mean survival time decreases as transition rates $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 increases and gives higher values with higher value of recovery rates β_1 and β_2. Further, the mean survival time (T_0) in case of the treatment-2 is higher than that in case of treatment-1.

10. Sensitivity and Relative Sensitivity Analysis

Sensitivity analysis is performed to find out how the variation in involved parameters affect the specific mean survival time under certain specific conditions. Since, there is significance difference between the values of parameters, therefore to compare their effects on mean survival time (MST), relative sensitivity function is used. The sensitivity and relative sensitivity functions for mean survival time (MST) are formulated as under:

$$\pi_k = \frac{\partial(\text{MST})}{\partial k} \text{ and } \delta_k = \pi_k \left(\frac{k}{\text{MST}} \right),$$

where $k = \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \beta_1, \beta_2$

Table1: *Sensitivity and relative sensitivity analysis of MST* (T_0) *with transition rate* (λ_1) *for* $\lambda_2 = 0.00092$, $\lambda_3 = 0.0025, \lambda_4 = 0.0015, \lambda_5 = 0.0035, \beta_1 = 0.7, \beta_2 = 0.9, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

λ_1	$\pi_{\lambda_1} = \frac{\partial(\mathrm{MST})}{\partial\lambda_1}$	$\delta_{\lambda_1} = \pi_{\lambda_1}\left(\frac{\lambda_1}{\mathrm{MST}}\right)$
0.00176	- 2.4884	- 0.2953
0.00178	- 2.4678	- 0.2972
0.00180	- 2.4475	- 0.2991
0.00182	- 2.4274	- 0.3009
0.00184	- 2.4076	- 0.3027
0.00186	- 2.3880	- 0.3045
0.00188	- 2.3686	- 0.3063
0.00190	- 2.3495	- 0.3081
0.00192	- 2.3306	- 0.3098

Table 2: *Sensitivity and relative sensitivity analysis of MST* (T_0) *with transition rate* (λ_2) *for* $\lambda_1 = 0.00184$, $\lambda_3 = 0.0025, \lambda_4 = 0.0015, \lambda_5 = 0.0035, \beta_1 = 0.7, \beta_2 = 0.9, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

λ_2	$\pi_{\lambda_2} = \frac{\partial(\mathrm{MST})}{\partial\lambda_2}$	$\delta_{\lambda_2} = \pi_{\lambda_2}\left(\frac{\lambda_2}{\mathrm{MST}}\right)$
0.00084	- 1.1087	- 0.6019
0.00086	- 1.0776	- 0.6075
0.00088	- 1.0478	- 0.6130
0.00090	- 1.0192	- 0.6183
0.00092	- 9.9172	- 0.6235
0.00094	- 9.6537	- 0.6285
0.00096	- 9.4006	- 0.6334
0.00098	- 9.1574	- 0.6382
0.00100	- 8.9234	- 0.6428

Table 3: *Sensitivity and relative sensitivity analysis of MST* (T_0) *with transition rate* (λ_3) *for* $\lambda_1 = 0.00184$, $\lambda_2 = 0.00092,\ \lambda_4 = 0.0015, \lambda_5 = 0.0035, \beta_1 = 0.7, \beta_2 = 0.9, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

λ_3	$\pi_{\lambda_3} = \frac{\partial(\mathrm{MST})}{\partial\lambda_3}$	$\delta_{\lambda_3} = \pi_{\lambda_3}\left(\frac{\lambda_3}{\mathrm{MST}}\right)$
0.0017	- 1.0919	- 0.0013
0.0019	- 1.0910	- 0.0014
0.0021	- 1.090	- 0.0016
0.0023	- 1.0891	- 0.0017
0.0025	- 1.0881	- 0.0019
0.0027	- 1.0872	- 0.0020
0.0029	- 1.0863	- 0.0022
0.0031	- 1.0853	- 0.0023
0.0033	- 1.0844	- 0.0024

Table 4: *Sensitivity and relative sensitivity analysis* of $MST(T_0)$ *with transition rate* (λ_4) for $\lambda_1 = 0.00184$, $\lambda_2 = 0.00092,\ \lambda_3 = 0.0025, \lambda_5 = 0.0035, \beta_1 = 0.7, \beta_2 = 0.9, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

λ_4	$\pi_{\lambda_4} = \frac{\partial(\mathrm{MST})}{\partial\lambda_4}$	$\delta_{\lambda_4} = \pi_{\lambda_4}\left(\frac{\lambda_4}{\mathrm{MST}}\right)$
0.0007	- 567.6037	- 2.7142
0.0009	- 567.2638	- 3.4879
0.0011	- 566.9242	- 4.2608
0.0013	- 566.5849	- 5.0329
0.0015	- 566.246	- 5.8041
0.0017	- 565.9073	- 6.5746
0.0019	- 565.5689	- 7.3442
0.0021	- 565.2309	- 8.1131
0.0023	- 564.8931	- 8.0011

Table 5: *Sensitivity and relative sensitivity analysis of* $MST(T_0)$ *with transition rate* (λ_5) *for* $\lambda_1 = 0.00184$, $\lambda_2 = 0.00092,\ \lambda_3 = 0.0025,\ \lambda_4 = 0.0015, \beta_1 = 0.7, \beta_2 = 0.9, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

λ_5	$\pi_{\lambda_5} = \frac{\partial(MST)}{\partial\lambda_5}$	$\delta_{\lambda_5} = \pi_{\lambda_5}\left(\frac{\lambda_5}{MST}\right)$
0.0027	- 5.1318	- 0.0927
0.0029	- 4.4481	- 0.0868
0.0031	- 3.8926	- 0.0817
0.0033	- 3.4351	- 0.0771
0.0035	- 3.0537	- 0.0730
0.0037	- 2.7325	- 0.0694
0.0039	- 2.4595	- 0.0660
0.0041	- 2.2254	- 0.0630
0.0043	- 2.0232	- 0.0603

Table 6: *Sensitivity and relative sensitivity analysis of* $MST(T_0)$ *with recovery rate* (β_1) $\lambda_1 = 0.0018$, $\lambda_2 = 0.00092, \lambda_3 = 0.0025, \lambda_4 = 0.0015, \lambda_5 = 0.0035, \beta_2 = 0.9, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

β_1	$\pi_{(\beta 1)} = \frac{\partial(MST)}{\partial\beta_1}$	$\delta_{\beta_1} = \pi_{\beta_1}\left(\frac{\beta_1}{MST}\right)$
0.4	8.5702	0.0023
0.6	3.8331	0.0016
0.8	2.1629	0.0012
1.0	1.3869	9.4735
1.2	0.9643	7.9033
1.4	0.7091	6.7796
1.6	0.5433	5.9357
1.8	0.4295	5.2786

Table 7: *Sensitivity and relative sensitivity analysis of MST* (T_0) *with recovery rate* (β_2) *for* $\lambda_1 = 0.0018$, $\lambda_2 = 0.00092, \lambda_3 = 0.0025, \lambda_4 = 0.0015, \lambda_5 = 0.0035, \beta_1 = 0.7, p = 0.7, q = 0.3,\ p_1 = 0.6, q_1 = 0.4$

β_2	$\pi_{\beta_2} = \frac{\partial(MST)}{\partial\beta_2}$	$\delta_{\beta_2} = \pi_{\beta_2}\left(\frac{\beta_2}{MST}\right)$
0.4	2.5847	7.0676
0.6	1.1526	4.7265
0.8	0.6494	3.5504
1.0	0.4161	2.8430
1.2	0.2891	2.3707
1.4	0.2125	2.0329
1.6	0.1628	1.7794
1.8	0.1286	1.5821

The sensitivity and relative sensitivity analyses of mean survival time are carried out with involved parameters $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \beta_1, \beta_2)$. The sensitivity and relative sensitivity analyses of mean survival time with these parameters are tabulated in table 1 to table 7. Tables 1-5 show that signs of the sensitivity of mean survival time with parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 are negative which implies that increase in these parameters decline the value of MST. Tables 6-7 show that signs of the sensitivity of mean survival time with parameters β_1 and β_2 are positive which lead to the conclusion that increase in these parameters improve the value of mean survival time. As transition rates $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 increase, sensitivity function increases whereas relative sensitivity function decreases and whenever recovery rates β_1 and β_2 increase, sensitivity function decreases.

11. Conclusion

Breast cancer is indeed one of the most common cancers diagnosed in women globally. It is a major public health concern and significant impact on women's health. In the paper, the evaluated expressions for mean sojourn time in the different states of the model gives estimates of the times for patient remains in a particular stage. The investigation through the stochastic analysis of the model on breast cancer considering two types of treatments in various progression stages concludes that the mean survival time lessens with the rise in the rates of transition. It has been observed that mean survival time from tamoxifen plus radiation is higher than tamoxifen only. It is concluded that tamoxifen plus radiation is more effective and useful than only tamoxifen for treatment of early-stage breast cancer.

References

[1] Anthony, W.F., David, R.M., Lee, A.M., Maureen, E.T., Patricia, M., Melania, P., Lorna, M.W. and Ivo, A.O. (2004). Tamoxifen with or without Breast irradiation in Women 50 Years of Age or Older with Early Breast Cancer. *Journal of Medicine*, 351(10):963-970.

[2] Bayer, P., Brown, J.S., Dubbeldam, J. and Broom, M. (2021). A Markov chain model of cancer treatment. doi: https://doi.org/10.1101/2021.06.16.448669.

[3] Chesney, T.R., Jennifer, X. Y., Rajaee, N., Andrea, C. T. and Anthony, W. (2017). Tamoxifen with radiotherapy compared with Tamoxifen alone in elderly women with early-stage breast cancer treated with breast conserving surgery: A systematic review and meta-analysis.

Journal of radiotherapy and oncology,127:1-9.

[4] Cong, C. and Tsokos, C. P. (2009). Markov Modelling of Breast Cancer. *Journal of Modern Applied Statistical Methods*, 8(2):626-631.

[5] Dey, B. and Arun, K. (2018). A Review Article on Breast Cancer. *International Journal of Pharmacy and Pharmaceutical Research*, 11(2):284-298.

[6] Duffy, S.W., Day, N.E. Taba´r, L., Chen, H. and Smith, T. C. (1997). Markov Models of Breast Tumor Progression: Some Age-Specific Results. *Journal of the National Cancer Institute*, 22:93-97.

[7] Grover, G., Prafulla, K. S., Komal, G. and Vikas, S. (2018). Multistate Markov Modelling for Disease Progression of Breast Cancer Patients Based on CA 15-3 Marker. *Thailand Statistician*,16(2):129-139.

[8] Huang, Y., Li, Q., Torres-Rueda, S. and Li, J. (2020). The Structure and Parameterization of the Breast Cancer Transition Model Among Chinese Women. *Values in Health Regional Issues*, 21(C):29-38.

[9] Johnstone, P.A.S., Northon, M. S. and Riffenburgh, R. H. (2000). Survival of Patients with Untreated Breast Cancer. *Journal of Surgical Oncology*, 73:273-277.

[10] Mubarik, S., Sharma, R., Hussain, S.R., Iqbal, M., Nawsherwan, Liu, X. and Yu, C. (2022). Breast Cancer Mortality Trends and predictions to 2030 and its Attributable Risk Factors in East and South Asian Countries. *Original Research*, 9:1-15.

[11] Newman, L.A. (2022). Breast cancer screening in low and middle-income countries. *Best Practice & Research Clinical Obstetrics and Gynecology*, 83:15-23.

[12] Ruiz-Castro, J. E. and Zenda, M. (2020). A General Piecewise Multi-State Survival Model: Application to Breast Cancer. *Statistical Methods & Applications*, 29:813-843.

[13] Schairer, C., Mink, P. J., Carroll, L. and Devesa S.S (2004). Probabilities of Death from Breast Cancer and Other Causes among Female Breast Cancer Patients. *Journal of National Cancer Institute*, 96(17):1311-1321.

[14] Taghipour, S., Banjevic, D., Miller, A. B., Montgomery, N., Jardine, A.K.S., and Harvey, B. J. (2013). Parameter estimates for invasive breast cancer progression in the Canadian National Breast Screening Study. *British Journal of Cancer*, 108:542–548, doi: 10.1038/bjc.2012.596.

[15] Ventura, L., Carreras, G., Puliti, D., Paci, E., Zappa, M. and Miccinesi, G. (2014). Comparison of multistate markov models for cancer progression with different procedures for parameter estimation. An Application to Breast Cancer. *Journal of Epidemiology and public Health*, 11:1-10.

SYNTHETIC RELIABILITY MODELING AND PERFORMANCE ENHANCEMENT FOR MULTI-UNIT SERIAL SYSTEMS: UNVEILING INSIGHTS VIA GUMBEL-HOUGARD FAMILY COPULA APPROACH

Ismail Muhammad Musa[1*] and Ibrahim Yusuf[2]

•

[1]Department of Mathematics and Statistics, Alqalam University, Katsina, Nigeria
[2]Department of Mathematical Sciences, Bayero University, Kano
beanec28@gmail.com
iyusuf.mth@buk.edu.ng

Abstract

This paper presents a comprehensive study of a series-parallel system comprising four interconnected subsystems: subsystem-1, subsystem-2, subsystem-3, and subsystem-4. Subsystem-1 stands as a single unit, subsystem-2 consists of three identical units in active parallel, subsystem-3 involves two identical units in series, while subsystem-4 incorporates two identical units in parallel. The system operates under good conditions, considering various failure rates and repair rates. The investigation employs Laplace transforms and Supplementary variable techniques to analyze the system's performance. Key reliability parameters, including Availability, Reliability, Mean Time to Failure (MTTF), Sensitivity, and Cost, are evaluated for specific values of failure and repair rates. The paper delves into the intricate analysis of a multi-unit series system, focusing on its reliability and performance evaluation. The study employs the Gumbel-Hougard Family Copula approach, a sophisticated and robust methodology to capture the interdependencies among system units. By utilizing this advanced technique, the paper provides a comprehensive understanding of the system's behavior under varying operating conditions. Various reliability and performance metrics, including Availability, Mean Time to Failure (MTTF), and Component Importance Measures, are thoroughly examined, offering valuable insights for optimizing the system's reliability and performance. The results are presented in a clear and visually appealing manner, utilizing tables and figures to aid in the comprehension of the findings.

Keywords: Availability; Reliability; Sensitivity; Mean time to system failure (MTTF); Cost Analysis;

I. Introduction

System reliability refers to the extent to which it can be relied upon to function correctly. Additionally, it encompasses critical aspects such as system usage, maintenance, and strategies for enhancing effectiveness by reducing failure occurrences and minimizing maintenance costs. By improving reliability, the risk of harm to maintenance personnel is reduced, as machine failures can result in significant injuries, revenue losses, reduced production output, and increased maintenance expenses. A key objective of system reliability analysis is to identify vulnerable components and assess the potential impacts of their failures.

The components of a serial or redundant system play a vital role in shaping its reliability and performance. The occurrence and nature of component failures within these systems significantly impact various key metrics, such as overall reliability, mean time to failure, dependability,

availability, and revenue generation. Consequently, it becomes crucial to identify and assess the most critical components with the highest reliability in terms of the aforementioned factors. In today's manufacturing industries, traditional production systems are no longer enough to ensure the continued existence and success of an organization. The focus on steady-state operations is insufficient; constant improvement in the reliability and performance of systems is now essential to meet the demands of enhanced productivity, adaptability, and the ever-changing competitive landscape. Systems engineering, improvement, and setup are increasingly driven by the consideration of performance as a crucial factor. Merely assuming that systems function properly is inadequate; their effectiveness is equally important. Numerous instances in domains like telecommunication systems, industrial, and manufacturing systems have shown that enhancing system reliability and performance can lead to significant savings in disaster mitigation, time, costs, labor, risks, and even human lives. Therefore, reliability and performance analysis surveys are conducted to evaluate existing or planned systems, explore different configurations, and strive for an optimum design setup. The ongoing pursuit of improved reliability and performance is vital for organizations seeking to thrive in today's dynamic industrial landscape.

Numerous researchers have employed diverse methodologies to explore the performance and reliability of various systems, and their findings have demonstrated notable enhancements in operational efficiency. For instance, Teslyuk et al. (2021) proposed models for reliably assessing metrics related to testing the performance of local area networks. Additionally, Rotar et al. (2021) introduced a mathematical approach to determine the reliability of solar tracking systems by considering fault coverage aware metrics. In the realm of reliability analysis for various network systems, several researchers have made significant contributions using diverse methodologies. Bisht et al. (2021) devised an algorithm to compute reliability metrics, component measures, and critical measures for communication networks. Arora et al. (2020) developed models specifically for determining reliability metrics in parallel systems with fault coverage. On the other hand, Bisht and Singh (2019a) focused on analyzing reliability metrics of complex networks using universal generating functions. In the context of distributed networks, Huang et al. (2020) introduced their models for reliability analysis. Furthermore, Bisht and Singh (2019b) employed Markov processes to analyze the reliability measures for enhancing transmission network systems. In a separate study, Bisht and Singh (2020) delved into the analysis of profit and reliability in transmission networks using artificial neural networks and Markov processes.These diverse approaches and methodologies have contributed to advancing the understanding of reliability and performance evaluation in various network systems, paving the way for more robust and efficient designs.

The research landscape on reliability assessment for various systems has witnessed several significant studies employing different methodologies. Ye et al. (2020) conducted an investigation on the reliability of a repairable machine, exploring its behavior under shocks and degradation caused by low-quality feedstocks. Sharifi et al. (2019) tackled a redundancy allocation problem, aiming to optimize the reliability and cost of weighted-k-of-n parallel systems. They employed a combination of a universal generating function, a non-dominated sorting genetic algorithm, and a non-dominated ranked genetic algorithm to determine the reliability and cost of each subsystem. In the field of power systems, Jia et al. (2020) introduced a multi-state decision diagram method for evaluating system reliability. Their approach incorporated a multi-state performance sharing mechanism and warm standby units. Lin et al. (2021) contributed to the reliability modeling domain by establishing a copula-based Bayesian model. This model effectively captures the interdependence between components in parallel systems and allows for the estimation of system failure rates. Additionally, Jia et al. (2021) presented a model for power systems that integrate warm standby and energy storage components. Their reliability assessment was calculated using the multi-valued decision diagram technique. These diverse research endeavors have enriched the understanding of

reliability assessment and paved the way for more robust and efficient designs across different systems and industries.

The realm of reliability analysis has seen diverse studies that explore different methodologies and applications. Pundir et al. (2021) conducted an analysis of reliability metrics for a system comprising two non-identical cold standby units, considering different types of priors for unknown parameters. Kumar et al. (2019a) introduced a novel approach inspired by the hierarchical and fishing behavior of gray wolves (Canis lupus) to enhance the technical specifications of a Nuclear Power Plant's safety system's residual heat removal system (RHRS). In another study, Kumar et al. (2019b) utilized a multi-objective gray wolf optimizer algorithm to optimize the reliability-cost trade-off for a space capsule's life support system. Fuzzy reliability evaluation was the focus of Kumar et al.'s (2020) investigation, specifically for series, parallel, and linear consecutive k-out-of-n: F systems. Hesitant fuzzy sets, triangular fuzzy numbers, and the Weibull distribution were employed to obtain fuzzy reliability measures for different types of systems. Mellal and Zio (2020) proposed a new cuckoo optimization algorithm to address reliability redundancy allocation problems, specifically with a cold-standby strategy. This innovative approach holds promise for optimizing the reliability and performance of redundant systems.

Numerous studies in the field of reliability engineering have shown that effective performance analysis can help to enhanced the reliability, avoid disasters and save time, money, or both. Xie et al. (2021) investigated and examined the performance of a safety system that is vulnerable to cascading failures that cause the appearance of further failures. In the paper, a unique technique for mitigating and preventing cascading failure is provided. Xie et al. (2019) suggested performance and an approximation approach for medium-frequency hazardous failures in safety instrumental systems prone to cascade failures. Yemane and Colledani (2019) offer a method for evaluating the performance of unstable manufacturing systems that takes into account unknown machine reliability predictions. Zhao et al. (2021) investigate and optimize the economic performance of a cold standby system susceptible to -shocks and imperfect repairs, proposing geometric process models to quantify the lifetime and repair time.
Numerous researchers have previously presented copula methods in the field of reliability and performance analysis of systems by examining system performance under various conditions. To name a few, Rawal et al. (2022) have concentrated on the reliability assessment of multi-computer systems consisting of n clients and k-out-of-n: G operational scheme with copula repair policy in the ongoing reliability investigations. Sha (2021) conducted research on a copula approach to reliability analysis for hybrid systems. Through the copula repair approach, Sanusi et al. (2022) estimate the dependability metrics of automated teller machine using Gumbel Hougaaard family copula.Yusuf et al. (2022) focus on reliability assessment and estimation of multi-unit of serial system. Maihulla et al. (2021) used the Gumbel-Hougaard family Copula to model and assess the dependability and performance of solar photovoltaic systems. Yusuf and Sanusi (2023) present copula technique in assessing and estimating the reliability characteristics of automated teller machine system. Maihulla and Yusuf (2022) analyzed the reliability of solar PV system through copula techniques. Yusuf et al. (2021) carried out a study on reliability analysis of distributed systems utilizing copula technique. Singh et al. (2021) examine the performance of a multi-unit k-out-of-n: G system through copula linguistic scheme. Singh et al. (2022) suggest a copula linguistic technique for analyzing the performance and effectiveness of a redundant k-out-of-n: G system with multiple successive state degradation. Abubakar and Singh (2019) analyses the Performance assessment of an industrial system (African Textile Manufactures Ltd.) through copula linguistic approach. Maihulla et al. (2021) used the Gumbel-Hougaard family Copula to model and assess the dependability and performance of solar photovoltaic systems.

The researchers mentioned above have made remarkable contributions in enhancing the reliability and performance of complex repairable systems using various techniques. However, there is a need for a new model that offers substantiated and comprehensive evaluations. With this in mind, this present paper focuses on the reliability and performance analysis of a serial system comprising five components. The paper introduces a novel technique called the copula repair technique to analyze the optimization of reliability and performance in this serial system. The main objective is to predict the system's performance optimization by employing two repair strategies. When the system experiences partial failure, the general repair technique is applied to fix it. On the other hand, if the system encounters complete failure, the copula repair technique is employed to fully recover from the failure.In pursuit of these objectives, the paper develops expressions for availability, reliability, mean time to failure (MTTF), sensitivity, and cost function. Through numerical analysis, the behavior of availability, reliability, and cost function over time is determined. This comprehensive approach aims to shed light on the dynamics of the serial system's performance and reliability, providing valuable insights for its optimization.

II. State Description, Notation, and assumptions

State descriptions

S_0 At the outset, the system is in an optimal operational state, where units B2, B3, and D2 are in hot standby mode, while units A1, B1, C1, C2, and D1 are actively functioning.

S_1 At this point, the system experiences a complete failure as subsystem 1 malfunctions.

S_2 At this moment, the system encounters a partial failure with units B3 and D2 in hot standby mode, units A1, B2, C1, C2, and D1 in working mode, and unit B1 undergoing repair.

S_3 At this point, the system experiences a complete failure as unit C1 malfunctions.

S_4 At this moment, the system faces a complete failure as unit C2 malfunctions.

S_5 At this point, the system encounters a partial failure with units B2 and B3 in hot standby mode, units A1, B1, C1, C2, and D2 in working mode, and unit D1 undergoing repair.

S_6 At this moment, the system experiences a partial failure with unit D2 in hot standby mode, units A1, B3, C1, C2, and D1 in working mode, and units B1 and B2 undergoing repair.

S_7 At this point, the system encounters a partial failure with unit B3 in hot standby mode, units A1, B2, C1, C2, and D2 in working mode, and units B1 and D1 undergoing repair..

S_8 At this moment, the system experiences a complete failure as all the units in subsystem 2 malfunction.

S_9 At this point, the system encounters a complete failure as all the units in subsystem 4 malfunction.

3.2 Notations

- t Stands for Time variable on a time scale.
- s Stands for Laplace transform variable for all expressions.
- λ_1 Stands for Failure rate of the unit in the subsystem 1.
- λ_2 Stands for Failure rate of any unit in subsystem 2.
- λ_3 Stands for Failure rate of the unit C_1 in subsystem 3.

- λ_4 Stands for Failure rate of the unit C_2 in subsystem 3.
- λ_5 Stands for Failure rate of any unit in subsystem 4.
- $\phi(x)$ Stands for Repair rates of the unit of subsystem 1.
- $\phi(y)$ Stands for Repair rate of units in the subsystem 2.
- $\phi(z)$ Stands for Repair rate of the unit C_1 in the subsystem 3.
- $\phi(m)$ Stands for Repair rate of the unit C_2 in the subsystem 3.
- $\phi(n)$ Stands for Repair rate of units in the subsystem 4.
- $P_i(t)$ Stands for the probability that the system is in S_i state at instants for i = 0 to 9.
- $\bar{P}(s)$ Stands for Laplace transformation of the state transition probability P(t).
- $P_1(x,t)$ Stands for the probability that a system is in state S_1 the system is running under repair and elapse repair time is (x, t) with repair variable x and time variable t.
- $E_p(t)$ Stands for Expected profit during the time interval [0, t).
- K_1, K_2 Stands for Revenue and service cost per unit time respectively.
- $P_i(y,t)$ Stands for Probability that the system is in state S_i for i=2, 6, 8, and the system is running under repair and elapse repair time is (y, t), with repair variable y and time variable t.
- $P_3(z,t)$ Stands for the probability that a system is in state S_3 the system is running under repair and elapse repair time is (z, t) with repair variable z and time variable t.
- $P_4(m,t)$ Stands for the probability that a system is in state S_4 the system is running under repair and elapse repair time is (m, t) with repair variable m and time variable t.
- $P_i(n,t)$ Stands for Probability that the system is in state S_i for i=5, 7, 9, and the system is running under repair and elapse repair time is (n, t), with repair variable n and time variable t.

1.3 Assumptions

- At the beginning, all the subsystems are in an ideal working mode
- One unit in subsystem 1, 2, 4 and all units in subsystem 3 are necessary for the system to be in operative mode
- All failure rates are unvarying and considered to undergo exponential distribution
- The repairs undergo a general distribution.
- It is considered that a repaired system performs like a new system and no damage seen during repair.
- Immediately the failed unit gets repaired, it is ready to undergo the task.

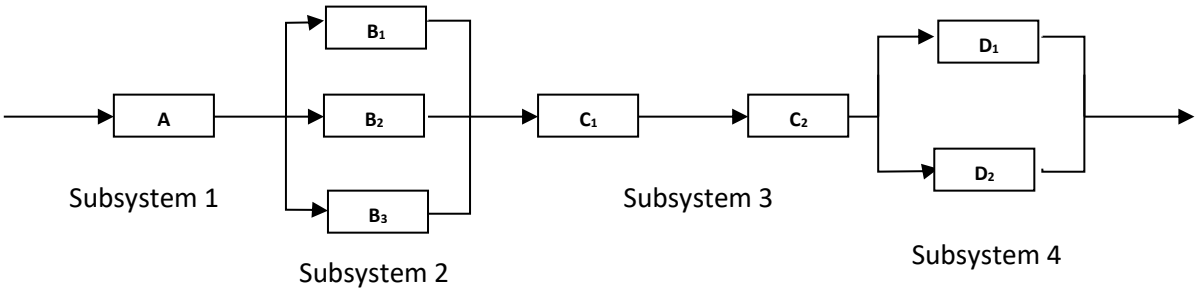


Figure 1 *Reliability Block Diagram of the System*

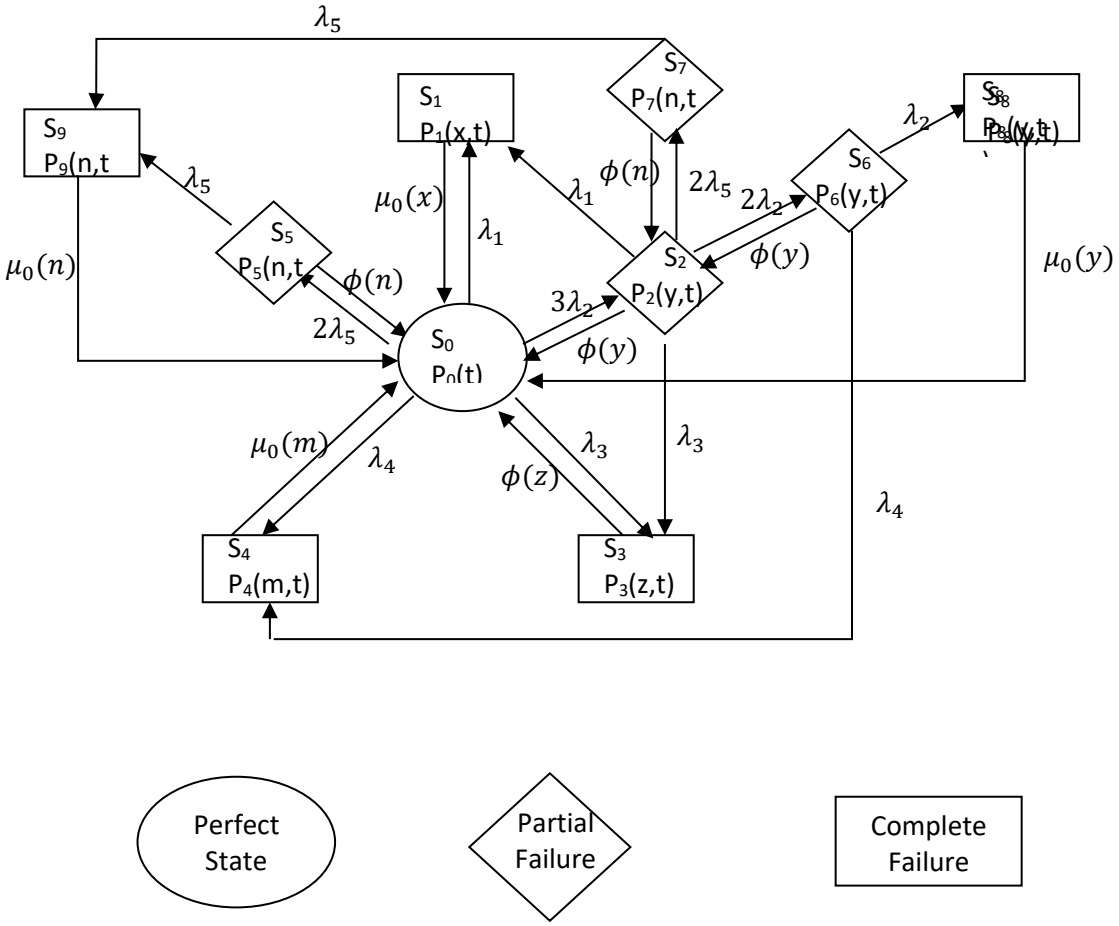


Figure 2 *State Transition diagram of Model*

III. Formulation of the Mathematical Model

With regards to the probability and continuity arguments, the train set of difference differential equations are lump together with the present mathematical model.

$$\left(\frac{\partial}{\partial t}+\lambda_1+3\lambda_2+\lambda_3+\lambda_4+2\lambda_5\right)p_0(t)=\int_0^\infty \phi(y)p_2(y,t)dy+\int_0^\infty \mu_0(z)p_3(z,t)dz+\int_0^\infty \mu_0(m)p_4(m,t)dm+\int_0^\infty \phi(n)p_5(n,t)dn+\int_0^\infty \mu_0(x)p_1(x,t)dx+\int_0^\infty \mu_0(y)p_8(y,t)dy+\int_0^\infty \mu_0(n)p_9(n,t)dn \quad (1)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\mu_0(x)\right)p_1(x,t)=0 \quad (2)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5+\phi(y)\right)p_2(y,t)=0 \quad (3)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial z}+\mu_0(z)\right)p_3(z,t)=0 \quad (4)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial m}+\mu_0(m)\right)p_4(m,t)=0 \quad (5)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial n}+\lambda_5+\phi(n)\right)p_5(n,t)=0 \quad (6)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\lambda_2+\phi(y)\right)p_6(y,t)=0 \quad (7)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial n}+\lambda_5+\phi(n)\right)p_7(n,t)=0 \quad (8)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\mu_0(y)\right)p_8(y,t)=0 \quad (9)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial n}+\mu_0(n)\right)p_9(n,t)=0 \quad (10)$$

Boundary Conditions

$$p_1(0,t)=\lambda_1 p_0(t)+\lambda_1 p_2(0,t) \quad (11)$$

$$p_2(0,t)=3\lambda_2 p_0(t) \quad (12)$$

$$p_3(0,t)=\lambda_3 p_0(t)+\lambda_3 p_2(0,t) \quad (13)$$

$$p_4(0,t)=\lambda_4 p_0(t)+\lambda_4 p_2(0,t) \quad (14)$$

$$p_5(0,t)=2\lambda_5 p_0(t) \quad (15)$$

$$p_6(0,t)=2\lambda_2 p_2(0,t) \quad (16)$$

$$p_7(0,t)=2\lambda_5 p_2(0,t) \quad (17)$$

$$p_8(0,t)=\lambda_2 p_6(0,t) \quad (18)$$

$$p_9(0,t)=\lambda_5 p_5(0,t)+\lambda_5 p_7(0,t) \quad (19)$$

Solution of the Model:

Using Laplace transformation of equations on (1) to (10) together with the initial condition, P0 (0) =1, one can attain.

$$(s+\lambda_1+3\lambda_2+\lambda_3+\lambda_4+2\lambda_5)\overline{p_0}(s)=1+\int_0^\infty \mu_0(x)\overline{p_1}(x,s)dx+\int_0^\infty \phi(y)\overline{p_2}(y,s)dy+\int_0^\infty \mu_0(z)\overline{p_3}(z,s)dz+\int_0^\infty \mu_0(m)\overline{p_4}(m,s)dm+\int_0^\infty \phi(n)\overline{p_5}(n,s)dn+\int_0^\infty \mu_0(y)\overline{p_8}(y,s)dy+\int_0^\infty \mu_0(n)\overline{p_9}(n,s)dn \quad (20)$$

$$\left(s+\frac{\partial}{\partial x}+\mu_0(x)\right)\overline{p_1}(x,s)=0 \quad (21)$$

$$\left(s+\frac{\partial}{\partial y}+\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5+\phi(y)\right)\overline{p_2}(y,s)=0 \quad (22)$$

$$\left(s+\frac{\partial}{\partial z}+\mu_0(z)\right)\overline{p_3}(z,s)=0 \quad (23)$$

$$\left(s+\frac{\partial}{\partial m}+\mu_0(m)\right)\overline{p_4}(m,s)=0 \quad (24)$$

$$\left(s+\frac{\partial}{\partial n}+\lambda_5+\phi(n)\right)\overline{p_5}(n,s)=0 \quad (25)$$

$$\left(s+\frac{\partial}{\partial y}+\lambda_2+\phi(y)\right)\overline{p_6}(y,s)=0 \tag{26}$$

$$\left(s+\frac{\partial}{\partial n}+\lambda_5+\phi(n)\right)\overline{p_7}(n,s)=0 \tag{27}$$

$$\left(s+\frac{\partial}{\partial y}+\mu_0(y)\right)\overline{p_8}(y,s)=0 \tag{28}$$

$$\left(s+\frac{\partial}{\partial n}+\mu_0(n)\right)\overline{p_9}(n,s)=0 \tag{29}$$

Laplace transform of boundary conditions

$$\overline{p_1}(0,s)=\lambda_1\overline{p_0}(s)+\lambda_1\overline{p_2}(0,s) \tag{30}$$

$$\overline{p_2}(0,s)=3\lambda_2\overline{p_0}(s) \tag{31}$$

$$\overline{p_3}(0,s)=\lambda_3\overline{p_0}(s)+\lambda_3\overline{p_2}(0,s) \tag{32}$$

$$\overline{p_4}(0,s)=\lambda_4\overline{p_0}(s)+\lambda_4\overline{p_2}(0,s) \tag{33}$$

$$\overline{p_5}(0,s)=2\lambda_5\overline{p_0}(s) \tag{34}$$

$$\overline{p_6}(0,s)=2\lambda_2\overline{p_2}(0,s) \tag{35}$$

$$\overline{p_7}(0,s)=2\lambda_5\overline{p_2}(0,s) \tag{36}$$

$$\overline{p_8}(0,s)=\lambda_2\overline{p_6}(0,s) \tag{37}$$

$$\overline{p_9}(0,s)=\lambda_5\overline{p_5}(0,s)+\lambda_5\overline{p_7}(0,s) \tag{38}$$

Solving (20) to (29), together with equations (30) to (38) one may attain;

$$\overline{P_0}(s)=\frac{1}{D(s)} \tag{39}$$

$$\overline{P_1}(s)=\frac{(\lambda_1+3\lambda_1\lambda_2)}{D(s)}\left\{\frac{1-\bar{S}_{\mu_0}(s)}{s}\right\} \tag{40}$$

$$\overline{P_2}(s)=\frac{3\lambda_2}{D(s)}\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5)}{s+\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5}\right\} \tag{41}$$

$$\overline{P_3}(s)\frac{=(\lambda_3+3\lambda_2\lambda_3)}{D(s)}\left\{\frac{1-\bar{S}_{\mu_0}(s)}{s}\right\} \tag{42}$$

$$\overline{P_4}(s)=\frac{(\lambda_4+3\lambda_2\lambda_4)}{D(s)}\left\{\frac{1-\bar{S}_{\mu_0}(s)}{s}\right\} \tag{43}$$

$$\overline{P_5}(s)=\frac{2\lambda_5}{D(s)}\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_5)}{s+\lambda_5}\right\} \tag{44}$$

$$\overline{P_6}(s)=\frac{6{\lambda_2}^2}{D(s)}\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_2)}{s+\lambda_2}\right\} \tag{45}$$

$$\overline{P_7}(s)=\frac{2\lambda_2\lambda_5}{D(s)}\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_5)}{s+\lambda_5}\right\} \tag{46}$$

$$\overline{P_8}(s) = \frac{6\lambda_2{}^3}{D(s)}\left\{\frac{1-\bar{S}_{\mu_0}(s)}{s}\right\} \quad (47)$$

$$\bar{P}_9(s) = \frac{(2\lambda_5{}^2+2\lambda_2\lambda_5{}^2)}{D(s)}\left\{\frac{1-\bar{S}_{\mu_0}(s)}{s}\right\} \quad (48)$$

Where $D(s) = s + \lambda_1 + 3\lambda_2 + \lambda_3 + \lambda_4 + 2\lambda_5 - \Big\{(\lambda_1 + 3\lambda_1\lambda_2)\{\bar{S}_{\mu_0}(s)\} + 3\lambda_2\{\bar{S}_{\phi}(s + \lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 + 2\lambda_5)\} + (\lambda_3 + 3\lambda_2\lambda_3)\{\bar{S}_{\mu_0}(s)\} + (\lambda_4 + 3\lambda_2\lambda_4)\{\bar{S}_{\mu_0}(s)\} + 2\lambda_5\{\bar{S}_{\phi}(s + \lambda_5)\} + 6\lambda_2{}^3\{\bar{S}_{\mu_0}(s)\} + (2\lambda_5{}^2 + 2\lambda_2\lambda_5{}^2)\{\bar{S}_{\mu_0}(s)\}\Big\}$

The Laplace transformations of the state transition probabilities that the system is in operative condition and failed condition at any time is as follows:

$$\overline{P_{up}}(s) = \overline{P_0}(s) + \overline{P_2}(s) + \overline{P_5}(s) + \overline{P_6}(s) + \overline{P_7}(s)$$

$$\overline{P_{up}}(s) = \overline{P_0}(s)\left(1 + 3\lambda_2\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5)}{s+\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5}\right\} + 2\lambda_5\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_5)}{s+\lambda_5}\right\} + 6\lambda_2{}^2\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_2)}{s+\lambda_2}\right\} + 2\lambda_2\lambda_5\left\{\frac{1-\bar{S}_{\phi}(s+\lambda_5)}{s+\lambda_5}\right\}\right) \quad (49)$$

$$\bar{P}_{down}(s) = 1 - \bar{P}_{up}(s)$$

IV. Analytical Computations

Availability Analysis

Applying $\bar{S}_{\phi}(s) = \frac{\phi}{s+\phi}$, $\bar{S}_{\mu_0}(s) = \frac{\mu_0}{s+\mu_0}$, $\frac{1-\bar{S}_{\phi}(s)}{s} = \frac{1}{s+\phi}$, $\frac{1-\bar{S}_{\mu_0}(s)}{s} = \frac{1}{s+\mu_0}$ and considering the values of different parameters as $\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_4 = 0.04, \lambda_5 = 0.05, \phi = \mu_0 = 1$ in (49), then with the inverse Laplace transform, the availability may be, obtained as:

$$\bar{P}_{up}(t) = 0.03439655083e^{-2.816787262t} - 0.01746442940e^{-1.293636565t} - 0.002630732952e^{-1.113039421t} + 0.9862252387e^{-0.004836752497t} - 0.0005266271282e^{-1.020000000t} \quad (50)$$

For different values of time variable t= 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, units of time, different values of Pup(t) with the help of (50) may be attained as shown in Table 1 and Figure 3.

Table1. *Variation of Availability with respect to time*

Time	0	1	2	3	4	5	6	7	8	9
Availability	1.0000	0.9772	0.9752	0.9715	0.9672	0.9626	0.9580	0.9534	0.9488	0.9442

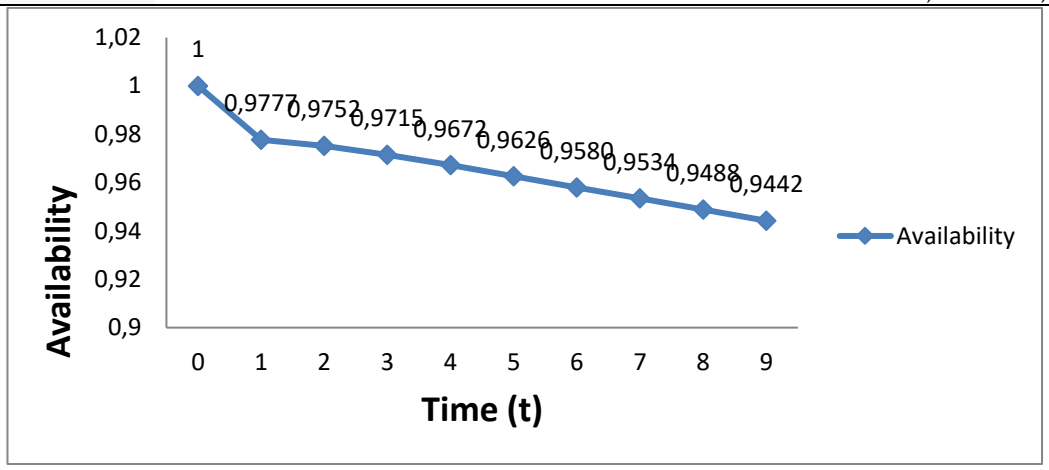


Figure 3 *Availability as function of time*

Reliability Analysis

Using all repair rates, ϕ, μ_0, in equation (49) to zero and for same values of failure rates as

$\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_4 = 0.04, \lambda_5 = 0.05$ And then computing inverse Laplace transform, the reliability for the system may be expressed as;

$$R(t) = 3.e^{-0.2200000000t} + 0.5368421053e^{-0.0500000000t} - 2.547751196e^{-0.2400000000t} + 0.01090909091e^{-0.0200000000t} \quad (51)$$

For, different values of time t= 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.., units of time, different values of Reliability may be attained as seen in Table 2 and graphical display in Figure 4.

