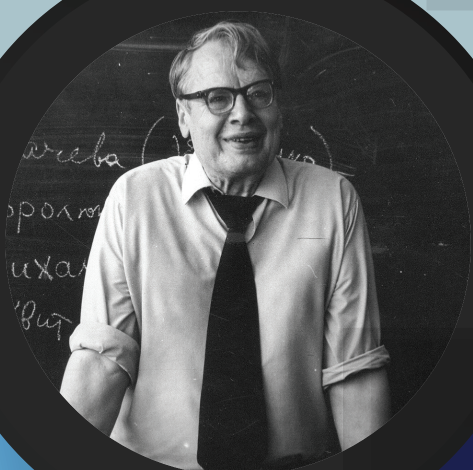
RTA

RELIABILITY: THEORY & APPLICATIONS

ISSN 1932-2321

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

INTERNATIONAL GROUP ON RELIABILITY



RELIABILITY

RISK ANALYSIS

MAINTENANCE

SAFETY

GNEDENKO FORUM PUBLICATIONS

#2

(68) VOL.17

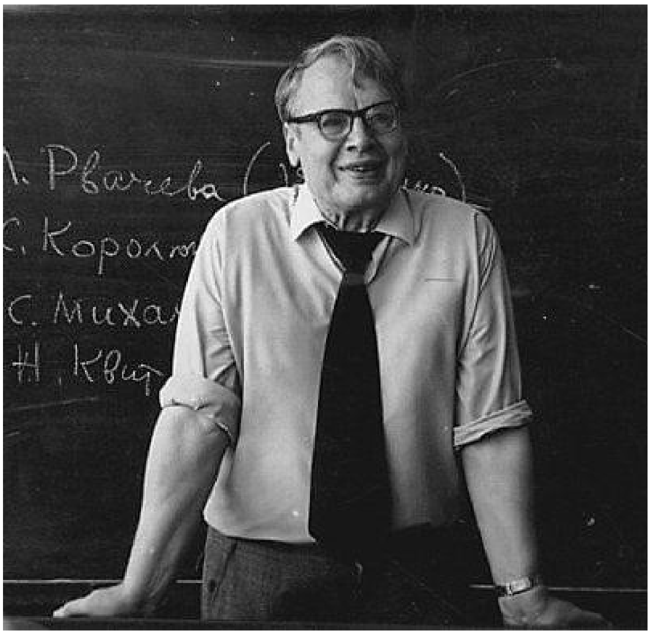
JUNE 2022

SAN DIEGO

ISSN 1932-2321

http://www.gnedenko.net/Journal/index.htm

Journal of International Group on Reliability



RELIABILITY:

THEORY & APPLICATIONS

Vol.17 No.2 (68),
June 2022

San Diego
2022

Editorial Board

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

Theoretical papers have to 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 have to 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 has to 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 taking into account 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

In this paper a parallel system has been discussed with the idea of priority to preventive maintenance over replacement. The system has two identical units and facility of inspection is given to the failed unit before repair/replacement. There is a single server who play four-in-one role of inspection, replacement, repair and preventive maintenance and comes immediately when required. Units are failed with constant rate whereas failure time is random. The distribution of time for repair activities is arbitrary and there rates follow exponential distribution. The random variable associate with different rates are stochastically independent. Mathematical expression for several reliability terms like MTSF, availability, busy period analysis for server , expected number of visits by the server and cost benefit are obtained by using semi-markov process and regenerative point technique. Graphs are drawn to find the effect of various parameters on MTSF, Availability and profit.

In real-life decisions, usually we happen to suffer through different states of uncertainties. In order to counter these uncertainties, in this paper, the author formulated a transportation problem in which availability, demand and costs are mixed terms of real, triangular intuitionistic fuzzy numbers. In this paper, a simple method for solving type-3 intuitionistic fuzzy transportation is applied. So, the proposed method gives the optimal solution directly. The solution procedure is illustrated with the help of numerical examples.

The roles of heavy – tailed distribution in modelling real life events, especially in financial and actuarial sciences, cannot be over – emphasized. In this paper, a new heavy right – tailed, three – parameter continuous distribution with increasing hazard rate called Zech distribution is developed. The proposed model is very suitable for modelling heavy right- tailed data. Zech distribution is the reciprocal of the random variable which follows Gompertz- Inverse – Exponential (GoIE) distribution and it does not involve addition of extra parameter, thereby removing the cumbersomeness in the estimation process posed by other methods involving additional extra parameters, especially where more than three parameters are involved. The statistical properties of the new distribution such as survival function, hazard function, cumulative hazard function, reversed hazard function, quantile function, order statistics, moments, mean, median, variance, skewness, and kurtosis were derived. The Linear representation of the pdf of the newly developed distribution revealed that its probability density function is a weighted exponential distribution. Also, method of maximum likelihood was used in estimating the model's parameters. The simulation results revealed that as the sample sizes increased, the root mean squared errors decreased which showed that the parameters of Zech distribution are stable. The proposed distribution was applied to two real life data sets. The results showed that Zech distribution performs better than Gompertz Inverse Exponential distribution, Weibull Exponential distribution and Gompertz Exponential distribution.

Purpose – This paper presents a sensitivity analysis of a urea fertilizer manufacturing system comprising several sub-systems of differing nature. Design/methodology/approach–A mathematical model is developed for the consistent general repair and disappointment rates for every subsystem. The framework is analyzed by utilizing regenerative point graphical technique; as a result, some recommendations are made for the optimized output. A state transition diagram of the system is developed to find mean time to busy period server, system failure and system availability. Findings – The present study suggests an approach to improve the system performance. The analysis and results outlined in this paper are useful to system managers, training supervisor, engineers and reliability analysts in the manufacturing industry. Originality/ value – The manufacturing system of Urea fertilizer consists of a complex structure with the high risk of machine failure. Machine/ Production failure leads to high risks of economic & environmental loss and worker's safety. To address this challenge effectively, sensitivity analysis of the urea fertilizer plant is discussed for minimizing the risk of machine failure.

This Fractional Transportation Problem arises when an enterprise has to face the issue of maintaining a good ratio of some critical parameters. These parameters are directly concerned with product(s) transportation from sources to destination. This paper considers a multi-objective Capacitated Transportation Problem with Fractional Objectives. A fuzzy goal programming approach with different membership functions is applied to generate a different set of solutions. We also use Chebyshev's Goal Programming for obtaining the solutions. Finally, a numerical illustration is provided to validate our proposed model.

Gas Insulated Substation (GIS) is essential for the transmission and control of power both in AC and DC electrical systems. Functionally Graded Material (FGM) technology is widely used for the design of the spacer material in the GIS to reduce the electric stress in the system. Optimal designing of the material of the spacer gradings with a particular attention to the number of gradings may prove to be very useful in reduction of the stress in the GIS at an effective cost. This paper deals with the design and development of an optimal dielectric material for the functionally graded material (FGM) spacer in a GIS. A novel optimization method has been proposed which is used for the optimization of the conductor material and the FGM epoxy spacer. The optimal value for each grading of functionally graded material spacer is determined by the proposed method. A dual-objective function is chosen for the optimization problem. The objective of the problem is to minimize the maximum field stress in addition to the standard deviation in the electric field. A post type spacer has been considered for the study. Initially, the optimization of the dielectric material is done only for 4 gradings. Gradually, the number of gradings in the FGM-spacer is increased to determine the optimal number of gradings suitable for the design.

Consider a multi-processors computer system consisting of a ready queue of different jobs to be executed/processed. Lottery scheduling is fair enough to schedule the resources for each and every job. The research idea assumes condition where one can observe some processes to be fully executed; some partially executed few blocked/suspended/ terminated, after sudden system breakdown. An estimation strategy has been designed for the estimation of the total time required to process all these types of processes (processed, partially processed and blocked processes). How much time is required to process the remaining in any hazardous situation? A regression type estimator of sampling theory is used to perform this task. This remaining time estimation technique deals with the backup cost and recovery management as well. Sampling techniques are used in proposed approach for the testing purpose and a simulation has been performed. Another tool adopted is the confidence intervals which are calculated and gives proper précised values in comparison to the true mean for the total remaining time. The linear, square root and square cost function model are adopted for the calculation of backup cost and recovery management. In addition some auxiliary information is also incorporated in the form of size measure of the processes which is an effective approach to calculate the complete remaining time of the processes in multiprocessor environment. The purpose of the proposed research has been served effectively as one can observe the results of disaster and recovery management of the computer system.

Improving the management of the technical condition of equipment, devices and installations, the service life of which exceeds the standard value, is one of the most important problems of state security. Today, the relative number of such equipment already exceeds 60%. The results of the analysis of literature data on this problem presented, which confirm its relevance and significance. It is important to note that these findings apply to not only electrical power systems, but many other production systems as well. The main difficulties in solving the analyzed problem, first of all, the paucity of statistical data characterizing the reliability of work, their multidimensional and random nature. The authors propose to solve this problem by moving from average annual reliability indicators to average monthly indicators of operational reliability. A brief description of the solution of individual tasks of this problem for overhead power lines is given, which together represent a new methodology for managing the technical condition of distributed type objects. Science-intensive, cumbersome and labor-intensive calculation algorithms determine the expediency of the transition to intelligent systems. At the same time, the management of the electric power system and its individual production enterprises will monthly receive specialized forms indicating recommendations that optimize the increase in the reliability of overhead power transmission lines by restoring wear and tear.

In recent past, Online Social Networks (OSN) has emerged as a platform for sharing information, thoughts, and activities. In the real-world network, method of considering the appropriate samples is most frequently used for network analysis. Graph sampling is a procedure used for computing unknown parameters. Many sampling algorithms exist in literature such as Random node, Random edge sampling, Rank degree, etc. can be used for estimation. This paper presents a comparison of clique based procedure (CBP) and shortest path based procedure (SPP) to estimate the average degree of a vertex in a social network using an overlapping cluster sampling. A comparative procedure is used to obtain the lower and upper limit of confidence intervals with the help of multiple samples. Ogive based simulation is also used for single value computation of limits of CI. The results, obtained from simulation, show that clique based sampling algorithm (CBP) is more efficient than the shortest path based sampling algorithm (SPP). The estimated confidence intervals can be used for monitoring the reliability of a social network in terms of control over average network degree.

The objective of this study is to develop an extension of the Burr-III distribution which is achieved by adopting the inverse Weibull-G family of distribution and is referred as inverse Weibull-Burr III distribution (IWB-III) to evaluate complicated data. Different structural characteristics of the suggested distribution have been determined and analysed. Distinct plots depict the behaviour of the probability density function (pdf) and the cumulative distribution function (cdf). The maximum likelihood estimation method is applied to estimate the stated distribution parameters. To assess and investigate the efficacy of estimators in terms of bias, variance, and mean square error (MSE), a simulation study was conducted. Lastly, the effectiveness of the stated distribution is proven by an actual data set relevant to survival rates in animals.

This paper reported the performance evaluation of Captive Power plant working in the fertilizer industry with possible production capacities. The idea of reduced capacity and load sharing to use the available system optimally is analyzed. The system works on two STG ' s (steam turbine generators) and one gridline. Gridline can bear the load of one or both STG ' s on failures. At the breakdown in gridline and STG, the system work at reduced capacity. Gridline repaired on a priority basis. The semi-Markov processses and regenerative point technique are used to evaluate the reliability and economic measures such as availability, busy period of repairman, and expected no. of repairs. The graphical study shows the relationships between these measures with the failure rates of STG and gridline.

A Discrete Parametric Markov-Chain Model of a Two Non-Identical Units Warm Standby Repairable System with Two Types of Failure

Pradeep Chaudhary, Anika Sharma, Rakesh Gupta
•
Department of Statistics
Ch. Charan Singh University, Meerut – 250004 (India)
E-mail: pc25jan@gmail.com; ash27sharma@gmail.com; smprgccsu@gmail.com

Abstract

The paper deals with the stochastic analysis of two non-identical units (unit-1 and unit-2). Initially, one unit is operative and other is kept into warm standby. Each unit of the system has two possible modes-Normal (N) and Total Failure (F). A single repairman is always available with the system to repair a failed unit. The operative unit is non-repairable, hence upon failure it goes for replacement. The system failure occurs when both the units are in total failure mode. Failure and repair times of a unit are taken as independent random variables of discrete nature having geometric distributions with different parameters.