Table 2. Computation of reliability for different values of time

Time	0	1	2	3	4	5	6	7	8	9
Reliability	1.0000	0.9248	0.8518	0.7828	0.7184	0.6592	0.6052	0.5561	0.5118	0.4718

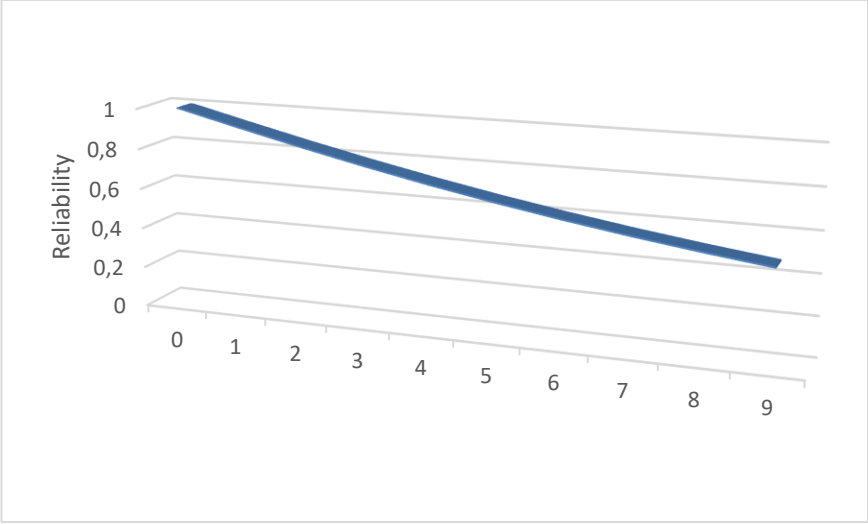


Figure 4 *Reliability as function of Time*

Mean Time to Failure (MTTF) Analysis

Mean time to failure (MTTF) analysis is an important tool in system reliability theory. It provides a measure of the expected time between failures of a system or component, and is often used to assess the reliability and performance of various systems, including mechanical, electrical, and software

systems. MTTF analysis plays a crucial role in system reliability theory. It provides a quantitative measure of a system's expected performance, which can be used to guide design improvements, maintenance activities, and safety procedures.

There are several reasons why MTTF analysis is important and necessary in system reliability theory:

1. Predictive Maintenance: MTTF analysis allows us to predict when a system or component is likely to fail. This information can be used to schedule maintenance activities and prevent costly and unexpected downtime.
2. Design Improvement: MTTF analysis can be used to identify weaknesses in a system's design or components, and guide design improvements to increase reliability.
3. Cost-Effective: MTTF analysis can help companies identify the most cost-effective approach to maintaining and repairing their systems. By prioritizing maintenance activities based on the expected MTTF, companies can optimize their maintenance budget and reduce overall costs.
4. Safety: MTTF analysis is crucial for ensuring the safety of critical systems, such as those used in aviation, healthcare, and nuclear power. By understanding the expected failure rate of these systems, we can design appropriate safety protocols and procedures.

Making all repairs to zero in equation (49), and then considering limit as s tends to zero, the MTTF may be expressed as:

$$MTTF = \lim_{s\to 0} \overline{P_{up}}(s) = \frac{1}{\lambda_1+3\lambda_2+\lambda_3+\lambda_4+2\lambda_5}\left(3 + 6\lambda_2 + \frac{3\lambda_2}{\lambda_1+2\lambda_2+\lambda_3+\lambda_4+2\lambda_5}\right) \quad (52)$$

Setting $\lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_4 = 0.04, \lambda_5 = 0.05$ and varying λ_1 one by one respectively as $0.01, 0.02, 0.03, 0.04, 005, 0.06, 0.07, 0.08, 0.09$, $\lambda_1 = 0.01, \lambda_3 = 0.03, \lambda_4 = 0.04, \lambda_5 = 0.05$ and varying λ_2 one by one respectively as $0.01, 0.02, 0.03, 0.04, 005, 0.06, 0.07, 0.08, 0.09$, $\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_4 = 0.04, \lambda_5 = 0.05$ and varying λ_3 one by one respectively as $0.01, 0.02, 0.03, 0.04, 005, 0.06, 0.07, 0.08, 0.09$, $\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_5 = 0.05$ and varying λ_4 one by one respectively as $0.01, 0.02, 0.03, 0.04, 005, 0.06, 0.07, 0.08, 0.09$, and $\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_4 = 0.04$ and varying λ_5 one by one respectively as $0.01, 0.02, 0.03, 0.04, 005, 0.06, 0.07, 0.08, 0.09$ in (52), the variation of MTTF with respect to failure rates may be attained as seen in Table 3 and corresponding Figure 5

Table 3. *Computation of MTTF corresponding to the various values of failure rates*

Failure rate	MTTF $\boldsymbol{\lambda_1}$	MTTF $\boldsymbol{\lambda_2}$	MTTF $\boldsymbol{\lambda_3}$	MTTF $\boldsymbol{\lambda_4}$	MTTF $\boldsymbol{\lambda_5}$
0.01	13.7576	15.0476	15.0909	15.8596	21.2857
0.02	13.1757	13.7576	14.3934	15.0909	18.7222
0.03	12.641	12.7037	13.7578	14.3934	16.7111
0.04	12.1481	11.8256	13.1757	13.7576	15.0909
0.05	11.6923	11.0823	12.641	13.1757	13.7576
0.06	11.2695	10.4444	12.1481	12.641	12.641
0.07	10.8762	9.891	11.6923	12.1481	11.6923
0.08	10.5095	9.4062	11.2695	11.6923	10.8762
0.09	10.1667	8.9778	10.8762	11.2695	10.1667

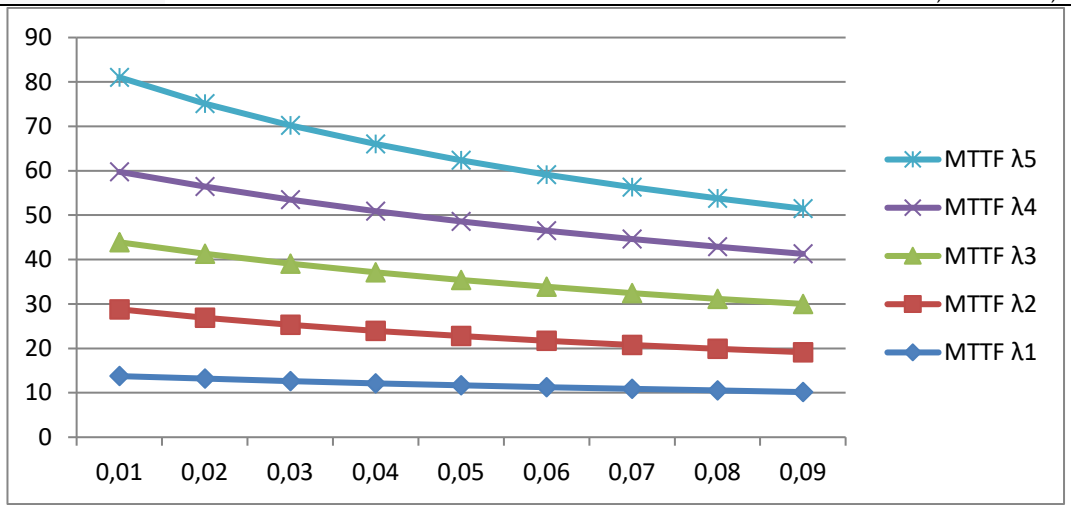


Figure 5 *MTTF as function of Failure rate*

Sensitivity Analysis corresponding to (MTTF)

Sensitivity analysis is an important tool in system reliability theory, as it can help identify critical components, optimize maintenance schedules, quantify uncertainty, and assess risk. Sensitivity analysis is an essential tool in system reliability theory for evaluating the impact of uncertainty and variability in the inputs of a system on the outputs. It involves varying the values of the inputs within a range and analyzing the corresponding changes in the outputs to determine how sensitive the outputs are to the inputs. Sensitivity analysis can help in several ways:

1. Identifying critical components: Sensitivity analysis can help identify the most critical components in a system, those whose failure has the most significant impact on the system's overall reliability. By varying the parameters associated with each component, sensitivity analysis can help determine which components are most sensitive to changes in their input values.
2. Optimal maintenance: Sensitivity analysis can also be used to determine the optimal maintenance schedule for a system. By varying the maintenance parameters and observing the corresponding changes in the system's reliability, it is possible to determine the maintenance schedule that maximizes the system's reliability while minimizing maintenance costs.
3. Uncertainty quantification: Sensitivity analysis can help quantify the uncertainty associated with a system's reliability estimates. By varying the input parameters and observing the corresponding changes in the output, it is possible to determine the range of variability in the system's reliability estimates.
4. Risk assessment: Sensitivity analysis can be used for risk assessment by identifying the most critical inputs in a system and quantifying their impact on the system's reliability. This information can be used to assess the risk associated with different scenarios and identify strategies to mitigate the risk.

Sensitivity in MTTF of the system may be calculated through the partial differentiation of MTTF with respect to the failure rates of the system. By executing the set of parameters as $\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_4 = 0.04, \lambda_5 = 0.05$ in the partial differentiation of MTTF, one may obtain the MTTF sensitivity as seen in Table 4 and corresponding graphs seen in Figure 6

Table 4 *MTTF sensitivity as function of time*

Failure rate	$\frac{\delta(MTTF)}{\lambda_1}$	$\frac{\delta(MTTF)}{\lambda_2}$	$\frac{\delta(MTTF)}{\lambda_3}$	$\frac{\delta(MTTF)}{\lambda_4}$	$\frac{\delta(MTTF)}{\lambda_5}$
0.01	-60.7668	-143.537	-73.1405	-80.7985	-291.581
0.02	-55.7272	-115.978	-66.5235	-73.1405	-225.386
0.03	-51.2903	-95.7819	-60.7668	-66.5235	-179.457

0.04	-47.3635	-80.5049	-55.7272	-60.7668	-146.281
0.05	-43.8715	-68.6513	-51.2902	-55.7272	-121.534
0.06	-40.7523	-59.2593	-47.365	-51.2903	-102.581
0.07	-37.9546	-51.6858	-43.8715	-47.3635	-87.7430
0.08	-35.4357	-45.4864	-40.7523	-43.8715	-75.9093
0.09	-33.1598	-40.3457	-37.9546	-40.7523	-66.3194

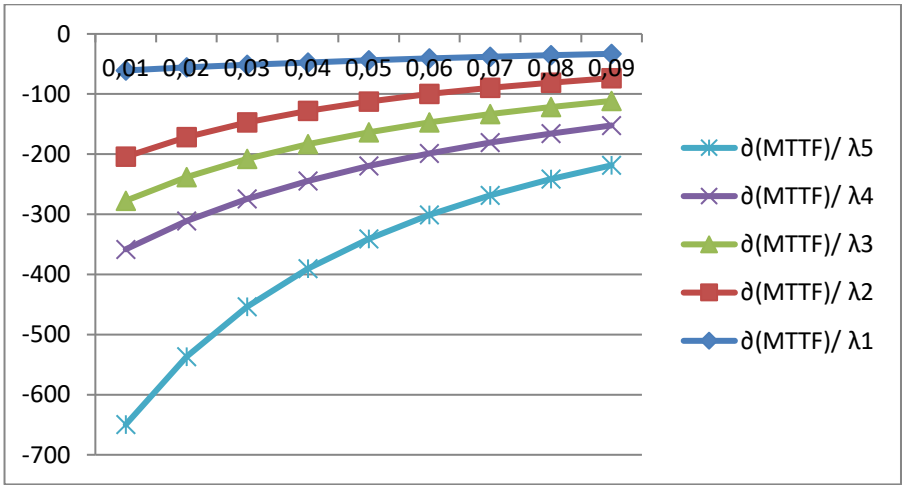


Figure 6 *MTTF Sensitivity with respect to time*

Cost Analysis

Conceding that the service facility be all the time available, then expected profit during the interval [0, t) of the system may be attained by the formula.

$$E_P(t) = K_1 \int_0^t P_{up}(t)dt - K_2 t \quad (53)$$

For the same set of a parameter of (49), one may attain (53). Therefore

$$E_P(t) = K_1(-0.01221127037e^{-2.816787262t} + 0.01350025956e^{-1.293636565t} + 0.002363557752e^{-1.113039421t} - 203.9023579e^{-0.004836752497t} + 0.0005163011061e^{-1.020000000t} + 203.898) - K_2 t \quad (54)$$

Setting $K_1 = 1$and $K_2 =$ 0.6, 0.5, 0.4, 0.3, 0.2 and 0.1 respectively and varying t =0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Units of time, the results for expected profit may be attained as seen in Figure 7.

Table 5. *Expected profit as function of time*

Time(t)	K2=0.6	K2=0.5	K2=0.4	K2=0.3	K2=0.2	K2=0.1
0	0	0	0	0	0	0
1	0.3834	0.4834	0.5835	0.6834	0.7834	0.8834
2	0.7599	0.9599	1.1599	1.3599	1.5599	1.7599
3	1.1333	1.4333	1.7333	2.0333	2.3333	2.6333
4	1.5027	1.9027	2.3027	2.7027	3.1027	3.5027
5	1.8677	2.3677	2.8677	3.3677	3.8677	4.3677
6	2.2280	2.8280	3.4280	4.0280	4.6280	5.2280
7	2.5837	3.2835	3.9837	4.6837	5.3837	6.0837
8	2.9348	3.7348	4.5348	5.3348	6.1348	6.9348

9	3.2813	4.1813	5.0813	5.9813	6.8813	7.7813

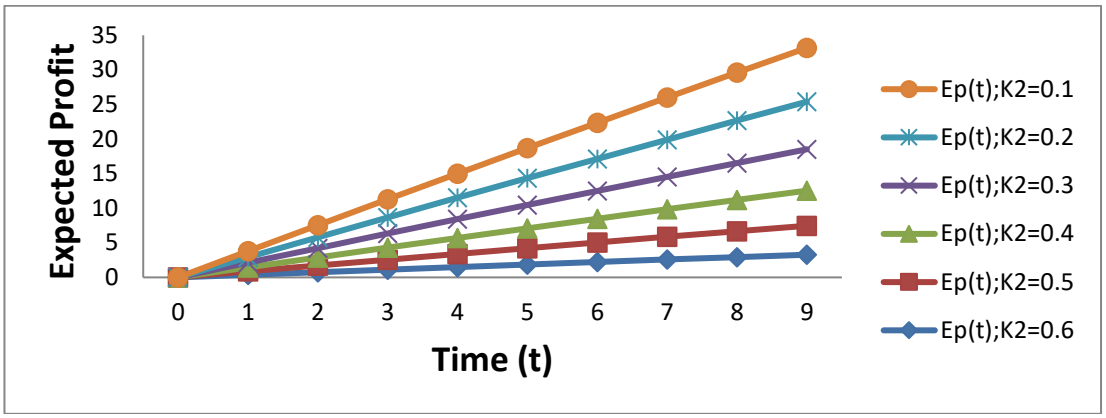


Figure 7 *Expected profit as function of time*

V. Discussion and Concluding Remark

The study aims to investigate and understand the behavior of the complex repairable system under different values of failure and repair rates. Figures 3 and 4 illustrate how the availability of the system changes over time with fixed failure rates at different values. When the failure rates are relatively low (e.g., $\lambda_1 = 0.01, \lambda_2 = 0.02, \lambda_3 = 0.03, \lambda_4 = 0.04, \lambda_5 = 0.05,$), both the availability and reliability of the system decrease gradually, while the probability of failure increases with time until it eventually stabilizes at zero after a sufficiently long period. By analyzing the graphical representation of the model, it becomes evident that the future behavior of the complex system can be reliably predicted for any given set of parametric values. Furthermore, from the observations in Table 3 and Table 4, it is evident that providing repair is more desirable for the system's performance compared to replacement, given the other parameters are held constant. Figure 5 presents the mean-time-to-failure (MTTF) of the system concerning variations in failure rates (. The reciprocal relationship between MTTF and failure rates indicates that these rates significantly influence the system's performance. The sensitivity analysis, as depicted in Figure 4, highlights how the system's sensitivity varies with changes in parameter values. In terms of profit analysis, the study considers a fixed revenue cost per unit (K1=1) and varying service costs (K2= 0.6, 0.5, 0.4, 0.3, 0.2, and 0.1). Figure 7 shows that as the service cost increases, the expected profit decreases.

In conclusion, this comprehensive investigation sheds light on the behavior of the complex repairable system under different conditions and parameter values. It provides valuable insights into the system's reliability, availability, MTTF, sensitivity, and profit optimization, offering practical guidance for decision-making and system performance enhancement.

References

[1] Abubakar, M.I. and Singh, V.V. (2019) 'Performance assessment of an industrial system (African Textile Manufactures Ltd.) through copula linguistic approach', Operation Research and Decisions, Vol. 29, No. 4, pp.1–18

[2] Bisht, S and Singh, S.B. (2019b). Reliability Analysis of Acyclic Transmission Network based on Minimal cuts using Copula in Repair. *Proc. Jangjeon Math. Soc.*, *22*, 163–173.

[3] Jia, H., Liua, D., Lia, Y., Dingb, Y., Liua, M and Peng, R. (2020). Reliability evaluation of power systems with multi-state warm standby and multi-state performance sharing mechanism, Reliability Engineering and System Safety, 204,https://doi.org/10.1016/j.ress.2020.107139

[4] Kumar,A., Pant, S and Ram, M. (2019a). Gray wolf optimizer approach to the reliability-cost optimization of residual heat removal system of a nuclear power plant safety system, Quality and

Reliability Engineering International, 35(7), 2228-2239. https://doi.org/10.1002/qre.2499

[5] Kumar, A., Singh, S.B and Ram, M. (2020). Systems reliability assessment using hesitant fuzzy set, International Journal of Operational Research,Vol.38 No.1, pp.1 – 18 DOI: 10.1504/IJOR.2020.106357

[6] Lin, S.-W. Matanhire, T.B and Liu, Y.-T.(2021). Copula-Based Bayesian Reliability Analysis of a Product of a Probability and a Frequency Model for Parallel Systems When Components Are Dependent. Appl. Sci. 2021, 11, 1697. https://doi.org/ 10.3390/app11041697

[7] Maihulla, A. S., Yusuf, I and Isa, M. S. (2021). Reliability Modeling and Performance Evaluation of Solar Photovoltaic System using Gumbel-Hougaard family Copula. International Journal of Quality & Reliability Management. Vol.1 (1)1- 17.

[8] Maihulla, A. S and Yusuf, I. (2022). Markov Modeling and Reliability analysis of solar photovoltaic system Using Gumbel Hougaard Family Copula. International Journal of Reliability, Risk and Safety: Theory and Application, Vol. 4/ Issue 2/ 2021/ pp. 47-58 DOI : 10.30699/IJRRS.4.2.6

[9] Mellal, M.A and Zio, E. (2020). System reliability-redundancy optimization with cold-standby strategy by an enhanced nest cuckoo optimization algorithm, Reliability Engineering & System Safety, 201, https://doi.org/10.1016/j.ress.2020.106973

[10] Mellal,M.A., Al-Dahidi, S and Williams, E.J. (2020). System reliability optimization with heterogeneous components using hosted cuckoo optimization algorithm, Reliability Engineering & System Safety, 203, https://doi.org/10.1016/j.ress.2020.107110

[11] Pundir, P.S., Patawa, R and Gupta, P.K. (2021). Analysis of Two Non-Identical Unit Cold Standby System in Presence of Prior Information, American Journal of Mathematical and Management Sciences, 40:4, 320-335, DOI: 10.1080/01966324.2020.1860840

[12] Rawal, D. K., Sahani, S. K., Singh, V. V. and Jibril, A. (2022) 'Reliability Assessment of Multi-Computer System Consisting N Clients and k-Out-of-n: G Operational Scheme with Copula Repair Policy; Life Cycle Reliability and safety Engineering, 11(1). pp. 163.

[13] Rotar, R.., Jurj, S.L., Opritoiu, F and Vladutiu, M. (2021). Fault Coverage-Aware Metrics for Evaluating the Reliability Factor of Solar Tracking Systems. *Energies*, *14*, 1074. https://doi.org/10.3390/en14041074

[14] Sha, N. (2021) 'A Copula Approach of Reliability Analysis for Hybrid System', Reliability: Theory and Application, 16(1), pp. 231-242.

[15] Sanusi, A., Yusuf, I and Yusuf, H. A. (2022). Evaluation of reliability characteristics of automated teller machine system using Gumbel–Hougaard family copula repair policies. Life Cycle Reliability and Safety Engineering, 11(4), 367-375

[16] Sharifi, M., Moghaddam, T.A and Shahriar, M. (2019). Multi-objective Redundancy Allocation Problem with weighted-k-out-of-n subsystems, Heliyon, 5, 1-8. https://doi.org/10.1016/j.heliyon.2019.e02346

[17] Singh, V.V., Mohammad, A. I., Ibrahim, K.H and Yusuf,I. (2021). Performance assessment of the complex repairable system with n-identical units under (k-out-of-n: G) scheme and copula linguistic repair approach. *Int. J. Qual. Reliab. Manag.* **38**(5), 10.1108/IJQRM-10-2020-0343.2150047

[18] Singh, V. V., Ismail, A.L., Chand, U and Maiti, S.S. (2022). Performance Assessment of Complex System Under the k-out-of-n: G Type Configuration with k Consecutive Degraded States through the Copula Repair Approach, International Journal of Reliability, Quality and Safety Engineering, Vol. 29, No. 02, 2150047, https://doi.org/10.1142/S0218539321500479

[19] Singh, V. V, Poonia, P. K and Rawal, D. K, (2020). Reliability analysis of repairable network system of three computer labs connected with a server under 2- out-of- 3: G configuration, Cycle Reliability, and Safety Engineering, DOI: 10.1007/s41872-020-00129-w

[20] Teslyuk, V.; Sydor, A.; Karoviˇc, V., ml.; Pavliuk, O.;mKazymyra, I. (**2021**). Modelling Reliability Characteristics of Technical Equipment of Local Area Computer Networks. *Electronics*, *10*, 955. https://doi.org/10.3390/ electronics10080955

[21] Xie, L., Lundteigen, M. A and Liu, Y. (2021). Performance analysis of safety instrumented systems against cascading failures during prolonged demands, Reliability Engineering and System Safety,2021, 216. https://doi.org/10.1016/j.ress.2021.107975

[22] Xie, L., Lundteigen, M.A., Liu, Y., Kassa, E and Zhu, S. (2019). Performance Assessment of Safety-instrumented Systems Subject to Cascading Failures in High-demand Mode, Proceedings of the 29th European Safety and Reliability Conference (ESREL), DOI: 10.3850/978-981-11-2724-3_0318-cd

[23] Yusuf, I., Usman, N. M. and Ismail, A. L. (2022) 'Reliability Estimation of a Serial System Subject to General and Gumbel-Hougaard Family Copula Repair Policies', Reliability: Theory and Application, 17(3), pp.69.

[24] Yusuf, I and Sanusi, A. (2023). Reliability Assessment and Profit Analysis of Automated Teller Machine System Under Copular Repair Policy. In Predictive Analytics in System Reliability (pp. 97-117). Springer, Cham.

[25] Yusuf, I., Ismail, A. L., Sufi, N. A., Ambursa, F. U., Samusi, A. and Isa, M. S (2021) 'Reliability Analysis of Distributed System for Enhancing Data Replication using Gumbel-Hougaard Family Copula Approach Joint Probability Distribution', Journal of Industrial Engineering International, 17(3), pp. 60-68.

A Bayesian Prediction for the Total Fertility Rate of Afghanistan Using the Auto-regressive Integrated Moving Average (ARIMA) Model

Sayed Rahmi Khuda Haqbin and Athar Ali Khan

•

Department of Statistics and Operations Research
Aligarh Muslim University, Aligarh-202002, India
haqbin2020@gmail.com,
atharkhan1962@gmail.com

Abstract

In this article, a simple methodology to predict the total fertility rate of Afghanistan via a Bayesian statistical analysis method has been applied. R- statistical analysis tool is used for data analysis. To forecast, the "bayesforecast" package is needed. It is a substitute package in R for the "forecast" package in the traditional (frequentest) statistical method. The Bayesian data analysis using the specific case of the general auto-regressive integrated moving average model (ARIMA) is processed as follows; As the first step, the stationarity of the given data-set is assessed, the time series has been made stationary by taking differences. After fitting several models, as the most appropriate fitted model, the ARIMA (1, 2, 1) model has been fitted to the data. The accuracy of the fitted model is examined, and thereafter, the developed model is analyzed. The posterior computation is done, using the Markov Chain Monte Carlo (MCMC) simulation method. The method ultimately focuses on drawing relevant inferences including the 16 years prediction, and the results are; in general, found to be satisfactory.

Keywords: Fertility, Total Fertility Rate, Age -Specific Fertility Rate, Auto-regressive Integrated Moving Average Model, Stationarity, Hamiltonian Monte Carlo Algorithm, Prediction.

1. Introduction

Demography is the "study of the human population" [1]. The size of families and children's numbers had a substantially decreased trend over the 20th century. Especially after 1960, the family size and the number of children (fertility) have been showing a broadly favorable pattern for sustainable development. These changes in the population resulted from a declining fertility rate. The economic development level of a nation could be related to its fertility. The developed countries have been showing a lower level of fertility trend, along with a better education system, a revolutionary change in urbanization, greater wealth, and other factors. On the other hand, undeveloped countries have been showing a higher fertility trend, the increased level of fertility associated with a desire for family, labor, and caregivers in old age, lack of education, unawareness of contraceptives, strict adherence to traditional religious beliefs, and lower rates of female employees. Population growth and dependency ratios are directly affected by fertility changes. Over the past century, fertility has been one of the most determinants, of the population growth rate. Infant and child survival will be increased to provide greater access to education and health services, especially for women. By improving the participation of women in the labor force and empowering and reducing the number of children, women have over a lifetime. Low levels of fertility also contributed by improving maternal health, reducing the mortality rate of children, poverty alleviation, and economic development. Our objective; in this article includes

the future prediction of the fertility of a population. Clearly, for this purpose, we must use some summary index of fertility. The Total Fertility Rate is a composite measure of the fertility of a population, it is the most significant component for Fertility projections.

1.1. Total Fertility Rate

The total fertility rate calculates how many children a cohort woman will have. so it represents the average **Family Size** of the population. The total fertility rate is the sum of the Age-Specific Fertility Rate (ASFR) over all ages of the childbearing period and is expressed as "per woman" i.e., the sum divided by 1000. The average number of children a woman would have over her lifetime of reproductive potential, which is between the ages of 15 and 49, is the population's total fertility rate.

Forecasting the fertility rate has a long history of being the interest of demographers. The total number of live births in Australia, which is significant for population growth has been forecasted by using an auto-regressive moving average model [2]. Another most important demo-graphical forecasted characteristic could be the mortality forecast of the U.S. through the application of a time series method [3].

Fitted a bi-variate auto-regressive model which acted as a transform function model for forecasting the total fertility rate and mean ages of childbearing in the U.S annual data [4]. Age-specific mortality in the United States is predicted for the long term, with confidence intervals, using time series methods from 1990 to 2065. [5]. In a different publication, a parameterized model used the age profiles of fertility to quickly and simply describe age-specific rates. Future vital rate profiles are projected using time series models of the parameters, which reflect the temporal patterns of the age profiles [6].

A mixed two methods; statistical time series and mathematical demography and applied to predict stochastic models of fertility rates and the mortality rates of the U.S. dataset then by using the random-matrix product theory forecasted various measures of demography [7]. In a well-done study, a range of non-stationary dynamic factor models were used to jointly estimate the total fertility rate changes within the rather homogeneous clusters of Southern European countries in order to investigate the viability of the forecast. [8]. In this research, three methods for forecasting demographic processes are used. The population forecasting for the demographic data from 1980 to 2005 is particularly intriguing. Exogenous variable structural modelling based on theory and expectation [9]. A recent work applied ARIMA model for the total fertility rate of Pakistan is similar work of the paper from a classical perspective [10]. One another work lately has been conducted by using the two-time series models (ARIMA and ARIMAX) for live birth forecasting in Nigeria, which is a measure of population changes [11]. Numerous methods for predicting fertility have been mentioned; the ones that have been discussed thus far were mostly developed for wealthy nations that have already passed the fertility transition stage and currently exhibit low fertility patterns. Another important consideration is that the developed methods are used from traditional viewpoints. Pre-transition (high fertility), the fertility transition, and post-transition are the three phases of the changes in total fertility rate in a Bayesian projection model for the total fertility rate of country-specific forecasts of all countries (low fertility). The model is based on the most recent data from the United Nations Population Division [12].

In the same way, we aim to model the TFR of Afghanistan via the ARIMA model and achieve it is 10 years forecast by harnessing the attributes of the Bayesian paradigm. The procedures of this article are as follows:

- A real dataset of the total fertility rate (TFR) of Afghanistan has been provided.
- By taking two time differences the non-stationary time series came into the form of stationary series.
- The given data set has been analyzed from the Bayesian perspective using the "Bayesforecast" package, in R statistical language and software.

- "bayesforecast" package in R is a package provides many functions for data analysis in the Bayesian prospective.
- "bayesforecast" package is a Stan-endded pack.
- Stan is probabilistic programming language, which is made easy the data analysis, when there is complex probabilistic model.
- Stan work by taking sample of the equivalent to the posterior distribution fo the model, by different algorithms.
- NUTS(MCMC) algorithm, which is Stan sampling algorithm is applied in this article for the sampling purpose of the posterior distributions.
- The model fitting procedure is done by the help of Bayesian model fitting criteria "WAIC" and "LOOIC".
- The ACF and PACF plot of the original data and differenced data has been given, it could be a proper judge property for model fitting.
- After all the fitted and proper model which is decided for the given data is; ARIMA (1,2,1) model.

2. Methods

In this section, very briefly the method which is applied for this study are expressed. The Total Fertility Rate data-set in Afghanistan has been given. The assigned object is Bayesian prediction of Total Fertility rate in Afghanistan at least for 16 future years. For further analysis, we have been applied Bayesian method of forecasting. The Bayesian substitute **R forecast package** (in classical statistics), is **bayesforecast R package**.The stationarity of the data tested, in a Bayesian standpoint, about the model decided. ARIMA(1,2,1) model selected for further analysis. After all analysis series, we reach to a satisfactory result.

2.1. ARIMA(p,d,q) Model

A time series Y_t is said to follows an auto-regressive integrated moving average (ARIMA) model if the d^{th} difference $Z_t = \nabla^d Y_t$ is stationary and follows (ARMA) process. If $Z_t \sim ARMA(p,q)$ model, we say that Y_t is an ARIMA (p, d, q) process. Where the three parameters; "p" in the Auto-regressive terms represents the current values depending on it's $p - previous$ values (lag values), while "q" in the moving average terms represents the current deviation from the mean depending on q's previous deviations and "d" is the number of differences, by taking, it makes a non-stationary time series in the format of stationary. When stationary is not an issue, then we can define an auto-regressive moving average (ARMA) model as follows;

$$Z_t = e_t + \phi_1 Z_1 + \phi_2 Z_2 + ... + \phi_p Z_t - \theta_1 e_1 - \theta_2 e_2 - ..... - \theta_q e_t \tag{1}$$

$$Z_t = e_t + \sum_{i=1}^{p} \phi_i Z_{t-i} - \sum_{j=1}^{q} \theta_j e_{t-j} \tag{2}$$

Where; e_t is a white noise process, and ϕ_i parameter which is the coefficient of the authoritative terms Z_{t-i}, and θ_j the parameter, which is the coefficient of the moving average e_{t-j} terms. Using the backshift operator, we can write this more succinctly, and $\phi_1, ..., \phi_p$ are the auto-regressive parameters to be estimated, $\theta_1, ., \theta_q$ are the moving average parameters to be estimated and $e_1, ., e_t$ a series of unknown random errors (residuals) that are assumed to follow a normal distribution. The (ARMA) model can be simplifies when Box-Jenkins back-shift operator applied, simplified (ARMA) model as follows:
The auto-regressive AR(p) model;

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - ..... - \phi_p Z_{t-p} = e_t$$

Where;

$\phi(B) = 1 - \phi_1 B - ..... - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - ..... - \theta_q B^q$ are polynomials in B of degrees of p and q respectively. For the process to be stationary the root of $\phi(B) = 0$ must lie outside the unit circle, if the root of $\theta(B) = 0$ is also outside the unit circle, it is called the invertible, there is a unique model corresponding to the likelihood in that case.

we can write Equation 1 in the form of;

$$\phi(B)Z_t = \theta(B)e_t \tag{3}$$

The Equation 3 is the Backshift form of the ARMA (p, q) and the experience suggests in data analysis models with AR and MA components often fit better data than the AR and MA pure. at the same method from Equation 2 we can write;

$$(1 - \sum_{i=1}^{p} \phi_i B^i)Z_t = (1 - \sum_{i=1}^{q} \theta_j B^j)e_t \Rightarrow (\sum_{i=1}^{p} \phi_i B^i)Z_t = (\sum_{i=1}^{q} \theta_j B^j)e_t$$

$$\phi_p(B)Z_t = \theta_q(B)e_t \tag{4}$$

The above (ARMA) model can be extended and written using differences;

$$BZ_t = Z_{t-1}$$

The difference operator $\nabla = 1 - B$ is;

$$BZ_t = (1 - B)Z_t = Z_t - Z_{t-1}.$$

Power in the Backshift operator is equal to the possible difference that a time series can take;

$$B^2 Z_t = B(BZ_t) = B(Z_{t-1}) = Z_{t-2}$$

for the second difference we can write;

$$\nabla^2 Z_t = (1 - B)(1 - B)Z_t = (1 - 2B + B^2)Z_t = Z_t - 2Z_{t-1} + Z_{t-2}$$

eventually the "d" difference can be write;

$$Y_t - \sum_{k=1}^{d} Y_t = (1 - B)^d Z_t \tag{5}$$

"d"is the order of differencing, in that case. The formal ARIMA (p, d, q) model in the Backshif form is produced by replacing "Z_t" in the ARMA model with the differences described above [15].

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B)e_t \tag{6}$$

The Equation 6 is the Backshift format of the ARIMA(p,d,q) model, with "d" difference.

3. Data and Model Specification

There is a yearly data-set, the total fertility rate of Afghanistan, from the year of 1983-2022 reported by the; (World bank open data, free and open access to global development data), the data-set is taken from the; **https://www.macrotrends.net/countries/AFG/afghan/fertility-rate**. For a better accuracy the total fertility rate data-set in Afghanistan for the year of 1983-2022 are presented as bellow;

Table 1: *Afghanistan - Historical Total Fertility Rate form 1982 to 2022.*

1982-1992	7.450	7.454	7.458	7.461	7.465	7.469	7.472	7.474	7.477	7.479	7.482
1993-2003	7.516	7.551	7.585	7.620	7.654	7.560	7.465	7.371	7.276	7.182	7.041
2004-2014	6.900	6.760	6.619	6.478	6.272	6.066	5.859	5.653	5.447	5.269	5.090
2015-2022	4.912	4.733	4.555	4.414	4.273	4.133	3.992				

The allocated projection of Afghanistan's total fertility rate for the next 16 years is the **objective** of this article. Graphical illustration is the first step in characterising a data-set for the analysis technique; thus, as a time series data-set has been provided, the most appropriate plot for the given data-set is time plot, as shown in Figure 1. From a plotted data, the below tips could be found;

- Patterns
- Unusual observations
- Change over time
- Relationship between variables

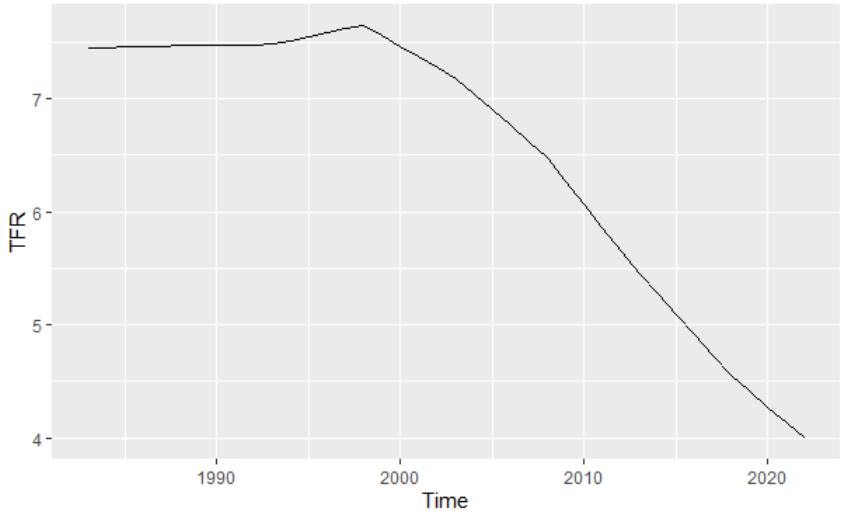


Figure 1: *Displays the time series plot of the Total Fertility Rate of Afghanistan.*

As we mentioned above, Figure 1 shows a constant pattern, from 1982 to 1998, a very short increase, afterwards, then the total fertility rate of the country has been showing a decreasing trend till 2022. Over all Figure 1 is visualized a decreasing trend, therefore, for fitting ARIMA model, testing the stationarity is the versatile condition. In a stationary time series, the statistical properties (mean, variance...) are independent of time, it does not depend on the time at which the series is observed. Time series with seasonality and trend patterns are not stationary, the trend and seasonality affect the value of the time series at different point of time. A series with the same pattern in any point of time is a stationary time series, (e.g: white noise series). Stationary time series will have not predictable pattern in the long- term [14].

3.1. Differencing to eliminate a trend

Differencing is the way to remove trends (non-stationarity) in a time series data. A difference operator is defined as;

$$\nabla y_t = y_t - y_{t-1}, \tag{7}$$

as well as, more broadly, for order d

$$\nabla^d y_t = (1 - \mathbf{B})^d y_t, \tag{8}$$

Where **B** denotes the backshift operator (i.e., $\mathbf{B}^k y - t = y_{t_k} fork \geq 1$) A random walk, for example, is one of the most basic and extensively used time series models, although it is not stationary. We may create a random walk model like follows:

$$y_t = y_{t-1} + w_t, w_t \sim N(0, q). \tag{9}$$

When the difference operator is applied to Equation 9, it produces a time series of Gaussian white noise errors w_t:

$\nabla(y_t = y_{t-1} + w_t)$
$y_t - y_{t-1} = y_{t-1} - y_{t-1} + w_t$
$y_t - y_{t-1} = w_t$
[16].

3.1.1 Making use of the 'diff()' function

There are many methods for making a non-stationary (time series with trend, seasonality, cyclic,...) data-set in the form of stationary. For the total fertility rate data-set, we applied two times differencing to making the time series stationary. Let's use R's diff() function to remove the trend from the total fertility rate time series. We will give some factors for identifying the stationarity of time series data; the correlation between a variable and itself at various time lags is known as the auto-correlation function (ACF). With the linear dependency of other $y_{t-1}, y_{t-2}, ..., y_{t-k}$ eliminated, the partial autocorrelation function (PACF) assesses the linear correlation between a series y_t and a lagged version of itself $y_t + k$. In this part, for the very important point, we will

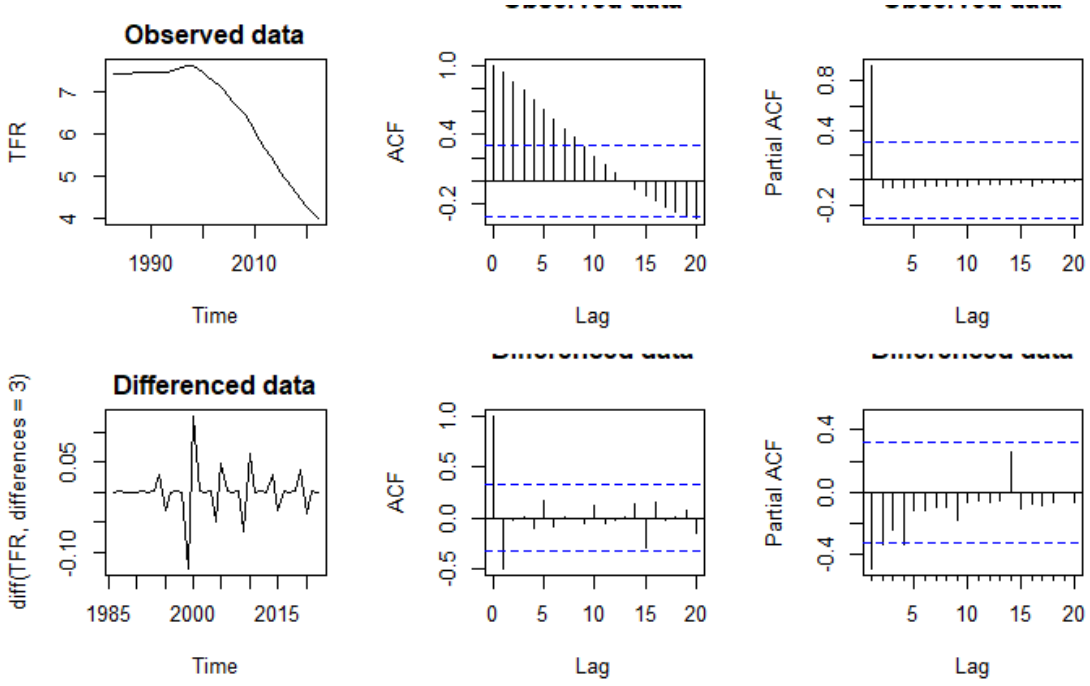


Figure 2: *Displays the observed and differenced Total Fertility Rate of Afghanistan.*

consider the Figure 2, and assesses it for further understanding of the process. In Figure 2 in the first row and first column is the time plot of the observed data, and the rest two other plots are the correlogram of the observed total fertility rate of Afghanistan, ACF and PACF respectively. The things we are going to learn from the correlogram are listed in bellow;

- The ACF at lag 0 is plotted as a
- reference point because it defaults to 1 (i.e., the correlation of a time series with itself).
- The blue horizontal lines represent the CIs for approximately 95%; and
- The auto-correlation is really high.

Figure 2 shows that although the partial autocorrelation at lag-1 is quite strong (equaling the ACF at $lag-1$), the other values at $lags > 1$ are comparatively low, in contrast to what we observed for the ACF. Indicies for the time lag are again real-valued in the PACF plot, but there is no value for lag-0 since it is impossible to eliminate any intermediate autocorrelation between t and $t-k$ when $k=0$, and as a result, the PACF does not exist at lag-0. Now the second row of the Figure 2 shows a differenced and stationary of the total fertility rate of Afghanistan, for model identifying as per ACF and PACF which are given the second row of the Figure 2 second and third columns, the ACF show auto-correlation at lag-1 and PACF shows also correlation to lag-1 and decreasing correlation toward zero. Therefore, we will go to the other model fitting criteria of Bayesian approaches.

3.2. Model Specification

In a next step, after visualizing the time series data, checking stationarity, and removing trend of the data, we will approximate a model for the given data for further statistical analysis. The model used ought to call for the fewest parameters, whose values may be calculated from the observable series. Everything should be made as standardized but not simpler, according to a quotation from Albert Einstein in Parzen (1982, p. 68). For fitting a model to the data, few more steps are very important.

3.2.1 Bayesian Model Fitting Criteria

There are various methods for assessing and contrasting Bayesian models. The posterior predictive checks can be used to evaluate the model's fit to the data. It can be instructive to assess the prediction accuracy of each model under consideration, compare them, and decide what to do next if all of the models under consideration have discrepancies with the data [17].

As the total fertility rate data set is in use,we can do a WAIC and LOOIC criteria test for model selection.

Watanable-Akaike Infromation Criteria (WAIC): It has also been suggested to utilise the log pointwise posterior predictive density, with the posterior variance of the likelihood being used to derive the effective number of parameters.

$$lppd = \sum_{i=1}^{n} log \int p(y_i|\theta)p(\theta|y)d\theta. \tag{10}$$

For more details; [18].

$$el\hat{p}pd_{waic} = l\hat{p}pd - \hat{p}_{waic} \tag{11}$$

$$WAIC = \sum_{i=1}^{n} logE[p(y_i|\theta,y)]) + 2\sum_{i=1}^{n} var_{post}(logp(y_i|\theta)). \tag{12}$$

For more details;[19].

Where the posterior mean of the likelihood of the ith observation is $E[p(y_i|theta,y)])$ and $\sum_{i=1}^{n} var_{post}(logp(y_i|\theta))$ the variance of each variable in the log predictive density summed over n data points. Prior knowledge is taken into account by the WAIC, and the posterior distribution is used in a non-normalway[16].

Leave-One-Out Cross Validation: When using the leave-one-out information criterion (LOO-IC), N-1 observations are used as the validation sample, the process is repeated N times, resulting

in a different observation being predicted each time, and the prediction results are added up to produce an estimate of expected log predictive density (**elpd**) that closely resembles the outcomes that would be obtained by gathering new data and performing the validation. We put the calculations into a R package named **"loo"**.