Keywords: Transition probabilities, mean sojourn time, geometric distribution, regenerative point technique, reliability, MTSF, availability, expected busy period of repairman, net expected profit.

1. Introduction

The Two non-identical units warm standby system have been widely studied in the literature of reliability as they are frequently used in modern business and industries. It is obvious that the standby unit is switched to operate when the operating unit fails and the switching device which is used to put the standby unit into operation may be perfect at the time of need. Some authors including [5, 9, 10 and 13] analyzed two unit warm standby and two non-identical units warm standby redundant system models using different concepts. All the above system models have been analyzed by considering continuous distributions of all the random variables involved.

In many realistic situations, some writers [4 and 7] analyzed a two identical unit and two non-identical units cold standby system with two types of failure and later on [3] analyzed two non-identical units parallel system subject to two types of failure and correlated life times. Some authors [11, 12] analyzed the deferent concepts of assumptions. So in case of discrete random variable, discrete distribution is considered to be appropriate for obtaining the effectiveness of different reliability measures.

In the area of reliability using discrete distribution had given their ideas by analyzing two non-identical unit parallel system with geometric failure and repair time distribution. Since there is always a possibility for failure of any system during in its operative conditions in different measures. So to detect the type of failure inspection is very much required which had been always ignored by the researchers, whether using continuous or discrete distributions.

This system model is based on discrete parametric Markov-chain. Moreover, [1, 2, 6 and 8] introduced the concept of discrete parametric Markov-chain in analyzing the system models in the field of reliability modeling. The following economic related measures of system effectiveness are obtained by using regenerative point technique-

i. Transition probabilities and mean sojourn times in various states.
ii. Reliability and mean time to system failure.
iii. Point-wise and steady-state availability of the system during time (0, t-1).
iv. Expected busy period of repairman during time (0, t-1).
v. Net expected profit incurred by the system during a finite and steady-state are obtained.

2. System Description and Assumptions

1. The system comprises of two non-identical units. Initially, one unit is operative and other is kept into warm standby.
2. Each unit of the system has two modes- Normal (N) and total failure (F).
3. A single repairman is always available with the system to repair a failed unit.
4. The operative unit is non-repairable, hence upon failure it goes for replacement.
5. The system failure occurs when both the units are in total failure mode.
6. The repaired unit works as good as new.
7. Failure and repair times of the units follow independent geometric distributions with different parameters.

3. Notations and States of the System

3.1 Notations :

$p_i q_i^x$: p.m.f. of failure time of type-1 and type-2 respectively for i=1,2,3 and $p_i + q_i = 1$.

$r_i s_i^x$: p.m.f. of repair time by repairman of type-1 and type-2 respectively for i=1, 2 and $r_i + s_i = 1$.

$p' q'^x$: p.m.f. of repair time of first unit; $p' + q' = 1$.

$\theta, \bar{\theta}$: probability that the replacement of a second unit respectively; $\theta + \bar{\theta} = 1$

$q_{ij}(\cdot), Q_{ij}(\cdot)$: p.m.f. and C.d.f. of one step or direct transition time from state S_i to S_j.

p_{ij} : steady state transition probability from state S_i to S_j.

$$p_{ij} = Q_{ij}(\infty)$$

$Z_i(t)$: probability that the system sojourn in state S_i up to epoch (t-1).

ψ_i : Mean sojourn time in state S_i.

$*, h$: symbol and dummy variable used in geometric transform e. g.

$$GT\left[q_{ij}(t)\right] = q_{ij}^*(h) = \sum_{t=0}^{\infty} h^t q_{ij}(t)$$

3.2 Symbols for the states of the system

N_o^i : unit-i is in normal mode(N) and operative; i=1,2

N_{ws}^i : unit-i is in normal mode(N) and warm standby.; i=1,2

F_R^2 / F_{wR}^2 : unit-2 is in total failure (F) mode and under replacement/waits for replacement.

F_r^1 / F_{2r}^1 : unit-1 is in total failure (F) mode and under repair.

The transition diagram of the system model is shown in Figure. 1.

TRANSITION DIAGRAM

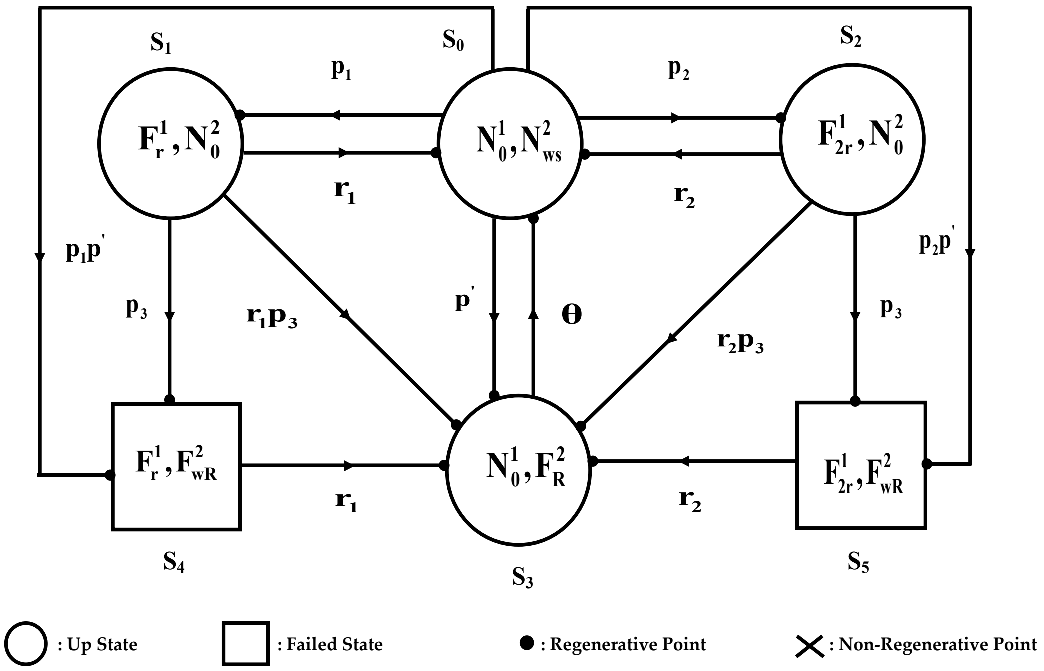


Figure.1

With the help of above symbols the possible states of the system are:

$S_0 \equiv \left(N_o^1, N_{ws}^2\right)$, $\quad S_1 \equiv \left(F_r^1, N_o^2\right)$, $\quad S_2 \equiv \left(F_{2r}^1, N_o^2\right)$

$S_3 \equiv \left(N_o^1, F_R^2\right)$, $\quad S_4 \equiv \left(F_r^1, F_{wR}^2\right)$, $\quad S_5 \equiv \left(F_{2r}^1, F_{wR}^2\right)$

The states S_0, S_1, S_2, S_3 are up states; S_4, S_5 are failed states.

4. Transition Probabilities and Sojourn Times

Let $Q_{ij}(t)$ be the probability that the system transits from state S_i to S_j during time interval (0, t) i.e., if T_{ij} is the transition time from state S_i to S_j then

$$Q_{ij}(t) = P\left\lfloor T_{ij} \le t \right\rfloor$$

By using simple probabilistic arguments we have,

$$Q_{01}(t) = \frac{p_1 q' q_2}{1 - q' q_1 q_2}\left[1 - \left(q' q_1 q_2\right)^{t+1}\right], \qquad Q_{02}(t) = \frac{p_2 q' q_1}{1 - q' q_1 q_2}\left[1 - \left(q' q_1 q_2\right)^{t+1}\right]$$

$$Q_{03}(t) = \frac{p' q_1 q_2}{1 - q' q_1 q_2}\left[1 - \left(q' q_1 q_2\right)^{t+1}\right], \qquad Q_{04}(t) = \frac{p' p_1 q_2}{1 - q' q_1 q_2}\left[1 - \left(q' q_1 q_2\right)^{t+1}\right]$$

$$Q_{05}(t) = \frac{p' p_2 q_1}{1 - q' q_1 q_2}\left[1 - \left(q' q_1 q_2\right)^{t+1}\right], \qquad Q_{10}(t) = \frac{r_1 q_3}{1 - s_1 q_3}\left[1 - \left(s_1 q_3\right)^{t+1}\right]$$

$$Q_{13}(t) = \frac{r_1 p_3}{1 - s_1 q_3}\left[1 - \left(s_1 q_3\right)^{t+1}\right], \qquad Q_{14}(t) = \frac{s_1 p_3}{1 - s_1 q_3}\left[1 - \left(s_1 q_3\right)^{t+1}\right]$$

$$Q_{20}(t)=\frac{r_2q_3}{1-s_2q_3}\left[1-(s_2q_3)^{t+1}\right], \qquad Q_{23}(t)=\frac{r_2p_3}{1-s_2q_3}\left[1-(s_2q_3)^{t+1}\right]$$

$$Q_{25}(t)=\frac{s_2p_3}{1-s_2q_3}\left[1-(s_2q_3)^{t+1}\right], \qquad Q_{30}(t)=\left[1-(\bar{\theta})^{t+1}\right]$$

$$Q_{43}(t)=\left[1-(s_1)^{t+1}\right], \qquad Q_{53}(t)=\left[1-(s_2)^{t+1}\right] \tag{1-14}$$

The steady state transition probabilities from state S_i to S_j can be obtained from (1-14) by taking t $\to\infty$, as follows:

$$p_{01}=\frac{p_1q^{'}q_2}{1-q^{'}q_1q_2}, \quad p_{02}=\frac{p_2q^{'}q_1}{1-q^{'}q_1q_2}, \quad p_{03}=\frac{p^{'}q_1q_2}{1-q^{'}q_1q_2}, \quad p_{04}=\frac{p^{'}p_1q_2}{1-q^{'}q_1q_2}$$

$$p_{05}=\frac{p'p_2q_1}{1-q^{'}q_1q_2}, \quad p_{10}=\frac{r_1q_3}{1-s_1q_3}, \quad p_{13}=\frac{r_1p_3}{1-s_1q_3}, \quad p_{14}=\frac{s_1p_3}{1-s_1q_3}$$

$$p_{20}=\frac{r_2q_3}{1-s_2q_3}, \quad p_{23}=\frac{r_2p_3}{1-s_2q_3}, \quad p_{25}=\frac{s_2p_3}{1-s_2q_3}$$

We observe that the following relations hold-

$$p_{30}=p_{43}=p_{53}=1, \qquad p_{01}+p_{02}+p_{03}+p_{04}+p_{05}=1$$

$$p_{10}+p_{13}+p_{14}=1, \qquad p_{20}+p_{23}+p_{25}=1, \tag{15-18}$$

5. Mean Sojourn Time

Let T_i be the sojourn time in state S_i (i=0-5) then ψ_i mean sojourn time in state S_i is given by

$$\psi_i=E(T_i)=\sum_{t=1}^{\infty}P\left[T_i\geq t-1\right]$$

In particular,

$$\psi_0=\frac{q^{'}q_1q_2}{1-q^{'}q_1q_2}, \quad \psi_1=\frac{s_1q_3}{1-s_1q_3}, \quad \psi_2=\frac{s_2q_3}{1-s_2q_3}$$

$$\psi_3=\frac{\bar{\theta}}{\theta}, \quad \psi_4=\frac{s_1}{r_1}, \quad \psi_5=\frac{s_2}{r_2} \tag{19-24}$$

6. Methodology for Developing Equations

In order to obtain various interesting measures of system effectiveness we developed the recurrence relations for reliability, availability and busy period of repairman as follows-