Table 2: *Displays the WAIC and LOOIC.*

ARIMA Models	LOOIC	WAIC
ARIMA(2,2,2)	-153.5	-155.5
ARIMA(0,2,2)	-155.8	-156.0
ARIMA(2,2,1)	-155.0	-156.8
ARIMA(2,2,0)	-156.0	-156.3
ARIMA(1,2,0)	-156.0	-155.5
ARIMA(0,2,1)	-156.6	-156.9
ARIMA(1,2,1)	**-157.0**	**-156.9**
ARIMA(1,2,2)	-154.8	-156.1

The WAIC and LOOIC are given for several ARIMA models, as per Table 2 ; ARIMA (1,2,1) Model has been showing the proper and best model among all other models, with lesser WAIC and LOOIC.

4. Bayesian Formulation

A Bayesian model is a parametric model as classical(or frequentest) model, but in Bayesian an addition prior probability distribution for the parameters of the model have to be defined. And the parameters of the model treat as a random variable rather to unknown constants. For a Bayesian model, the components are listed as bellow:

- the data, denoted by Z.
- the parameters, denoted by the Greek letters.
- the distributions of the model, given by a specification $f(Z|\theta)$ or $F(Z|\theta)$
- the prior distribution, specification of $f(\theta)$ or $F(\theta)$ or the distribution of the θ .

f is the probability distribution (cdf), F is a symbol which denotes the cumulative distribution density(cdf), θ is the prior.

4.1. Bayes' theorem

As a starting point for the Bayesian analysis, we would like to bring the Bayes' Theorem which is the important formula for Bayesian. Bayes' theorem is based on the conditional probability distribution:

$$f(\theta \mid Z) = \frac{f(Z \mid \theta) f(\theta)}{f(Z)} \tag{13}$$

f(x) is the unconditional (or prior) pdf of x. The proportionality format of the Equation 13 is;

$$f(\theta \mid Z) \propto f(Z \mid \theta) f(\theta) \tag{14}$$

In the Equation 14 the $f(\theta \mid Z)$ is the posterior; the posterior is proportional to the likelihood times to the prior. The model is emphasized that the proportionality of the posterior is specifically related to the θ. As the demonstrator is independent of θ we easily ignore it.

4.2. Choosing the Prior Distribution of the Parameters

For the four parameters the location, scale, auto-regressive, and moving average parameters, , we must describe a four-dimensional joint prior distribution in the general case. However, it is common practise to assume the parameters' prior independence, which means that before examining the data, we are unaware of whether the parameters are positively or negatively connected. We can just define four distinct priors for the four parameters with independence. We investigate weakly informative prior distributions.

$$TFR \sim ARIMA(1,2,1)$$
$$mu_0 = \epsilon_0 \sim t(0,2.5,7)$$
$$\sigma_0 \sim t(0,1,7)$$
$$ar_1 \sim N(0,0.5)$$
$$ma_1 \sim N(0,0.5)$$

we can write the proportional prior density of the defined parameters;

$\pi_1(\epsilon_0) \propto (1+\frac{\epsilon_0^2}{7})^{-4}$
$\pi_2(\phi_1) \propto exp^{-\frac{\phi_1^2}{0.5}}$
$\pi_3(\theta_1) \propto exp^{-\frac{{\theta_1}^2}{0.5}}$
$\pi_4(\sigma_0) \propto (1+\frac{\sigma_0^2}{7})^{-4}$

4.3. Likelihood Function for the Model

$Z : Z_1, Z_2, ..., Z_{T-d}$ are the observations, from Equation 2, the conditional density of Z_t over Z_{t-1}, Z_{t-2},
..., Z_{t-p} is given by;

$$f(Z_t \mid Z_{t-1}, Z_{t-2}, ..., Z_{t-p}; \epsilon_0, \phi_i, \theta_j, \sigma^2) \propto (\frac{1}{\sigma^2}) exp^{\frac{-1}{2\sigma^2}(Z_t - \epsilon_0 - \sum_{i=1}^{p}\phi_i Z_{t-i} - \sum_{j=1}^{q}\theta_j \epsilon_{t-j})^2} \quad (15)$$

From Equation 15 the likelihood function of the given density, can be approximated by its conditional forms;

$$L(Z \mid \epsilon_0, \phi_i, \theta_j, \sigma^2) \propto \prod_{t=p+1}^{T-d} f(Z_t \mid Z_{t-1}, Z_{t-2}, ..., Z_{t-p}; \epsilon_0, \phi_i, \theta_j), \quad (16)$$

For more details; [17].

For simplification we can write;

$$L(Z \mid \epsilon_0, \phi_i, \theta_j, \sigma^2) \propto (\frac{1}{\sigma^2})^{\frac{(T-d-p)}{2}} exp^{-\frac{1}{2\sigma^2}\sum_{t=p+1}^{T-d}(Z_t - \epsilon_0 - \sum_{i=1}^{p}\phi_i Z_{t-i} - \sum_{j=1}^{q}\theta_j \epsilon_{t-j})^2} \quad (17)$$

where; $\phi_i = \phi_1, \phi_2, ..., \phi_p$, $\theta_j = \theta_1, \theta_2, ..., \theta_q$. If one has a sample size of T, and would like to estimate a ARMA(p,q), there are many ways to estimate the parameters of the model: For the two times differenced Total Fertility Rate series in Afghanistan, the Likelihood can be defined from Equation 17 as;

$$f(Z \mid \epsilon_0, \phi_1, \theta_1, \sigma^2) \propto (\frac{1}{\sigma^2})^{\frac{(T-2-1)}{2}} exp^{(-\frac{1}{2\sigma^2})\sum_{t=2}^{T-2}(Z_t - \epsilon_0 - \phi_1 Z_{t-1} - \theta_1 \epsilon_{t-1})^2} \quad (18)$$

where Z_t obviously ∇y_t [18].

4.4. Posterior Distribution Function of the Model

As, it has been discussed in Section 4.1, Bayesian approach is the implication of the Bayes'theorem for further details Section 4.1. The prior and likelihood for the given data-set (total fertility rate) have been defined, now by applying the Bayes' theorem we can reach to the posterior distribution

of the Bayesian distribution of the data-set, once the posterior defined, it will be easy to apply any probabilistic programming language for fitting the posterior in a model and estimate the parameters. From Section 4.1, the posterior;

posterior $\propto$ **likelihood** $\times$ **prior**
log(posterior) = **log(prior)** + **log(likelihood)**

$$p(\epsilon_0, \phi_1, \theta_1, \sigma^2 \mid Z) \propto (1 + \frac{\epsilon_0^2}{7})^{-4} exp^{-\frac{\phi_1^2}{0.5}} exp^{-\frac{\theta_1^2}{0.5}} (1 + \frac{\sigma^2}{7})^{-4} \times (\frac{1}{\sigma^2})^{\frac{(T-3)}{2}} exp^{(-\frac{1}{2\sigma^2}) \sum_{t=2}^{T-2} (Z_t - \epsilon_0 - \phi_1 Z_{t-1} - \theta_1 \epsilon_{t-1})^2} \quad (19)$$

From Equation 19, it is obvious which eventually we reach to a joint joint posterior distribution of all parameters, we can find the marginal density function of each parameter and then generate sample. As we obtained a joint posterior distribution of all parameters, we can find the marginal density function of each parameter and then generate sample. In this article, which is the total fertility rate data-set, in a Bayesian approach, we are going to analyse the data-set and estimate the parameters. As we know the most important step for a Bayesian data analysis is; to create a joint posterior function of all the parameters, now we have the joint posterior function of the parameters. There are many computational tools for the Bayesian data analysis, as we will discuss in Section 5, we only run the **Code** in **R**, in **"bayesforecast" package' fucntion** $stan_sarima()$ fucntion, which is a Bayesian **forecast'package** in **R**, based of **Stan**, and using the **MCMC** simulation method, the sampling algorithm in this **function** is **NUTS(MCMC).**

4.5. Markov Chain Monte Carlo (MCMC)

The modern method for approximating complex forms of the posterior distribution is the Markov Chain Monte Carlo (MCMC) simulation method. The concept is akin to treating the posterior distribution as a population and then drawing samples from it repeatedly. When you draw a large enough sample (say 1,000), the sample distribution should be extremely close to the population distribution, as you learnt in basic statistics. The samples drawn in the above comparison are correlated, thus if the first sample is high, the second one is more likely to be high as well. This is necessary since there is no direct way to take samples from the posterior distribution, which typically has a very complex structure; instead, we have various procedures that can lead us to the posterior indirectly. Correlation between samples is usually not a big deal, except that we need to draw extra samples to compensate. In Sections 5.1 and 5.2 are given the values of the parameters and in Section 5.2 plots are showing the trance plot and the sample distribution of 4,000 samples using MCMC. The shape of the MCMC sample distribution approximates the shape of the posterior distribution [19].

5. Bayesian Inference Using NUTS(MCMC)

NUTS is an extension of Hamiltonian Monte Carlo (HMC) which is a type of MCMC. NUTS can even sample from the model with a greater number of parameters, it is a strength of NUTS sampler algorithm. Stan is a very good probabilistic programming language, the only disadvantage is that, Stan cannot sample for discrete parameters. NUTS uses a recursive algorithm to build a set of likely candidate points that spans a wide swath of the target distribution, stopping automatically when it starts to double back and retrace its steps. Hamiltonian Monte Carlo is a Markov Chain Monte Carlo (MCMC) algorithm that avoids the random walk behavior and sensitivity to correlated parameters that plague many MCMC methods by taking a series of steps informed by first order gradient information [20].

We will describe how to give the data to the model created by the R package "bayesforecast" in this part. For forecasting, the "bayesforecast" package in R is a Stan end-packed package. We must use R to give the data as a certain type to the "bayesforecast" package's function for ARMA

model fitting in order to accomplish this and obtain the outcome. In "bayesforecast" package, the model can be estimated using the guidelines below:

```
stan_sarima(
ts,
order = c(1, 0, 0),
seasonal = c(0, 0, 0),
xreg = NULL,
period = 0,
chains = 4,
iter = 2000,
warmup = floor(iter/2),
adapt.delta = 0.9,
tree.depth = 10,
stepwise = TRUE,
prior_mu0 = NULL,
prior_sigma0 = NULL,
prior_ar = NULL,
prior_ma = NULL,
prior_sar = NULL,
prior_sma = NULL,
prior_breg = NULL,
series.name = NULL,
...)
```

Using the function *stan_sarima*()'s from "bayesforecast" package in R the default setting are; *ts* =, The ARIMA(p,d,q) order, chains is equaling 4, iterations per chain equaling 1000, and 1000 warmups. used as a warmup to eliminate dependencies in the initial settings. We fit the selected model to the total fertility rate; by the function above.

5.1. The Estimated Results

The given Table 3: shows the numerical estimated values for the parameters of the fitted model ARIMA (1,2,1) on the total fertility rate of Afghanistan.

Table 3: *Displays the estimated posterior distribution of the parameters*

	Mean	Se	%5	% 95	N_ess	Rhat
Intercept	-0.0103	0.0001	-0.0215	-0.0007	4282.407	1.0000
σ_0	0.0270	0.0001	0.0221	0.0329	3940.089	1.0015
ϕ_1	0.9209	0.0008	0.8339	0.9872	4034.682	1.0000
θ_1	0.0658	0.0024	-0.1909	0.3134	4109.169	1.0005
loglik	77.7385	0.0323	73.8071	80.3939	3808.628	1.0034

Table 3 provides a summary of the statistics for the marginalized posterior distribution as well as the MCMC convergence diagnostics. In line 3, for instance, the marginalized posterior distribution for parameter ϕ_1 is presented. Now, let's examine each column. The name of the parameter is displayed in the first column. The posterior mean, also known as the mean of the marginalized posterior distribution, is the value in the second column. The arithmetic mean of all the **draws** in this case, a total of **4000** is used to calculate this. For instance, the posterior mean of θ_1 is **0.0658** , and this finding shows that, on average, the baseline annual income grows by **0.0658**with each extra year of work experience. The Monte Carlo standard errors of the Mean are shown in the third column as Se. By dividing the SD by the square root of $N - eff$, it is calculated. The quantiles of the posterior marginalized distributions are displayed in columns

four and five. These are calculated using the draw's quantiles. Stan determines the effective number of drawings as N_eff in column 6 using the autocorrelation of the draws. We believe this value should be at least 100 in order to estimate distributions and perform other statistical calculations. Additionally, a small value for this parameter suggests that it has probably not yet converged; this information could be used to improve the model. An indication of MCMC convergence, $R - hat$, is computed for each parameter and is shown in column seven. Typically, it compares the sample variance of each chain using findings from many chains. According to Section 11.4 [21].

The MCMC is said to have converged if it meets the condition $R_hat < 1.1$ for all the parameters. We must confirm that the MCMC is converged before evaluating the results. Beginners frequently skip over this stage in favour of moving on to the analysis' next step, when they could use the posterior mean, plot the histogram of draws, or quickly interpret the findings using an MCMC that hasn't even begun to converge. Without a doubt, avoid this. The $try - and - error$ procedure must be repeated numerous times before the MCMC converges. Eventually the ARIMA (1,2,1) model for the time series data as per Equation 1 can be written:

$$Y_t = -0.0103 + 0.92y_{t-1} + 0.66e_{t-1} \tag{20}$$

From Table 3, the simulation estimate for $E(\phi_1|y)$ is 0.92009, according the 90% credible interval which is (0.8339 , 0.9872) of the Table 3 statistically significant. in the same way, the simulation estimate for $E(\theta_1|y)$ is 0.06258, it is not statistically significant 90% probability interval for the variable(parameter) θ_1 is (-0.1909 , 0.3134) . Intercept and the posterior standard deviation values are given respectively. In Bayesian statistics, the credible interval is analogous to the confidence interval, but with radically distinct implications. A 90% credible interval, for example, is an interval that has a 90% chance of containing the true value of the parameter, an interpretation that is frequently and incorrectly connected with the confidence interval. Because the population parameter is fixed, frequentists cannot use probability for the parameter; instead, only the sample is probabilistic, and a 90% confidence interval must be interpreted in the sense that 90% of the intervals constructed with repeated sampling will contain the true parameter. As we can see in the Table 3 a 90% credible interval is given for each parameter of the model.

5.2. The estimated results figure

The sported plot of the estimated values are given as Figure 3;

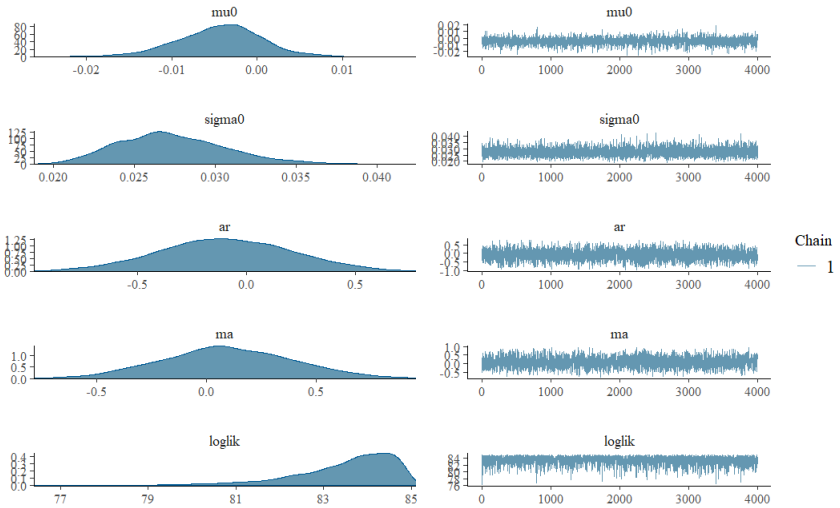


Figure 3: *Displays the supported of the estimated parameters.*

The probability density functions of the marginalised posterior distributions, which are

calculated from the draws after warm-up in each chain, are shown on the left column side. The right column, which additionally displays the trace plots following warm-up. These plots show that all of the chains are oscillating around a set of values after the warm-up, which indicates that the MCMC has converged.

6. Model Checking

A common procedure in Bayesian statistics for validating your model is the posterior predective check [22]. We must first talk about the posterior predictive distribution in order to comprehend the posterior predictive check. The posterior predictive distribution in a mathematical notation.

$$p(\hat{y}|y) = \int_{\theta} p(\hat{y}|\theta, y)p(\theta|y)d\theta \tag{21}$$

From the Equation 21, θ is the parameter, in our posterior distribution which is discussed in the Section 4.4, θ is the set of all parameters θ=(ϵ, σ, ϕ_1, θ_1). The formula is merely provided for your convenience because we won't be using it to make posterior predictions. As an alternative, we will employ simulations to roughly approximate the posterior predictive distribution of $\hat{y}$, which is really an extension of the posterior distribution of θ's approximation. The posterior predictive check simply makes a comparison between the observed data and the model's prediction ẇe proceed to review the the graphical posterior predictive distribution of the total fertility rate.

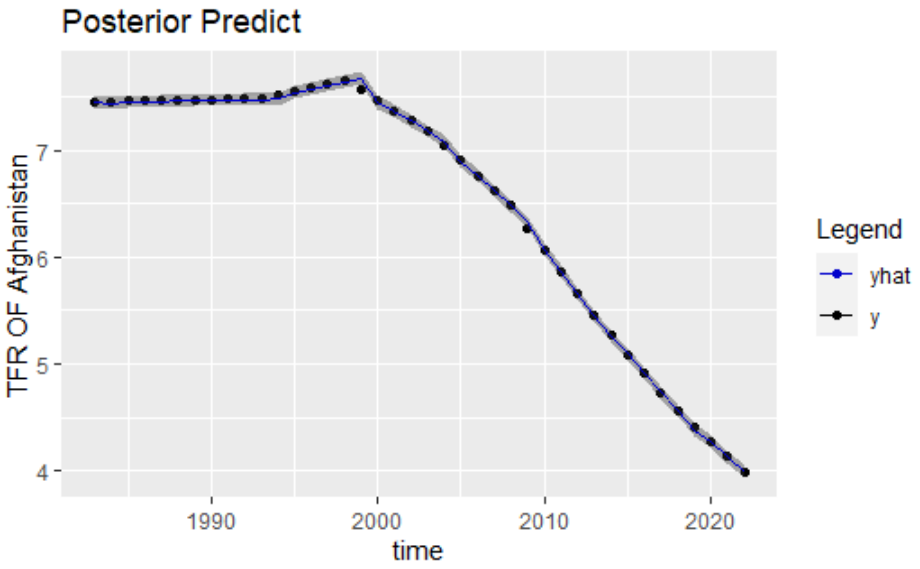


Figure 4: *Displays the posterior predicted.*

From Figure 4, we can see that; the fitted model on the generated draws $y - hat$ is very good fit. Therefore, we can results, the fitted model is a proper model for the given data-set.

6.1. Check Residuals

We may examine the residuals' autocorrelation using "$check_residuals()$" in the forecast package will automatically run a test and display a few common diagnosis charts. Figure 5 is the common diagnostics charts of the total fertility rate.

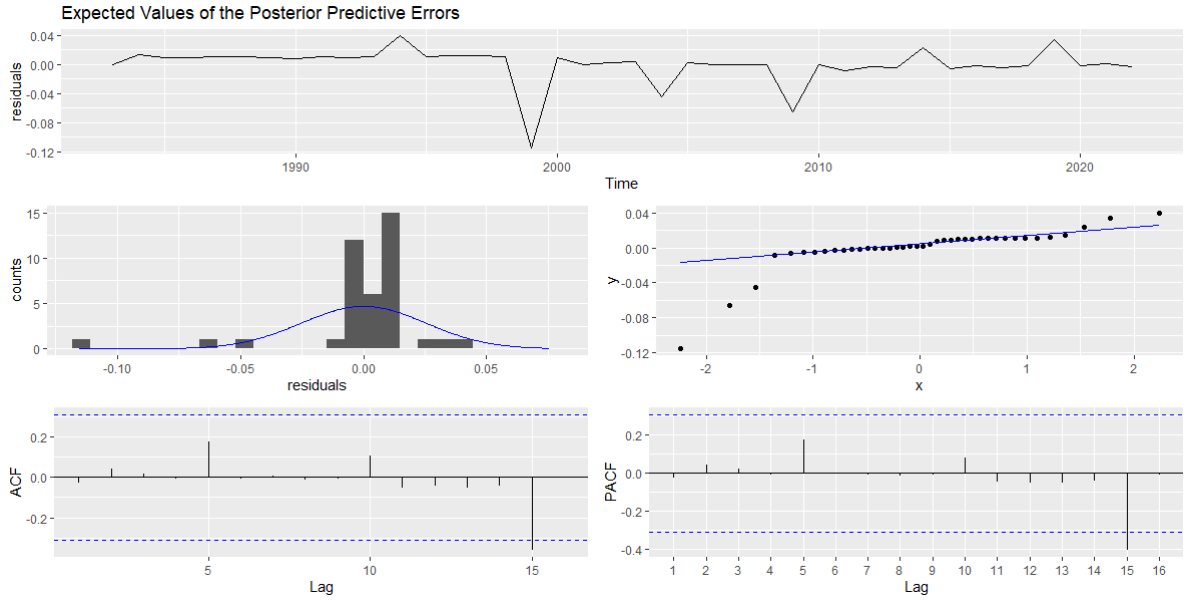


Figure 5: *expected values of the posterior predictive errors.*

The residuals series (upper part) shows that the given series is white noise. The dash blue lines are indicated the correlations are significantly not zero. In a white noise series, we expect each correlation to be zero. For a white noise series, we expect 95% of the spikes to lie within the $\frac{2}{\sqrt{T}}$, where T is the length of the series. If more than 5% of the spikes are not in the blue bounds, then the series is not white noise probably. The histogram and quantile graph (middle part) show that the model is around normally distributed. Finally, the residual auto-correlation presents have showed that there is no correlation between the series segments. A summary result from the above interpretation is; we can step ahead and do our prediction based, as all the evidences present a good and fair series.

7. Prediction

Forecasting can be used in many situations: for making a schedule to the staff of a call center for some times in the future, we need to do a forecast as per call going to be received. For an inventory or stocking there is need for forecast of stock requirement. Forecast can be for few minutes beforehand till several years in advance. For making an efficient and effective plan, need to do forecasting. Forecasting can be done very easy and sometimes very tough. predictability for an event or quantity related to several factors;

- Understanding the factors to contribute on the event.
- How much data are accessible.
- Is there any effect of forecasting to the things are going to be forecasted [23].

Forecasting fertility; as the demographic characteristics (fertility, mortality, migration) of the human population is significant for socio-economics planning. There are mentioned lots of fertility forecasting methods, the methods discussed on the Section 1.1. Mostly, they have developed for the countries which are rich today and already passed their fertility transition, and at the current time they have low fertility patterns. Another point, is the developed methods are applied from the classical perspectives. A Bayesian projection model for the total fertility rate of country-specified projections of all countries have been done, the developed model consisted of three phases for the changes in total fertility rate, pre-transition (high fertility), the fertility

transition, and post-transition (low fertility). The model has been built on the United Nations Population Division's current [24].

In the same way, we aim to model the Total Fertility Rate of Afghanistan via the ARIMA model and achieve it is 16 years forecast, by harnessing the attributes of the Bayesian paradigm.

7.1. Forecasted Results

In the Table 4, which is numerical forecasted years, from 2022-2037 the 16 years of the total fertility rate of Afghanistan.

Table 4: *Future forecasting*

	Point Forecast	Lo.0.9	Hi.0.9	Lo.0.95	Hi.0.95
2023	3.853670	3.802970	3.903317	3.794032	3.915283
2024	3.718556	3.668241	3.768550	3.656223	3.776726
2025	3.586596	3.533241	3.639957	3.520394	3.650888
2026	3.455740	3.396801	3.510490	3.386231	3.520109
2027	3.332134	3.261010	3.388971	3.242005	3.400166
2028	3.207198	3.120882	3.278480	3.100771	3.290015
2029	3.085855	2.983132	3.162357	2.957278	3.173630
2030	2.967033	2.836914	3.055491	2.812259	3.068283
2031	2.851968	2.693900	2.959655	2.657657	2.972596
2032	2.739253	2.551682	2.859017	2.500849	2.878114
2033	2.626087	2.400823	2.762809	2.352485	2.776872
2034	2.519611	2.262385	2.677497	2.209586	2.693992
2035	2.411963	2.088992	2.588108	2.044437	2.604896
2036	2.311417	1.955690	2.510772	1.883435	2.533944
2037	2.210127	1.821591	2.432696	1.725132	2.452562
2038	2.114983	1.655969	2.371578	1.576980	2.388008

The variable's values we are are interested to forecast is unknown, therefore, it is a random variable. The random variable can take a range of random values. In Table 4, the first column is the years of forecast, the second column is the forecasted point or mean of the future values of the total fertility rate (which is the random variable). From columns 3-6 are the credible interval of 90% to 95%. For more details about the credible interval Section 5.1. The value for the total fertility rate of Afghanistan which is under consideration is given in Table 1, form 1982-2022. Afterwards, the value of the fertility of the country is random variable, because it is unknown and we are going to do forecast. In a Bayesian prediction we reach to the forecasting point which is the Table 4 is the numerical values of the forecasted amount.

7.2. Forecasted Results Figure

From Table 4, which is a 16 years Bayesian prediction of the total fertility rate of Afghanistan. Year of 2038 is showing a very good results. 2.1 is the Replacement fertility level: Total fertility of approximately 2.1 children per woman. This value shows the average number of children a woman would need to have in order to reproduce herself by having a daughter who reaches reproductive age. Now it is time to go to the graphical representation of the results. Figure 6, is going ot support the numerical values of Table 4, which is the forecasted values of the total fertility rate of Afghanistan from 2023-2038.

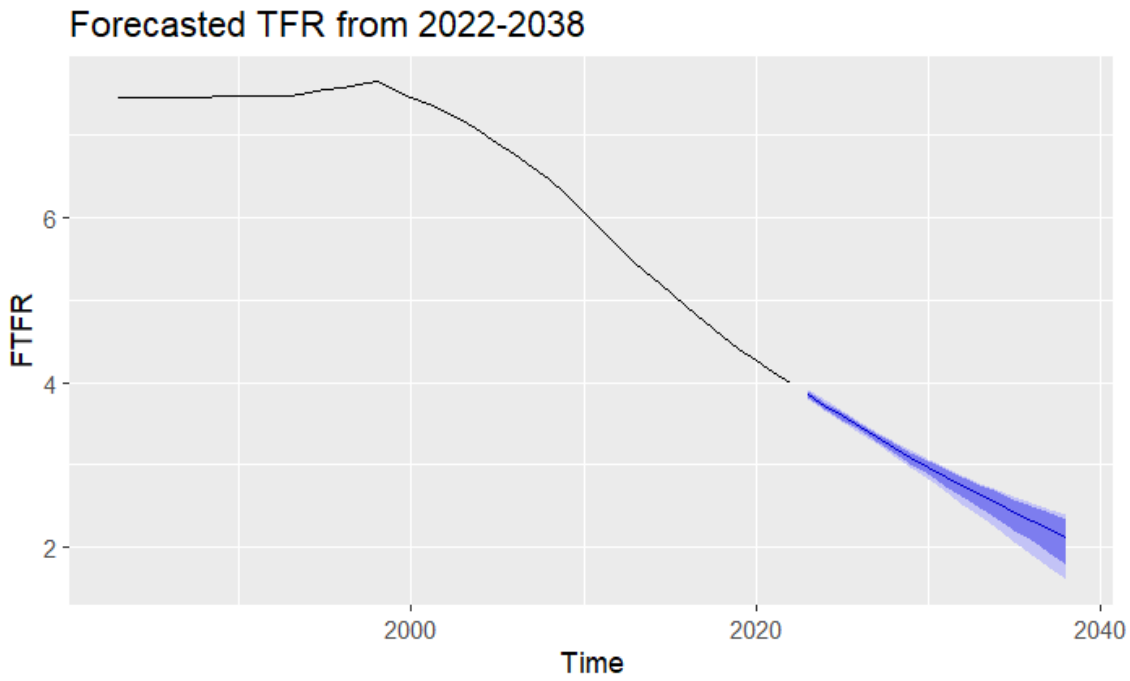


Figure 6: *Future prediction*

Figure 6 are consisted of two parts, from 1982-2022 the observed data, and 2023-2038 is the forecasted values of the total fertility rate of Afghanistan. The blue line in the Figure 6 shows the forecasted part of the value of the total fertility rate, which is the averages possible future value, and numerically was shown in the Table 4. In the Figure 6 it is obvious, there is two types of shadowed area, the 90% and the 95% credible intervals of the forecasted values. Forecast accuracy also can be analyzed form the shadowed areas, as much as, our forecasted time increase the credible interval or probability interval increase. And on the same way as much as years of forecast is near to the observed value the credible interval is more narrow and it shows the certainty or accuracy of the forecast.

8. Conclusion

In this article, we were started the work for doing Bayesian data analysis, using the total fertility rate of Afghanistan data-set. We discussed selecting a statistical model that is assumed to have generated the observed data, defining the prior and likelihood for the data, summarizing the posterior distribution, and using the posterior predictive check to assess the model's fit to the observed data. The model For the model ARIMA (1,2,1) has been chosen to the data-set. Model checking conducted, and 16 years forecasting has been done to the data. The total fertility rate of the country is projected to decline, and specifically it has been expected that approximately 2.1 (average number of children per woman) for the year 2038, this results is a very fundamental results that is the normal and standard measure of the fertility rate which is called replacement level of fertility. The importance of the demographic changes, obviously is a concern for the countries, for the effects on all affairs of human activity, economically, culturally, socially, and politically, therefore; the predicted value of the total fertility could be a good source for the government of the country for taken effective steps to control and decrease the fertility rate.

Need of the study: Afghanistan is a deve1oping country. The country has been suffering for a long time, for a long-term war perspective, awareness of this type of issue could be significant for the government and policymakers, and non-government organizations to plan in order to monitor the fertility rate of the country. There was no such study, so there is a great need for the study, it

can be a good hint for those, who are going to work on the demographic area of Afghanistan.

One important tip is needed to mention is that we have done a prediction for the total fertility rate of the country for a specific period of future time, based on the series period of past time, no prediction can be a real future, but it is a motivation work for those, who are concern in such study.

References

[1] Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F., Buettner, T., and Heilig, G. K. (2011). Probabilistic projections of the total fertility rate for all countries. Demography, 48(3):815–839.

[2] Bolanle, A. (2017). Forensic accounting as a tool for fighting corruption and terrorism: Case study of nigeria. Int. J. Soc. Sci. Econ. Res, 2(1):2210–2225.

[3] Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. International journal of forecasting, 22(3):547–581.

[4] Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley, Sons.

[5] Carmichael, G. A. et al. (2016). Fundamentals of demographic analysis: Concepts, measures, and methods. Springer.

[6] Gamerman, D. and Lopes, H. F. (2006). Markov chain Monte Carlo: stochastic simulation for Bayesian inference. CRC press.

[7] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). Bayesian data analysis. CRC press.

[8] Gelman, A., Hwang, J., and Vehtari, A. (2014). Understanding predictive information criteria for bayesian models. Statistics and computing, 24(6):997–1016.

[9] Gelman, A., Meng, X.-L., and Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. Statistica sinica, pages 733–760.

[10] Hoffman, M. D., Gelman, A., et al. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. J. Mach. Learn. Res., 15(1):1593–1623.

[11] Holmes, E. E., Scheuerell, M. D., and Ward, E. (2020). Applied time series analysis for fisheries and environmental data. NOAA Fisheries, Northwest Fisheries Science Center, Seattle, WA.

[12] Hyndman, R. J. and Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.

[13] Knudsen, C., McNown, R., and Rogers, A. (1993). Forecasting fertility: An application of time series methods to parameterized model schedules. Social Science Research, 22(1):1–23.

[14] Korner-Nievergelt, F., Roth, T., Von Felten, S., Gu?elat, J., Almasi, B., and Korner-Nievergelt, P. (2015). Bayesian data analysis in ecology using linear models with R, BUGS, and Stan. Academic Press.

[15] Lee, R. D. (1993). Modeling and forecasting the time series of us fertility: Age distribution, range, and ultimate level. International Journal of Forecasting, 9(2):187–202.

[16] Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting us mortality. Journal of the American statistical association, 87(419):659–671.

[17] Lee, R. D. and Tuljapurkar, S. (1994). Stochastic population forecasts for the united states: Beyond high, medium, and low. Journal of the American Statistical Association, 89(428):1175–1189.

[18] Marriott, J., Ravishanker, N., and Gelfand, A. E. (1993). Bayesian analysis of arma processes: Complete sampling based inference under full likelihoods. Technical report, STANFORD UNIVCA DEPT OF STATISTICS.

[19] McDonald, J. (1979). A time series approach to forecasting australian total live births. Demography, 16(4):575–601.

[20] McNown, R. and Rogers, A. (1989). Forecasting mortality: A parameterized time series approach.Demography, 26(4):645–660.

[21] Ortega, J. A. and Poncela, P. (2005). Joint forecasts of southern european fertility rates with non-stationary dynamic factor models. International Journal of Forecasting, 21(3):539–550.

[22] Raziq, A., Sheikh, N., Rafique, M., Murtaza, G., and Muhammad, A. (2021). Forecasting fertility and child mortality rates of pakistan using time series model. Linguistica Antverpiensia, 2021(2):1194–1213.

[23] Tripathi, P. K., Mishra, R. K., and Upadhyay, S. (2018). Bayes and classical prediction of total fertility rate of india using autoregressive integrated moving average model. Journal of Statistics Applications , Probability, 7(2):233–244.

[24] Watanabe, S. and Opper, M. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. Journal of machine learning research, 11(12).

Exponentiated Discrete Hypo Exponential Distribution and its Generalizations

Krishnakumari.K[1] and Dais George[2]

[1]Research Scholar, St.Thomas College, Palai
(SAS SNDP YOGAM College, Konni), Kerala, India.
[2]Catholicate College, Pathanamthitta, Kerala, India.
krishnavidyadharan@gmail.com, daissaji@rediffmail.com

Abstract

Generalizations of standard probability distributions is a thought-provoking concept in statistical literature and was inspired by many researchers in recent days. This is because the addition of parameters may increase the flexibility of the new models. Now a days various generalization techniques are available in literature. In this work, we proposed a generalization of discrete hypo exponential distribution and studied its various properties. A real data analysis is carried out and check the flexibility of the new model by comparing it with other standard distributions. Two generalizations of the proposed distribution are introduced.

Keywords: Discrete hypo exponential distribution, Estimation, Generalization, Moments, Stress-strength analysis.

1. Introduction

Over the last few decades, there has been growing interest in adding supplementary parameters to the baseline distributions to broaden generalized families of distributions. The addition of parameters may increase the flexibility of the new models. So generalization of the standard distributions are attracted by many researchers and are prominent in recent days. In literature there exists various generalization techniques and for a detailed review, see Tahir and Nadarajah[32]. These techniques resulted in the generalizations of various standard distributions. For details see, Gupta and Kundu [14], Eugene et.al [11], Zografos and Balakrishnan [33], Gomez-Deniz [13], Mahmoudi and Zakerzadeh [19], Cordeiro and Castro [4], Nadarajah [24], Nadarajah et.al [26], Cordeiro et al. [5], Ristic and Balakrishnan [30], Lemonte et.al [17], Liyanage and Pararai [18], Merovci and Elbatal [22], Merovci and Sharma [23], Nadarajah and Bakar [25], Ahmad and Ghazal [1], Sulami [2] etc. Recently exponentiated family of distributions due to Lehman [16] has got special attention and various standard distributions were generalized. The most prominent distribution introduced in the 20^{th} century is the exponentiated exponential distribution and inspired by this many existing distributions were generalized and for details see, Pal et al. [28], Nekoukhou and Bidram [27], Morshedy et al. [8], El-Bassiouny et al. [7], Morshedy et al.[10], Morshedy et al. [9], Mashhadzadeh and Mirmostafaee [21] and Baharith and Alamoudi [3].
The layout of this article is in this way. In Section 2, we introduced exponentiated discrete hypo exponential distribution and studied its various properties. In Section 3 the parameters of the distribution is done through non linear maximization method. To evaluate the performance of the nlm estimator a simulation study is done in Section 4. A real data analysis is done in Section 5. In Section 6 some generalizations of the proposed distribution are introduced. Some concluding remarks are recorded in Section 7.

2. Exponentiated Discrete Hypo Exponential Distribution

Consider the discrete hypo exponential (DHE) distribution having model parameters ϕ_1, $\phi_2 > 0$, $\phi_1 \neq \phi_2$ with the distribution function

$$F(x;\phi_1,\phi_2) = \frac{\phi_2}{\phi_2-\phi_1}(1-e^{-\phi_1 x}) - \frac{\phi_1}{\phi_2-\phi_1}(1-e^{-\phi_2 x}) \tag{1}$$

By inserting (1) into the resilience parameter family of distributions, the distribution function of the resulting distribution is given by

$$\begin{aligned} G(x;\phi_1,\phi_2,\alpha) &= [F(x;\phi_1,\phi_2)]^\alpha \\ &= \frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha} \end{aligned} \tag{2}$$

where

$$V(x;\phi_1,\phi_2,\alpha) = [\phi_2(1-e^{-\phi_1 x}) - \phi_1(1-e^{-\phi_2 x})]^\alpha \tag{3}$$

We call such a random variable X, having distribution function (2), is an exponentiated DHE distribution with parameters ϕ_1, $\phi_2 > 0$, $\phi_1 \neq \phi_2$, $\alpha > 0$ and denote it as EDHE (ϕ_1, ϕ_2, α).
The probability mass function(pmf) of EDHE distribution is given by

$$\begin{aligned} P(X=x) &= v(x;\phi_1,\phi_2,\alpha) \\ &= \frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}; x = 0,1,2,.... \end{aligned} \tag{4}$$

The plots of pmf of EDHE distribution is given in Figure 1.

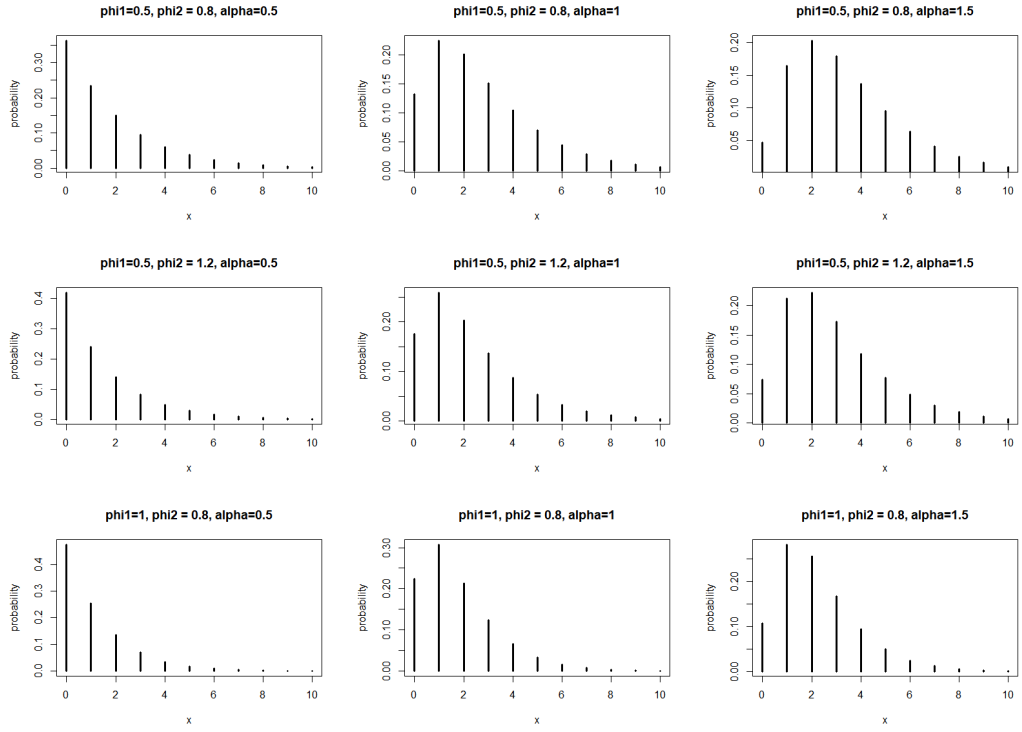


Figure 1: *Plots of pmf of EDHE distribution*

From Figure 1 it is understood that the EDHE distribution is unimodel. Since every log-concave density is unimodel, it is also inferred that EDHE distribution is log-concave.

2.1. Reliability characteristics

$$\text{Survival function, } S(x) = 1 - \frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\ ; x = 0,1,2..., \quad \text{and}$$

$$\text{hazard rate, } r(x) = \frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha - V(x;\phi_1,\phi_2,\alpha)}$$

The plots of hazard rate of EDHE distribution is given in Figure 2.

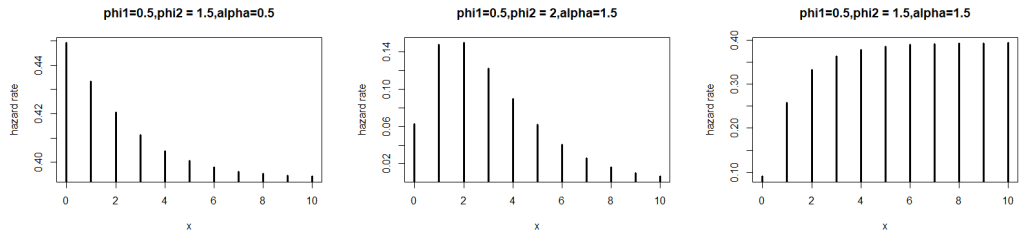


Figure 2: *Plots of hazard rate of EDHE distribution.*

From Figure 2, it is evident that for various model parameters, the hazard rate functions can be decreasing, increasing and increasing-decreasing, which makes the EDHE distribution more flexible and can model different types of data sets such as count data, failure time data etc.

2.2. Moments

Let X $\sim$ EDHE(ϕ_1,ϕ_2,α), then for $n \geq 1$,

$$E(X^n) = \sum_{x=0}^{\infty} x^n \left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right].$$

In particular

$$E(X) = \sum_{x=0}^{\infty} x \left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right].$$

and

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right].$$

The expression for V(X) can be obtained using he relation

$$V(X) = E(X^2) - [E(X)]^2$$

2.3. Infinite Divisibility

According to Steutel and Van Harn [31], a necessary condition for infinite divisibility of a discrete distribution P_y is that $P_0 > 0$. For EDHE distribution this condition is satisfied for all values of the parameters. Hence it is infinitely divisible.

2.4. Theorem

If X follows an exponentiated hypo exponential distribution with parameters ϕ_1 , ϕ_2 and α then the random variable W=[X] follows a exponentiated discrete hypo exponential distribution with parameters ϕ_1, ϕ_2 and α .

Proof:
Consider w=0, 1, 2,... then using Lemma1 of Krishna and Pundir (2009), we have

$$\begin{aligned} P(W \geq w) &= P([X] \geq w) \\ &= P(X \geq w) \\ &= 1 - \left[\frac{\phi_2(1-e^{-\phi_1 x}) - \phi_1(1-e^{-\phi_2 x})}{\phi_2 - \phi_1}\right]^{\alpha} \end{aligned}$$

which is the survival function of EDHE distribution and hence the theorem.

2.5. Order Statistics

Order statistics are sample values placed in ascending order. It is a very useful concept in statistical sciences. It has a far reaching applications especially in modeling auctions, car races, insurance policies and estimating parameters of distributions etc.
Let $X_{(1:n)} \leq X_{(2:n)} \leq X_{(3:n)} \leq ... \leq X_{(n:n)}$ represents the order statistics obtained from the i.i.d. EDHE(ϕ_1, ϕ_2, α) distribution of size n. Then probability mass function of first order statistics is given by

$$f_{X_{(1:n)}}(x) = \left[1 - \frac{V(x-1;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^n - \left[1 - \frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^n$$

and the distribution function is

$$F_{X_{(1:n)}}(x) = 1 - \left[1 - \frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^n.$$

The probability mass function of n^{th} order statistics is given by

$$f_{X_{(n:n)}}(x) = \left[\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^n - \left[\frac{V(x-1;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^n$$

and the distribution function is

$$F_{X_{(n:n)}}(x) = \left[\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^n$$

where $V(x;\phi_1,\phi_2,\alpha)$ is given by (3).

2.6. Entropy

The Shannon's entropy of random variable X having probability mass function P(x) is given by

$$H(X) = E(-logP(x)).$$

For EDHE distribution, H(X) is obtained as

$$H(X) = -\sum_{x=0}^{\infty}\left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right] log\left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right].$$

Renyi's entropy of order $\beta > 0$ $(\beta \neq 1)$ is

$$H_{\beta}(p) = \frac{1}{1-\beta} log \sum_{x=0}^{\infty}\left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]^{\beta}.$$

2.7. Stress-strength Analysis

Stress-strength analysis is a mechanism and a topic used in reliability engineering. When the probability of stress exceeding the strength of an item, that item fails. Hence the expected reliability (R) is calculated as

$$R = P(stress \leq strength) = \sum_{x=0}^{\infty} f_{stress}(x) R_{strength}(x)$$

where the strength and stress are in the positive domain. When stress $\sim$ EDHE$(\phi_1, \phi_2, \alpha_1)$ and strength$\sim$ EDHE$(\phi_3, \phi_4, \alpha_2)$, the expected reliability is

$$R = \sum_{x=0}^{\infty} \frac{V(x+1;\phi_1,\phi_2,\alpha_1) - V(x;\phi_1,\phi_2,\alpha_2)}{(\phi_2-\phi_1)^{\alpha_1}} \left[1 - \frac{V(x;\phi_3,\phi_4,\alpha_2)}{(\phi_4-\phi_3)^{\alpha_2}}\right].$$

Tables 1-4 show the numerical values of R for different values of stress-strength parameters.

Table 1: *Values of R for* $\phi_1 = 0.1$, $\phi_2 = 0.3$, $\phi_3 = 0.3$, $\phi_4 = 0.6$ *and different values of* α_1 *and* α_2

α_1 \ α_2	0.2	0.6	1	1.5	2	2.5
0.2	0.5185	0.6342	0.6962	0.7409	0.7685	0.7874
0.6	0.1707	0.2952	0.3708	0.4319	0.4735	0.5040
1	0.0735	0.1588	0.2166	0.2676	0.3050	0.3338
1.5	0.0336	0.0834	0.1207	0.1564	0.1843	0.2068
2	0.0181	0.0479	0.0718	0.0960	0.1157	0.1321

Table 2: *Values of R for* $\phi_1 = 0.3$, $\phi_2 = 0.5$, $\phi_3 = 0.6$, $\phi_4 = 0.8$ *and different values of* α_1 *and* α_2

α_1 \ α_2	0.2	0.6	1	1.5	2	2.5
0.2	0.6338	0.7240	0.7777	0.8185	0.8441	0.8617
0.6	0.2799	0.4162	0.5027	0.5733	0.6210	0.6557
1	0.1430	0.2657	0.3487	0.4210	0.4730	0.5127
1.5	0.0753	0.1690	0.2373	0.3011	0.3499	0.3890
2	0.0469	0.1169	0.1712	0.2249	0.2680	0.3037

Table 3: *Values of R for* $\phi_1 = 0.5$, $\phi_2 = 0.8$, $\phi_3 = 0.5$, $\phi_4 = 0.8$ *and different values of* α_1 *and* α_2

α_1 \ α_2	0.2	0.6	1	1.5	2	2.5
0.2	0.7372	0.8258	0.8748	0.9092	0.9293	0.9421
0.6	0.4272	0.5967	0.6952	0.7686	0.8138	0.8442
1	0.2720	0.4598	0.5747	0.6649	0.7235	0.7646
1.5	0.1759	0.3550	0.4713	0.5685	0.6352	0.6838
2	0.1277	0.2886	0.3991	0.4962	0..5658	0.6183

Table 4: *Values of R for α_1=$\alpha_2 = 0.6$, and different values of ϕ_1, ϕ_2, ϕ_3, ϕ_4 $\phi_2 = \phi_4$*

$\phi_1 = \phi_3$	0.6	0.7	0.8	1	1.5
0.1	0.4940	0.4974	0.50000	0.5037	0.5091
0.2	0.5437	0.5470	0.54972	0.5540	0.5611
0.3	0.5627	0.5668	0.5704	0.5761	0.5863
0.4	0.5751	0.5803	0.5848	0.5923	0.6058
0.5	0.5850	0.5912	0.5967	0.6059	0.6226

From tables 1-3 it is clear that for fixed values of ϕ_1, ϕ_2, ϕ_3, ϕ_4 and α_1 reliability increases as α_2 tends to infinity. But the reliability decreases with α_1 tends to infinity for fixed values of ϕ_1, ϕ_2, ϕ_3, ϕ_4 and α_2 . Table 4 shows that reliability increases with increasing values of ϕ_1, ϕ_3 for fixed values of ϕ_2, ϕ_4, α_1 and α_2. Also for fixed values of ϕ_1, ϕ_3, α_1 and α_2, reliability increases with increasing values of ϕ_2 , ϕ_4.

2.8. Estimation

In this section we estimate the parameters ϕ_1, ϕ_2 and α of EDHE distribution using the method of maximum likelihood. Let us take a random sample$X_1, X_2...X_n$ of size n from EDHE distribution. Then the logarithm of likelihood function is

$$logL = \sum_{x=0}^{\infty} log\left[\frac{V(x+1;\phi_1,\phi_2,\alpha) - V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^{\alpha}}\right]$$

The maximum likelihood estimators of ϕ_1, ϕ_2 and α are obtained by solving the equations

$$\frac{\partial logL}{\partial \phi_1} = 0, \frac{\partial logL}{\partial \phi_2} = 0, \frac{\partial logL}{\partial \alpha} = 0.$$

But these equations cannot be solved analytically. So we use Non Linear Maximization (nlm) method for estimating the parameters ϕ_1, ϕ_2 and α.

3. Simulation Study

In this section, we use Monte-Carlo simulation method to illustrate the performance of the nlm estimator of the parameters ϕ_1 and ϕ_2 and α. We generate 5000 random samples of sizes n=20, 30, 75 and 100 from the $HE(\phi_1, \phi_2)$ distribution for some true values of the parameter set (ϕ_1, ϕ_2) = (15,18), (15,21), (16,18) and (16,21). We discretize the generated data and find out 5000 estimates of ϕ_1 and ϕ_2 and α using (4)for each sample sizes. The estimate of the parameter, average bias and mean square error of the estimate (MSE) are computed and it is given in Table 5 to Table 12.

Table 5: *Values of estimates, average bias and average MSE for ϕ_1=15 and different values of, ϕ_2, α and n=20.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
15	18	0.4	14.5668	-0.4331	0.1876	17.8307	-0.1693	0.0287	0.3775	-0.0224	0.00005
		0.8	14.4120	-0.5879	0.3457	17.7375	-0.2624	0.0689	0.7133	-0.0866	0.0075
15	21	0.4	14.8996	-0.1003	0.0101	20.3668	-0.6331	0.4009	0.3370	-0.0629	0.004
		0.8	14.3655	-0.6344	0.4025	20.6269	-0.3730	0.1392	0.6763	-0.1236	0.0153

Table 6: *Values of estimates, average bias and average MSE for ϕ_1=15 and different values of ϕ_2, α and n=30.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
15	18	0.4	14.5689	-0.4310	0.1858	17.8316	-0.1683	0.0283	0.3807	-0.0192	0.00004
		0.8	14.5509	-0.4490	0.2016	17.5699	-0.4300	0.185	0.7356	-0.0643	0.0041
15	21	0.4	14.9042	-0.0957	0.0092	20.6514	-0.3485	0.1215	0.3487	-0.0512	0.0026
		0.8	14.5340	-0.4659	0.2171	20.6291	-0.3708	0.1375	0.6883	-0.1116	0.0125

Table 7: *Values of estimates, average bias and average MSE for ϕ_1=15 and different values of ϕ_2, α and n=75.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
15	18	0.4	14.6128	-0.3871	0.1499	17.8372	-0.1627	0.0265	0.4090	0.0099	0.00001
		0.8	14.4120	-0.5879	0.3457	17.7375	-0.2624	0.0689	0.7133	-0.0866	0.0075
15	21	0.4	14.9614	-0.0380	0.0014	20.8440	-0.1550	0.0241	0.3502	-0.0497	0.0024
		0.8	14.6453	-0.3546	0.1258	20.6317	-0.3683	0.1356	0.6979	-0.1020	0.0104

Table 8: *Values of estimates, average bias and average MSE for ϕ_1=15 and different values of ϕ_2, α and n=100.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
15	18	0.4	14.6154	-0.3845	0.1478	17.8376	-0.1623	0.0263	0.4089	0.0089	0.00001
		0.8	14.7504	-0.2494	0.0622	17.8349	-0.2494	0.0273	0.7988	-0.0011	0.0000
15	21	0.4	15.0249	0.0249	0.00006	20.8483	-0.1516	0.0230	0.37770	-0.0226	0.00005
		0.8	15.0320	0.0320	0.0010	20.65213	-0.3478	0.1210	0.7445	-0.0554	0.0031

Table 9: *Values of estimates, average bias and average MSE for ϕ_1=16 and different values of ϕ_2, α and n=20.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
16	18	0.4	15.4568	-0.5431	0.2950	17.8664	-0.1335	0.0178	0.3409	-0.0590	0.0035
		0.8	15.6653	-0.3346	0.1120	17.4611	-0.5388	0.2903	0.6452	-0.1547	0.0240
16	21	0.4	15.1495	-0.8504	0.7233	20.4267	-0.5732	0.3286	0.3003	-0.0996	0.0099
		0.8	15.7852	-0.2147	0.0461	20.4893	-0.5101	0.2603	0.6690	-0.1309	0.0172

Table 10: *Values of estimates, average bias and average MSE for ϕ_1=16 and different values of ϕ_2, α and n=30.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
16	18	0.4	15.4593	-0.5406	0.2923	17.8662	-0.1337	0.0179	0.3452	-0.0547	0.0030
		0.8	15.8042	-0.1957	0.0383	17.4674	-0.5325	0.2836	0.6708	-0.1291	0.0167
16	21	0.4	15.1507	-0.8492	0.7212	20.4282	-0.5717	0.3269	0.3008	-0.0991	0.0098
		0.8	15.8184	-0.1815	0.0330	20.4918	-0.5081	0.2583	0.7341	-0.0658	0.0043

Table 11: *Values of estimates, average bias and average MSE for ϕ_1=16 and different values of ϕ_2, α and n=75.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
16	18	0.4	15.4822	-0.5177	0.2681	17.8708	-0.1291	0.0167	0.3791	-0.0208	0.00004
		0.8	15.8212	-0.1787	0.0319	17.5036	-0.4963	0.2463	0.7450	-0.0549	0.0030
16	21	0.4	15.1631	-0.8368	0.7002	20.4289	-0.5710	0.3261	0.3171	-0.08281	0.0069
		0.8	15.9054	-0.0945	0.0089	20.5039	-0.4961	0.2461	0.7946	0.0053	0.0012

Table 12: *Values of estimates, average bias and average MSE for ϕ_1=16 and different values of ϕ_2, α and n=100.*

ϕ_1	ϕ_2	α	$\hat{\phi_1}$	Bias(ϕ_1)	MSE(ϕ_1)	$\hat{\phi_2}$	Bias(ϕ_2)	MSE(ϕ_2)	$\hat{\alpha}$	Bias(α)	MSE(α)
16	18	0.4	15.4951	-0.5048	0.2548	17.8731	-0.1268	0.0161	0.3989	-0.0010	0.0000
		0.8	15.9119	-0.0880	0.0078	17.9695	-0.0304	0.00009	0.7942	-0.0057	0.0000
16	21	0.4	15.1645	-0.8354	0.6979	20.4290	-0.5710	0.3261	0.3183	-0.0816	0.0067
		0.8	15.9066	-0.0933	0.0087	20.5044	-0.4956	0.2456	0.8010	0.0010	0.0000

From tables 5-12, it is clear that as sample size increases, the average bias and average MSE becomes very small for different choices of the values of the parameters. This indicates the consistency of the estimators.

4. Real Data Analysis

For studying the efficiency of EDHE distribution we consider the data set used by Krishna and Pundir [15] and it represents the total number of carious teeth among the four deciduous molars in a sample of 100 children 10 and 11 years of old. The data are given in Table 13.

Table 13: *Observed data*

X	0	1	2	3	4
f	64	17	10	6	3

Figure 3 shows the observed data.

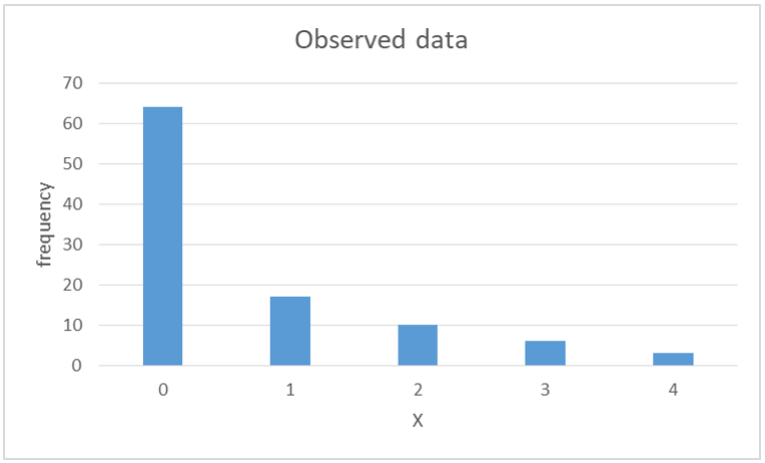


Figure 3: *Observed data.*

We fit the EDHE distribution using the empirical data set and the embeded figure is given in Figure 4.

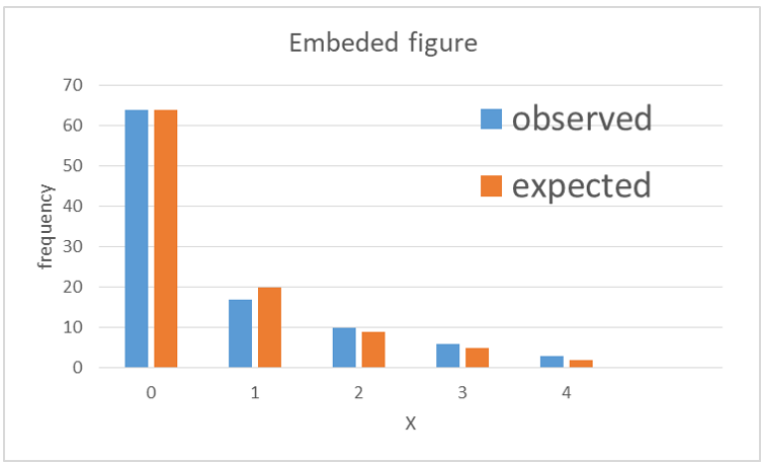


Figure 4: *Embeded figure.*

In order to assess the suitability of the proposed model, we use chi-square test of goodness of fit. Also, we compare the EDHE distribution with discrete Lindley (DL) distribution discrete Pareto (DP) distribution and the values of Log-likelihood, AIC, BIC are computed and is shown in Table 14.

Table 14: *MLE's, Chi-square value, -Log-likelihood value, AIC values, BIC values and P values for the observed data.*

Distribution fitted	estimators	Chi-square	-LL value	AIC value	BIC value	p value
DLD	$\hat{\theta} = 0.275$	6.637	113.68	229.36	229.36	0.036
DPD	$\hat{\beta} = 0.1837$	3.226	116.83	235.66	235.66	0.199
EDHED	$\hat{\phi_1} = 0.9824779$ $\hat{\phi_2} = 0.9824794$ $\hat{\alpha} = 0.3346$	1.2611	111.54	229.08	229.08	0.8679

From Table 14, it is inferred that the EDHE distribution is a better fit than discrete Lindley

and discrete Pareto distributions.

5. Generalizations

5.1. Transmuted exponentiated discrete hypo exponential (TEDHE) distribution

Many transmuted distributions are proposed and studied in literature. For details see Rahman et al. [29] and Dey et al. [6]. In this section we present a generalization of (4) called the transmuted exponentiated discrete hypo exponential distribution. A random variable X is said to have transmuted distribution if its distribution function and probability mass functions are respectively given by

$$F(x) = G(x)[1 + \beta - \beta G(x)]; |\beta| \leq 1 \tag{5}$$

and

$$P(X = x) = g(x)[1 + \beta - 2\beta G(x)] \tag{6}$$

where G(x), g(x) are the distribution function and probability mass function of the baseline distribution. Also if $\beta = 0$, we will get the baseline distribution. By using equations (5) and (6), the distribution function and probability mass function of the TEDHE distribution is obtained as

$$F(x;\phi_1,\phi_2,\alpha,\beta) = \frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\left[1+\beta-\beta\left(\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right)\right] \tag{7}$$

and

$$f(x;\phi_1,\phi_2,\alpha,\beta) = \frac{V(x+1;\phi_1,\phi_2,\alpha)-V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\left[1+\beta-2\beta\left(\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right)\right] \tag{8}$$

The plot of pmf of TEDHE distribution is given in Figure 5.

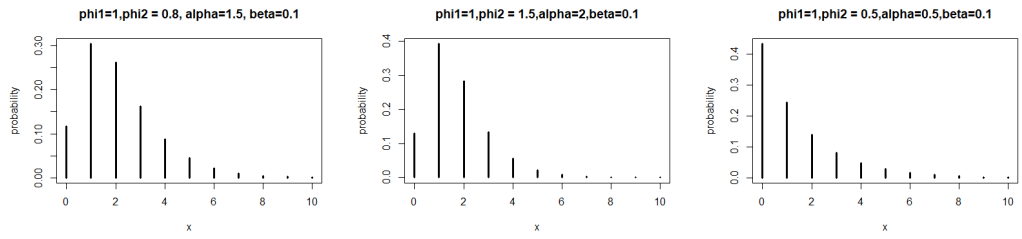


Figure 5: *Plot of pmf of TEDHE distribution.*

The survival function and hazard rate functions are given by the expressions

$$S(x) = 1 - \frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\left[1+\beta-\beta\left(\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right)\right]$$

and

$$h(x) = \frac{\frac{V(x+1;\phi_1,\phi_2,\alpha)-V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\left[1+\beta-2\beta\left(\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right)\right]}{1-\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\left[1+\beta-\beta\left(\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}\right)\right]}$$

The hazard plots of TEDHE distribution is given in Figure 6.

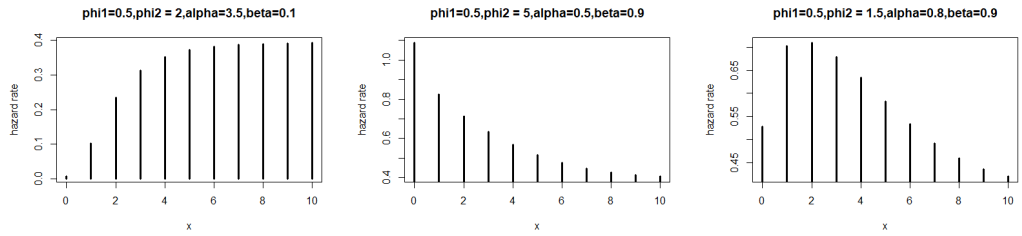


Figure 6: *Plot of hazard function of TEDHE distribution.*

From Figure 6, it is understood that for different model parameters, the hazard rate function can be decreasing, increasing and increasing-decreasing, which makes the TEDHE distribution more flexible and can model different types of data sets.

5.2. Marshall-Olkin exponentiated discrete hypo exponential (MOEDHE) distribution

Marshall and Olkin [20] introduced a new method for adding a parameter $\theta(>0)$ to the baseline distribution in order to generalize it. Using this method many generalized distributions are proposed and for a detailed review see Gillariose et al. [12]. If $\overline{F}(x)$ is the survival function of a distribution, then, by Marshall-Olkin method, another survival function $\overline{G}(x)$ is obtained as

$$\overline{G}(x,\theta) = \frac{\theta\overline{F}(x)}{1-(1-\theta)\overline{F}(x)}; -\infty < X < \infty, \theta > 0.$$

The corresponding distribution function, probability mass function and hazard rate is obtained as

$$\begin{aligned} G(x,\theta) &= \frac{F(x)}{1-(1-\theta)\overline{F}(x)} \\ g(x,\theta) &= G(x,\theta) - G(x-1,\theta) \\ &= \frac{\theta f(x)}{[1-(1-\theta)\overline{F}(x)][1-(1-\theta)\overline{F}(x-1)]} \\ h(x) &= \frac{g(x)}{\overline{G}(x)}. \end{aligned}$$

where f(x) is the probability mass function corresponding to the distribution function F(x). Using Marshall-Olkin method the survival function of MOEDHE distribution is

$$\overline{G}(x,\theta) = \frac{\theta(1-\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})}{1-[(1-\theta)(1-\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})]}.$$

The corresponding distribution function, probability mass function and hazard rate are respectively given by

$$\begin{aligned} G(x,\theta) &= \frac{\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}}{1-[(1-\theta)(1-\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})]} \\ g(x,\theta) &= \frac{\theta\frac{V(x+1;\phi_1,\phi_2,\alpha)-V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}}{(1-[(1-\theta)(1-\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})])(1-[(1-\theta)(1-\frac{V(x-1;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})])} \end{aligned}$$

and

$$h(x) = \frac{\frac{V(x+1;\phi_1,\phi_2,\alpha)-V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha}}{(1-[(1-\theta)(1-\frac{V(x-1;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})])(1-\frac{V(x;\phi_1,\phi_2,\alpha)}{(\phi_2-\phi_1)^\alpha})}.$$

The plot of probability mass function of MOEDHE distribution is given in Figure 7.

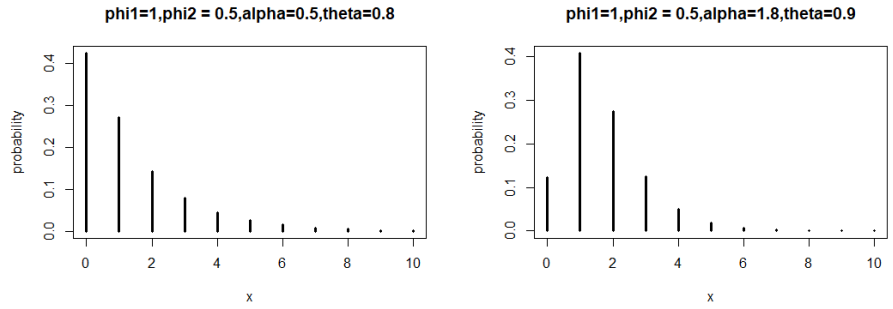


Figure 7: *Plot of pmf of MOEDHE distribution.*

The hazard plots are given in Figure 8.

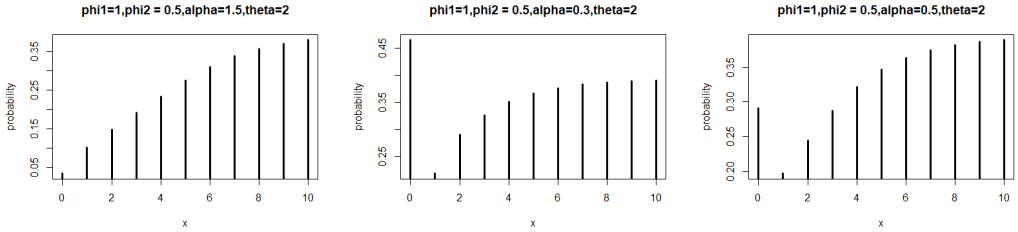


Figure 8: *Hazard plots of MOEDHE distribution.*

Figure 8 shows different shapes of hazard rate functions and so we can conclude that the MOEDHE distribution is a flexible model in modeling different types of data sets.

6. Summary

Recently, there has been thriving interest in developing new families of distributions by adding one or more additional parameters to the baseline distributions. The existence of various generalization techniques were attracted by many researchers and using one among them we proposed and studied a new distribution called exponentiated discrete hypo exponential distribution. Various distributional and structural properties of this distribution are studied. Also stress-strength analysis is carried out. To evaluate the performance of the nlm estimator, we conducted a simulation study and found that the nlm estimator is consistent. A real data application is carried out and inferred that our proposed distribution is better model than discrete Lindley and discrete Pareto distribution. Two generalizations of the proposed distribution namely transmuted exponentiated discrete hypo exponential distribution and Marshall-Olkin exponentiated discrete hypo exponential distribution are introduced.

References

[1] Ahmad, A. A. and Ghazal, M.G. M. (2020). Exponentiated additive Weibull distribution. *Reliability Engineering and System Safety*, 193, 106663.

[2] Al-Sulami, D. (2020). Exponentiated Exponential Weibull Distribution. *American Journal of Applied Sciences*, **17**, 188-195.

[3] Baharith, L. A and Alamoudi, H. H. (2021). The exponentated Frechet Generator of distributions with applications. *Symmetry*, **13(4)**, 572.

[4] Cordeiro, G. M and Castro, M. (2011). A new family of generalized distributions. *J.Stat.Comput.Simul.*, **81**, 883-898.

[5] Cordeiro, G. M. ,Ortega, E. M. M and Silva, G. (2012). The Beta extended Weibull family. *J.Prob.Stat.Sci.*, **10**, 15-40.

[6] Dey, S., Kumar, D.,Anis, M. Z. and Nadarajah, S. (2021). A review of transmuted distributions. *Journal of the Indian Society for Probability and Statistics*, **22**, 47-111.

[7] El-Bassiouny, A. H., EL-Damcese, M., Abdelfattah, M. and Eliwa, M. S. (2017). Exponentiated generalized Weibull-Gompertz distribution with application in survival analysis. *Journal of Statistics Applications and Probability*, **6**, 7-16.

[8] El-Morshedy, M., El-Bassiouny, A. H. and El-Gohary, A. (2017). Exponentiated Inverse flexible Weibull extension distribution. *Journal of Statistics Applications and Probability*, **6**, 169-183.

[9] El-Morshedy, M., Eliwa, M. S., El-Gohary, A and Khalil, A. A. (2020). Bivariate exponentiated discrete Weibull distribution: Statistical Properties,estimation, Simulation and Applications. *Mathmatical Sciences* , **14(1)**.

[10] El-Morshedy, M., Eliwa, M. S. and Nagy, H. (2019). A new two-parameter exponentiated discrete Lindley distribution: Properties,estimation and Applications. *Journal of Applied Statistics* .

[11] Eugene, N., Lee, C and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, **31**, 497-512.

[12] Gillariose, J., Tomy, L., Chesneau, C. and Jose, M. (2020). Recent developments in Marshall-Olkin distributions. *Contributions to Mathematics*,**2**, 71-75.

[13] Gomez-Deniz, E. (2010). Another Generalization of the geometric distribution. *Test 19*, 399-415

[14] Gupta, R. D and kundu, D. (1999). Generalized exponential distribution, *Aust.N.Z.J.Stat.*, **41**, 173-188.

[15] Krishna, H and Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, **6**, 177-188.

[16] Lehmann, E. L. (1953). The power of rank tests. *Ann. Math. Statistics*, **24**, 23-43.

[17] Lemonte, W., Souza, w. B and Cordeiro, G. M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazillian Journal of Probability and Statistics*, **27**, 31-53.

[18] Liyange, W and Pararai, M. (2014). A generalized power lindley distribution with applications. *Asian Journal of Math. Applications and Probability*, **18**, 1-23.

[19] Mahmoudi, E and Zakerzadeh, H. (2010). Generalized Poisson-Lindley distribution. *Communications in Statistics-Theory and Methods*, **39**, 1785-1798.

[20] Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrica*, **84(3)**, 641-652.

[21] Mashhadzadeh, Z. H and Mirmostafaee, S. M. T. K. (2020). The exponentiated discrete inverse Rayleigh distribution. *Journal of Hyperstructures*, **9(1)**, 54-61.

[22] Merovci, F. and Elbatal, I. (2014). Transmuted Lindley - Geometric distribution and its applications. *Journal of Statistics Applications and Probability*, **3**, 77-91.

[23] Merovci, F. and Sharma, V. K. (2014). The Beta-Lindley distribution: Properties and Applications. *Journal of Applied Mathematics*.

[24] Nadarajah, S. (2011). The exponentiated exponential distribution:asurvey. *AStA Adv. Stat. Anal.* , **95**,219-251.

[25] Nadarajah, S. and Bakar, S. S. A. (2015). An exponentiated geometric distribution. *Applied Mathematical Modelling* , 1-10.

[26] Nadarajah, S. Bakouch, H. S. and Tahmasbi, R. (2011). A generalized Lindley distribution. *Sankhya B 73* , 331-359.

[27] Nekoukhou, V. and Bidram, H. (2015a). The exponentiated discrete Weibull distribution. *SORT 39*, 127-146.

[28] Pal, M., Ali, M. M and Woo, J. (2006). Exponentiated Weibull Distribution. *Statistica*, **66(2)**, 139-147.

[29] Rahman, M. M, Zahrani, B. L.,Shahbaz, S. H. and Shahbaz, M. Q. (2020). Transmuted probability distributions: A Review. *Pakistan Journal of Statistics and Operations Research*, **16(1)**, 83-94.

[30] Ristic, M. M and Balakrishnan, N. (2012). The gamma exponentiated exponential distribution. *J.Stat.Comput.Simul.*, **82**, 1191-1206.

[31] Steutel, F. W and Van Harn, K. Infinite divisibility of Probability Distributions on the Real Line. Vol.259, New York, Marcel Dekker, 2004.

[32] Tahir, M. H and Nadarajah, S. (2015). Parameter induction in continuous univariate distributions: Well established G-families. *Anais da Academia Brasileira de Cincias* , **87(2)**, 539-568.

[33] Zografos, K and Balakrishnan, N. (2009). On families of Beta and generalized Gamma generated distributions and associated Inference. *Statistical Methodology*, 344-362.

An Innovative Approach for Reliability Modeling of HVDC Converter Station

Aditya Tiwary, R. S. Mandloi
•
Department of Fire Technology & Safety Engineering, IPS Academy, IES, Indore (M.P) India
raditya2002@gmail.com

Department of Electrical Engineering, Shri G. S. Institute of Technology & Science, Indore (M.P) India

Abstract

Assessment of reliability indices is important when availability and unavailability of the system or systems or components or group of components are to be assessed. There are various reliability indexes which are very important for overall performance of any complex engineering system. Reliability block diagram modeling is required to be formulated for evaluating different essential and important reliability parameters of any complex engineering system. In view of above, in this paper, reliability block diagram modeling of HVDC converter station is represented and formulated. The schematic diagram of the HVDC converter station is available in literature and based on that schematic diagram the modeling of HVDC converter station is formulated in this paper. After the reliability block diagram modeling of HVDC converter station, the mean time to failure (MTTF) of each and every components of HVDC converter station are also evaluated and represented in the result and discussion section. The reliability of each and every component of the HVDC converter station is evaluated and expressed in result section. Assessment of unavailability is also obtained and shown in result section.

Keywords: Reliability, Mean time to failure, HVDC converter station, availability, Reliability indices.

I. Introduction

Reliability evaluation of a system or component or element is very important in order to predict its availability and other relevant indices. Reliability is the parameter which tells about the availability of the system under proper working conditions for a given period of time. A Markov cut-set composite approach to the reliability evaluation of transmission and distribution systems involving dependent failures was proposed by Singh et al. [1]. The reliability indices have been determined at any point of composite system by conditional probability approach by Billinton et al. [2]. Wojczynski et al. [3] discussed distribution system simulation studies which investigate the effect of interruption duration distributions and cost curve shapes on interruption cost estimates. New indices to reflect the integration of probabilistic models and fuzzy concepts was proposed by Verma et al. [4]. Zheng et al. [5] developed a model for a single unit and derived expression for

availability of a component accounting tolerable repair time. Distributions of reliability indices resulting from two sampling techniques are presented and analyzed along with those from MCS by Jirutitijaroen and Singh [6]. Dzobe et al. [7] investigated the use of probability distribution function in reliability worth analysis of electric power system. Bae and Kim [8] presented an analytical technique to evaluate the reliability of customers in a micro grid including distribution generations. Reliability network equivalent approach to distribution system reliability assessment is proposed by Billinton and Wang [9].

Customer and energy based indices consideration for reliability enhancement of distribution system using Improved Teaching Learning based optimization is discussed [10]. An Innovative Self-Adaptive Multi-Population Jaya Algorithm based Technique for Evaluation and Improvement of Reliability Indices of Electrical Power Distribution System, Tiwary et al. [11]. Determination of reliability indices for distribution system using a state transition sampling technique accounting random down time omission, Tiwary et al. [12]. Tiwary et al. [13] proposed a methodology based on Inspection-Repair-Based Availability Optimization of Distribution System Using Bare Bones Particle Swarm Optimization. Bootstrapping based technique for evaluating reliability indices of RBTS distribution system neglecting random down time was evaluated [14].

Volkanavski et al. [15] proposed application of fault tree analysis for assessment of the power system reliability. Li et al. [16] studies the impact of covered overhead conductors on distribution reliability and safety. Self-Adaptive Multi-Population Jaya Algorithm based Reactive Power Reserve Optimization Considering Voltage Stability Margin Constraints was obtained in Tiwary et al. [17]. A smooth bootstrapping based technique for evaluating distribution system reliability indices neglecting random interruption duration is developed [18]. Tiwary et al. [19] have developed an inspection maintenance based availability optimization methodology for feeder section using particle swarm optimization. The impact of covered overhead conductors on distribution reliability and safety is discussed [20]. Tiwary et al. [21] has discussed a methodology for reliability evaluation of an electrical power distribution system, which is radial in nature. Sarantakos et al. [22] introduced a method to include component condition and substation reliability into distribution system reconfiguration. Tiwary et al. [23] has discussed a methodology for evaluation of customer orientated indices and reliability of a meshed power distribution system. Reliability evaluation of engineering system is discussed [24]. Battu et al. [25] discussed a method for reliability compliant distribution system planning using Monte Carlo simulation. Application of non-parametric bootstrap technique for evaluating MTTF and reliability of a complex network with non-identical component failure laws is discussed [26]. Tiwary and Tiwary [27] have developed an innovative methodology for evaluation of customer orientated indices and reliability study of electrical feeder system. Tiwary and Tiwary [28] proposed the evaluation of reliability indices of Roy Billinton Test System (RBTS) Bus-2 Distribution System.

Tiwary and Tiwary [29] have proposed a methodology for reliability block diagram representation of electric traction system and identification of various reliability indices. In view of the above, in this paper reliability block diagram modeling of HVDC converter station is represented and formulated. The reliability of each and every component of the HVDC converter station is obtained. Mean time to failure (MTTF) of the components are also evaluated. Assessment of overall reliability of the system and overall mean time to failure is calculated and presented.

II. Reliability block diagram representation of HVDC converter station

Reliability block diagram which is a diagrammatic method for showing how different components are connected in a system is obtained for the HVDC converter station. The schematic diagram of the typical HVDC converter station is given by [30]. The HVDC converter station

consists of 12 pulse converter, transformer, smoothing reactors, DC filters, tuned AC filters, HP AC filters. The reliability block diagram of HVDC converter station is shown in Fig. 1. From Fig. 1, it is clear that HP AC filter and tuned AC filters are connected in a parallel manner. While the 12 pulse converter, smoothing reactors and DC filters are connected in series configuration as shown.

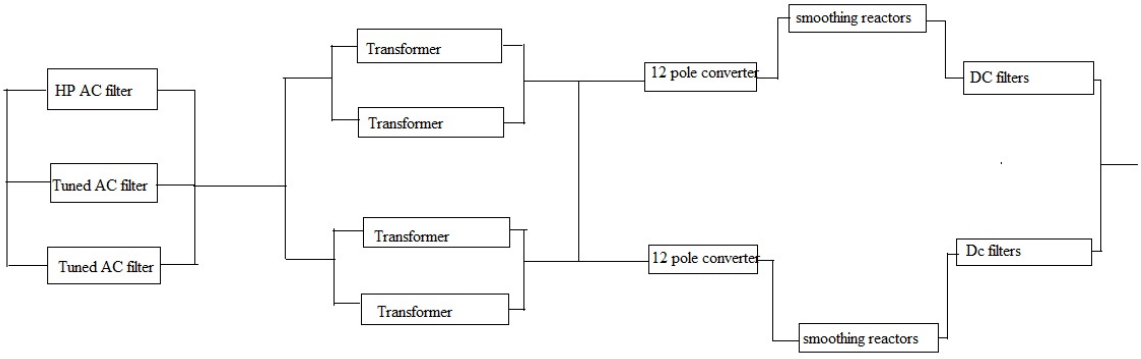


Figure 1: *Reliability block diagram of HVDC converter station*

III. Evaluation of reliability and its various indices of HVDC converter station

The system is having a constant failure rate and therefore the reliability of the system having constant failure rate is evaluated by using the following relation.

$$R(t) = e\text{^}(-\lambda t) \quad (1)$$

Where $R(t)$ represents the reliability of each and every component. λ represents the failure rate per year and t represents time period which is taken as one year.

The mean time to failure (MTTF) can be obtained as follows:

$$MTTF = 1/\lambda \quad (2)$$

A series system is that system in which one component fails, the complete system will fail and for working of the whole system it is mandatory that all the component are in working condition. If one assumes time independent reliability r1, r2…rn, then reliability of series system is given as:

$$R_s = \prod_{i=1}^{n} r_i \quad (3)$$

In series configuration combined failure rate is calculated as follows.

$$\lambda_{Total} = \sum \lambda \quad (4)$$

Unavailability of series configuration is calculated by using following relation.

$$U_{Total} = \sum \lambda r \quad (5)$$

Total repair rate of the components connected in series manner is obtained as follows.

$$r_{Total} = \sum U/\lambda \quad (6)$$

The system or network fails, if all components fail and the system will perform its function even if a single component is working, such a system or network is known as parallel reliability system or network.

The reliability of parallel system (R_p) is given as [23]

$$R_p = 1 - \prod_{i=1}^{n}(1 - r_i) \quad (7)$$

Where ri represents the reliability of components from i=1….n.

Following relation are used to evaluate indices if two components are connected in parallel.

$$\lambda_{para} = \frac{\lambda_1 \lambda_2 (r_1 + r_2)}{8760} \quad (8)$$

$$r_{para} = \frac{r_1 . r_2}{r_1 + r_2} \quad (9)$$

$$U_{para} = \lambda_{para} . r_{para} \quad (10)$$

IV. Result and Discussion

Initial data for different components of the HVDC converter station is shown in Table 1 [30]. There are six components in the HVDC converter station 12 pulse converter, transformer, smoothing reactors, DC filters, tuned AC filters, HP AC filters having mean time to failure (MTTF) as 13.7, 16.1, 76.8, 19.7, 12.6 and 12.6 respectively [30]. The values of mean time to repair (MTTR) for the six components in hours are 6.1, 1700.0, 1700.0, 7.9, 9.3, 9.3 respectively.

Table 1: *Initial data for different components of the HVDC converter station [30].*

Components	MTTF (years)	MTTR (hours)
12 pulse converter (c1)	13.7	6.1
Transformer (c2)	16.1	1700.0
smoothing reactors (c3)	76.8	1700.0
DC filters (c4)	19.7	7.9
tuned AC filters (c5)	12.6	9.3
HP AC filters (c6)	12.6	9.3

Table 2 provides the evaluated values of the failure rate as obtained from equation (2), of the six components of the HVDC converter station, which are obtained as 0.0730, 0.0621, 0.0130, 0.0508, 0.0794 and 0.0794 respectively.

Table 2: *Evaluated failure rate of different components of the HVDC converter station.*

Component	failure rate
c1	0.0730
c2	0.0621
c3	0.0130
c4	0.0508
c5	0.0794
c6	0.0794

The evaluated values of the reliability as obtained from equation (1) for the HVDC converter station components are obtained as 0.9296, 0.9398, 0.9871, 0.9505, 0.9237 and 0.9237 respectively, are shown in Table 3.

Table 3: *Evaluated Reliability of each component of the HVDC converter station.*

Component	Reliability
c1	0.9296
c2	0.9398
c3	0.9871
c4	0.9505
c5	0.9237
c6	0.9237

Table 4 shows the values of evaluated unavailability which can be obtained from equation (5) for each and every individual component only is obtained as 0.4453, 105.57, 22.1, 0.4013, 0.7384 and 0.7384 respectively.

Table 4: *Evaluated unavailability for each and every component of the HVDC converter station.*

component	c1	c2	c3	c4	c5	c6
Unavailability	0.4453	105.57	22.1	0.4013	0.7384	0.7384

Fig. 2 shows the magnitude of failure rate of each component of the HVDC converter station. Magnitude of evaluated reliability of each component of the HVDC converter station is shown in Fig. 3. Fig. 4 and Fig. 5 shows the magnitude of evaluated Unavailability of all components of the HVDC converter station.