6.1 Reliability of the system

Here we define $R_i(t)$ as the probability that the system does not fail up to epochs 0, 1, 2,.., (t-1) when it is initially started from up state S_i. To determine it, we regard the failed states S_4,S_5 as absorbing state. Now, the expression for $R_i(t)$; i=0, 1, 2, 3; we have the following set of convolution equations.

$$\begin{aligned}R_0(t)&=q^{'t}q_1^tq_2^t+\sum_{u=0}^{t-1}q_{01}(u)R_1(t-1-u)+\sum_{u=0}^{t-1}q_{02}(u)R_2(t-1-u)\\&=Z_0(t)+q_{01}(t-1)©R_1(t-1)+q_{02}(t-1)©R_2(t-1)+q_{03}(t-1)©R_3(t-1)\end{aligned}$$

Similarly,

$$R_1(t)=Z_1(t)+q_{10}(t-1)©R_0(t-1)+q_{13}(t-1)©R_3(t-1)$$
$$R_2(t)=Z_2(t)+q_{20}(t-1)©R_0(t-1)+q_{23}(t-1)©R_3(t-1)$$
$$R_3(t)=Z_3(t)+q_{30}(t-1)©R_0(t-1) \qquad (25\text{-}28)$$

Where,

$$Z_1(t)=s_1^t q_3^t, \qquad Z_2(t)=s_2^t q_3^t, \qquad Z_3(t)=q'^t$$

6.2 Availability of the System

Let $A_i(t)$ be the probability that the system is up at epoch (t-1), when it initially started from state S_i. Then, by using simple probabilistic arguments, as in case of reliability the following recurrence relations can be easily developed for $A_i(t)$; i=0 to 5.

$$A_0(t)=q'^t q_1^t q_2^t+\sum_{u=0}^{t-1}q_{01}(u)©A_1(t-1-u)+\sum_{u=0}^{t-1}q_{02}(t-1)©A_2(t-1)+\sum_{u=0}^{t-1}q_{03}(t-1)©A_3(t-1)$$
$$+\sum_{u=0}^{t-1}q_{04}(t-1)©A_4(t-1)+\sum_{u=0}^{t-1}q_{05}(t-1)©A_5(t-1)$$
$$=Z_0(t)+q_{01}(t-1)©A_1(t-1)+q_{02}(t-1)©A_2(t-1)+q_{03}(t-1)©A_3(t-1)$$
$$+q_{04}(t-1)©A_4(t-1)+q_{05}(t-1)©A_5(t-1)$$

Similarly,

$$A_1(t)=Z_1(t)+q_{10}(t-1)©A_0(t-1)+q_{13}(t-1)©A_3(t-1)+q_{14}(t-1)©A_4(t-1)$$
$$A_2(t)=Z_2(t)+q_{20}(t-1)©A_0(t-1)+q_{23}(t-1)©A_3(t-1)+q_{25}(t-1)©A_5(t-1)$$
$$A_3(t)=Z_3(t)+q_{30}(t-1)©A_0(t-1)$$
$$A_4(t)=q_{43}(t-1)©A_3(t-1)$$
$$A_5(t)=q_{53}(t-1)©A_3(t-1) \qquad (29\text{-}34)$$

Where the values of $Z_i(t)$; i=0 to 3 are same as given in section 6.1.

6.3 Busy Period of Repairman

Let $B_i^r(t)$ and $B_i^R(t)$ be the probability that the repairman is busy in the repair and replacement of a failed unit at epoch t-1, when it initially started from state S_i. Then, by using simple probabilistic arguments, as in case of reliability the following recurrence relations can be easily developed for $B_i^r(t)$ and $B_i^R(t)$; i=0 to 5.

$$B_0^r(t)=q_{01}(t-1)©B_1^r(t-1)+q_{02}(t-1)©B_2^r(t-1)+q_{03}(t-1)©B_3^r(t-1)+q_{04}(t-1)©B_4^r(t-1)$$
$$+q_{05}(t-1)©B_5^r(t-1)$$
$$B_1^r(t)=Z_1(t)+q_{10}(t-1)©B_0^r(t-1)+q_{13}(t-1)©B_3^r(t-1)+q_{14}(t-1)©B_4^r(t-1)$$
$$B_2^r(t)=Z_2(t)+q_{20}(t-1)©B_0^r(t-1)+q_{23}(t-1)©B_3^r(t-1)+q_{25}(t-1)©B_5^r(t-1)$$
$$B_3^r(t)=q_{30}(t-1)©B_0^r(t-1)$$
$$B_4^r(t)=Z_4(t)+q_{43}(t-1)©B_3^r(t-1)$$
$$B_5^r(t)=Z_5(t)+q_{53}(t-1)©B_3^r(t-1) \qquad (35\text{-}40)$$

Where,

The values of $Z_1(t)$ and $Z_2(t)$ are same as given in section 6.1, $Z_4(t)=s_1^t$ and $Z_5(t)=s_2^t$.

$$\begin{aligned}
B_0^R(t) &= q_{01}(t-1)©B_1^R(t-1)+q_{02}(t-1)©B_2^R(t-1)+q_{03}(t-1)©B_3^R(t-1)+q_{04}(t-1)©B_4^R(t-1)\\
&\quad +q_{05}(t-1)©B_5^R(t-1)\\
B_1^R(t) &= q_{10}(t-1)©B_0^R(t-1)+q_{13}(t-1)©B_3^R(t-1)+q_{14}(t-1)©B_4^R(t-1)\\
B_2^R(t) &= q_{20}(t-1)©B_0^R(t-1)+q_{23}(t-1)©B_3^R(t-1)+q_{25}(t-1)©B_5^R(t-1)\\
B_3^R(t) &= Z_3(t)+q_{30}(t-1)©B_0^R(t-1)\\
B_4^R(t) &= q_{43}(t-1)©B_3^R(t-1)\\
B_5^R(t) &= q_{53}(t-1)©B_3^R(t-1)
\end{aligned} \tag{41-46}$$

Where,

The value of $Z_3(t)$ is same as given in section 6.1.