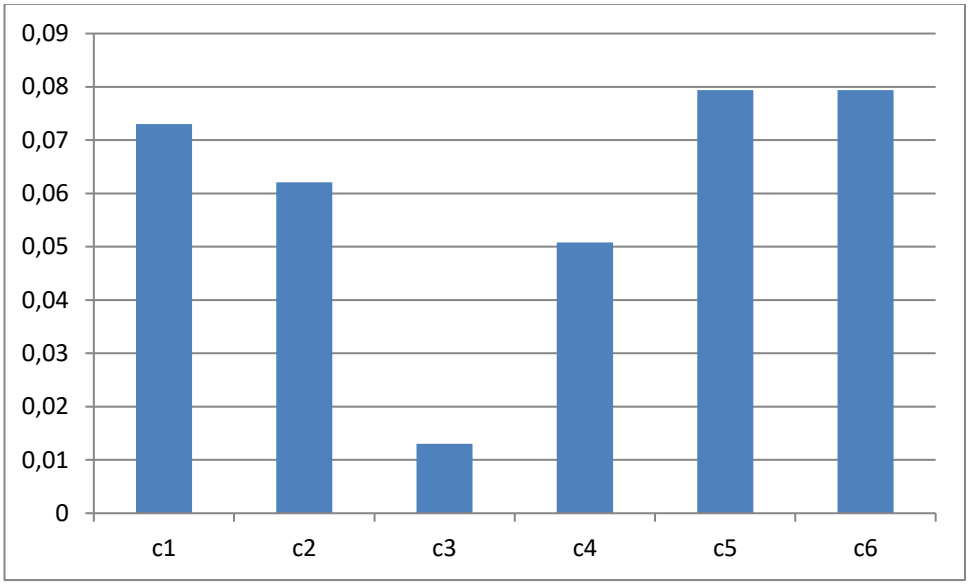


Figure 2: *Magnitude of failure rate of each component of the HVDC converter station.*

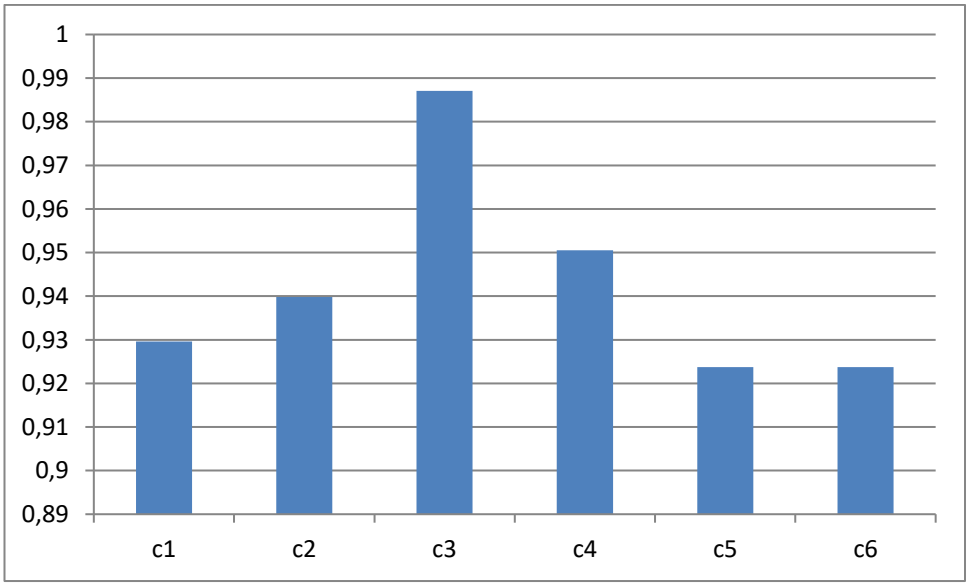


Figure 3: *Magnitude of evaluated reliability of each component of the HVDC converter station.*

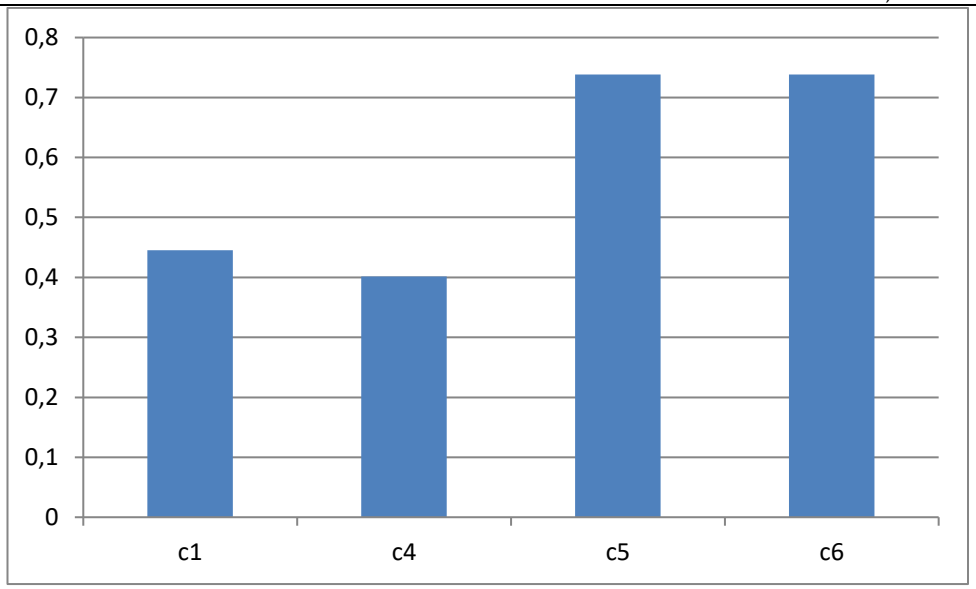


Figure 4: *Magnitude of evaluated Unavailability of components c1, c4, c5, c6 of the HVDC converter station.*

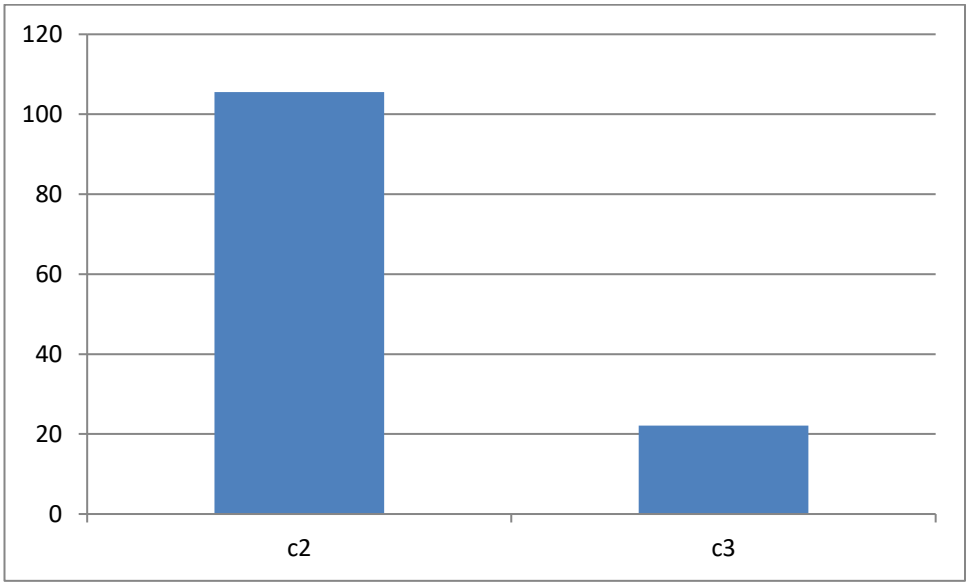


Figure 5: *Magnitude of evaluated Unavailability of components c2, c3 of the HVDC converter station.*

V Conclusion

Evaluation of reliability of engineering system is important and necessary for the overall impact of the system. The reliability of the system depends on its different components and the manner of their existence. This paper proposes the reliability block diagram representation of HVDC converter station. The HVDC converter station whose modeling has been done consists of six components, namely 12 pulse converter, transformer, smoothing reactors, DC filters, tuned AC filters and HP AC filters, which are connected in a manner which is shown in the block diagram. Reliability parameters such as failure rate and reliability are obtained for each component and are discussed in the result section. Unavailability for each and every component of the HVDC converter station is also obtained and discussed in result section. Result section also shows the magnitude of the parameters which are evaluated.

References

[1] Singh, C. (1981). Markov cut-set approach for the reliability evaluation of transmission and distribution systems. *IEEE Trans. on Power Apparatus and Systems*, 100: 2719-2725.

[2] Billinton, R. (1969). Composite system reliability evaluation. *IEEE Trans. on Power Apparatus and Systems*, 88: 276-281.

[3] Wojczynski, E. and Billinton, R. (1985). Effects of distribution system reliability index distributions upon interruption cost/reliability worth estimates. *IEEE Trans. on Power Apparatus and Systems*, 11: 3229-3235.

[4] Verma, A. K., Srividya, A., Kumar, H. M. R. (2002). A framework using uncertainties in the composite power system reliability evaluation. *Electric Power Components and Systems*, 30: 679-691.

[5] Zheng, Z., Cui, L., Hawkes, A. G. (2006). A study on a single-unit Markov repairable system with repair time omission. *IEEE Trans. on Reliability*, 55: 182-188.

[6] Jirutitijaroen, P. and singh, C. (2008). Comparison of simulation methods for power system reliability indexes and their distributions. *IEEE Trans. on Power Systems*, 23: 486-493.

[7] Dzobe, O., Gaunt, C. T., Herman, R. (2012). Investigating the use of probability distribution functions in reliability-worth analysis of electric power systems. *Int. J. of Electrical Power and Energy Systems*, 37: 110-116.

[8] Bae, I. S. and Kim, J. O. (2008). Reliability evaluation of customers in a microgrid. *IEEE Trans. on Power Systems*, 23: 1416-1422.

[9] Billinton, R. and Wang, P. (1998). Reliability-network-equivalent approach to distribution-system-reliability evaluation. *IEE Proc. generation, transmission and distribution*, 145: 149-153.

[10] Tiwary, A. (2017). Reliability enhancement of distribution system using Teaching Learning based optimization considering customer and energy based indices. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 3: 58-62.

[11] Tiwary, A. (2018). An Innovative Self-Adaptive Multi-Population Jaya Algorithm based Technique for Evaluation and Improvement of Reliability Indices of Electrical Power Distribution System. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 4: 299-302.

[12] Tiwary, A., Arya, R., Choube, S. C., Arya, L. D. (2013). Determination of reliability indices for distribution system using a state transition sampling technique accounting random down time omission. *Journal of The Institution of Engineers (India): series B (Springer)*, 94: 71-83.

[13] Tiwary, A. (2019). Inspection-Repair-Based Availability Optimization of Distribution System Using Bare Bones Particle Swarm Optimization. Computational Intelligence: Theories, Applications and Future Directions – Volume II, *Advances in Intelligent Systems and computing*, 799.

[14] Tiwary, A., Arya, R., Arya, L. D., Choube, S. C. (2017). Bootstrapping based technique for evaluating reliability indices of RBTS distribution system neglecting random down time. *The IUP Journal of Electrical and Electronics Engineering*, X: 48-57.

[15] Volkanavski, Cepin, M., Mavko, B. (2009). Application of fault tree analysis for assessment of the power system reliability. *Reliab Eng Syst Safety*, 94: 1116–27.

[16] Li, B.M., Su, C.T., Shen, C.L. (2010). The impact of covered overhead conductors on distribution reliability and safety. *Int J Electr Power Energy Syst*, 32: 281–9.

[17] Tiwary, A. (2018). Self-Adaptive Multi-Population Jaya Algorithm based Reactive Power Reserve Optimization Considering Voltage Stability Margin Constraints. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 4: 341-345.

[18] Arya, R., Tiwary, A., Choube, S. C., Arya, L. D. (2013). A smooth bootstrapping based technique for evaluating distribution system reliability indices neglecting random interruption duration. *Int. J. of Electrical Power and Energy System*, 51: 307-310.

[19] Tiwary, A. (2018). Inspection–Maintenance-Based Availability Optimization of Feeder Section Using Particle Swarm optimization. *Soft Computing for Problem Solving-Advances in Intelligent Systems and Computing*, 816: 257-272.

[20] BinLi, M., TzongSu, C., LungShen, C. (2010). The impact of covered overhead conductors on distribution reliability and safety. *Int. J. of Electrical Power and Energy System*, 32: 281-289.

[21] Tiwary, A. (2019). Reliability evaluation of radial distribution system – A case study. *Int. J. of Reliability: Theory and Applications*, 14, 4(55): 9-13.

[22] Sarantakos, I., Greenwood, D. M., Yi, J., Blake, S. R., Taylor, P. C. (2019). A method to include component condition and substation reliability into distribution system reconfiguration. *Int. J. of Electrical Power and Energy System*, 109: 122-138.

[23] Tiwary, A. (2020). Customer orientated indices and reliability evaluation of meshed power distribution system. *Int. J. of Reliability: Theory and Applications*, 15, 1(56): 10-19.

[24] Tiwary, A. and Patel, P. (2020). Reliability Evaluation of Hose Reel System - A Practical Approach. *Journal of Industrial Safety Engineering*, 7: 30-34.

[25] Battu, N. R., Abhyankar, A. R., Senroy, N. (2019). Reliability Compliant Distribution System Planning Using Monte Carlo Simulation. *Electric power components and systems*, 47: 985-997.

[26] Tiwary, A. (2020). Application of Non-Parametric Bootstrap Technique for evaluating MTTF and Reliability of a Complex Network with Non-Identical Component Failure Laws. *Reliability: Theory and Applications*, 15: 62-69.

[27] Tiwary, A. and Tiwary, S. (2020). Evaluation of Customer Orientated Indices and Reliability Study of Electrical Feeder System. *Reliability: Theory and Applications*, 15: 36-43.

[28] Tiwary, A. and Tiwary, S. (2021). Evaluation of Reliability Indices of Roy Billinton Test System (RBTS) Bus-2 Distribution System for EducationalPurpose. *Reliability: Theory and Applications*, 16: 54-61.

[29] Tiwary, A. and Tiwary, S. (2021). An Innovative Methodology for Evaluation of Reliability Indices of Electric Traction System. *Reliability: Theory and Applications*, 16: 13-21.

[30] K. R. Padiyar. (2010). HVDC Power Transmission Systems. *2nd edition, New Age International (P) Limited*.

TIME DEPENDENT BEHAVIOUR OF A SINGLE SERVER QUEUEING SYSTEM WITH DIFFERENTIATED WORKING VACATIONS SUBJECT TO SYSTEM DISASTER

V Karthick, V Suvitha

•

Department of Mathematics,
College of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulthur – 603 203, Tamilnadu, India.

kv0994@srmist.edu.in, suvithav@srmist.edu.in

Abstract

This study investigates the time dependent behaviour of the single server queue with differentiated working vacations. The model also takes into account the possibility of a disaster happening during busy periods and working vacations, with the repair procedure starting right away. The time-dependent probabilities of system size are described in terms of modified Bessel functions in the paper using explicit equations that were generated using generating functions. Numeric instances have been added to support the theoretical findings even more.

Keywords: Transient Analysis, Differentiated Vacations, Quasi birth death process, Disaster, Repair

1. Introduction

Many queueing systems allow servers to go offline when the system is empty for any period of time. This random period of server absence, known as a server vacation, could indicate the server can take a break or perform an additional task during this period. In 1975, Levy and Yechiali first presented the vacation queuing model. Numerous researchers have worked on queues with vacationing servers during the past few decades. Doshi [5], Takagi [19], Upadhyaya [23], Tian and Zhang [20] and Ke et al. [9] have conducted comprehensive surveys on vacation queueing models, considering various contexts. The queueing model can be applied to a variety of real-world stochastic service systems since server vacations are especially advantageous for systems where the server can use idle time for other activities. Recent research on vacation queueing models has also been done by Sapkota [13] and Tian et al. [21]. The working vacation queue is the queue in which the server serves customers at a rate that is lower than the busy time service rate. This kind of technology has a wide range of practical applications, including the rate at which employees perform their official work both in the office and at home. Servi and Finn [14] introduced the M/M/1 queueing model with working vacations, where a customer isserved at a lower service rate instead of stopping the service completely. M/M/1 queueing model with working vacation and two types of server failure was discussed by Agrawal et al. [1]. Recently, Tian et al. (2021) conducted an analysis of Markovian queues with Bernoulli interruptions and single working

vacations. Recently, Kumar et al. [11] presented a transient analysis of working vacation queueing system.

Differentiated vacations are one of the various types of queueing vacation types. In this case, whenever the system becomes empty, the server initiates a type I vacation, which has a random length. When a server returns from vacation and discovers there is no queue, a new type II random vacation is started. When the server returns from either type I or type II vacations and there are customers in the system, the server starts providing services to them right away and continues doing so until there are no more users in it. The queueing model with differentiated queueing vacations were initially proposed by Ibe and Isijola [6]. Since then, a number of scholars have examined differentiated vacation queueing systems, including Suranga Sampath et al. [17], Suranga Sampath and Jicheng [18], and Jain and Sigman [7]. Vijayashree and Janani [24] analyse a single server differentiated working vacation queueing model's transient behaviour. Recently, a transient analysis of a single server differentiated queuing system is given by Azhagappan and Deepa [3].

In computer systems, telecommunication networks, and other queueing systems, congestion and blocking are frequently predicted using queuing theory and network analysis. Disasters can happen as a result of unforeseen situations, and these systems are frequently prone to unreliability. All jobs sitting in the buffer, including the one being processed by the server, are lost when a calamity happens since the system is rendered inoperable. Towsley and Tripathi [22] were the first to analyse queueing systems that were subject to disasters. This phenomenon was examined by Chen et al. [4], who called it a "mass exodus". As part of their "stochastic clearing" research, Artaljo and Gomez-Coral [2] examined queueing systems with catastrophes. Single-server queueing systems with disasters have been the subject of transient analysis by Kumar et al. [10], Sudhesh and Vairthiyanathan (2019), and Jain and Singh [8]. Recently, Sudhesh et al. [15] gave the transient analysis of single server queue with disaster. We take into account an M/M/1 queue with differentiated working vacations subject to system disaster and server repair. In this sense, we have seen that the service rate is different, but arrival rate is same for all the states.

The proposed queueing model is motivated real-time application with power-saving features in a smart home automation system.The smart home automation system monitors various sensors and devices within a home, controlling tasks such as lighting, temperature, security, and appliances. It continuously processes sensor data and user commands to maintain an optimal and comfortable environment. When no user commands are received and there are no sensor-triggered events for a certain period, the system switches to an type I power save mode to conserve power. In this state, non-essential components are turned off or put into low-power mode.After the type I power save mode duration expires, the system periodically wakes up to check for any new user commands or sensor-triggered events. If there are pending actions or events, it resumes normal operation and executes the necessary tasks. If there are no pending commands or events, the system enters a deeper power-saving state known as type II power saving mode. In this mode, only essential components remain active to maintain basic system functionality and listen for any incoming commands or events.The smart home automation system may be susceptible to security attacks, such as unauthorized access, data breaches, or control manipulation. These attacks can compromise the integrity and privacy of the system and disrupt its normal operation. The repair process starts immediately.The power-saving features in this smart home automation system help reduce energy consumption during periods of inactivity, contributing to energy efficiency and cost savings. The security considerations highlight the importance of safeguarding the system against potential attacks to protect the privacy, safety, and functionality of the smart home environment.

2. Model description

In this research, a $M/M/1$ queuing model with differentiated working vacations and the possibility of disastrous breakdown and repair is taken into account. These are the main

presumptions that underlie this model:

1. Customer Arrivals: Customers arrive according to a Poisson process with a rate of λ. Customers join a single waiting queue based on the sequence in which they arrive. The capacity of the system is similarly predicated on an endless number of potential clients.
2. Service Process: A single server offers the service, and service times during a normal busy period follow an exponential distribution with parameter μ.
3. Vacation Policy : The servers take a type I vacation after fully serving every customer in the system. Once the servers have attended to at least a single customer, this vacation starts. If the system is empty when the servers return after a vacation, a new random vacation of type II is started. If there are still customers in the system when the server returns from either a Type I or type II vacation, it begins serving them right away until the system is completely empty once more.
4. The servers type I and type II vacation times follow exponential distributions, and their vacation rates are indicated by the symbols θ and γ respectively.
5. Arriving customers are served at rates of $\mu_{v1}(\mu_{v2})$ during type-I (II) vacations.
6. Disastrous Breakdown and Repair: There is a chance that a disastrous breakdown will occur when the servers are either away on working vacation or busy serving customers. The breakdowns occurrence follows exponential distribution with a rate of α . When a server fails, the repair procedure begins right away at a rate of β, enabling the servers to function again as quickly as possible.

2.1. The Quasi-Birth-and-Death (QBD) process

At the time t the number of customers in the systems is consider as $H(t)$ and let $I(t)$ be the servers state, where

$$I(t) = \begin{cases} 0, & \text{the server is in busy} \\ 1, & \text{the server is in type-I vacation} \\ 2, & \text{the server is in type-II vacation} \\ 3, & \text{the server in disaster} \end{cases}$$

Then $X(t) = \{H(t), I(t)\}$, is a Continuous time Markov chain with a state space denoted by Ω as follows:
$\Omega = \{\{(i,j), i \geq 0, j = 0,1,2,3\}$.

3. Transient Analysis

Let $P_{n,j}(t)$ be the time-dependent probability for the system to be in state j with n customers at time t.

$$P'_{0,0}(t) = -(\lambda+\alpha)P_{0,0}(t) + \beta P_{0,3}(t) \tag{1}$$

$$P'_{n,0}(t) = -(\lambda+\alpha+\mu)P_{n,0}(t) + \mu P_{n+1,0}(t) + \gamma_1 P_{n,1}(t) + \gamma_2 P_{n,2}(t) + \beta P_{n,3}(t) + \lambda P_{n-1,0} \text{ for } n \geq 1 \tag{2}$$

$$P'_{0,1}(t) = -(\lambda+\alpha+\gamma_1)P_{0,1}(t) + \mu_{v1} P_{1,1}(t) + \mu P_{1,0}(t) \tag{3}$$

$$P'_{n,1}(t) = -(\lambda+\alpha+\gamma_1+\mu_{v1})P_{n+1,1}(t) + \lambda P_{n-1,1}(t) \text{ for } n \geq 1 \tag{4}$$

$$P'_{0,2}(t) = -(\lambda+\alpha)P_{0,2}(t) + \gamma_1 P_{0,1}(t) + \mu_{v2} P_{1,2}(t) \tag{5}$$

$$P'_{n,2}(t) = -(\lambda+\alpha+\mu_{v2}+\gamma_2)P_{n,2}(t) + \lambda P_{n-1,2}(t) + \mu_{v2} P_{n+1,2}(t) \text{ for } n \geq 1 \tag{6}$$

$$P'_{0,3}(t) = -(\lambda+\beta)P_{0,3}(t) + \alpha\big(1 - \sum_{n=0}^{\infty} P_{n,3}(t)\big) \tag{7}$$

$$P'_{n,3}(t) = -(\lambda+\beta)P_{n,3}(t) + \lambda P_{n-1,3}(t) \text{ for } n \geq 1 \tag{8}$$

We assume the initial condition as,
$P_{0,1}(0) = 1$, $P_{0,i}(0) = 0$ for $i = 0, 2, 3,$, $P_{n,i}(0) = 0$ for $n \geq 1, i = 0, 1, 2, 3,$
Taking laplace on equations (1), (3), (5), (7), (8).

$$\hat{P}_{0,0}(s) = \frac{\beta}{(s+\lambda+\alpha)}\hat{P}_{0,3}(s) \tag{9}$$

$$\hat{P}_{0,1}(s) = \frac{1}{s+\lambda+\alpha+\gamma_1} + \frac{\mu_{v1}}{s+\lambda+\alpha+\gamma_1}\hat{P}_{1,1}(s) + \frac{\mu}{(s+\lambda+\alpha+\gamma_1)}\hat{P}_{1,0}(s) \tag{10}$$

$$\hat{P}_{0,2}(s) = \frac{\gamma_1}{s+\lambda+\alpha}\hat{P}_{0,1}(s) + \frac{\mu_{v2}}{(s+\lambda+\alpha)}\hat{P}_{1,2}(s) \tag{11}$$

$$\hat{P}_{0,3}(s) = \frac{\alpha}{s(s+\lambda+\beta)} - \frac{\alpha}{(s+\lambda+\beta)}\sum_{n=0}^{\infty}\hat{P}_{n,3}(s) \tag{12}$$

$$\hat{P}_{n,3}(s) = \frac{\lambda}{(s+\lambda+\beta)}\hat{P}_{n-1,3}(s) \tag{13}$$

The above equation (13) recursively yields

$$\hat{P}_{n,3}(s) = \frac{\lambda^n}{(s+\lambda+\beta)^n}\hat{P}_{0,3}(s) \text{ for } n \geq 1 \tag{14}$$

Define
$Q_1(z,t) = \sum_{n=0}^{\infty} P_{n,1}(t)z^n$ then $\frac{\partial Q_1(z,t)}{\partial t} = \sum_{n=0}^{\infty} P'_{n,1}(t)z^n$
Multiplying the equations (3) and (4) by the appropriate powers of z and summing over $n \geq 1$ we obtain,

$$\frac{\partial Q_1(z,t)}{\partial t} + \left((\lambda+\alpha+\gamma_1+\mu_{v1}) - \left(\frac{\mu_{v1}}{z} + \lambda z\right)\right)Q_1(z,t) = \mu_{v1}P_{0,1}(t) - \frac{\mu_{v1}}{z}P_{0,1}(t) + \mu P_{1,0}(t)$$

Upon integrating the above linear differential equation with respect to t, we get

$$Q_1(z,t) = \int_0^t \left(\mu_{v1}P_{0,1}(t) + \frac{\mu_{v2}}{z}P_{0,1}(t) + \mu P_{0,1}(t)\right)(e^{-(\lambda+\alpha+\gamma_1+\mu_{v1})(t-y)} \times e^{((\mu_{v1}/z)+\lambda z)(t-y)})dy \tag{15}$$

If $a_i = 2\sqrt{\lambda\mu_{vi}}$ and $b_i = \sqrt{\lambda/\mu_{vi}}$ then $e^{(\mu_{vi}/z+\lambda z)t} = \sum_{-\infty}^{\infty}(b_i z)^n I_n(a_i t)$ for $i = 1, 2$ where $I_n(a_i t)$ is a bessel funtion of order n. Using that fact in equation (15) and comparing the terms coeffients of z^n for $n = 1, 2, 3, ...$

$$P_{n,1}(t) = \int_0^t \left(\big(\mu_{v1}P_{0,1}(t) + \mu P_{1,0}(t)\big)b_1^n I_n(.)e^{-k_1(t-y)}\right)dy + \int_0^t \left(\mu_{v1}P_{1,0}(t)b_1^{n+1}I_{n+1}(.)e^{-k_1(t-y)}\right)dy \tag{16}$$

Equating the coeffients of z^{-n} for $n = 1, 2, ...$ and applying $I_{-n}(.) = I_n(.)$ we get

$$0 = \int_0^t \left(\big(\mu_{v1}P_{0,1}(t) + \mu P_{1,0}(t)\big)b_1^{-n} I_n(.)e^{-k_1(t-y)}\right)dy + \int_0^t \left(\mu_{v1}P_{1,0}(t)b_1^{-n+1}I_{-n+1}(.)e^{-k_1(t-y)}\right)dy \tag{17}$$

where $k_1 = \lambda + \alpha + \gamma_1 + \mu_{v1}$ and $I_n(.) = I_n(a(t-y))$

Multiply equation (17) by b_1^{2n} and subtract from equation (16)

$$P_{n,1}(t) = \int_0^t \left(\mu_{v1} P_{0,1}(t) b_1^{n+1} \left[I_{n-1}(.) - I_{n+1}(.) \right] e^{-k_1(t-y)} \right) dy$$

Taking laplace transform on both sides

$$\hat{P}_{n,1}(s) = 2\mu_{v1} \frac{b_1^{n+1}}{a_1} \hat{\psi}(s)^n \hat{P}_{0,1}(s) \quad \text{for } n \geq 0 \tag{18}$$

In similar way using the equations (5), (6) we get

$$\hat{P}_{n,2}(s) = 2\mu_{v2} \frac{b_2^{n+1}}{a_2} \hat{\psi}(s)^n \hat{P}_{0,2}(s) \quad \text{for } n \geq 0 \tag{19}$$

Define
$Q_3(z,t) = \sum_{n=0}^{\infty} P_{n,0}(t) z^n$ then $\frac{\partial Q_3(z,t)}{\partial t} = \sum_{n=0}^{\infty} P'_{n,0}(t) z^n$
Multiplying the equations (1) and (2) by the appropriate powers of z and summing over $n \geq 0$ we obtain,

$$\begin{aligned} \frac{\partial Q_3(z,t)}{\partial t} + \left((\lambda + \alpha + \mu) - \left(\frac{\mu}{z} + \lambda z \right) \right) Q_3(z,t) &= \mu P_{0,0}(t) - \frac{\mu}{z} P_{0,0}(t) + \mu P_{1,0}(t) \\ &+ \gamma_1 \sum_{n=1}^{\infty} P_{n,1}(t) z^n + \gamma_2 \sum_{n=1}^{\infty} P_{n,2}(t) + \beta \sum_{n=0}^{\infty} P_{n,3}(t) z^n \end{aligned} \tag{20}$$

If $a_3 = 2\sqrt{\lambda\mu}$ and $b_i = \sqrt{\lambda/\mu}$ then $e^{(\mu/z + \lambda z)t} = \sum_{-\infty}^{\infty} (b_3 z)^n I_n(a_3 t)$ where $I_n(a_3 t)$ is a bessel funtion of order n. Using that fact in equation (20) and comparing the terms coeffients of z^n for $n = 1, 2, 3, ...$

$$\begin{aligned} P_{n,0}(t) = \int_0^t \left(\left(\mu P_{0,1}(t) - \mu P_{1,0}(t) \right) b_3^n I_n(.) e^{-k_3(t-y)} \right) dy - \int_0^t \left(\mu P_{0,0}(t) b_3^{n+1} I_{n+1}(.) \right. \\ \left. \times e^{-k_3(t-y)} \right) dy + \int_0^t \left(\gamma_1 \sum_{m=1}^{\infty} \left(\gamma_1 P_{m,1}(t) z^m + \gamma_2 P_{n,2}(t) z^m \right. \right. \\ \left. \left. + \beta P_{m,3}(s) \right) b_3^{n-m} I_{n-m}(.) e^{-k_3(t-y)} \right) dy \end{aligned} \tag{21}$$

Equating the coeffients of z^{-n} for $n = 1, 2, ...$ and applying

$$\begin{aligned} 0 = \int_0^t \left(\left(\mu P_{0,1}(t) - \mu P_{1,0}(t) \right) b_3^{-n} I_n(.) e^{-k_3(t-y)} \right) dy - \int_0^t \left(\mu P_{0,0}(t) b_3^{-n+1} I_{-n+1}(.) \right. \\ \left. \times e^{-k_3(t-y)} \right) dy + \int_0^t \left(\gamma_1 \sum_{m=1}^{\infty} \left(\gamma_1 P_{m,1}(t) z^m + \gamma_2 P_{n,2}(t) z^m \right. \right. \\ \left. \left. + \beta P_{m,3}(s) \right) b_3^{-(n+m)} I_{n+m}(.) e^{-k_3(t-y)} \right) dy \end{aligned} \tag{22}$$

Multiply equation (22) by b_1^{2n} and subtract from equation (21)

$$\begin{aligned} P_{n,0}(t) = \int_0^t \left(\mu P_{0,0}(t) b_3^{n+1} \left[I_{n-1}(.) - I_{n+1}(.) \right] e^{-k_3(t-y)} \right) dy + \int_0^t \left(\gamma_1 \sum_{m=1}^{\infty} \left(\gamma_1 P_{m,1}(t) z^m \right. \right. \\ \left. \left. + \gamma_2 P_{n,2}(t) z^m + \beta P_{m,3}(s) \right) b_3^{n-m} \left[I_{n-m}(.) - I_{n+m}(.) \right] e^{-k_3(t-y)} \right) dy \end{aligned}$$

Taking laplace transform on both sides

$$\hat{P}_{n,0}(s) = \frac{1}{\sqrt{\omega_3^2 - a_3^2}}\left[\sum_{m=1}^{\infty}\left(\gamma_1\hat{P}_{m,1}(s) + \gamma_2\hat{P}_{m,2}(s)\right) + \sum_{m=0}^{\infty}\beta\hat{P}_{m,3}(s)\right] + 2\mu\frac{b_3^{n+1}}{a_3}\hat{\psi(s)}^n\hat{P}_{0,0}(s) \text{ for } n \geq 0 \tag{23}$$

substitute $n = 1$ in (18) and (19)

$$\hat{P}_{1,1}(s) = 2\mu_{v1}b_1^2\frac{\hat{\psi(s)}^s}{a1}\hat{P}_{0,1}(s) \tag{24}$$

$$\hat{P}_{1,2}(s) = 2\mu_{v2}b_2^2\frac{\hat{\psi(s)}^s}{a2}\hat{P}_{0,2}(s) \tag{25}$$

Substitute (24) and (25) in $\hat{P}_{0,1}(s)$

$$\hat{P}_{0,1}(s) = \frac{1}{(s+\lambda+\alpha+\gamma_1)}\left[1 + 2\mu_{v1}^2b_1^2\frac{\hat{\psi(s)}}{a_2}\hat{P}_{0,1}(s) + \mu\hat{P}_{1,0}(s)\right]$$

$$\hat{P}_{0,1}(s) = \left[\sum_{j=0}^{\infty}\left(\frac{2\mu_{v1}^2b_1^2\hat{\psi(s)}}{(s+\lambda+\alpha+\gamma_1)a_1}\right)^j\right]\frac{1}{(s+\lambda+\alpha+\gamma_1)} + \left[\sum_{j=0}^{\infty}\left(\frac{2\mu_{v1}^2b_1^2\hat{\psi(s)}}{(s+\lambda+\alpha+\gamma_1)a_1}\right)^j\right]\frac{\mu}{(s+\lambda+\alpha+\gamma_1)}\hat{P}_{1,0}(s)$$

$$\hat{P}_{0,1}(s) = \hat{A}_1(s) + \hat{A}_1(s)\mu\hat{P}_{1,0}(s) \tag{26}$$

Substitue (25) and (26) in $\hat{P}_{0,2}(s)$

$$\hat{P}_{0,2}(s) = \frac{\gamma_1}{(s+\lambda+\alpha)}\hat{A}_1(s) + \frac{\gamma_2}{(s+\lambda+\alpha)}\hat{A}_2(s)\hat{P}_{1,0}(s) + \frac{2\mu_{v2}^2b_2^2}{(s+\lambda+\alpha)a_2}\hat{\psi}(s)\hat{P}_{0,2}(s)$$

$$\hat{P}_{0,2}(t) = \left[\sum_{j=0}^{\infty}\frac{2\mu_{v2}^2b_2^2}{(s+\lambda+\alpha)a_2}\hat{\psi}(s)\right]^j\frac{\gamma_1}{(s+\lambda+\alpha)}\hat{A}_1(s) + \left[\sum_{j=0}^{\infty}\frac{2\mu_{v2}^2b_2^2}{(s+\lambda+\alpha)a_2}\hat{\psi}(s)\right]^j\frac{\gamma_2}{(s+\lambda+\alpha)}\hat{A}_2(s)\hat{P}_{1,0}(s)$$

$$\hat{P}_{0,2}(s) = \gamma_1\hat{A}_2(s) + \gamma_2\hat{A}_2(s)\hat{P}_{1,0}(s) \tag{27}$$

Substitue (26) and (27) in (18) and (19)

$$\hat{P}_{n,1}(s) = 2\mu_{v1}b_1^{n+1}\frac{\hat{\psi}(s)^n}{a_1}\left(\hat{A}_1(s) + \mu\hat{A}_1(s)\hat{P}_{1,0}(s)\right) \tag{28}$$

$$\hat{P}_{n,2}(s) = 2\mu_{v2}b_2^{n+1}\frac{\hat{\psi}(s)^n}{a_2}\left(\gamma_1\hat{A}_2(s) + \gamma_2\hat{A}_2(s)\hat{P}_{1,0}(s)\right) \tag{29}$$

Substitute (14), (28), (29) in (23)

$$\hat{P}_{n,0}(s) = \frac{1}{\sqrt{\omega_3^2 - a_3^2}}\left[\sum_{m=1}^{\infty}\left(2\gamma_1\mu_{v1}b_1^{m+1}\frac{\hat{\psi}(s)}{a_1}\left(\hat{A}_1(s) + \mu\hat{A}_1(s)\hat{P}_{1,0}(s)\right) + 2\gamma_2\mu_{v2}b_2^{m+1}\frac{\hat{\psi}(s)}{a_2}\left(\gamma_1\hat{A}_2(s) + \gamma_2\hat{A}_2(s)\hat{P}_{1,0}(s)\right)\right) + \sum_{m=0}^{\infty}\frac{\beta\lambda^m}{(s+\lambda+\beta)^m\hat{P}_{0,3}(s)}\right] \times\hat{X}_3(s) + 2\frac{\mu b_3^{n+1}}{a_3}\hat{\psi}(s)^n\frac{\beta}{(s+\lambda+\alpha)}\hat{P}_{0,3}(s) \tag{30}$$

Substitute $n = 1$ in the above equation

$$\hat{P}_{1,0}(s) = \frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \left[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \hat{A}_1(s) + 2\gamma_2 \mu_{v2} b_2^{m+1} \frac{\hat{\psi}^m(s)}{a_2} \gamma_1 \hat{A}_2(s) \right] \hat{X}_3(s)$$
$$+ \frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \left[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \hat{A}_2(s) + 2\gamma_2 \mu_{v2} b_2^{m+1} \frac{\hat{\psi}^m(s)}{a_2} \gamma_2 \hat{A}_2(s) \right]$$
$$\times \hat{X}_3(s) \hat{P}_{1,0}(s) + \sum_{m=0}^{\infty} \frac{\beta \lambda^m}{(s+\lambda+\beta)^m \hat{P}_{0,3}(s)} \Big] \hat{X}_3(s)$$
$$+ 2\frac{\mu b_3^2}{a_3} \hat{\psi}(s) \frac{\beta}{(s+\lambda+\alpha)} \hat{P}_{0,3}(s)$$

$$\hat{P}_{1,0}(s) = \Big[\sum_{j=0}^{\infty} \Big(\frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \Big[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \mu \hat{A}_1(s) + 2\gamma_2 \mu_{v2} b_2^{m+1} \frac{\hat{\psi}^m(s)}{a_2} \gamma_2$$
$$\times \hat{A}_2(s) \Big] \hat{X}_3(s) \Big)^j \Big] \Big[\frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \Big[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \hat{A}_1(s) + 2\gamma_2 \mu_{v2} b_2^{m+1}$$
$$\times \frac{\hat{\psi}^m(s)}{a_2} \gamma_1 \hat{A}_2(s) \Big] \hat{X}_3(s) + \Big(\sum_{m=0}^{\infty} \frac{\beta \lambda^m}{(s+\lambda+\beta)^m} \hat{X}_3(s) + 2\frac{\mu b_3^2}{a_3} \hat{\psi}(s)$$
$$\times \frac{\beta}{(s+\lambda+\alpha)} \Big) \hat{P}_{0,3}(s) \Big]$$
$$\hat{P}_{1,0}(s) = \hat{A}_3(s) \hat{P}_{0,3}(s) + \hat{A}_4(s) \tag{31}$$

where,

$$A_1(s) = \Big[\sum_{j=0}^{\infty} \Big(\frac{2\mu_{v1}^2 b_1^2 \hat{\psi(s)}}{(s+\lambda+\alpha+\gamma_1) a_1} \Big)^j \Big] \frac{1}{(s+\lambda+\alpha+\gamma_1)}$$
$$A_2(s) = \Big[\sum_{j=0}^{\infty} \frac{2\mu_{v2}^2 b_2^2}{(s+\lambda+\alpha) a_2} \hat{\psi}(s) \Big]^j \frac{1}{(s+\lambda+\alpha)}$$
$$A_3(s) = \Big[\sum_{j=0}^{\infty} \Big(\frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \Big[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \mu \hat{A}_1(s) + 2\gamma_2 \mu_{v2} b_2^{m+1}$$
$$\times \frac{\hat{\psi}^m(s)}{a_2} \gamma_2 \hat{A}_2(s) \Big] \hat{X}_3(s) \Big)^j \Big] \Big[\frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \Big[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \hat{A}_1(s)$$
$$+ 2\gamma_2 \mu_{v2} b_2^{m+1} \frac{\hat{\psi}^m(s)}{a_2} \gamma_1 \hat{A}_2(s) \Big] \hat{X}_3(s)$$
$$A_4(s) = \Big[\sum_{j=0}^{\infty} \Big(\frac{1}{\sqrt{\omega_3^2 - a_3^2}} \sum_{m=1}^{\infty} \Big[2\gamma_1 \mu_{v1} b_1^{m+1} \frac{\hat{\psi}^m(s)}{a_1} \mu \hat{A}_1(s) + 2\gamma_2 \mu_{v2} b_2^{m+1} \frac{\hat{\psi}^m(s)}{a_2} \gamma_2$$
$$\times \hat{A}_2(s) \Big] \hat{X}_3(s) \Big)^j \Big] \Big(\sum_{m=0}^{\infty} \frac{\beta \lambda^m}{(s+\lambda+\beta)^m} \hat{X}_3(s) + 2\frac{\mu b_3^2}{a_3} \hat{\psi}(s) \frac{\beta}{(s+\lambda+\alpha)} \Big)$$
$$\hat{X}_3(s) = b_3^{n-m} \big[I_{n-m}(.) - I_{n+m}(.) \big] e^{-k_3(t-y)}$$

Substitute (31) in (30)

$$\hat{P}_{n,0}(s) = \Bigg[\frac{1}{\sqrt{\omega_3^2 - a_3^2}}\Bigg[\sum_{m=1}^{\infty}\left(2\gamma_1\mu_{v1}b_1^{m+1}\frac{\hat{\psi}(s)^m}{a_1}\right)\mu\hat{A}_1(s)\hat{A}_3(s) + \left(2\gamma_2\mu_{v2}b_2^{m+1}\frac{\hat{\psi}(s)^m}{a_2}\right)\gamma_2\hat{A}_2(s)\hat{A}_3(s)\Bigg]\hat{X}_3(s) + \sum_{m=0}^{\infty}\frac{\beta\lambda^m}{(s+\lambda+\beta)^m} + 2\mu\frac{b_3^{n+1}}{a_3}\hat{\psi}^n(t)\frac{\beta}{(s+\lambda+\alpha)}\Bigg]\hat{P}_{0,3}(s) + \Bigg[\frac{1}{\sqrt{\omega_3^2 - a_3^2}}\sum_{m=1}^{\infty}\Bigg(2\gamma_1\mu_{v1}b_1^{m+1}\frac{\hat{\psi}(s)^m}{a_1}\big(\hat{A}_1(s) + \mu\hat{A}_1(s)\hat{A}_4(s)\big) + 2\gamma_2\mu_{v2}b_2^{m+1}\frac{\hat{\psi}(s)^m}{a_2}\big(\gamma_1\hat{A}_2(s) + \gamma_2\hat{A}_2\hat{A}_4(s)\big)\Bigg)\Bigg]\hat{X}_3(s) \tag{32}$$

Substitute (31) in (28) and (29)

$$\hat{P}_{n,1}(s) = 2\mu_{v1}b_1^{n+1}\frac{\hat{\psi}(s)}{a_1}\big[\hat{A}_1(s) + \mu\hat{A}_1(s)\hat{A}_4(s) + \mu\hat{A}_1(s)\hat{A}_3(s)\hat{P}_{0,3}(s)\big] \tag{33}$$

$$\hat{P}_{n,2}(s) = 2\mu_{v2}b_2^{n+1}\frac{\hat{\psi}(s)}{a_2}\big[\gamma_1\hat{A}_2(s) + \gamma_2\hat{A}_2(s)\hat{A}_4(s) + \gamma_2\hat{A}_2(s)\hat{A}_6(s)\hat{P}_{0,3}(s)\big] \tag{34}$$

Inverting (14), (32)-(34)

$$P_{n,5}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!}e^{-(\lambda+\beta)t} * P_{0,3}(t) \text{ for } n \geq 1 \tag{35}$$

$$P_{n,0}(t) = \Bigg[I_0(t)\Bigg(\sum_{m=1}^{\infty}2\gamma_1\mu_{v1}\frac{b_1^{m+1}}{a_1}\psi(t)^m * \mu A_1(t) * A_3(t) + 2\gamma_2\mu_{v2}\frac{b_2^{m+1}}{a_2}\psi(t)^m * \gamma_2 A_2(t) * A_3(t)\Bigg) * X_3(t) + \sum_{m=0}^{\infty}\frac{\lambda^m t^{m-1}}{(m-1)!}e^{-(\lambda+\beta)t} + 2\mu\frac{b_3^{n+1}}{a_3}\psi(t) * \beta e^{-(\lambda+\beta)t}\Bigg] * P_{0,3}(t) + \Bigg[I_0(t)\Bigg(\sum_{m=1}^{\infty}2\gamma_1\mu_{v1}\frac{b_1^{m+1}}{a_1}\psi(t)^m * \big(\mu A_1(t) + \mu A_1(t) * A_4(t)\big) + 2\gamma_2\mu_{v2}\frac{b_2^{m+1}}{a_2}\psi(t)^m * \big(\gamma_1 A_2 + \gamma_2 A_2(t) * A_4(t)\big)\Bigg) * X_3(t)\Bigg] \text{ for } n \geq 0 \tag{36}$$

$$P_{n,1}(t) = 2\mu_{v1}\frac{b_1^{n+1}}{a_1}\psi(t)\big[A_1(t) + \mu A_1(t) * A_4(t) + \mu A_1(t) * A_3(t)P_{0,3}(t)\big] \text{ for } n \geq 0 \tag{37}$$

$$P_{n,2}(t) = 2\mu_{v2}\frac{b_2^{n+1}}{a_2}\psi(t)\big[\gamma_1 A_2(t) + \gamma_2 A_2(t) * A_4(t) + \gamma_2 A_2(t) * A_4(t)P_{0,3}(t)\big] \text{ for } n \geq 0 \tag{38}$$

Here all the probabilities are purely expressed in terms of $P_{0,3}(t)$. Using (12) we can find $P_{0,3}(t)$ in the following manner

$$\hat{P}_{0,3}(s) = \Bigg[\sum_{j=0}^{\infty} -\Bigg(\frac{\alpha}{(s+\lambda+\alpha+\beta)}\sum_{n=1}^{\infty}\frac{\lambda^n}{(s+\lambda+\beta)^n}\Bigg)^j\Bigg]\Bigg[\frac{\alpha}{s(s+\lambda+\alpha+\beta)}\Bigg] \tag{39}$$

Inverting the above

$$P_{0,3}(t) = \left[\sum_{j=0}^{\infty} \alpha e^{-(\lambda+\alpha+\beta)t} * \sum_{n=1}^{\infty} \frac{\lambda^n t^{n-1} e^{-(\lambda+\beta)t}}{(n-1)!}\right] * \left[\delta(t) * \alpha e^{-(\lambda+\alpha+\beta)t}\right] \tag{40}$$

4. Numerical Analysis

In this section, graphs show the system's transient probabilities in various states, including busy state, type-I and type-II vacation states, as well as disaster state. Additionally, the system's mean is recorded over time. The following parameter values were used to generate the graphs: $\lambda = 2$, $\mu = 3$, $\mu_{v1} = 0.8$, $\mu_{v2} = 0.7$, $\alpha = 0.6$ and $\beta = 0.5$. Figure 2 dipicts the behaviour of the transient probability of busy period against time t for different values of n. The probability curves start at 0 and converge to a steady state over time, as shown by this graph. Figure 3 displays the behaviour of transient probabilities during the type-I vacation period, demonstrating that all probability curves begin at 0 and progressively rise to a certain extent as t rises before stabilising.