7. Analysis of Reliability and MTSF

Taking geometric transform of (25-28) and simplifying the resulting set of algebraic equations for $R_0^*(h)$ we get

$$R_0^*(h)=\frac{N_1(h)}{D_1(h)} \tag{47}$$

Where,

$$N_1(h)=Z_0^*+hq_{01}^*Z_1^*+hq_{02}^*Z_2^*+\left\lfloor h^2q_{01}^*q_{13}^*+h^2q_{02}^*q_{23}^*+hq_{03}^*\right\rfloor Z_3^*$$

$$D_1(h)=1-h^2q_{01}^*q_{10}^*-h^2q_{02}^*q_{20}^*-h^2q_{03}^*q_{30}^*-h^3q_{01}^*q_{13}^*q_{30}^*-h^3q_{02}^*q_{23}^*q_{30}^*$$

Collecting the coefficient of h^t from expression (47), we can get the reliability of the system $R_0(t)$. The MTSF is given by-

$$E(T)=\lim_{h\to 1}\sum_{t=1}^{\infty}h^tR(t)=\frac{N_1(1)}{D_1(1)}-1 \tag{48}$$

Where,

$$N_1(1)=\psi_0+p_{01}\psi_1+p_{02}\psi_2+[p_{01}p_{13}+p_{02}p_{23}+p_{03}]\psi_3$$

$$D_1(1)=1-p_{01}p_{10}-p_{02}p_{20}-p_{03}-p_{01}p_{13}-p_{02}p_{23}$$

8. Availability Analysis

On taking geometric transform of (29-34) and simplifying the resulting equations for we get,

$$A_0^*(h)=\frac{N_2(h)}{D_2(h)} \tag{49}$$

Where,

$$N_2(h)=\begin{vmatrix}
Z_0^* & -hq_{01}^* & -hq_{02}^* & -hq_{03}^* & -hq_{04}^* & -hq_{05}^*\\
Z_1^* & 1 & 0 & -hq_{13}^* & -hq_{14}^* & 0\\
Z_2^* & 0 & 1 & -hq_{23}^* & 0 & -hq_{25}^*\\
Z_3^* & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & -hq_{43}^* & 1 & 0\\
0 & 0 & 0 & -hq_{53}^* & 0 & 1
\end{vmatrix}$$

and

$$D_2(h)=\begin{vmatrix} 1 & -hq_{01}^* & -hq_{02}^* & -hq_{03}^* & -hq_{04}^* & -hq_{05}^* \\ -hq_{10}^* & 1 & 0 & -hq_{13}^* & -hq_{14}^* & 0 \\ -hq_{20}^* & 0 & 1 & -hq_{23}^* & 0 & -hq_{25}^* \\ -hq_{30}^* & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -hq_{43}^* & 1 & 0 \\ 0 & 0 & 0 & -hq_{53}^* & 0 & 1 \end{vmatrix}$$

The steady state availabilities of the system due to operation of unit –

$$A_0 = \lim_{t\to\infty} A_0(t) = \lim_{h\to 1}(1-h)\frac{N_2(h)}{D_2(h)}$$

But $D_2(h)$ at h=1 is zero, therefore by applying L. Hospital rule, we get

$$A_0 = -\frac{N_2(1)}{D_2'(1)} \tag{50}$$

Where,

$$N_2(1) = \psi_0 + p_{01}\psi_1 + p_{02}\psi_2 + \left[p_{01}p_{13} + p_{01}p_{14} + p_{02}p_{23} + p_{02}p_{25} + p_{03} + p_{04} + p_{05}\right]\psi_3$$

and

$$D_2'(1) = \psi_0 + p_{01}\psi_1 + p_{02}\psi_2$$

Now the expected uptime of the system due to operative unit upto epoch (t-1) are given by

$$\mu_{up}(t) = \sum_{x=0}^{t-1} A_0(x)$$

So that

$$\mu_{up}^*(h) = \frac{A_0^*(h)}{(1-h)} \tag{51}$$

9. Busy Period Analysis

On taking geometric transforms of (35-40) and (41-46), simplifying the resulting equations, we get

$$B_0^{r*}(h) = \frac{N_3(h)}{D_2(h)} \quad \text{and} \quad B_0^{R*}(h) = \frac{N_4(h)}{D_2(h)} \tag{52-53}$$

Where,

$$N_3(h) = Z_1^* hq_{01}^* + Z_2^* hq_{02}^* + Z_4^* h^2 q_{01}^* q_{14}^* + Z_4^* hq_{04}^* + Z_5^* h^2 q_{02}^* q_{25}^* + Z_5^* hq_{05}^*$$

and

$$N_4(h) = Z_3^* \left\lfloor h^2 q_{01}^* q_{13}^* + h^3 q_{01}^* q_{14}^* q_{43}^* + h^2 q_{02}^* q_{23}^* + h^3 q_{02}^* q_{25}^* q_{53}^* + hq_{03}^* + h^2 q_{04}^* q_{43}^* + h^2 q_{05}^* q_{53}^* \right\rfloor$$

and $D_2(h)$ is same as in availability analysis.

In the long run the respective probabilities that the repairman is busy in the repair and replacement of a failed unit are given by-

$$B_0^r = \lim_{t\to\infty} B_o^r(t) = \lim_{h\to 1}(1-h)\frac{N_3(h)}{D_2(h)} \quad \text{and} \quad B_0^R = \lim_{t\to\infty} B_o^R(t) = \lim_{h\to 1}(1-h)\frac{N_4(h)}{D_2(h)}$$

But $D_2(h)$ at h=1 is zero, therefore by applying L. Hospital rule, we get

$$B_0^r = -\frac{N_3(1)}{D_2'(1)} \quad \text{and} \quad B_0^R = -\frac{N_4(1)}{D_2'(1)} \tag{54-55}$$

Where,

$$N_3(1) = p_{01}\psi_1 + p_{02}\psi_2 + (p_{01}p_{14} + p_{04})\psi_4 + (p_{02}p_{25} + p_{05})\psi_5$$

and

$$N_4(1) = \psi_3\left[p_{01}p_{13} + p_{01}p_{14} + p_{02}p_{23} + p_{02}p_{25} + p_{03} + p_{04} + p_{05}\right]$$

and $D_2'(1)$ is same as in availability analysis.