Figure 4 exhibits the graph of type-II vacation period transient probability over time t. The probability curves in that graph start at 0 and move towards a steady state over time. Furthermore, it is clear that a type-I vacation has a higher probability of having more customers than a type-II vacation does at any given time instance t. This discrepancy results from the servers in type-II vacation mode quickly switching to busy mode after the vacation is over, whereas type-I vacation takes a longer period of time. As a result, during type-II vacation, customers do not need to wait for additional processing time.

The behaviour of $P_{n,3}(t)$ appears in Figure 5, where all probability curves initially start at 0 and gradually grow to some extent as t increases, finally reaching a steady state. The mean behaviour for various disaster and repair rate values is shown in Figures 6 and 7. According to these data, an increase in repair rate results in an increase in the mean size. Similar to this, a rise in the disaster rate causes a fall in the mean size.

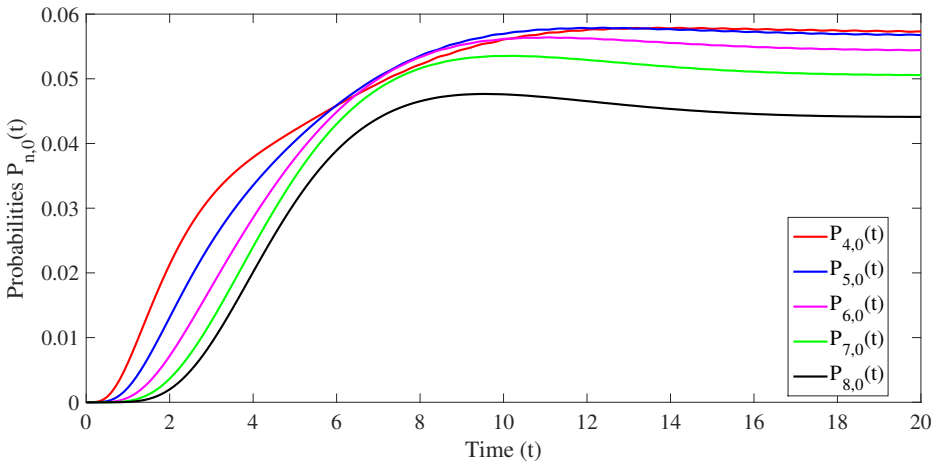


Figure 1: *Probabilities $P_{n,0}(t)$ Vs Time*

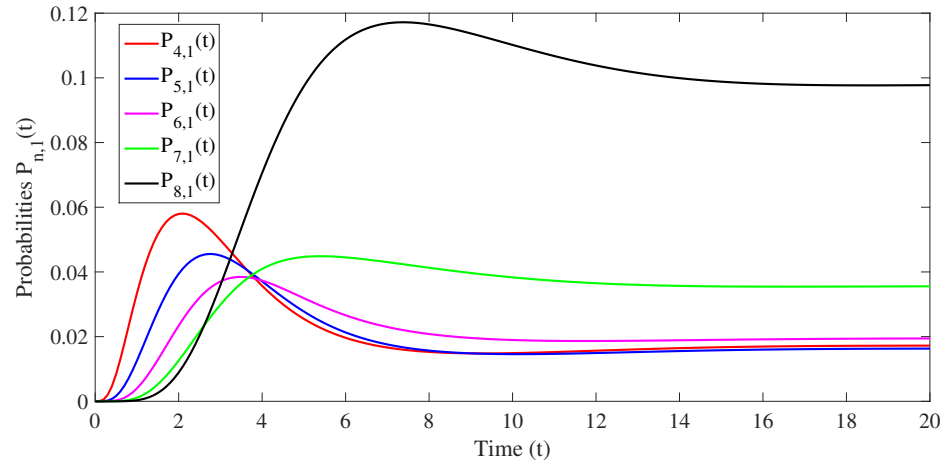


Figure 2: *Probabilities $P_{n,1}(t)$ Vs Time*

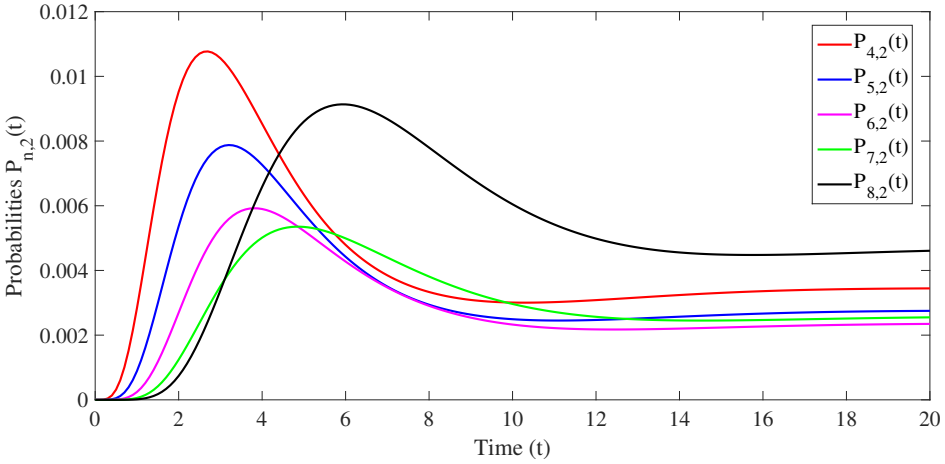


Figure 3: *Probabilities $P_{n,2}(t)$ Vs Time*

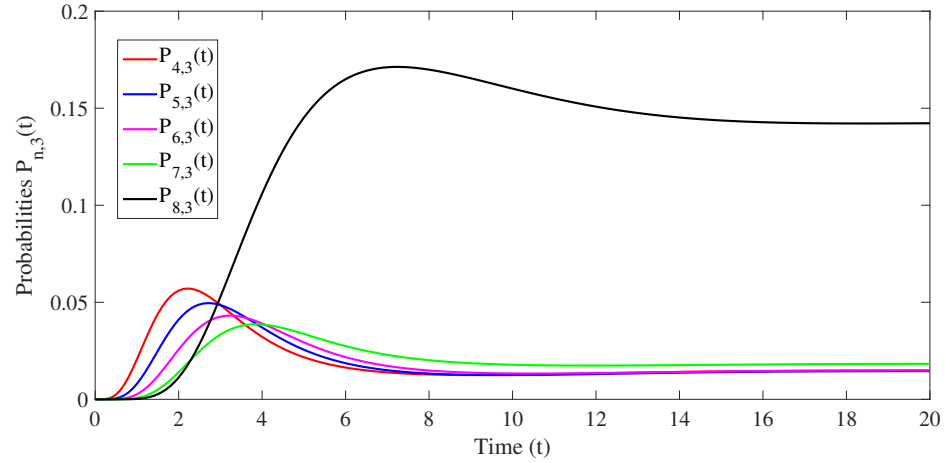


Figure 4: *Probabilities $P_{n,3}(t)$ Vs Time*

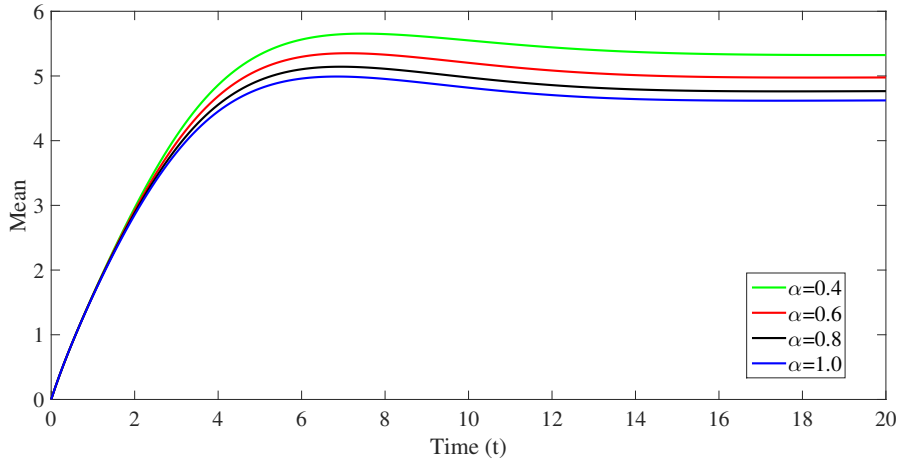


Figure 5: *Probabilities Mean Vs Time*

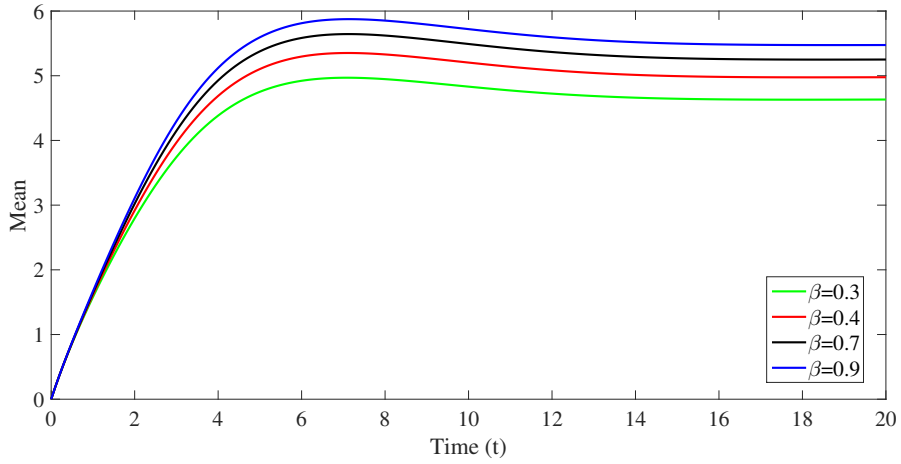


Figure 6: *Mean Vs Time*

5. Conclusion and Future work

In this study, a single server queueing system with multiple differentiated vacation, disaster, and repair periods was investigated. The modified Bessel function of the first kind was used to derive the time-dependent probability of the system size. The proposed model's numerical results indicate that the time-dependent probabilities eventually reach their respective steady-state probabilities.

By taking into account multi-server differentiated vacation queueing systems with disaster and repair, future research can build on this study. Analysing such systems would provide us with a more complete understanding of their performance and behaviour. It would also be advantageous to investigate stochastic decomposition for this model since it can provide insightful information about the dynamics of the system and aid in improving performance. These avenues of inquiry will advance queueing theory as a whole and improve our comprehension of intricate queueing systems in real-world situations.

References

[1] Agarwal, P. K., Jain, A. and Jain, M. (2021). M/M/1 queueing model with working vacation and two type of server breakdown, *Journal of Physics: Conference series*, 1849:012021.

[2] Artalejo, J. T. and Gomez-Corral, A. (1998). Analysis of a stochastic clearing system with repeated attempts, *Stochatic Models*, 14:623-645.

[3] Azhagappan, A. and Deepa, T. (2023). Transient analysis of various vacation interruption policies in a Markovian queue with differentiated multiple vacations, *International journal of operational research*, 45:492-510.

[4] Chen, J. J., Wu, S. L. and Wang, S. W. (2009). Power saving class management for energy saving in IEEE 802.16e wireless networks, *Tenth International Conference on Mobile Data Management: Systems, Services and Middleware*, 490-495.

[5] Doshi, B. T. (1986). Queueing systems with vacations-a survey, *Queueing systems*, 1:22-66.

[6] Ibe, O. C. and Isijola, O. A. (2014). $M/M/1$ Multiple vacation queueing systems with differentiated vacations, *Modelling and Simulation in Engineering*, DOI: 10.1155/2014/158247.

[7] Jain, G. and Sigman, K. (1996). A Pollaczek-Khintchine formula for $M/G/1$ queues with disasters, *Journal of Applied Probability*, 33:1191-1200.

[8] Jain, M. and Singh, M. (2020). Transient analysis of a Markov queueing model with feedback, discouragement and disaster, *International Journal of Applied and Computational Mathematics*, 6:1-14.

[9] Ke, J. C., Wu, C. H. and Zhang, Z. G. (2010). Recent developments in vacation queueing models: A short survey, *Operation Research*, 7:3-8.

[10] Kumar, B. K., Anantha Lakshmi, S. R., Anbarasu, S. and Pavai Madheswari, S. (2014). Transient and steady-state analysis of queueing systems with catastrophes and impatient customers, *International Journal of Mathematics in Operational Research*, 6:523-549.

[11] Kumar, P., Jain, M. and Meena, R. K. (2023). Transient analysis and reliability modeling of fault-tolerant system operating under admission control policy with double retrial features and working vacation, *ISA transactions*, 134:183-199.

[12] Levy, Y. and Yechiali, U. (1975). Utilization of idle time in an $M/G/1$ queueing system, *Management Science*, 22:202-211.

[13] Sapkota, G. and Ghimire, R. P. (2022). Mathematical analysis of $M/M/C$ vacation queueing model with a waiting server and impatient customers, *Journal of Mathematics and Statistics*, 18:36-48.

[14] Servi, L. D. and Finn, S. G. (2002). M/M/1 queues with working vacations (M/M/1/WV), *Performance Evaluation*, 50:41-52.

[15] Sudhesh, R., Mohammed Shapique, A. and Dharmaraja, S. (2022). Analysis of a multiple dual-stage vacation queueing system with disaster and repairable server, *Methodology and Computing in Applied Probability* 24:2485-2508.

[16] Sudhesh, R. and Vaithiyanathan, A. (2019). Analysis of state-dependent discrete-time queue with system disaster, *RAIRO-Operations Research*, 53:1915-1927.

[17] Suranga Sampath, M. I. G., Khalidass, K. and Jicheng, L. (2020). Transient analysis of an $M/M/1$ queueing system subjected to multiple differentiated vacations, impatient customers and a waiting server with application IEEE 802.16e power saving mechanism, *Indian Journal of Pure and Applied Mathematics*, 51:297-320.

[18] Suranga Sampath, M. I. G. and Jicheng, L. (2018). Impact of customers impatience on an M/M/1 queueing system subject to differentiated vacations with a waiting server, *Quality Technology and Quantative Management*, 17:125-148.

[19] Takagi, H. Queueing Analysis: A Foundation of Performance Analysis, Vacation and Priority Systems, Volume 1, Part 1, North-Holland, New York, USA, 1991.

[20] Tian, N. and Zhang, Z. G. Vacation queueing models: theory and Applications, Springer Science and Business Media, US, 2006.

[21] Tian, R. and Zhang, Z. G. and Su, S. (2022). On Markovian queues with single working vacation and bernoulli interruptions, *Probability in the Engineering and Informational Sciences*, 36:616-643.

[22] Towsley, D. and Tripathi, S.K. (1991) A single server priority queue with server failures and queue flushing, *Operation Research Letters* , 10:353-362.

[23] Upadhyaya, S. (2016). Queueing systems with vacations: An overview, *International Journal of Mathematics in Operational Research*, 9:167-213.

[24] Vijayashree, K. V. and Janani, B. (2017). Transient analysis of an $M/M/1$ queueing system subject to differentiated vacations, *Quality Technology and Quantitavtive Management*, 15:730-748.

ON THE PROPERTIES AND APPLICATIONS OF TOPP-LEONE GOMPERTZ INVERSE RAYLEIGH DISTRIBUTION

Sule Omeiza Bashiru[1] & O.Y. Halid[2]

•

[1]Department of Mathematical Sciences, Prince Abubakar Audu University, Anyigba, Kogi State, Nigeria

[2]Department of Statistics, Ekiti State University, Ado-Ekiti, Nigeria

Email: [1]bash0140@gmail.com, [2]omobolaji.halid@eksu.edu.ng

Abstract

In this study, we introduce a new four-parameter continuous probability distribution known as the Topp-Loene Gompertz Inverse Rayleigh (TLGoIRa) distribution. This novel model extends the Gompertz Inverse Rayleigh distribution. We present various mathematical properties of the distribution, including moments, moment generating functions, quantile functions, survival functions, hazard functions, reversed hazard functions, and odd functions. We also derive the distribution of order statistics, yielding both the maximum and minimum order statistics. This process of parameter estimation using the maximum likelihood estimation method is discussed. Furthermore, we present two real-life applications that illustrate the effectiveness and robustness of the TLGoIRa distribution when compared to several considered lifetime models. Our analysis reveals that the TLGoIRa distribution demonstrates superior robustness in comparison to the competing lifetime models. Additionally, the study highlights the distribution's efficacy in fitting biomedical datasets.

Keywords: Bladder cancer patients, goodness of fit, continuous probability, Gompertz Inverse Rayleigh, adequacy model.

I. Introduction

The Inverse Rayleigh distribution, originally introduced by [1], stands as a continuous probability distribution with significant utility in modeling the time until failure of a system. It emerges as a specialized variant within the broader framework of the inverse Weibull (IW) distribution, a potent tool for modeling lifetime data. The realm of statistical research has witnessed a fervent exploration of the inverse Rayleigh distribution, with scholars such as [2], [3], [4], [5], [6], [7], [8], and [9] delving into its intricacies, unraveling its nuances, and investigating its various generalizations and extensions. This study is inherently driven by a paramount objective: to fortify the existing model by incorporating additional shape parameters into the GoIRa distribution, as initially proposed by [9]. This augmentation is sought to yield a heightened degree of robustness, amplifying its aptness in fitting real-world datasets that abound within the realm of medical science. The proposed enhancement is envisaged to concomitantly bolster the goodness-of-fit of the model to such complex and diverse datasets, rendering it a more potent tool in the hands of researchers, clinicians, and decision-makers tasked with extracting meaningful insights and informed decisions from the intricate fabric of medical data. By infusing the GoIRa distribution with further shape parameters, the study seeks to attain a higher degree of flexibility, enabling the model to more accurately capture the multifaceted variability present in medical datasets. This endeavor is underscored by a firm statistical foundation, harnessing robust quantitative techniques to ensure the viability and efficacy of the proposed enhancement. The profound implications of such an augmented model resonate across medical research, aiding in predictive modeling, risk assessment, and optimal resource

allocation, ultimately advancing the frontiers of knowledge and facilitating improved healthcare outcomes. The GoIRa distribution will be combined with the Toppleone family of distributions, introducing additional skewness to the GoIRa distribution. This augmentation aims to enhance the baseline distribution's capacity to accurately model datasets that demonstrate a significant degree of skewness.

II. Methods

2.1 Topp-Leone Gompertz Inverse Rayleigh (TLGoIRa) distribution

In this section a new continuous probability distribution function (pdf) known as TLGoIRa distribution is derived. Also, some plots of its pdf and hazard rate function (hrf) were plotted in order to assess the shape of the new distribution.

The cdf and pdf of the family of distribution proposed by [10] are given as:

$$F\left(x;\theta,\boldsymbol{\xi}\right)=\left[1-\left[1-G\left(x;\boldsymbol{\xi}\right)\right]^{2}\right]^{\theta} \tag{1}$$

$$f\left(x;\theta,\boldsymbol{\xi}\right)=2\theta g\left(x;\boldsymbol{\xi}\right)\left[1-G\left(x;\boldsymbol{\xi}\right)\right]\left[1-\left[1-G\left(x;\boldsymbol{\xi}\right)\right]^{2}\right]^{\theta-1} \tag{2}$$

where ξ is the vector of parameters of the baseline distribution.

where $G(x;\xi)$ is the cumulative distribution function (cdf) of the baseline distribution with vector of parameter ξ.

for $x\geq 0,\theta,\xi>0$,

where equations (1) and (2) are the cdf and pdf of the family of distributions proposed by [10].

The cdf and pdf of the GoIRa distribution are given a

$$G\left(x;\alpha,\lambda,\beta\right)=1-e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta}\right]} \tag{3}$$

$$g\left(x;\alpha,\lambda,\beta\right)=2\alpha\lambda^{2}x^{-3}e^{-\left(\frac{\lambda}{x}\right)^{2}}\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta-1}e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta}\right]} \tag{4}$$

To obtain the cdf of the new model, equation (3) is inserted into equation (1) as

$$F\left(x;\theta,\alpha,\lambda,\beta\right)=\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta}\right]}\right]^{2}\right]^{\theta} \tag{5}$$

On differentiating equation (5), the pdf of TLGoIRa distribution is obtained which is given as

$$f\left(x;\theta,\alpha,\lambda,\beta\right)=4\alpha\theta\lambda^{2}x^{-3}e^{-\left(\frac{\lambda}{x}\right)^{2}}\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta-1}e^{2\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta}\right]}\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^{2}}\right]^{-\beta}\right]}\right]^{2}\right]^{\theta-1} \tag{6}$$

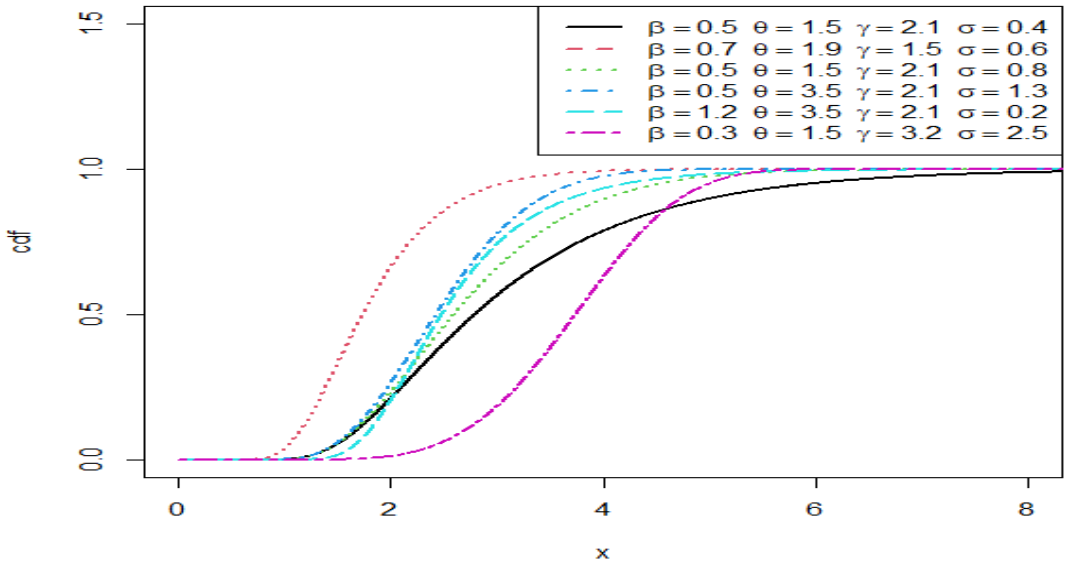


Figure 1: *Plots of cdf of TLGoIRa distribution for different parameter value*

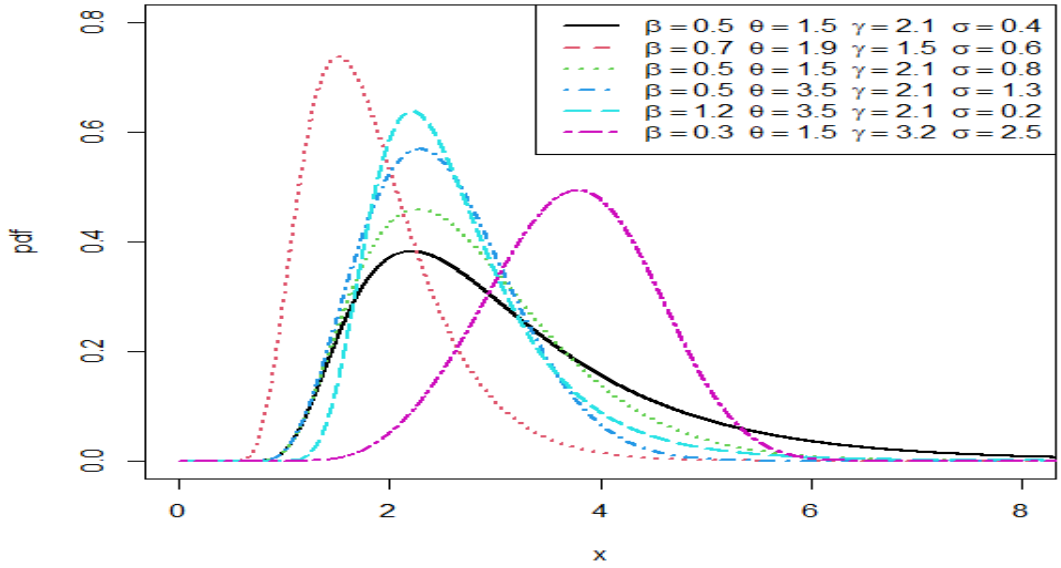


Figure 2: *Plots of pdf of TLGoIRa distribution for different parameter values*

Where $x \geq 0, \lambda > 0$ is the scale parameter and $\alpha, \theta, \beta > 0$ are the shape parameters respectively

2.1.1 Expansion of density

In this section the pdf in equation (6) is expanded using binomial expansion. Expanding the last term in equation (6), we have

$$e^{2\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]} = \sum_{i=0}^{\infty} \frac{\left(\frac{2\alpha}{\beta}\right)^i}{i!}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]^i$$

$$\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta-1}=\sum_{j=0}^{\infty}(-1)^j\binom{\theta-1}{j}\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^{2j}$$

$$\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^{2j}=\sum_{k=0}^{\infty}\frac{\left(\frac{2j\alpha}{\beta}\right)^k}{k!}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]^k$$

$$\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]^{k+1}=\sum_{m=0}^{\infty}(-1)^m\binom{k+1}{m}\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta m}$$

$$\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta(1+m)-1}=\sum_{w=0}^{\infty}(-1)^w\binom{-\beta(1+m)-1}{w}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^w$$

On substituting all the expansions into equation (6), we have

$$f(x)=4\alpha\theta\lambda^2\sum_{i,j,k,m,w=0}^{\infty}\frac{\left(\frac{2\alpha}{\beta}\right)^i}{i!}\frac{\left(\frac{2j\alpha}{\beta}\right)^k}{k!}(-1)^{j+m+w}\binom{\theta-1}{j}\binom{k+1}{m}\binom{-\beta(1+m)-1}{w}x^{-3}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{w+1} \tag{7}$$

Now

$$f(x;\theta,\alpha,\lambda,\beta)=\sum_{i,j,k,m,w=0}^{\infty}\gamma_q x^{-3}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{w+1} \tag{8}$$

where

$$\gamma_q=4\alpha\theta\lambda^2\frac{\left(\frac{2\alpha}{\beta}\right)^i}{i!}\frac{\left(\frac{2j\alpha}{\beta}\right)^k}{k!}(-1)^{j+m+w}\binom{\theta-1}{j}\binom{k+1}{m}\binom{-\beta(1+m)-1}{w}$$

Equation (7) is the expansion of equation (6) which will be used to derive some of the properties of the distribution.

Also, equation (5) is expanded as

$$\left[F(x;\theta,\alpha,\lambda,\beta)\right]^h=\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta h}$$

$$\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta h}=\sum_{z=0}^{h}(-1)^z\binom{\theta h}{z}\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^{2z}$$

$$\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^{2z}=\sum_{t=0}^{\infty}\frac{\left(\frac{2z\alpha}{\beta}\right)^t}{t!}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]^t$$

$$\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]^t=\sum_{d=0}^{\infty}(-1)^d\binom{t}{d}\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta d}$$

$$\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta d}=\sum_{p=0}^{\infty}(-1)^p\binom{-\beta d}{p}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^p$$

Substituting all the expansions into equation (5), we have

$$\left[F\left(x;\theta,\alpha,\lambda,\beta\right)\right]^h=\sum_{z,d,p=0}^{\infty}\sum_{t=0}^{h}\frac{\left(\frac{2z\alpha}{\beta}\right)^t}{t!}(-1)^{z+d+p}\binom{t}{d}\binom{-\beta d}{p}\binom{\theta h}{z}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^p \tag{9}$$

Now

$$\left[F\left(x;\theta,\alpha,\lambda,\beta\right)\right]^h=\sum_{t=0}^{h}\zeta_c\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^p \tag{10}$$

where

$$\zeta_c=\sum_{z,d,p=0}^{\infty}\frac{\left(\frac{2z\alpha}{\beta}\right)^t}{t!}(-1)^{z+d+p}\binom{t}{d}\binom{-\beta d}{p}\binom{\theta h}{z}$$

Equation (9) is the expansion of equation (5) which will be used to derive some of the properties of the distribution.

2.1.2 Properties of the TLGoIRa distribution

In this section, some of the mathematical and statistical properties of TLGoIRa distribution such as the quantile function, moments, moment generating function, reliability measure, odds function, reversed hazard function and order statistics are derived.

2.1.2.1 Moments

$$E\left(X^r\right)=\int_0^{\infty}x^r f\left(x\right)dx \tag{11}$$

$$=\sum_{i,j,k,m,w=0}^{\infty}\gamma_q\int_0^{\infty}x^{r-3}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{w+1}dx$$

$$\int_0^{\infty} x^{r-3}\left[e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{w+1} dx = \frac{\lambda^r \Gamma\left(1-\frac{r}{2}\right)}{\left(w+1\right)^{1-\frac{r}{2}}}$$

$$E\left(X^r\right) = \sum_{i,j,k,m,w=0}^{\infty} \gamma_q \frac{\lambda^r \Gamma\left(1-\frac{r}{2}\right)}{\left(w+1\right)^{1-\frac{r}{2}}} \tag{12}$$

Equation (12) is the moments of TLGoIRa distribution. To obtain mean, we set *r = 1* in equation (12).

2.1.2.2 Moment generating function (mgf)

$$M_{(x)}(t) = \int_0^{\infty} e^{tx} f\left(x\right) dx \tag{13}$$

the series expansion for e^{tx} is given as

$$e^{tx} = \sum_{v=0}^{\infty} \frac{\left(tx\right)^v}{v!} \tag{14}$$

$$M_{(x)}(t) = \sum_{v=0}^{\infty} \frac{t^v}{v!} \sum_{i,j,k,m,w=0}^{\infty} \gamma_q \frac{\lambda^v \Gamma\left(1-\frac{v}{2}\right)}{\left(w+1\right)^{1-\frac{v}{2}}} \tag{15}$$

2.1.2.3 Quantile function

Quantile function has a significant position in probability theory and it is the inverse of the cdf. The quantile function is obtained using

$$Q(u) = F^{-1}(u) \tag{16}$$

Using the inverse of equation (5), we have the quantile function of TLGoIRa distribution given as

$$x = Q\left(u\right) = \frac{\lambda}{\sqrt{-\log\left[1-\left[1-\left[\frac{\beta}{\alpha}\log\left[1-U^{\frac{1}{\theta}}\right]^{\frac{1}{2}}\right]\right]^{\frac{-1}{\beta}}\right]}} \tag{17}$$

The median is obtained by setting *u = 0.5* in equation (17) given as

$$x_{median} = Q\left(0.5\right) = \frac{\lambda}{\sqrt{-\log\left[1-\left[1-\left[\frac{\beta}{\alpha}\log\left[1-0.5^{\frac{1}{\theta}}\right]^{\frac{1}{2}}\right]\right]^{\frac{-1}{\beta}}\right]}} \tag{18}$$

2.1.2.4 Hazard function

Hazard function is given as

$$\tau(x;\alpha,\beta,\lambda,\theta) = \frac{f(x;\alpha,\beta,\lambda,\theta)}{R(x;\alpha,\beta,\lambda,\theta)} \tag{19}$$

The hazard function of the TLGoIRa distribution is given as

$$\tau(x)=\frac{4\alpha\theta\lambda^2x^{-3}e^{-\left(\frac{\lambda}{x}\right)^2}\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta-1}e^{2\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta-1}}{1-\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta}} \tag{20}$$

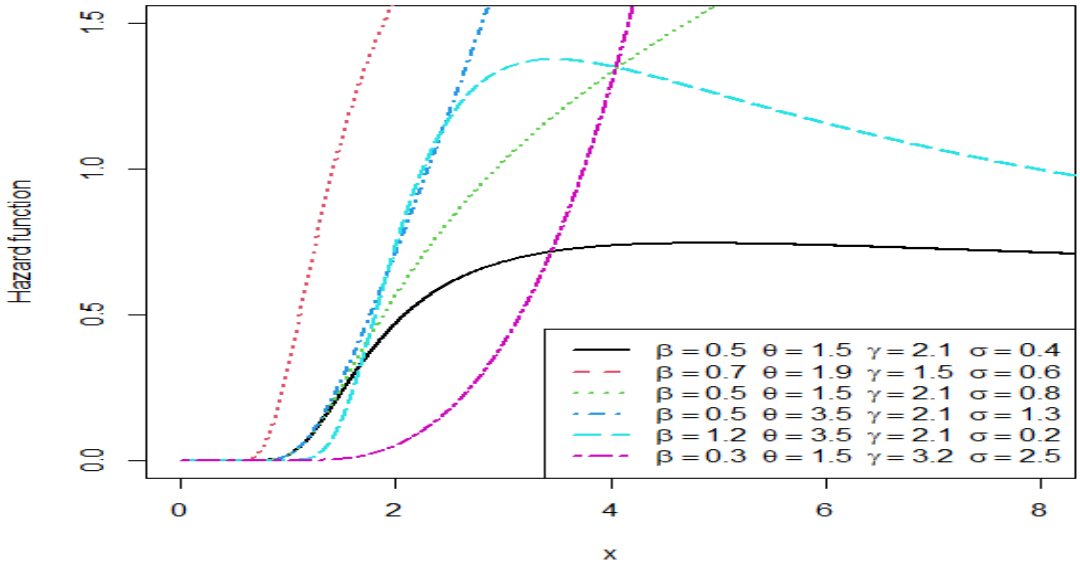


Figure 3: *Plots of hazard function of the TLGoIRa distribution for different parameter values*

2.1.2.5 Survival function

The reliability function is also known as survival function, which is the probability of an item not failing prior to some time. It can be defined as

$$R(x;\alpha,\beta,\lambda,\theta)=1-F(x;\alpha,\beta,\lambda,\theta) \tag{21}$$

The survival function of the TLGoIRa distribution is given as

$$\mathrm{R}(x)=1-\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta} \tag{22}$$

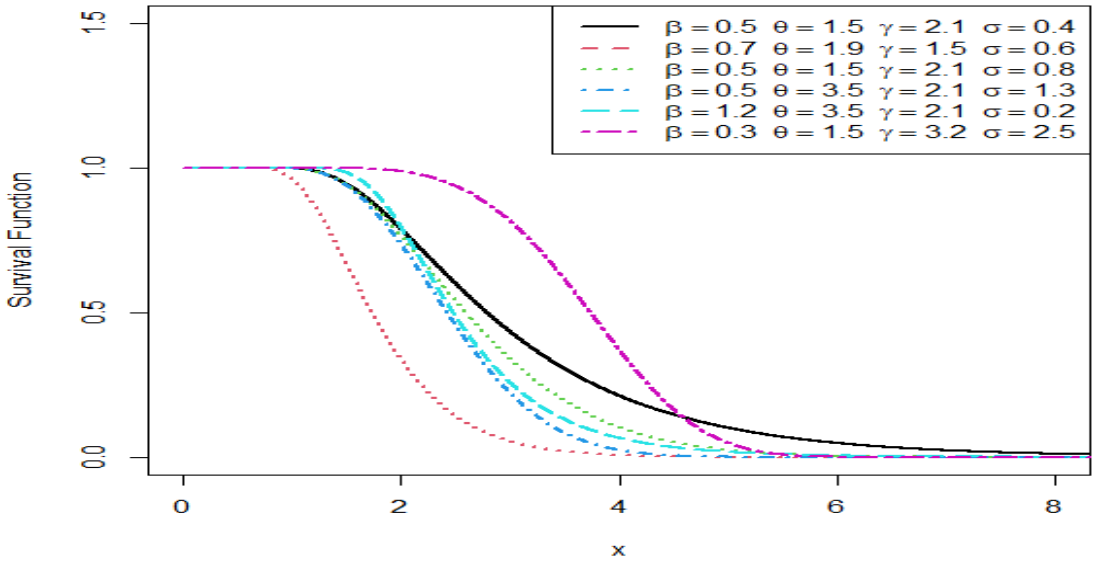


Figure 4: *Plots of survival function of the TLGoIRa distribution for different parameter values*

2.1.2.6 Reversed hazard function

Reversed hazard function of a random variable x is given as

$$\Re(x;\alpha,\beta,\lambda,\theta)=\frac{f(x;\alpha,\beta,\lambda,\theta)}{F(x;\alpha,\beta,\lambda,\theta)} \tag{23}$$

The reverse hazard rate function of the TLGoIRa distribution is given as

$$\Re(x)=\frac{4\alpha\theta\lambda^2x^{-3}e^{-\left(\frac{\lambda}{x}\right)^2}\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta-1}e^{2\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta-1}}{\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta}} \tag{24}$$

2.1.2.7 Odds function

The odds function of the TLGoIRa distribution is given as

$$O(x;\alpha,\beta,\lambda,\theta)=\frac{\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta}}{1-\left[1-\left[e^{\frac{\alpha}{\beta}\left[1-\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right]^{-\beta}\right]}\right]^2\right]^{\theta}} \tag{25}$$

2.2 Order Statistics

Let $X_1, X_2, ..., X_n$ be n independent random variable from the TLGoIRa distributions and let $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$ be their corresponding order statistic. Let $F_{r:n}(x)$ and $f_{r:n}(x)$, $r = 1, 2, 3, ...n$ denote the cdf and pdf of the r^{th} order statistics $X_{r:n}$ respectively. The pdf of the r^{th} order statistics of $X_{r:n}$ is given as

$$f_{r:n}(x) = \frac{f(x)}{B(r, n-r+1)} \sum_{v=0}^{n-r} (-1)^v \binom{n-r}{v} F(x)^{v+r-1} \tag{26}$$

The pdf of r[th] order statistic for distribution is obtained also by replacing h with v+r-1 in cdf expansion. We have

$$f_{r:n}(x) = \gamma_q \zeta_c \frac{1}{B(r, n-r+1)} \sum_{v=0}^{n-r} \sum_{i,j,k,m,w=0}^{\infty} \sum_{t=0}^{v+r-1} x^{-3} \left[e^{-\left(\frac{\lambda}{x}\right)^2} \right]^{w+p+1} \tag{27}$$

where

$$\gamma_q = 4\alpha\lambda^2 \frac{\left(\frac{2\alpha}{\beta}\right)^i}{i!} \frac{\left(\frac{2j\alpha}{\beta}\right)^k}{k!} (-1)^{j+m+w} \binom{\theta-1}{j} \binom{k+1}{m} \binom{-\beta(1+m)-1}{w}$$

and

$$\zeta_c = \sum_{z,d,p=0}^{\infty} \frac{\left(\frac{2z\alpha}{\beta}\right)^t}{t!} (-1)^{z+d+p} \binom{t}{d} \binom{-\beta d}{p} \binom{\theta(v+r-1)}{z}$$

The pdf of minimum order statistic of the distribution is obtained by setting r=1 in equation (27).

$$f_{1:n}(x) = \gamma_q \zeta_c n \sum_{v=0}^{n-1} \sum_{i,j,k,m,w=0}^{\infty} \sum_{t=0}^{v} x^{-3} \left[e^{-\left(\frac{\lambda}{x}\right)^2} \right]^{w+p+1} \tag{28}$$

Now ,

$$\zeta_c = \sum_{z,d,p=0}^{\infty} \frac{\left(\frac{2z\alpha}{\beta}\right)^t}{t!} (-1)^{z+d+p} \binom{t}{d} \binom{-\beta d}{p} \binom{\theta v}{z}$$

Also, the pdf of maximum order statistic of the distribution is obtained by setting r = n in equation (27)

$$f_{n:n}(x) = \gamma_q \zeta_c n \sum_{i,j,k,m,w=0}^{\infty} \sum_{t=0}^{v+n-1} x^{-3} \left[e^{-\left(\frac{\lambda}{x}\right)^2} \right]^{w+p+1} \tag{29}$$

where

$$\zeta_c = \sum_{z,d,p=0}^{\infty} \frac{\left(\frac{2z\alpha}{\beta}\right)^t}{t!} (-1)^{z+d+p} \binom{t}{d} \binom{-\beta d}{p} \binom{\theta(v+n-1)}{z}$$

2.3 Estimation method

The method of maximum likelihood estimation (MLE) is used in this section to estimate the parameters of the TLGoIRa distribution. For a random sample, $X_1, X_2, ..., X_n$ of size n from the TLGoIRa$(\alpha,\beta,\theta,\lambda)$, the log-likelihood function L$(\alpha,\beta,\theta,\lambda)$ of (6) is given as

$$\log L = n\log(4) + n\log(\theta) + n\log(\alpha) + 2n\log(\lambda) - 3\sum_{i=1}^{n}\log(x_i) - \lambda^2\sum_{i=1}^{n}\left(\frac{1}{x_i}\right)^2$$
$$-(\beta+1)\sum_{i=1}^{n}\log\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right] + \frac{\alpha}{\beta}\sum_{i=1}^{n}\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right] + (\theta-1)\sum_{i=1}^{n}\log\left[1-\left[e^{\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]}\right]^2\right] \tag{30}$$

Differentiating the log-likelihood with respect to $\lambda,\alpha,\theta,\beta$ and equating the result to zero, we have

$$\frac{\partial \log l}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n}\log\left[1-\left[e^{-\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]}\right]^2\right] = 0 \tag{31}$$

$$\frac{\partial \log l}{\partial\alpha} = \frac{n}{\alpha} + \frac{2}{\beta}\sum_{i=1}^{n}\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right] + 2(\theta-1)\sum_{i=1}^{n}\frac{e^{\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]}e^{\left(\frac{1}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]}}{\left[1-e^{\left(\frac{\beta}{\sigma}\right)\left[1-\left[1-e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right]}\right]} = 0 \tag{32}$$

$$\frac{\partial \log l}{\partial\lambda} = \frac{2n}{\lambda} - 2\lambda\sum_{i=1}^{n}\left(\frac{1}{x_i}\right)^2 - (\beta+1)\sum_{i=1}^{n}\frac{\lambda e^{-\left(\frac{\lambda}{x}\right)^2}}{x^2\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]} + 4\alpha\lambda\sum_{i=1}^{n}e^{-\left(\frac{\lambda}{x_i}\right)^2}\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta-1}$$
$$+2\beta(\theta-1)\sum_{i=1}^{n}\frac{e^{\left(\frac{\alpha}{\beta}\right)}e^{-\left(\frac{\lambda}{x_i}\right)^2}e^{\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]}\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta-1}}{\left[1-\left[e^{\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]}\right]^2\right]} = 0 \tag{33}$$

$$\frac{\partial \log l}{\partial \beta} = -\sum_{i=1}^{n} \log\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right] - \frac{2\alpha}{\beta^2}\sum_{i=1}^{n}\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta} \log\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right] e^{\left(\frac{\alpha}{\beta}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]} e^{\left(\frac{\alpha}{\beta^2}\right)\left[1-\left[1-e^{-\left(\frac{\lambda}{x_i}\right)^2}\right]^{-\beta}\right]} e^{\left(\frac{\alpha}{\beta}\right)\left[1-e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\beta}} \log\left[1-e^{-\left(\frac{\lambda}{x}\right)^2}\right] + 2(\theta-1)\sum_{i=1}^{n} \frac{}{\left[1-e^{-\left(\frac{\beta}{\sigma}\right)\left[1-\left[1-e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right]}\right]} = 0 \quad (34)$$

Now, equations (31), (32), (33) and (34) do not have a simple analytical form and are therefore not tractable. As a result, we have to resort to non-linear estimation of the parameters using iterative method.

III. Results

3.1 Applications

In this section, we present two applications of the TLGoIRa distribution using different datasets from the biomedical field. These applications are intended to demonstrate the flexibility of the distribution in modeling real-life datasets. The data are fitted to the TLGoIRa distribution, as well as four other comparator distributions: Gompertz Inverse Rayleigh (GoIRa) distribution, Generalized Gompertz (GGo) distribution, Exponentiated Exponential (EtEx) distribution, and Inverse Rayleigh (IRa) distribution. This fitting process is carried out to test the new distribution's flexibility against these comparators. We utilized the Adequacy Model package within the R software to perform the analysis and produce the results. To evaluate the performance of the distribution, we employed the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). These criteria were used to compare the performance of the TLGoIRa distribution with other existing distributions that align with the baseline distribution in terms of providing a good parametric fit to the dataset.

$$AIC = -2ll + 2k \quad (30)$$

$$BIC = -2ll + k\log(n) \quad (31)$$

The model selection is carried out using the AIC and the BIC. Where ll denotes the log-likelihood function evaluated at the maximum likelihood estimates, k is the number of parameters, and n is the sample size from the data. The model with minimum value of AIC and BIC is chosen as the best model to fit the data set.

Data set 1 has been used by [11] and represents the sum of skin folds in 202 athletes collected at the Australian Institute of Sports as.

28.0, 98, 89.0, 68.9, 69.9, 109.0, 52.3, 52.8, 46.7, 82.7, 42.3, 109.1, 96.8, 98.3, 103.6, 110.2, 98.1, 57.0, 43.1, 71.1, 29.7, 96.3, 102.8, 80.3, 122.1, 71.3, 200.8, 80.6, 65.3, 78.0, 65.9, 38.9, 56.5, 104.6, 74.9, 90.4, 54.6, 131.9, 68.3, 52.0, 40.8, 34.3, 44.8, 105.7, 126.4, 83.0, 106.9, 88.2, 33.8, 47.6, 42.7, 41.5, 34.6, 30.9, 100.7, 80.3, 91.0, 156.6, 95.4, 43.5, 61.9, 35.2, 50.9, 31.8, 44.0, 56.8, 75.2, 76.2,101.1, 47.5, 46.2, 38.2, 49.2, 49.6, 34.5, 37.5, 75.9, 87.2, 52.6, 126.4, 55.6, 73.9, 43.5, 61.8, 88.9, 31.0, 37.6,52.8, 97.9, 111.1, 114.0, 62.9, 36.8, 56.8, 46.5, 48.3, 32.6, 31.7, 47.8, 75.1, 110.7, 70.0, 52.5, 67, 41.6, 34.8, 61.8, 31.5, 36.6, 76.0, 65.1, 74.7, 77.0, 62.6, 41.1, 58.9, 60.2, 43.0, 32.6, 48, 61.2, 171.1, 113.5, 148.9, 49.9, 59.4, 44.5, 48.1, 61.1, 31.0, 41.9, 75.6, 76.8, 99.8, 80.1, 57.9, 48.4, 41.8, 44.5, 43.8, 33.7, 30.9, 43.3, 117.8, 80.3, 156.6, 109.6, 50.0, 33.7, 54.0, 54.2, 30.3, 52.8, 49.5, 90.2, 109.5, 115.9, 98.5, 54.6, 50.9, 44.7, 41.8, 38.0, 43.2, 70.0, 97.2, 123.6, 181.7, 136.3, 42.3, 40.5, 64.9, 34.1, 55.7, 113.5, 75.7, 99.9, 91.2, 71.6, 103.6, 46.1, 51.2, 43.8, 30.5, 37.5, 96.9, 57.7, 125.9, 49.0, 143.5, 102.8, 46.3, 54.4, 58.3, 34.0, 112.5, 49.3, 67.2, 56.5, 47.6, 60.4, 34.9.

Data set 2 has been used by [12] and it represents the remission times (in months) of a random sample

of one hundred and twenty-eight (128) bladder cancer patients.
0.08, 2.09, 3.48, 4.87, 6.94 , 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46 , 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

Table 1: *The models' MLEs and performance requirements based on data set 1*

Models	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\beta}$	ll	AIC	BIC
TLGoIR	0.8992	0.2125	121.0713	1.2501	-949.7464	1907.4930	1920.7260
GoIR	0.0031	-	0.0000	0.8601	-987.5204	1981.0410	1990.9660
GGo	-0.0052	15.4031	-	0.0597	-956.0865	1918.1730	1928.9200
EtEx	0.0406	8.5786	-	-	-958.0065	1920.0130	1926.6300
IR	52.6054	-	-	-	-966.4625	1934.9250	1938.2330

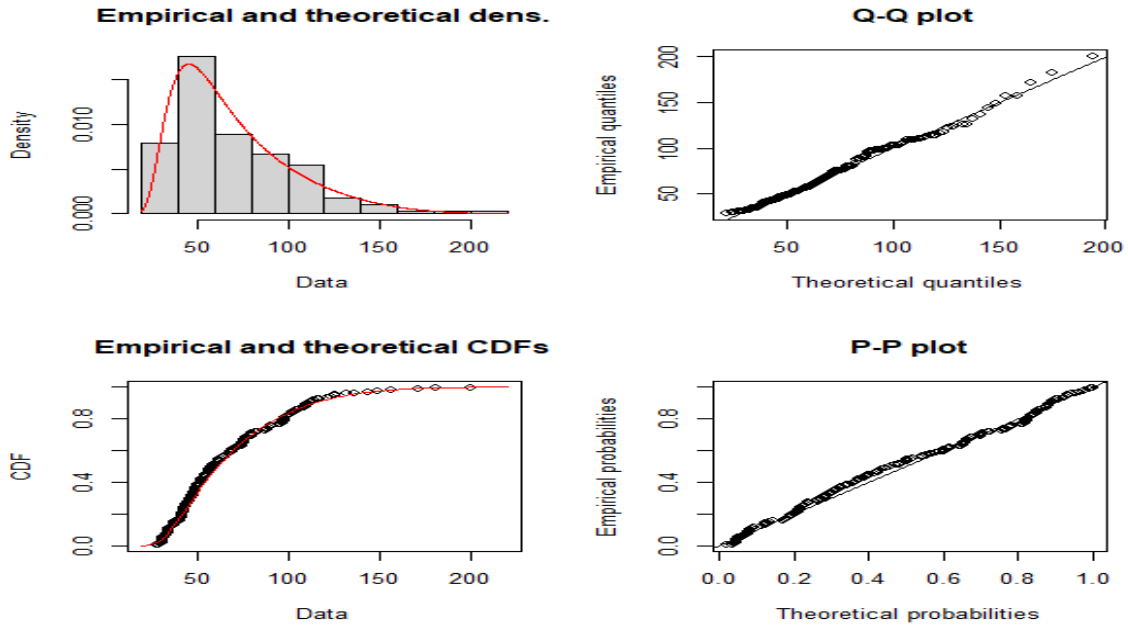


Figure 5: *Density plots for data set 1*

Table 2: *The models' MLEs and performance requirements based on data set 2*

Models	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\beta}$	ll	AIC	BIC
TLGoIR	0.0002	2.6653	0.0001	0.3345	-410.6935	829.3871	834.7952
GoIR	0.0839	-	0.0041	0.5129	-413.5753	833.1505	836.1377
GGo	-0.0224	1.5034	-	0.1678	-413.1834	832.3668	835.3539
EtEx	0.1213	1.2180	-	-	-413.0776	830.1552	834.8592
IR	2.2612	-	-	-	-774.3416	1550.6830	1553.5350

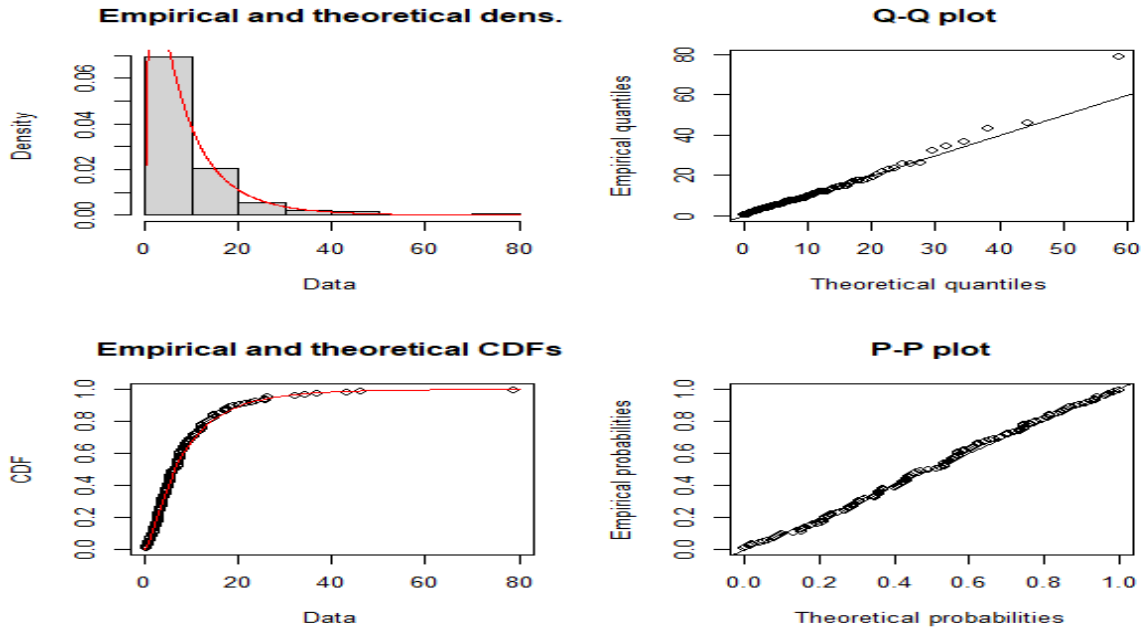


Figure 6: *Density plots for data set 2*

IV. Discussion

In this study, we have introduced and thoroughly examined a novel probability distribution, coined the Topp Leone Gompertz Inverse Rayleigh distribution. Drawing from the framework of distributions proposed by [10], we have both theoretically established its foundations and conducted practical investigations. A comprehensive exploration of the distribution's mathematical properties has been undertaken, encompassing crucial characteristics such as moments, moment generating function, quantile function, odd function, reverse hazard function, order statistics, and reliability analysis. These analyses have contributed to a robust comprehension of the distribution's behaviors and traits. The method of maximum likelihood estimation has been employed to effectively determine the distribution's parameters, lending statistical rigor to our subsequent analyses. In a bid to assess the distribution's practical utility, it has been subjected to rigorous testing using two distinct datasets from the realm of medical sciences. This assessment has entailed a comparative study against several competing distributions, including the GoIRa, Generalized Gompertz, Exponentiated Exponential, and Inverse Rayleigh distributions, all sharing common baseline distributions. The outcomes of our meticulous analyses indicate that the Topp-Leone Gompertz Inverse Rayleigh distribution stands out as the most adept candidate for fitting both datasets. The distribution's apparent superior performance among the alternatives highlights its potential to adeptly capture the underlying data-generation mechanisms. This research introduces not only a novel distribution but also offers practical insights into its capacity to effectively represent intricate real-world phenomena.

References

[1] Trayer, V. N. Doklady Acad, Nauk, Belorus, U.S.S.R, 1964.

[2] Gharraph, M.K. (1993). Comparison of Estimators of Location Measures of an Inverse Rayleigh Distribution. *The Egyptian Statistical Journal*, 37:295–309.

[3] Mukarjee, S.P. and Maitim, S.S. (1996) A Percentile Estimator of the Inverse Rayleigh Parameter, *IAPQR Transactions*, 21, 63–65.

[4] Voda, V. G. (1972). On the Inverse Rayleigh Random Variable, *Pep. Statist. App. Res.*, JUSE, 19, 13-21.

[5] Mohsin, M. and Shahbaz, M.Q (2005). Comparison of Negative Moment Estimator with Maximum Likelihood Estimator of Inverse Rayleigh Distribution, *PJSOR*, 1(1): 45–48.

[6] Khan, M.S. (2014). Modified Inverse Rayleigh Distribution. *International Journal of Computer Applications*, 87(13) : 28–33.

[7] Khan, M. S. and King, R. (2015). Transmuted Modified Inverse Rayleigh Distribution, *Austrian Journal of Statistics*, 44(3): 17–29

[8] Malik, A. S. and Ahmad, S. P. (2018). A New Inverse Rayleigh Distribution: Properties and Application, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 5(5): 92-96.

[9] Halid, O.Y. and Sule, O.B. (2022). A classical and Bayesian estimation techniques for Gompertz inverse Rayleigh distribution: properties and application, *Pakistan Journal of Statistics*, 38(1): 49-76

[10] Al-Shomrani, A., Arif, O., Shawky, A., Hanif, S. and Shahbaz, M. Q., (2016). Topp-Leone family of distributions: Some properties and application, *Pakistan Journal of Statistics and Operation Research*. XII, 3, 443-451.

[11] Hosseini, B., Afshari, M. and Alizadeh M. (2018). The generalized odd gamma-G family of distributions: properties and applications. *Austrian Journal of Statistics*, 47(2):69-89

[12] Shabeer, A.M., Krishnana, B. and Jayakumar, K. (2023). A Generalization of Lindley Distribution: Characterizations, Methods of Estimation and Applications. *Reliability: Theory & Applications*, 1(72): 575-588.

SOME APPLICATIONS OF TRANSMUTED LOG-UNIFORM DISTRIBUTION

Ashin K Shaji, Rani Sebastian

•

Department of Statistics
St.Thomas' college (Autonomous)
Thrissur, Kerala, 680001, India
ashinkidangan@gmail.com
ranikanjiram@gmail.com

Abstract

As a generalization of the Log Uniform distribution, Transmuted Log - Uniform distribution is introduced and its properties are studied.We obtained graphical representations of its pdf, cdf, hazard rate function and survival function. We have derived statistical properties such as moments, mean deviations, and the quantile function for the Transmuted Log-Uniform distribution. We also obtained the order statistics of the new distribution. Method of maximum likelihood is used for estimating the parameters. Estimation of stress strength parameters is also done. We applied the Transmuted Log-Uniform distribution to a real data set and compared it with Transmuted Weibull distribution and Transmuted Quasi-Akash distribution. It was found that the Transmuted Log-Uniform distribution was a better fit than the Transmuted Weibull distribution and Transmuted Quasi-Akash distribution distributions based on the values of the AIC, CAIC, BIC, HQIC, the Kolmogorov-Smirnov (K-S) goodness-of-fit statistic and the p-values.

Keywords: Transmuted distribution, Transmuted Log- Uniform distribution,Stress- strength parameters

1. Introduction

The Transmuted family of distributions was first introduced by Shaw and Buckley (2007) and has since been widely used in various fields including finance, engineering, and environmental sciences. According to quadratic rank transmutation map (QRTM) technique approach, a random variable X is said to have a Transmuted distribution, if its cdf is given by,

$$D(x) = (1+\lambda)G(x) - \lambda(G(x))^2; \quad -1 \leq \lambda \leq 1 \tag{1}$$

where G(x) is the c.d.f of the base distribution.

The corresponding probability density function (p.d.f) with parameter λ is given by:

$$d(x) = g(x)(1+\lambda-2\lambda G(x)); \quad -1 \leq \lambda \leq 1 \tag{2}$$

where λ is a scale parameter.

There are different families of distributions which are useful for developing flexible compound distributions for solving real life problems. Transmuted distributions have emerged as the superior option, surpassing their standard counterparts in terms of flexibility and performance. Some of the models studied were Transmuted Exponential Lomax distribution by Abdullahi and Ieren [1], Transmuted complementary Weibull Geometric distribution by Afify [2], the Transmuted Weibull Lomax distribution by Afify [3], the Transmuted Weibull distribution by Aryaland Tsokos

[5], Transmuted Additive Weibull distribution by Elbatal and Aryal [6], Transmuted Quasi Akash distribution by Hassan [8], Transmuted Exponentiated Gamma distribution by Hussian [9], Transmuted modified Weibull distribution by Khanand King [11], Transmuted Inverse Weibull distribution by Khan [10], Transmuted Gompertz distribution by Khan [12], Transmuted Lindley distribution by Mansour [13], Transmuted Rayleigh distribution by Merovci [15], Transmuted Pareto distribution by Merovci and Puka [14], Transmuted Inverse Exponential distribution by Oguntunde and Adejumo [16] and Transmuted generalized Uniform distribution by subramanian [18] etc.

In this article we present a new generalization of Log-Uniform distribution called the Transmuted Log-Uniform distribution.

2. Transmuted Log-Uniform Distribution

A Log-Uniform distribution is a probability distribution where the logarithm of the random variable is uniformly distributed.

A random variable X is said to have the Log-Uniform distribution with parameter λ if its probability density function is defined as,

$$g(x) = \begin{cases} \frac{1}{x\left[ln(b)-ln(a)\right]}; & \text{if }, a \leq x \leq b, 0 < a < b, a, b \in R \\ 0; & \text{otherwise} \end{cases} \tag{3}$$

where a and b are the parameters of the distribution and they are location parameters that define the minimum and maximum values of the distribution on the original scale and ln is the natural Log function (the logarithm to base e).

The corresponding Cumulative distribution function (c.d.f.) is,

$$G(x) = \begin{cases} \frac{ln(x)-ln(a)}{ln(b)-ln(a)}; & \text{if }, a \leq x \leq b, 0 < a < b, a, b \in R \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

The cdf of a Transmuted Log-Uniform distribution,

$$F(x) = \begin{cases} (1+\lambda)\frac{ln(\frac{x}{a})}{ln(\frac{b}{a})} - (\lambda)\left[\frac{ln(\frac{x}{a})}{ln(\frac{b}{a})}\right]^2; & \text{if } |\lambda| \leq 1, a \leq x \leq b, \\ & 0 < a < b, \quad a, b \in R \\ 0; & \text{otherwise.} \end{cases} \tag{5}$$

The pdf of Transmuted Log-Uniform distribution is

$$f(x) = \begin{cases} \frac{(1+\lambda)}{(x)ln(\frac{b}{a})} - \frac{(2\lambda)ln(\frac{x}{a})}{(x)(ln(\frac{b}{a}))^2}; & \text{if } |\lambda| \leq 1, a \leq x \leq b, \\ & 0 < a < b, \quad a, b \in R \\ 0; & \text{otherwise} \end{cases} \tag{6}$$

The survival function of Transmuted Log-Uniform distribution is given by:

$$S(x) = \frac{(ln(\frac{b}{a}))^2 - (1+\lambda)ln(\frac{x}{a})ln(\frac{b}{a}) - (\lambda)(ln(\frac{x}{a}))^2}{(ln(\frac{b}{a}))^2} \tag{7}$$

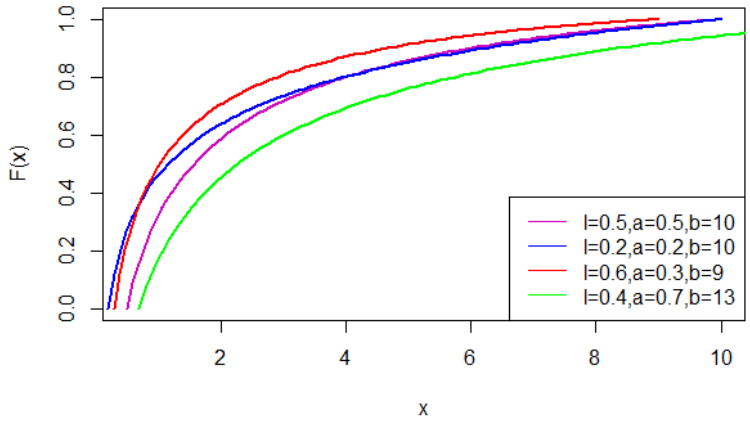


Figure 1: *Plot of cdf of the Transmuted Log-Uniform distribution*

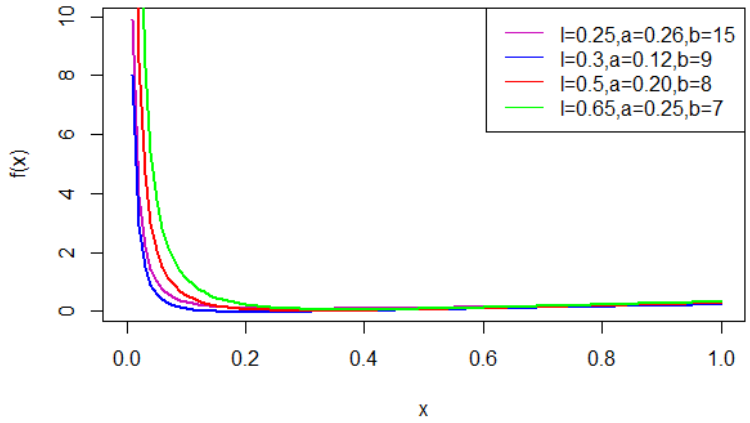


Figure 2: *Plot of pdf of the Transmuted Log-Uniform distribution*

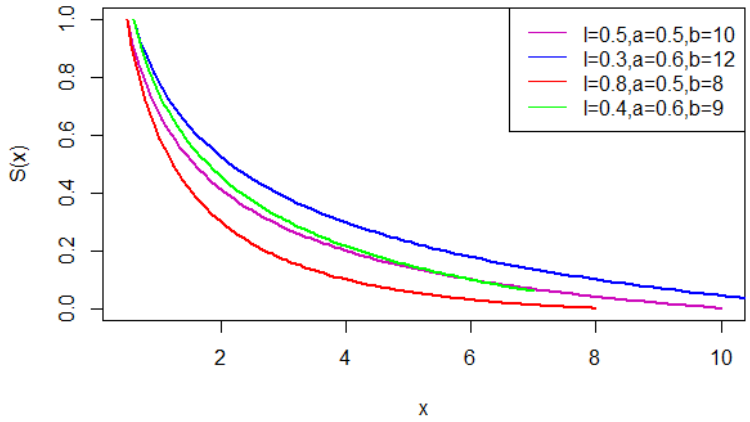


Figure 3: *Plot of survival function of the Transmuted Log-Uniform distribution*

The failure rate function or hazard function of Transmuted Log-Uniform distribution is:

$$h(x) = \frac{(1+\lambda)ln(\frac{b}{a}) - (2\lambda)ln(\frac{x}{a})}{x[(ln(\frac{b}{a}))^2 - (1+\lambda)ln(\frac{x}{a})ln(\frac{b}{a}) - (\lambda)(ln(\frac{x}{a}))^2]} \tag{8}$$

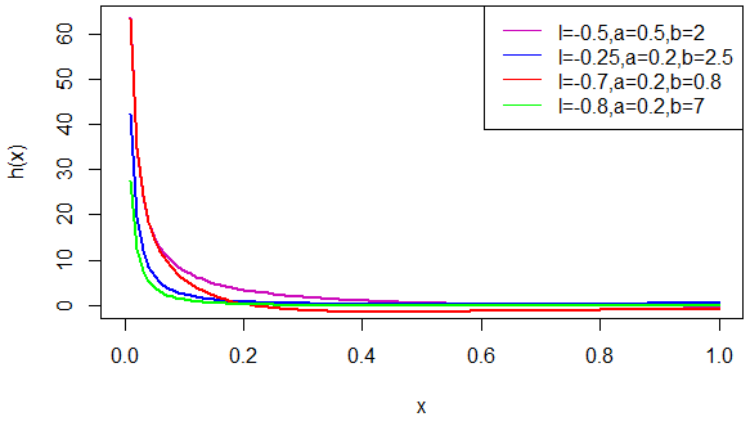


Figure 4: *Plot of hazard rate function of the Transmuted Log-Uniform distribution*

- Special Case:
  If we put$\lambda = 0$, then Transmuted Log-Uniform distribution reduces to Log-Uniform distribution.

3. Statistical properties

3.1. Moments

Let X is a random variable following Transmuted Log-Uniform distribution with parameters a,b,λ and then the r^{th} moment for a given probability distribution is given by:

$$\mu_r^{'} = \int_a^b x^r \left[\frac{(1+\lambda)}{x\ln\frac{b}{a}} - \frac{2\lambda\ln\frac{x}{a}}{x(\ln\frac{b}{a})^2}\right]dx$$

$$E(X^r) = \mu_r^{'} = \frac{(1+\lambda)(b^r - a^r)}{r\ln(\frac{b}{a})} - \frac{2\lambda a^r}{(\ln(\frac{b}{a}))^2}\left[(\frac{b}{a})\ln(\frac{b}{a}) - (\frac{b}{ar}) + (\frac{1}{r^2})\right] \tag{9}$$

Mean of the Transmuted Log-Uniform distribution is obtained as:
$\mu_1^{'} = \frac{(1+\lambda)(b-a)}{ln(\frac{b}{a})} - \frac{2\lambda a}{(ln(\frac{b}{a}))^2}\left[(\frac{b}{a})ln(\frac{b}{a}) - (\frac{b}{a}) + 1\right]$.

3.2. Quantile function

The Quantile function of Transmuted Log-Uniform distribution is obtained by inverting distribution function.

$$p = (1+\lambda)\frac{ln(\frac{x}{a})}{ln(\frac{b}{a})} - \lambda\left[\frac{ln(\frac{x}{a})}{ln(\frac{b}{a})}\right]^2$$

$$x = a \exp\left(\frac{\left[(1+\lambda) + \sqrt{\lambda^2 + 2\lambda + 1 - 4\lambda p}\right] \ln \frac{b}{a}}{2\lambda} \right) \tag{10}$$

The median, 2^{nd} quartile is obtained by putting $p = \frac{1}{2}$ in (10)

$$x = a \exp\left(\frac{\left[(1+\lambda) + \sqrt{\lambda^2 + 1}\right] \ln \frac{b}{a}}{2\lambda} \right) \tag{11}$$

3.3. Mean deviation

Let X follows Transmuted Log-Uniform distribution with mean μ and median M.

- Mean Deviation from the Mean is given by:

$$\delta_1(x) = \int_a^b |x - \mu| f(x) dx \ = 2\mu(F(\mu) - 1) + 2T(\mu) \tag{12}$$

where μ is the mean of the distribution and

$$T(\mu) = \int_\mu^b x f(x) dx$$

$$T(\mu) = \int_\mu^b x \left[\frac{(1+\lambda)}{(x) ln(\frac{b}{a})} - \frac{(2\lambda) ln(\frac{x}{a})}{(x)(ln(\frac{b}{a}))^2} \right] dx$$

$$T(\mu) = \frac{(1+\lambda)(b-\mu)}{ln(\frac{b}{a})} - \frac{(2\lambda)}{(ln(\frac{b}{a}))^2} \left[b\left(\ln \frac{b}{a} - 1 \right) - \mu\left(\ln \frac{\mu}{a} - 1 \right) \right] \tag{13}$$

- Similarly, the Mean Deviation about Median is:

$$\delta_2(x) = \int_a^b |x - M| f(x) dx \ = 2T(M) - \mu \tag{14}$$

where M is the median of the distribution and μ is the mean of the distribution and

$$T(M) = \int_M^b x f(x) dx$$

$$T(M) = \frac{(1+\lambda)(b-M)}{ln(\frac{b}{a})} - \frac{(2\lambda)}{(ln(\frac{b}{a}))^2} \left[b\left(\ln \frac{b}{a} - 1 \right) - M\left(\ln \frac{M}{a} - 1 \right) \right] \tag{15}$$

The mean deviations about mean is obtained by substituting the mean, cdf and $T(\mu)$ in (12). The mean deviations about median is obtained by substituting the mean, cdf and $T(M)$ in (14).

3.4. Order Statistics

Let $X_{(1)}$,$X_{(2)}$,$X_{(3)}$,...,$X_{(n)}$ denote the order statistics of a random sample X_1,X_2,X_3,...,X_n drawn from the continuous distribution with pdf $f_X(x)$ and cdf $F_X(x)$, then the pdf of r^{th} order statistics $X_{(r)}$ is given by:

$$f_{X_{(r)}}(x) = \frac{n!}{(r-1)!(n-r)!} f(x) [F(x)]^{(r-1)} [1 - F(x)]^{(n-r)} \tag{16}$$

Using the equations (5) and (6) the probability density function of r^{th} order statistics $X_{(r)}$ of Transmuted Log-Uniform distribution is given by:

$$f_{x_{(r)}}(x,a,b,\lambda) = \frac{n!}{(r-1)!(n-r)!}\left[\frac{(1+\lambda)}{x\ln(\frac{b}{a})} - \frac{(2\lambda)\ln(\frac{x}{a})}{x(\ln(\frac{b}{a}))^2}\right]\left[(1+\lambda)\frac{\ln(\frac{x}{a})}{\ln(\frac{b}{a})} - \lambda\left[\frac{\ln(\frac{x}{a})}{\ln(\frac{b}{a})}\right]^2\right]^{(r-1)} \left[1-((1+\lambda)\frac{\ln(\frac{x}{a})}{\ln(\frac{b}{a})} - (\lambda)\left[\frac{\ln(\frac{x}{a})}{\ln(\frac{b}{a})}\right]^2)\right]^{(n-r)} \tag{17}$$

4. Parameter estimation

In this section, we discuss the method of maximum likelihood(ML) for the estimation of the unknown parameters a, b, λ of Transmuted Log-Uniform distribution. Let X_1,X_2,X_3,...,X_n be the random sample of size n drawn from Transmuted Log-Uniform distribution, then the likelihood function is given by:

$$L(x_i;a,b,\lambda) = \prod_{i=1}^{n}\frac{1}{x_i}\prod_{i=1}^{n}\left[\frac{1+\lambda}{ln(\frac{b}{a})} - \frac{2\lambda ln(\frac{x_i}{a})}{(ln(\frac{b}{a}))^2}\right] \tag{18}$$

The log-likelihood function is given by:

$$lnL(x_i;a,b,\lambda) = ln\left[\prod_{i=1}^{n}\frac{1}{x_i}\right] + ln\prod_{i=1}^{n}\left[\frac{1+\lambda}{ln(\frac{b}{a})} - \frac{2\lambda ln(\frac{x_i}{a})}{(ln(\frac{b}{a}))^2}\right] \tag{19}$$

Therefore, the maximum likelihood estimator of a, b and λ which maximize equation (19), must satisfy the following normal equations given by

$$\frac{\partial lnL}{\partial a} = \sum_{i=1}^{n}\frac{\ln(\frac{b}{a})(1+\lambda) - 4\lambda\ln(\frac{x_i}{a}) + 2\lambda\ln(\frac{b}{a})}{[(1+\lambda)\ln(\frac{b}{a}) - 2\lambda\ln(\frac{x_i}{a})]\ln(\frac{b}{a})} = 0 \tag{20}$$

$$\frac{\partial lnL}{\partial b} = \sum_{i=1}^{n}\frac{4\lambda\ln(\frac{x_i}{a}) - \ln(\frac{b}{a})(1+\lambda)}{[(1+\lambda)\ln(\frac{b}{a}) - 2\lambda\ln(\frac{x_i}{a})]b\ln(\frac{b}{a})} = 0 \tag{21}$$

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{n}\frac{\ln(\frac{b}{a}) - 2\ln(\frac{x_i}{a})}{(1+\lambda)\ln(\frac{b}{a}) - 2\lambda(\ln\frac{x_i}{a})} = 0 \tag{22}$$

Solving this system of equations, in a, b, β gives the MLEs of a, b, λ as $\hat{a}$,$\hat{b}$,$\hat{\lambda}$.

5. Estimation of Stress-Strength parameter

In this section, the procedure of estimating reliability of $R = P(X_2 < X_1)$ model is considered. The expression $R = P(X_2 < X_1)$ measures the reliability of a component in terms of probability and the random variables X_1 representing the stress experienced by the component does not exceed X_2 which represents the strength of the component. If stress exceeds strength, the component would fail and vice-versa.

In order to estimate the stress-strength parameter, considering two random variables X and Y with Transmuted Log-Uniform (λ_1, a, b) and Transmuted Log-Uniform (λ_2, a, b) distributions respectively. We assume that X and Y are independent random variables and the stress-strength parameter is obtained in the form:

$$R = P(Y < X) = \int_{X<Y} f(x,y)dxdy = \int_0^\infty f(x;\lambda_1,a,b)F(x;\lambda_2,a,b)dx$$

where $f(x,y)$ is the joint probability density function of random variables X and Y, having Transmuted Log-Uniform distribution so that:

$$R = \int_a^b \left[(1+\lambda_1)\left(\frac{\ln(x/a)}{\ln(b/a)}\right) - \lambda_1\left(\frac{\ln(x/a)}{\ln(b/a)}\right)^2\right]\left[\frac{(1+\lambda_2)}{x\ln(b/a)} - \frac{2\lambda_2\ln(x/a)}{x(\ln(b/a))^2}\right]dx$$

On simplification we get:

$$R = \frac{(\lambda_1 - \lambda_2 + 3)}{6} \tag{23}$$

To compute the maximum likelihood estimate of R, we need to compute the maximum likelihood estimate of λ_1 and λ_2.Suppose $X_1, X_2, ..., X_n$ is random sample of size n from the Transmuted Log-Uniform (λ_1, a, b) and $Y_1, Y_2, ..., Y_m$ is an independent random sample of size m from Transmuted Log-Uniform (λ_2, a, b). Then the likelihood function of the combined random sample can be obtained as follows:

$$L = \prod_{i=1}^{n} f(x_i;\lambda_1,a,b)\prod_{i=1}^{m} f(y_i;\lambda_2,a,b)$$

$$L = \prod_{i=1}^{n}\frac{1}{x_i}\prod_{i=1}^{n}\left[\frac{1+\lambda_1}{ln(\frac{b}{a})} - \frac{2\lambda_1 ln(\frac{x_i}{a})}{(ln(\frac{b}{a}))^2}\right]\prod_{i=1}^{m}\frac{1}{x_i}\prod_{i=1}^{n}\left[\frac{1+\lambda_2}{ln(\frac{b}{a})} - \frac{2\lambda_2 ln(\frac{x_i}{a})}{(ln(\frac{b}{a}))^2}\right] \tag{24}$$

The log-likelihood function is given by:

$$lnL = ln\left[\prod_{i=1}^{n}\frac{1}{x_i}\right] + ln\prod_{i=1}^{n}\left[\frac{1+\lambda_1}{ln(\frac{b}{a})} - \frac{2\lambda_1 ln(\frac{x_i}{a})}{(ln(\frac{b}{a}))^2}\right] + ln\left[\prod_{i=1}^{m}\frac{1}{x_i}\right] + ln\prod_{i=1}^{m}\left[\frac{1+\lambda_2}{ln(\frac{b}{a})} - \frac{2\lambda_2 ln(\frac{x_i}{a})}{(ln(\frac{b}{a}))^2}\right] \tag{25}$$

The maximum likelihood estimate (MLE) of λ_1 and λ_2 can be obtained by solving the following equations:

$$\frac{\partial lnL}{\partial\lambda_1} = \sum_{i=1}^{n}\frac{\ln(\frac{b}{a}) - 2\ln(\frac{x_i}{a})}{(1+\lambda_1)\ln(\frac{b}{a}) - 2\lambda_1(\ln\frac{x_i}{a})} = 0 \tag{26}$$

$$\frac{\partial lnL}{\partial\lambda_2} = \sum_{i=1}^{m}\frac{\ln(\frac{b}{a}) - 2\ln(\frac{x_i}{a})}{(1+\lambda_2)\ln(\frac{b}{a}) - 2\lambda_2(\ln\frac{x_i}{a})} = 0 \tag{27}$$

From the equations (26) and (27), we can obtain the ML estimates $\hat{\lambda_1}$ and $\hat{\lambda_1}$. The corresponding ML estimate of R is computed from (23) by replacing λ_1 and λ_2 by their ML estimates

$$\hat{R} = \frac{(\hat{\lambda_1} - \hat{\lambda_2} + 3)}{6} \tag{28}$$

This can be used in estimation of stress-strength for the given data.

6. Simulation study and Data analysis

6.1. Simulation study

Simulation studies are an important tool for statistical research. These help researchers assess the performance of a model, understand the different properties of statistical methods and compare them. Here we take distinct combinations of parameters a, b, λ with sample size as bias and the mean square error(MSE) of the parameter estimates.

The simulation is done by using different true parameter values. The chosen true parameter values are as follows:

- $a = 7.5$,$b = 16$,$\lambda = -0.25$
  As the n increases, MSE decreases for the selected parameter values given in table 1. Moreover, the bias is close to zero as the sample size increases. Thus, as the sample size increases the estimates tend to be closer to the true parameter values.

Table 1: *Simulation study at a = 7.5, b = 16,$\lambda = -0.25$*

n	Parameter	Estimate	Bias	MSE
	a	8.5628	1.0628	1.1296
250	b	14.8466	1.1533	1.3302
	λ	0.3504	0.6004	0.3604
	a	8.5470	1.0470	1.0962
350	b	15.8661	0.9338	0.1790
	λ	0.1411	0.3912	0.1530
	a	8.3061	0.8061	0.6499
500	b	16.2075	0.2075	0.0430
	λ	0.0465	0.2965	0.0879
	a	7.4086	0.0913	0.0083
600	b	16.2010	0.2010	0.0404
	λ	-0.2993	0.0493	0.0024
	a	7.4955	0.0044	0.0000193
750	b	16.0647	0.0647	0.004182
	λ	-0.2533	0.00331	0.0000109

6.2. Data analysis

The data set given in Table 2 represents the relief times (in minutes) of twenty patients receiving an analgesic Gross and Clark (1975). We fit the Transmuted Log-Uniform distribution to a real life data set and compare the results with the Transmuted Quasi Akash distribution and Transmuted Weibull distribution.

Table 2: *Relief times of 20 patients receiving an analgesic*

0				
1.1	1.4	1.3	1.7	1.9
1.8	1.6	2.2	1.7	2.7
4.1	1.8	1.5	1.2	1.4
3.0	1.7	2.3	1.6	2.0

Table 3: *AIC, CAIC, BIC,and HQIC statistics of the fitted model in data set*

Distribution	AIC	CAIC	BIC	HQIC
Transmuted Log-Uniform Distribution	6.0016	8.1258	5.9790	8.1834
Transmuted Quasi Akash Distribution	49.79	51.78	50.18	50.50
Transmuted Weibull Distribution	63.3218	65.446	63.299	65.503

From the table 3, it has been observed that the Transmuted Log-Uniform Distribution possesses the lesser AIC, CAIC, BIC,and HQIC values as compared to Transmuted Quasi Akash distribution and Transmuted Weibull distribution. To check the model goodness of fit we had considered the Kolmogorov-Smirnov (K-S) test (goodness-of-fit) statistics for the relief times of patients receiving an anelgesic data.
To determine the Goodness of fit of the models, the magnitude of K-S Statistic is obtained. Since

the p-value of fitted model is highest than the other distributions we have considered.Therefore the results indicate, that the proposed model performed better than other models.

7. Conclusion

In this study, we introduced a new distribution called the Transmuted Log-Uniform distribution. The distribution was generated using the Transmuted technique and the Log-Uniform distribution as the base distribution with two parameters. We obtained graphical representations of its pdf, cdf, hazard rate function and survival function. We have derived statistical properties such as moments, mean deviations, and the quantile function for the Transmuted Log-Uniform distribution. We also obtained the order statistics of the new distribution.

We used the maximum likelihood method to estimate the parameters of the distribution and the stress strength parameters. We performed a simulation study to validate the estimates of the model parameters, and it was observed that the distribution showed the least bias, with the values of mean square error decreasing as the sample size increased. Finally, we applied the Transmuted Log-Uniform distribution to a real data set and compared it with Transmuted Weibull distribution and Transmuted Quasi-Akash distribution. It was found that the new distribution was a better fit than these distributions based on the values obtained for the AIC, CAIC, BIC, HQIC, the Kolmogorov-Smirnov (K-S) goodness-of-fit statistic, and the p-values obtained for the models.

References

[1] Abdullahi, Umar Kabir and Ieren, Terna Godfrey (2018). On the inferences and applications of transmuted exponential Lomax distribution. *International Journal of Advanced Statistics and Probability*, 6(1):30–36.

[2] Afify, Ahmed Z and Nofal, Zohdy M and Butt, Nadeem Shafique (2014). Transmuted complementary Weibull geometric distribution. *Pakistan Journal of Statistics and Operation Research*, 435–454.

[3] Afify, Ahmed Z and Nofal, Zohdy M and Yousof, Haitham M and El Gebaly, Yehia M and Butt, Nadeem Shafique (2015). The transmuted Weibull Lomax distribution: properties and application. *Pakistan Journal of Statistics and Operation Research*, 135–152.

[4] Al-Kadim, Kareema Abed and Mahdi, Ashraf Alawi (2018). Exponentiated transmuted exponential distribution. *Journal of University of Babylon for Pure and Applied Sciences*, 26(2):78–90.

[5] Aryal, Gokarna R and Tsokos, Chris P (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution. *European Journal of pure and applied mathematics*, 4(2):89–102.

[6] Elbatal, Ibrahim and Aryal, Gokarna (2013). On the transmuted additiveweibull distribution. *Austrian Journal of statistics*, 42(2):117–132.

[7] Hamming, Richard Wesley (1970). On the distribution of numbers. *The bell system technical journal*, 49(8):1609–1625.

[8] Hassan, Anwar and Wani, Sameer Ahmad and Shafi, Sumeera and Dar, Showkat Ahmad (2022). Transmuted Quasi Akash Distribution Applicable to Survival Times Data. *Thailand Statistician*, 20(2):338–356.

[9] Hussian, Mohamed A (2014). Transmuted exponentiated gamma distribution: A generalization of the exponentiated gamma probability distribution. *Applied Mathematical Sciences*, 8(27):1297–1310.

[10] Khan, Muhammad Shuaib and King, Robert and Hudson, Irene (2013). Characterizations of the transmuted inverse Weibull distribution. *Anziam Journal*, 55:C197–C217.

[11] Khan, Muhammad Shuaib and King, Robert (2013). Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. *European Journal of pure and applied mathematics*, 6(1):66–88.

[12] Khan, Muhammad Shuaib and Robert, K and Irene, LH (2016). Transmuted Gompertz distribution: Properties and estimation. *Pak. J. Statist*, 32(3):161–182.
[13] Mansour, Mahmoud M and Mohamed, Salah M (2015). A new generalized of transmuted Lindley distribution. *Applied Mathematical Sciences*, 9(55):2729–2748.
[14] Merovci, Faton and Puka, Lukan (2014). Transmuted pareto distribution. *ProbStat Forum*, 1:1–11.
[15] Merovci, Faton (2013). Transmuted rayleigh distribution. *Austrian Journal of statistics*, 42(1):21–31.
[16] Oguntunde, Pelumi E and Adejumo, Adebowale Olusola and Owoloko, Enahoro A (2017). On the flexibility of the transmuted inverse exponential distribution. *World Congress on Engineering*.
[17] Shaw, William T and Buckley, Ian RC (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*.
[18] Subramanian, C and Rather, AA (2018). Transmuted generalized uniform distribution. *Int. J. Sci. Res. in Mathematical and Statistical Sciences Vol*, 5:5.
[19] Varghese, Akhila K and Chacko, VM (2022). Estimation of stress-strength reliability for Akash distribution *Reliability: Theory & Applications*, 17(3(69)):52–58.

MEASUREMENTS OF BRIDGE STRUCTURES USING NON-DESTRUCTIVE TESTING METHODS AND THEIR STABILITY IN WIND GUSTS

Alena Rotaru
•
All-Russian Scientific Research Institute for Civil Defence
and Emergencies of the EMERCOM of Russia
alenarotaru@mail.ru

Abstract

As bridge structures become older and older, they are subject to wear and tear due to ageing, weather conditions or environmental effects, as well as due to surprise structural modifications substantially affecting the condition of the structures. Therefore, the condition assessment of bridge structures is a must for the safety and absence of risk. The condition assessment of bridge structures is also necessary for the maintenance and repair of existing structures having been in service for more than 30 years, in order to avoid breakdowns and save human lives. This paper states the condition assessment performed with the use of various nondestructive test methods.

Keywords: bridge structures, nondestructive, test, methods, risk.

I. Introduction

Assessment of structures has been a top subject of studies in recent years. The studies in this field encompass signal processing, data management, and measurements. Topical issues of bridge structures assessment are currently defined as distributed and integrated data control and storage, data analysis and knowledge discovery, diagnostic techniques and provision of useful and reliable information to bridge owners/managers for making service and control decisions.

II. Methods

The structure condition is assessed with the use of nondestructive test methods to obtain information on the condition of the structures. This paper contains some exemplary bridge structure studies with the use of the following methods:
1) Visual inspection;
2) Ultrasonic testing;
3) Radio-wave inspection (structure penetrating radar survey);
4) Radio-wave inspection (ground penetrating radar survey);
5) Vibroacoustic inspection (dynamic testing).
Visual inspection is a nondestructive test method. Visual monitoring and field tests are used to determine the numerical indicator of damage between 0.5 and 10, as shown in Table 1.

Table 1: *Degree of damage during measuring control*

Range of measured value, mm	Measurement error, mm
Up to 0,5 inclusive	0,1
Above 0,5 to 1,0 inclusive	0,2
Above 1,0 to 1,5 inclusive	0,3
Above 1,5 to 2,5 inclusive	0,4
Above 2,5 to 4,0 inclusive	0,5
Above 4,0 to 6,0 inclusive	0,6
Above 6,0 to 10,0 inclusive	0,8
Above 10,0	1,0

The visual inspection is conducted with the use of measuring tools such as: tape measure, photo camera, calipers.