Now the expected busy period of the repairman in repair of a failed unit up to epoch (t-1) are respectively given by-

$$\mu_b^r(t) = \sum_{x=0}^{t-1} B_0^r(x) \qquad \text{and} \qquad \mu_b^R(t) = \sum_{x=0}^{t-1} B_0^R(x) \tag{56-57}$$

10. Profit Function Analysis

We are now in the position to obtain the net expected profit incurred up to epoch (t-1) by considering the characteristics obtained in earlier section. Let us consider,

$K_0 =$ revenue per-unit time by the system due to operative unit.

$K_1 =$ cost per-unit time when repairman is busy in the repair of failed unit.

$K_2 =$ cost per-unit time when repairman is busy in the replacement of a failed unit.

Then, the net expected profit incurred up to epoch (t-1) is given by,

$$P(t) = K_0\mu_{up}(t) - K_1\mu_b^r(t) - K_2\mu_b^R(t) \tag{58}$$

The expected profit per unit time in steady state is given by-

$$\begin{aligned} P &= \lim_{t\to\infty}\frac{P(t)}{t} = \lim_{h\to 1}(1-h)^2 P^*(h) \\ &= K_0 \lim_{h\to 1}(1-h)^2\frac{A_0^*(h)}{(1-h)} - K_1\lim_{h\to 1}(1-h)^2\frac{B_0^{r*}(h)}{(1-h)} - K_2\lim_{h\to 1}(1-h)^2\frac{B_0^{R*}(h)}{(1-h)} \\ &= K_0A_0 - K_1B_0^r - K_2B_0^R \end{aligned} \tag{59}$$

11. Graphical Representation

The curves for MTSF and profit function have been drown for different values of failure parameters. Fig.2 depicts the variation in MTSF with respect to failure rate (p_1) for different values of repair rate (p_2) of a unit and constant repair rate (p') when values of other parameters are kept fixed as $p_3 = 0.001$, $r_1 = 0.5$, $r_2 = 0.7$ and $\theta = 0.01$. From the curves we conclude that expected life of the system decrease with increase in p_1. Further, increases as the values of p_2 and p' increases.

Similarly, Fig.3 reveals the variations in profit (P) with respect to p_1 for varying values of p_2 and p', when other parameters are kept fixed as $p_3 = 0.01$, $r_1 = 0.92$, $r_2 = 0.99$ and $\theta = 0.01$, $K_0 = 100$, $K_1 = 100$, $K_2 = 400$ and $K_3 = 300$. From the figure it is clearly observed from the smooth curves, that the system is profitable if the value of parameter p_1 is greater than 0.2, 0.33 and 0.5 respectively for $p_2 = 0.4$, 0.6 and 0.8 for fixed value of $p' = 0.15$. From dotted curves, we conclude that system is profitable only if value of parameter p_1 is greater than 0.27, 0.39 and 0.6 respectively for $p_2 = 0.4$, 0.6 and 0.8 for fixed value of $p' = 0.3$.