1) Visual inspection of a bridge (Pictures 1-6).

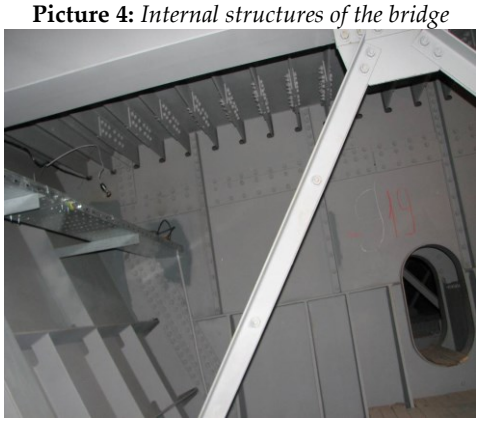
Picture 1: *View of the bridge from the right bank of the river*

Picture 2: *View of the roadway and pedestrian part of the bridge*

Picture 3: *Roadway fence construction*

Picture 4: *Internal structures of the bridge*

Picture 5: *Transition bays between bridge sections at the piers*



Picture 6: *Internal construction of bridge span adjacency to piers*

2) Ultrasonic testing

The structures are measured by means of UK-1401 instrument designed for measuring the time and velocity of propagation of longitudinal ultrasonic oscillations in solid materials at the surface sounding on a fixed base, in order to determine the strength and integrity of materials and structures. The test results of metalwork (mw) of a bridge crossing a river are given in Table 2.

Table 2: *The results of strength calculation*

Segment	*Value*			*Arithmetic mean*	*Sigma*	*0,95*	*Strength, MPa*
	1	*2*	*3*				
Strength of metal structures between 9 and 10 supports right side							
1	5700	5700	5700	5700,00	0,00	5700	39,75
	5680	5700	5680	5686,67	11,55	5667	39,52
	5700	5660	5680	5680,00	20,00	5646	***39,37***
	5680	5680	5700	5686,67	11,55	5667	39,52
	5700	5700	5730	5710,00	17,32	5681	39,61
	5700	5680	5700	5693,33	11,55	5674	39,57
	5700	5680	5700	5693,33	11,55	5674	39,57
Mean value							**39,56**
Strength of metal structures between 9 and 10 centerpiece							
2	6050	6170	6050	6090,00	69,28	5973	41,66
	6070	6100	6170	6113,33	51,32	6027	***42,04***
	6120	6000	6070	6063,33	60,28	5962	41,58
	5910	5930	5980	5940,00	36,06	5879	41,00
	5950	5860	6000	5936,67	70,95	5817	40,57
	5910	6000	5930	5946,67	47,26	5867	40,92
Mean value							**35,39**
Strength of metal structures between 9 and 10 support left side							
3	5700	5750	5810	5753,33	55,08	5660	39,47
	5840	5810	5810	5820,00	17,32	5791	40,38
	5810	5880	5840	5843,33	35,12	5784	40,34
	5770	5840	5840	5816,67	40,41	5749	40,09
	5860	5860	5860	5860,00	0,00	5860	40,87
	5840	5910	5880	5876,67	35,12	5817	40,57
Mean value							**34,53**

3) Structure penetrating radar survey

Structures are scanned by means of the OKO-M1 penetrating radar (Picture 7) with AB-1700 antenna unit, as designed for detecting various objects (metal and non-metal articles, e.g., reinforcement bars) in concrete structures, cavities and for studying homogeneity of structures.

Picture 7: *GPR with antenna unit AB-1700*

The penetrating radar operates based on the property of radio waves to reflect from the interface of media with different dielectric permeability. The radar with central frequency of 1700 MHz, sounding pulse duration of 1 to 5 ns and scanning resolution of 5 to 10 cm has a maximum sounding depth of 0.8 to 1.0 m. The penetrating radar sounding was made on the columns in axes X, Y, and Z. The radar survey diagram is shown in Figure 1. The scanning radargrams are shown in Figure 2. The radargrams have been obtained by computer-aided processing with the GeoScan-32 software.

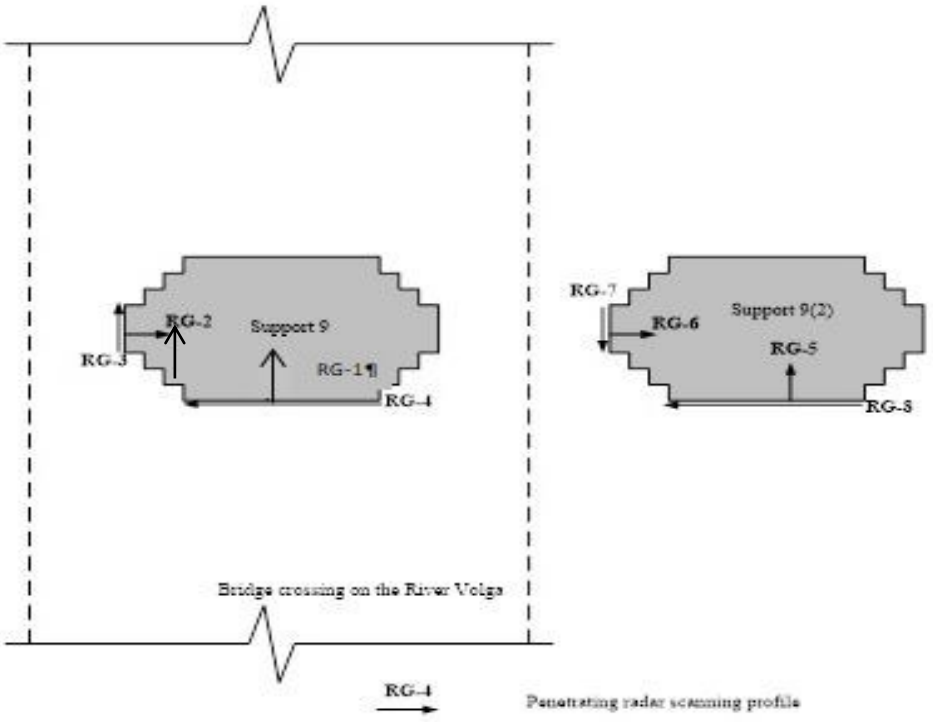


Figure1: *Study diagram of supports No. 9*

Radargrams of support 4000x3000 mm

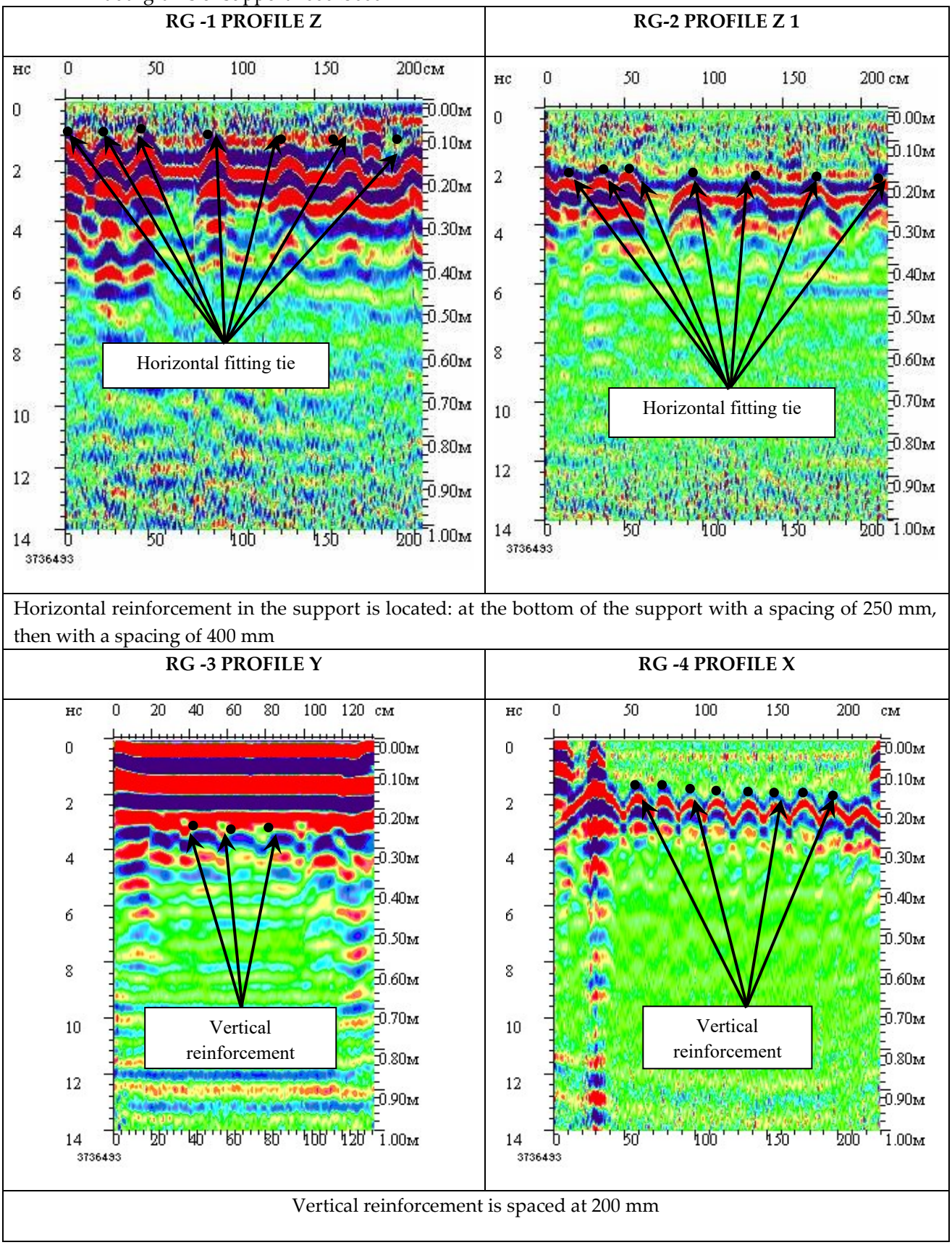


Figure 2: *Radargrams of support No. 9*

Finding: The scanning of the supports located on the right bank of the Volga River has yielded the following results:

Supports No. 9 – the column section is 4,000x3,000 mm, the horizontal in-column reinforcement is arranged as follows: with 200-250 mm spacing at the bottom of the column, with subsequent 400 mm spacing and the vertical reinforcement to have been arranged with 200 mm spacing. No cavities or discontinuities have been found in the concrete. Generally, the structural members of the supports are consistent with the project design.

4) Ground penetrating radar survey

The OKO-M1 penetrating radar (Picture 8) with AB-400 antenna unit is designed for detecting various objects (metal and nonmetal articles underground, underwater and in cavities) and for studying the homogeneity of soils. The penetrating radar operates based on the property of radio waves to reflect from the interface of media with different dielectric permeability. The radar with central frequency of 400 MHz, sounding pulse duration of 1 to 5 ns and scanning resolution of 5 to 10 cm has a maximum sounding depth of 4.0 to 5.0 m for dry soils.

Picture 8: *GPR «OKO-M1» with antenna unit AB-400*

The scanning was performed all round the supports and in the coastal area near the supports to determine the homogeneity of soils, the ground water level, and the presence of cavities. The survey diagram is shown in Figure 3. The scanning profiles (radargrams) of the survey area are shown in Figure 4. The radargrams have been obtained by computer-aided processing with the GeoScan-32 software.

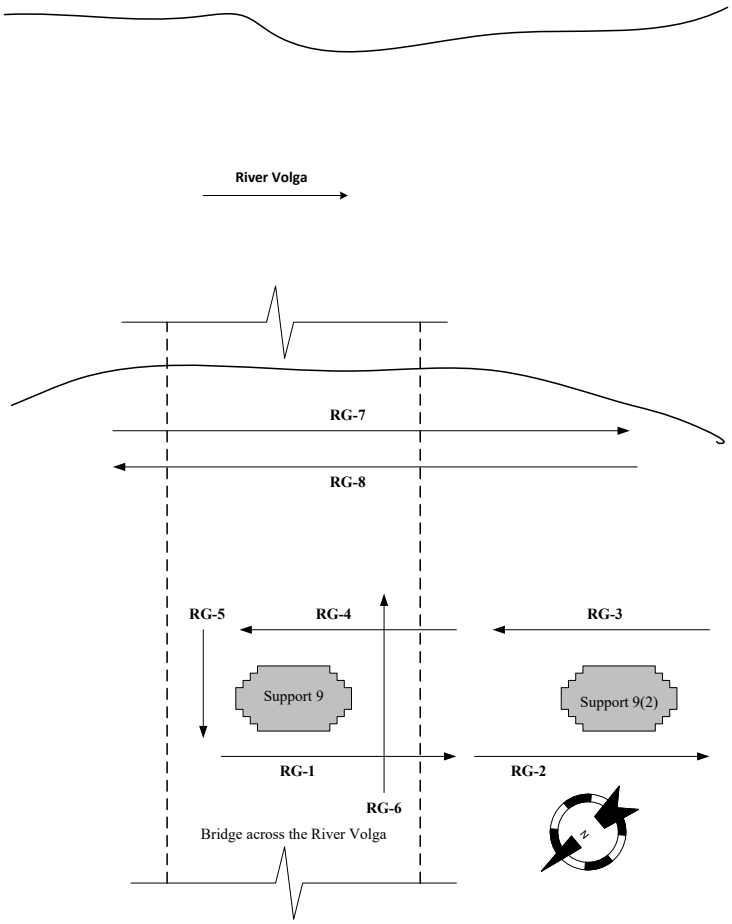
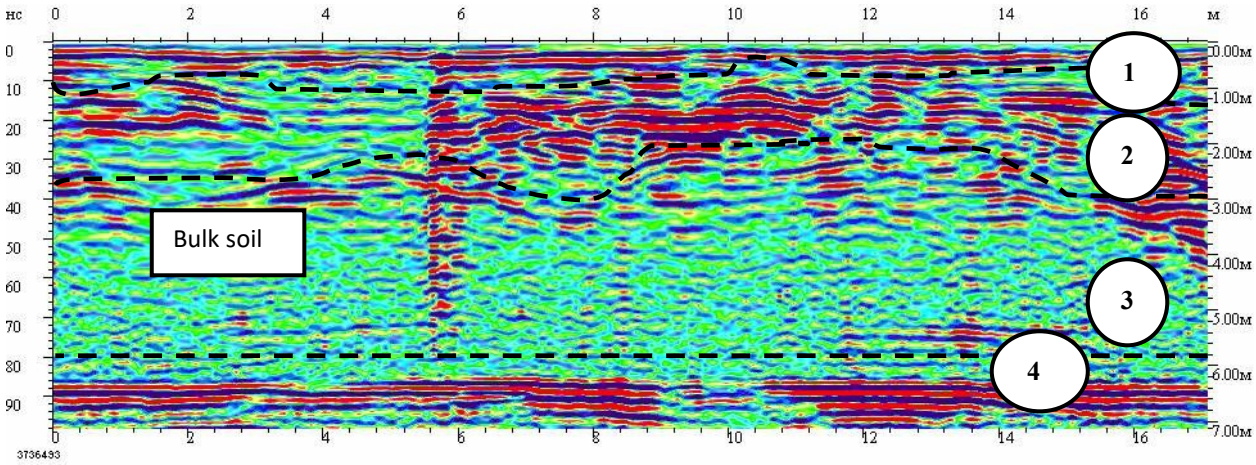


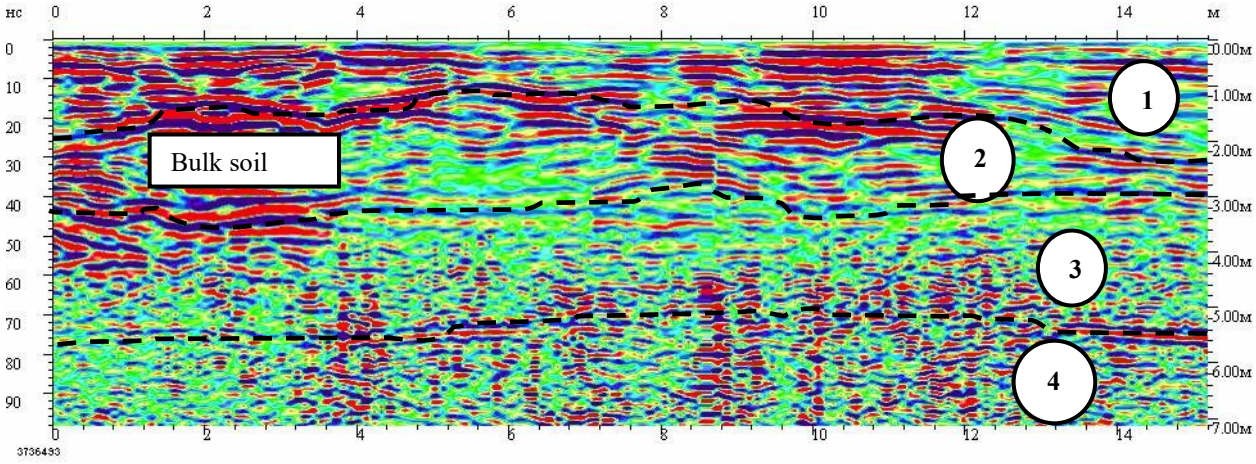
Figure 3: *Ground penetrating radar survey near supports No. 9*

RG-1 Surface scanning profile



There is a uniform occurrence of soils without sharp differences in depth. The bulk soil consists of three layers. The thickness of the bulk soil is 6.2-6.3 m, then the bedrock lies.

RG-2 Surface scanning profile



There is a uniform occurrence of soils without sharp differences in depth. The bulk soil consists of three layers. The thickness of the bulk soil is 5.2-5.8 m, then the bedrock lies.

Figure 4: *Radargrams of soil on profile 1 u 2*

5) Dynamic testing of structures

The identification of the bridge dynamic parameters and the bridge structure hidden defects required to use loading devices; three-component seismic detectors; connecting cables; multi-channel analog-to-digital converters; computers with a software package for seismic vibration signal analysis [2].

The dynamic tests are conducted to determine dynamic and stiffness properties, load-bearing capacity of structural members of buildings and structures, and to detect any hidden defect (Figures 5, 6).

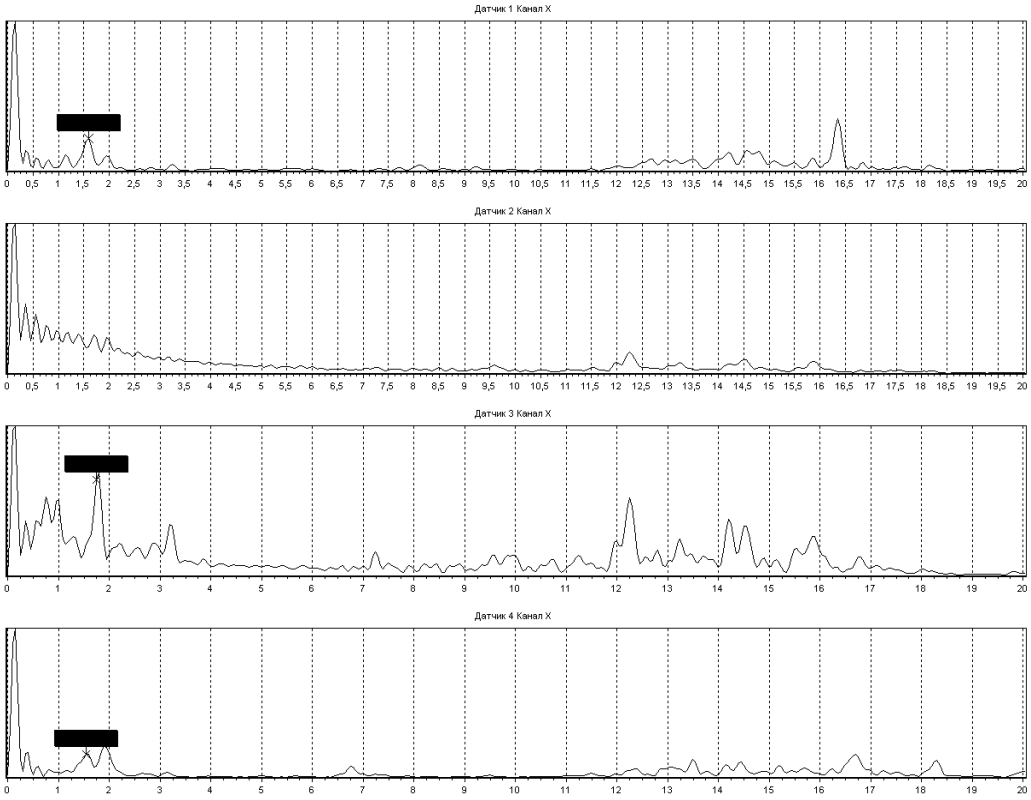


Figure 5: *The natural vibrations of the bridge along the X axis*

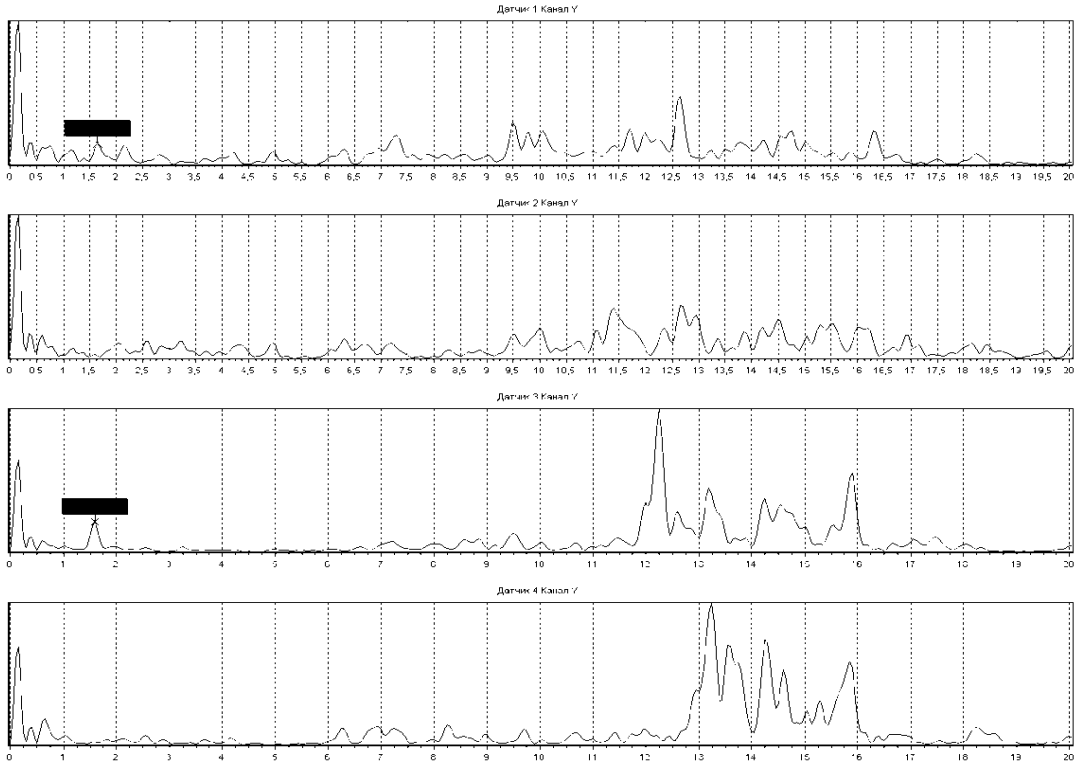


Figure 6: *Natural vibrations of building block 1 along the Y axis*

Measurements during wind gusts have been conducted (Figures 7).

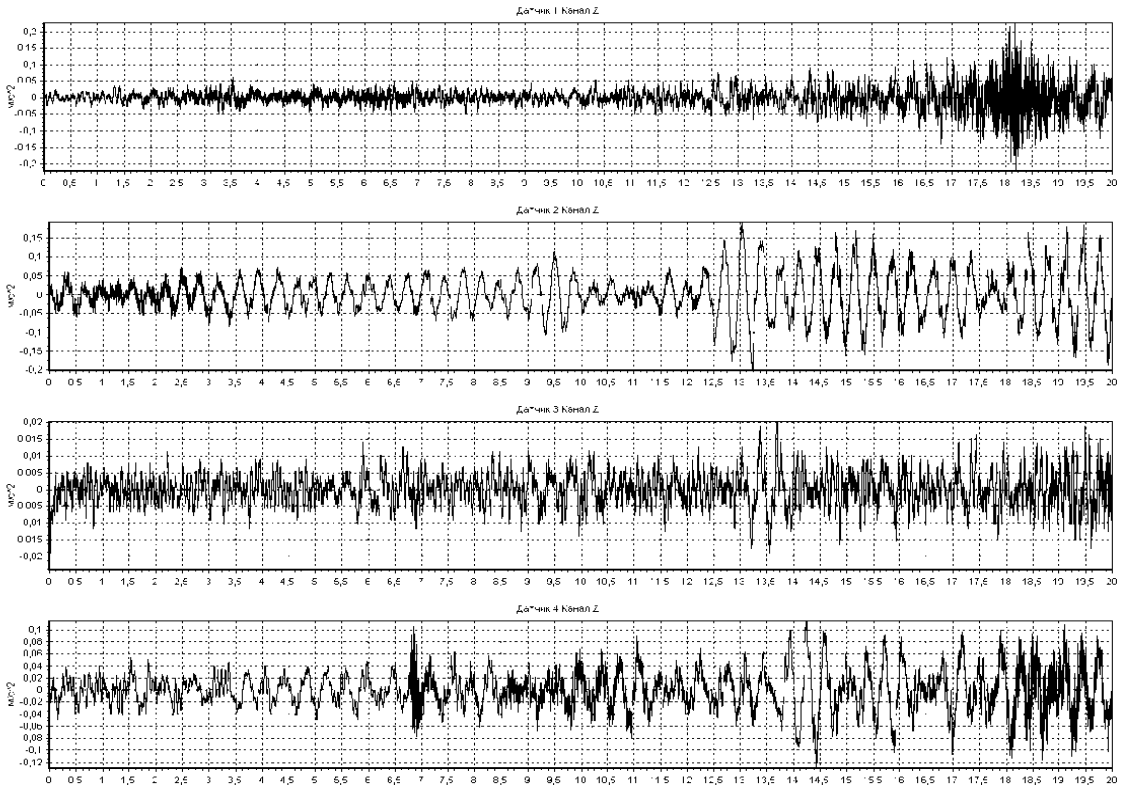


Figure 7: *Accelerations along the X-axis in the middle part of the bridge in small wind gusts*

Instruments have been used for the dynamic testing of the ground-to-structure system with the dynamic parameters of the structures obtained. The natural period of structural system oscillations as the basic dynamic parameter is known to relate to *EJ* stiffness thereof. Therefore, the dynamic test results to increase the natural period of oscillations *T1* of a structure are going to yield the value of reduction in the structure integral stiffness. Mathematically, the dependency of the natural period on the stiffness can be expressed as:

$$T_1 = \kappa\sqrt{\frac{m}{EJ}}$$

where
κ is a factor accounting for the type and dimensions of the structural layout,
m is mass of the structure,
E is modulus of elasticity,
J is moment of inertia.
Generally, the natural period of oscillations *T1* features the structural system's stiffness. Afterwards, the integral reduction in the inertia moment will show the potential defects in the sections of the structural members of the structure. A reduction in the modulus of elasticity shows that an integral reduction in the strength of the structural members occurs.
From the design data: *T1* norm. = 2.4 sec.
The survey and acceptance test results of the structure have shown that the natural periods of oscillations of the structure for spans 8-9 and 7-8, respectively, are equal to:
T = 0.37 sec;
T1y = 1.67 sec;
T2 = 0.4 sec.
For city bridges, the natural period of oscillations of the first 2 modes should not fall within the range of 0.45 to 0.6 sec.
The test results have shown the natural periods of oscillations of the structure to be, respectively, equal to:
T1x = 0.3-0.5-0.625 sec;
T1y = 0.3-0.5-0.625 sec;

I.e., the dynamic test results show the oscillation periods of the bridge to fall within the range of 0.45 to 0.6 sec with reference to car traffic movement and light wind gusts, and this proves that the bridge is subject to dynamic impacts.
In addition, the wind gusts were found to lead to a "play" phenomenon on the bridge, with 3-4 sec period. With the same period, but greater amplitudes, critical "play" oscillations were occurring, which led to the stopping of the movement [4].

III. Results

No visually identifiable defects have been found in the load-bearing structures of the bridgework. The ultrasonic testing yielded the following results:
- the strength (of mw between supports 9 and 10) is varying between 39.37 and 42.04 MPa (up to 6 % of the maximum), the average values being between 34.53 and 39.56 MPa. The ground scanning suggests the following conclusions:
- the backfill soil thickness in this geological area is 6.2 to 6.3 m, as followed by deeply formed alluvial soils;
- no substantial changes have been found in the soil mass.

Generally, the soils have been found to lay evenly, with no abrupt depth changes. In general, the geologic conditions of the area are considered favorable upon the ground penetrating radar scanning. The visual, strength, and geodetic surveys yielded no changes in the technical condition of the bridge. The wind gusts turned out to lead to a "play" phenomenon on the bridge with 3-4 sec period [2, 3]. With the same period, but greater amplitudes, critical "play" oscillations were occurring, which led to the stopping of the movement. During storm winds the bridge will be periodically subject to the "play" or critical "play" phenomena. This paper provides the measured data of the bridge, which have been obtained with the use of the methods applied. The application of such methods improves the specialist knowledge.

IV. Conclusions

The analysis shows that the monitoring of bridge structures from the point of view of the users of bridge structures aims at the following important tasks: ensuring the preservation of bridge structures; increasing durability bridges structures by timely detecting damage and repairing it in a timely manner; maintaining the load capacity of bridge structures by managing their behavior during operation; improving the efficiency of expenditure on repair increase efficiency of expenses for repair works by means of correct determination of time and types of necessary increase cost efficiency of repair works by correct timing and type of necessary repairs [2, 3, 4, 5].

References

[1] Kappos, A.J. (2001). Dynamic Loadind and Design of Structures. London and New York.

[2] Rotaru, A.N. and Savinov, A.M. (2021). Civil Security Technology 18(3) (69) 53-55.

[3] Savinov, A.M., Rotaru, A.N., Nigmetov, G.M. (2020). Determination of vulnerabilities in of multi-storey buildings by the method of dynamic-geophysical tests, Proceedings of Section No. 12 XXX International Scientific and Practical Conference PREVENTION. RESCUE. HELP" 113-118, Himki, Russian Federation.

[4] Rotaru, A.N., Nigmetov, G.M., Maklakov, A.S., Gaifullin, Z.G. (2021). Civil Security Technology 18(1) (67), 46 –53.

[5] Sardar Kashif Ur Rehman, Zainah Ibrahim, Shazim Ali Memon, Mohammed Jameel, (2016). Construction and Building Materials Volume 107, 58-86.

[6] Todorov, E.I. (2022). Encyclopedia of Materials: Metals and Alloys, Volume 3, 168-192.

STATISTICAL MODELS FOR FORECASTING NATURAL EMERGENCIES

Valery Akimov, Maxim Bedilo, Olga Derendiaeva
•
All-Russian Research Institute for Civil Defense and Emergencies of the Ministry of Emergency Situations of Russia
akimov@vniigochs.ru

Abstract

The article considers predictive-analytical solutions for natural hazards for urbanized areas, the mathematical basis of which is Bayesian classifiers. The result of the work is a formalized description of models for predicting forest fires, the consequences of earthquakes and floods resulting from floods.

Keywords: predictive and analytical solutions; Bayes method; emergency situations of a natural nature; Forest fires; earthquakes; floods.

I. Introduction

The subject of the study is natural hazards for urban areas with the aim of their formalized description using Bayesian classifiers. The main part of the work contains three sections: a model for predicting forest fires; model for predicting the consequences of earthquakes; model for forecasting floods due to floods.

II. Methods

Statistical Model for Forest Fire Prediction. As part of the creation of the hardware-software complex "Safe City", an approach to modeling forest fires (hereinafter referred to as LP) is proposed, which is based on Bayesian methods [1]. In the statistical model for predicting the LP, the following parameters are subject to probabilistic assessment using a Bayesian classifier: speed of the riding LP; area and speed of the downstream LP [2]. The calculation of the average speed of the front of the upper LP for various gradations of wind speed is calculated by the equation:

$$v_{frsr}^{j} = \frac{\sum_{i=1}^{N_j} v_{fri}}{N_j}, \quad (1)$$

Where:
j is the range of wind speed values corresponding to the j-th gradation (according to Table 1);

Nj is the number of upper LPs recorded at wind speeds that are in the detected values, detection of the j-th gradation, units;

νfri is the front speed of the i-th riding LP, m/min.

Table 1: *Directory of gradations of wind speed values*

№	Range of values, m/s
1.	Less than 0,3
2.	[0,3; 1,6)
3.	[1,6; 3,4)
4.	[3,4; 5,5)
5.	[5,5; 8)
6.	[8; 10,8)
7.	[10,8; 13,9)
8.	[13,9; 17,2)
9.	[17,2; 20,8)
10.	[20,8; 24,5)
11.	[24,5; 28,5)
12.	[28,5; 32,7)
13.	from 32,7 and more

The main indicators of the propagation of low-level LPs are calculated according to the Rothermel model, which is based on the position that the flame propagation speed is proportional to the ratio of the heat of combustion of the material to the heat of heating new portions of combustible material to the ignition temperature [3].

Determining the speed of the propagation front of the grassroots LP in this model is carried out as follows.

According to the data obtained from the results of collecting the characteristics of the forest area, the composition of the forest combustible material (FCM) for this area is determined, after which the surface area per 1 m2 of the territory is determined for each layer of the FCM according to the formula:

$$A_i = \frac{\sigma_i \omega_{0i}}{\rho_i}, \quad (1)$$

Where:

ω_0 is the average reserve of the layer of forest materials, kg/m²;

σ is the specific surface area of the FCM layer, m–1;

ϱ is the density of the LCM layer, kg/m^3.

The values of the parameters of formula (2) are determined according to table 2.

Table 2: *Average characteristics of the ground cover*

Compound	Designation	Lichen	Moss schreber	Needles	Leaves	Dry cereals	Bush	Logging waste
Average stock of a layer of LCM, kg/m²	ω_0	1,7	1,0	0,3	0,15		0,225	0,9
Specific surface area of the LCM layer, m^{-1}	σ	2000	2500	6000	11560	18170	6560	4920

Heat of combustion, kcal/kg	h	4300		4700	4500	4880	4200		4400			
Density LGM layer, kg/m^3	ρ	300			512	460	420		512			
Layer height, m	δ	0,12		0,1		0,075	0,5	0,6		0,7		
Critical moisture content, kg/kg	M_x	0,3		0,5	0,3	0,4			0,2			

The total surface area (AT) of all LCM layers is calculated by the formula:

$$A_T = \sum_{i=1}^{n} A_{i.} \quad (3)$$

Next, the wind speed is calculated at a level of 6 meters above the crowns of the forest stand (WZ6) according to the formula:

$$W_{Z6} = W_{Z10}\left[\frac{Z_d + 6}{10}\right]^{0,28}, \quad (4)$$

Where:

W_{Z10} - wind speed according to the nearest weather station;

$Z_d$is the average stand height.

Then the speed of propagation of the fire front in the absence of wind and slope (v_{fr}^0) can be calculated by the formula:

$$v_{fr}^0 = \frac{0{,}048\mathrm{h}_{sr}\eta_S\omega_{0sr}\eta_m r\xi}{(\rho_{sr} + \rho_{sr}S_T)Q\gamma\varepsilon{\sigma_{sr}}^{-0,8189}}, \quad (5)$$

Where:

η_S is a coefficient that takes into account the mineral composition of forest fuels (assumed to be 0.42);

SR - mineral content.

Based on the assumption that the lower LP in a homogeneous medium propagates in an ellipse, the area of the LP can be determined as follows.

First, the ratio of the length of the LP to its width (LB) is determined by the formula:

$$LB = 0{,}936\exp(0{,}2566W) + 0{,}461\exp(-0{,}1548W) - 0{,}397. \quad (6)$$

Next, the ratio of the leading edge of the LA to its trailing edge (HB) is determined by the formula:

$$HB = \frac{LB + \sqrt{LB^2 - 1}}{LB - \sqrt{LB^2 - 1}}. \quad (7)$$

The values a, b and c are calculated by the formulas:

$$b = \frac{v_{fr}(1 + HB)}{2HB}, \quad (8)$$

$$a = \frac{b}{LB}, \tag{9}$$

$$c = b - \frac{v_{fr}}{HB}. \tag{10}$$

The final value of a, b and c is determined by multiplying each of the given values by the forecast period in minutes.

Then the LP area (S_{lp}) is calculated by the formula:

$$S_{lp} = \pi ab. \tag{11}$$

Statistical model for predicting the consequences of earthquakes

The model for predicting the consequences of earthquakes is based on calculation methods for assessing the parameters of seismic impact, determining the degree of destruction of buildings and structures, elements of urban infrastructure, including theoretical approaches in the field of statistical data analysis based on the Bayesian method of interpreting probability.

The main initial data for the formation of the basic training set of the model for predicting the consequences of earthquakes are the following groups of parameters: characteristics of earthquakes; characteristics of damage to buildings and structures during earthquakes; characteristics of buildings and structures; characteristics of territories; parameters of the meteorological situation [4].

At the stage of training the model, the parameters of the hypotheses are determined based on the values of the parameter "Degree of damage to the building (structure)".

After training, the process of predicting events corresponding to hypotheses begins on new values of the observed parameters.

The order of collecting new values of the observed input data and processing them corresponds to similar processes when training the model, except that the parameters of the hypotheses (estimating the posterior probabilities of the hypotheses) are determined by the Bayesian classifier.

When forecasting using this model, each building or structure is considered separately. For these purposes, the characteristics of all buildings and structures in the controlled area are prepared in advance [5].

Statistical model for flood forecasting

The main input data for the formation of the base training set of the model for flood forecasting due to floods are flood data due to heavy rainfall floods; data characterizing sections of rivers with sections of their watersheds; data characterizing the watersheds of the observed sections of the rivers; data characterizing water management systems in the observed sections of rivers; data characterizing controlled settlements (hereinafter referred to as CPs) located in areas of probable flooding of the area; data characterizing the hydrological situation in river sections; data characterizing the meteorological situation in the observed areas of the terrain; data characterizing the prevailing landscape of the area, the types and composition of soils within the boundaries of the NP and the catchment areas of the observed rivers [6].

In this model, the most significant indicators are the calculated hydromorphological parameters during the flood period.

When determining the calculated hydromorphological characteristics of river points, the

following calculation methods are used [7]:

in the presence of data from hydrometric observations - directly from these data;

in case of insufficiency of hydrometric observation data, by reducing them to a multi-year period according to the data of analogue rivers with longer observation series;

in the absence of hydrometric observation data and the impossibility of comparison with the given analogous rivers, according to calculation formulas using maps based on the totality of observational data from the entire network of monitoring points for the hydrological situation in the corresponding area or a larger territory.

III. Results

Thus, this article considers predictive and analytical solutions for natural hazards for urban areas, the mathematical basis of which is Bayesian classifiers. The result of the work is a formalized description of models for predicting forest fires, the consequences of earthquakes and floods resulting from floods.

IV. Discussion

The discussion of the verbal and mathematical foundations of predictive modeling of natural emergencies is quite active in the scientific literature [8 - 10].

The scientific novelty of the developed models lies in a single scientific approach to their creation, namely, the use of a statistical processing method based on Bayes' theorem.

Bayesian probability is currently developing as the main method of forecasting in terms of building and training neural networks, in contrast to frequency probability - when the probability is determined by the relative frequency of occurrence of a random event with sufficiently long observations.

For scientific forecasting of crisis situations and incidents using the Bayesian method and Bayesian networks, a large amount of up-to-date data is required for modeling natural disasters, which is typical for often recurring negative events.

Due to the lack of statistical data, Bayesian methods are not applicable for predicting catastrophic natural disasters that occur rarely, but with significant damage.

References

[1] Predictive and analytical solutions for natural, man-made and biological and social threats of a unified system of information and analytical support for the safety of the environment and public order "Safe City" / V. A. Akimov, A. V. Mishurny, O. V. Yakimyuk [and etc.]. - Moscow: All-Russian Research Institute for Civil Defense and Emergency Situations of the Ministry of Emergency Situations of Russia, 2022. - 315 p. – ISBN 978-5-93970-278-2. – EDN MGXNYI.

[2] Akimov, V. A. Large natural fires as sources of natural emergencies / V. A. Akimov, M. V. Bedilo, E. O. Ivanova // Civil Security Technologies. - 2022. - T. 19, No. 2 (72). - S. 11-13. – DOI 10.54234/CST.19968493.2022.19.2.72.2.11. – EDN PGJCMG.

[3] Barovik, D. V. Computer model, examples of the analysis of the spread of ground fires / D.V. Barovik, V. B. Taranchuk // Problems of physics, mathematics and technology. - 2020. - No. 4 (45). - S. 113-120. – EDN DIJSBZ.

[4] Akimov, V. A. Hazardous geophysical phenomena and processes as sources of natural emergencies: a model of medium-term earthquake forecasting / V. A. Akimov, M. V. Bedilo, E. O. Ivanova // Civil Security Technologies. - 2022. - T. 19, No. 1 (71). - S. 20-23. – DOI 10.54234/CST.19968493.2022.19.1.71.4.20. – EDN QICTLF.

[5] Preliminary national standard of the Russian Federation PNST 763-2022 “Safety in emergency situations. Safe city. Earthquake forecasting. General requirements".

[6] Akimov, V. A. Hazardous hydrological phenomena and processes as sources of natural emergencies: a verbal model / V. A. Akimov, M. V. Bedilo, S. P. Sushchev // Civil Security Technologies. - 2021. - T. 18, No. 4 (70). - P. 4-8. – DOI 10.54234/CST.19968493.2021.18.4.70.1.4.

– EDN EVUUFA.

[7] Security of Russia. Legal, socio-economic and scientific and technical aspects. Analysis and provision of security from emergency situations / V. A. Akimov, A. A. Antyukhov, E. V. Arefieva [and others]; Security Council of the Russian Federation, Russian Academy of Sciences, EMERCOM of Russia, Rostekhnadzor, Russian Science Foundation, Rostec State Corporation, Rosatom State Corporation, Rosneft Oil Company PJSC, Russian Railways OJSC, Transneft PJSC, Gazprom PJSC. - Moscow: MGOF "Knowledge", 2021. - 500 p. – ISBN 978-5-87633-199-1. – EDN FXIJPZ.

[8] Akimov, V. A. Classical and post-non-classical methods for predicting natural emergencies / V. A. Akimov, M. V. Bedilo // Civil defense on guard of peace and security: Proceedings of the VII International scientific and practical conference dedicated to the World Day of Civil Defense in the year of the 90th anniversary of the formation of the Academy of the State Fire Service of the Ministry of Emergency Situations of Russia. At 5 a.m., Moscow, March 01, 2023. Volume Part I. - Moscow: Academy of the State Fire Service of the Ministry of the Russian Federation for Civil Defense, Emergencies and Disaster Relief, 2023. - P. 20-27. – EDN YAVGTG.

[9] Akimov, V. A. Scientific forecasting of natural emergencies / V. A. Akimov, E. O. Ivanova // Russia in the XXI century in the context of global challenges: problems of risk management and security of socio-economic and social political systems and natural-technogenic complexes: collection of materials of the All-Russian Scientific and Practical Conference, Moscow, April 26-27, 2022 / Russian Academy of Sciences, International Independent Environmental and Political University, State University of Management. Volume Issue 1. – Moscow: State University of Management, 2022. - P. 114-119. – EDNFFDBHY.

[10] Akimov, V. A. Modern methods of studying emergency situations of natural, technogenic and biological and social nature / V. A. Akimov, M. V. Bedilo // Civil defense on guard of peace and security: Materials of the VI International scientific and practical conference dedicated to World Civil Defense Day. In 4 parts, Moscow, March 01, 2022 / Comp. V.S. Butko, M.V. Aleshkov, S.V. Podkosov, A.G. Zavorotny [i dr.]. Volume Part I. - Moscow: Academy of the State Fire Service of the Ministry of the Russian Federation for Civil Defense, Emergencies and Disaster Relief, 2022. - P. 16-24. – EDN HSSYVB.