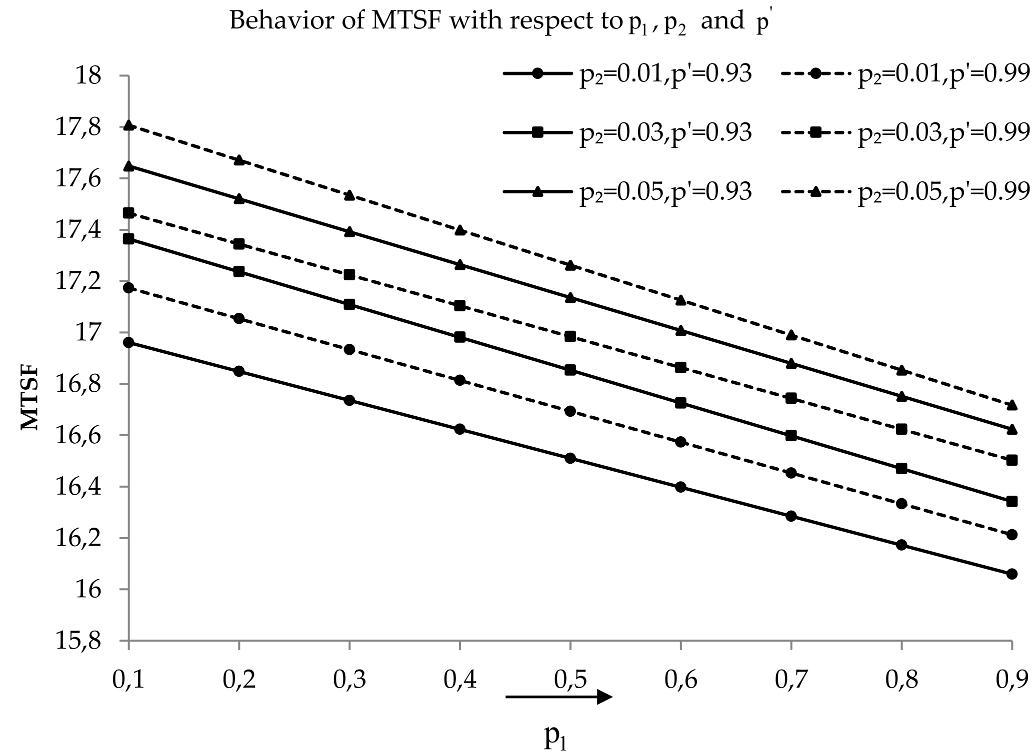


Figure.2

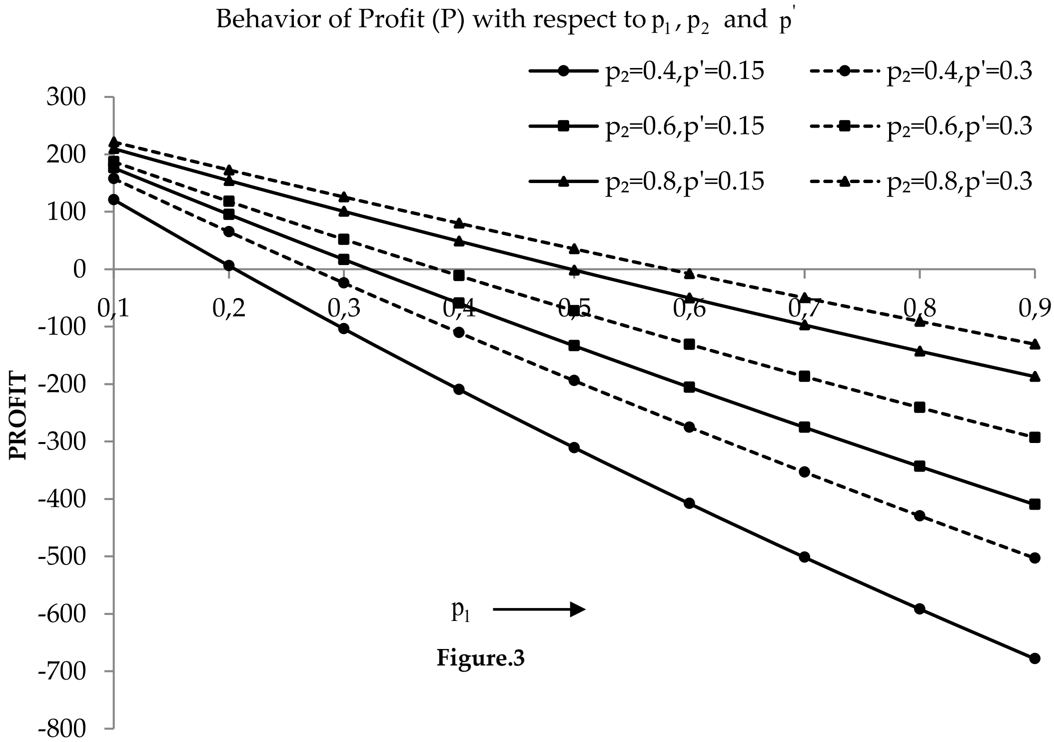


Figure.3

12. Conclusions

1. It is indicated in fig.2 that we can easily obtain the upper limit of "p_1" to achieve at least a particular value of MTSF. As an illustration to get at least MTSF 16.8 unit, the failure rate "p_1" must be less than 0.24, 0.56 and 0.79 respectively for repair rate p_2 = 0.01, 0.03 and 0.05 when

activation rate is kept fixed as $p' = 0.93$. Similarly, when $p' = 0.99$ is kept fixed as "p_1" must be less than 0.43, 0.68 and 0.87 corresponding to $p_2 = 0.01$, 0.03 and 0.05.

2. In fig. 3 it is reveled from the smooth curves, that the system is profitable if the value of parameter p_1 is greater than 0.2, 0.33 and 0.5 respectively for $p_2 = 0.4$, 0.6 and 0.8 for fixed value of $p' = 0.15$. From dotted curves, we conclude that system is profitable only if value of parameter p_1 is greater than 0.27, 0.39 and 0.6 respectively for $p_2 = 0.4$, 0.6 and 0.8 for fixed value of $p' = 0.3$.

References

[1] Bhatti, J., Chitkara, A.K. and Kakkar, M. (2013). Profit analysis of non-identical parallel system with two types of failure using discrete distribution. *Mathematical Journal of Interdisciplinary Sciences*, 1(2):27-40.

[2] Bhatti, J., Chitkara, A.K. and Bhardwaj, N. (2011). Profit analysis of two unit cold standby system with two types of failure under inspection policy and discrete distribution. *International Journal of Scientific and Engineering Research*, 2(12):1-6.

[3] Chaudhary, P. and Tyagi, L. (2021). A two non-identical unit parallel system subject to two types of failure and correlated life times. *Reliability: Theory and Applications*, 16(2):247-258.

[4] Chaudhary, P. and Tomar, R. (2019). A two identical unit cold standby system subject to two types of failure. *Reliability: Theory and Applications*, 14(1):34-43.

[5] Gupta, R., Jaiswal, S. and Chaudhary, A. (2015). Cost benefit analysis of a two unit warm standby system with correlated working and rest time of repairman. *International Journal of Statistics and Reliability Engineering*, 2(1):1-17.

[6] Gupta, R. and Varshney, G. (2006). A two non-identical unit parallel system with geometric failure and repair time distribution. *IAPQR*, 31(2):127-139.

[7] Gupta, P. and Vinodiya, P. (2018). Analysis of reliability of a two non-identical units cold standby repairable system has two types of failure. *International Journal of Computer Sciences and Engineering*. 6(11):907-913.

[8] Gupta, S., Chaudhary, P. and Vaishali (2020). A three unit warm and cold standby system model of discrete parametric markov chain. *Reliability: Theory and Applications*, 15(4):117-127.

[9] Kumar, A., Malik, S.C. and Pawar, D. (2019). Profit analysis of a warm standby non-identical units system with single server subject to preventive maintenance. *International Journal of Agricultural and Statistics Science*, 15(1):261-269.

[10] Kumar, A., Malik, S.C. and Pawar, D. (2018). Profit analysis of a warm standby non-identical units system with single server subject to priority. *International Journal of Future Revolution*, 4(10):108-112.

[11] Kozyrev, D., Kolev, N. and Rykov, V. (2018). Reliability function of renewable system under marshall-olkin failure model. *Reliability: Theory and Applications*, 13 No.1(48):39-46.

[12] Rykov, V., Efrosinin, D., Stepanova, N. and Sztrik, J. (2020). On reliability of a double redundant renewable system with a generally distributed life and repair times. *Mathematics*, 8(2):278.

[13] Wang, J., Xie, N. and Yang, N. (2021). Reliability of a two dissimilar-unit warm standby repairable system with priority in use. *Communications in Statistics-Theory and Methods*, 50(4):792-814.

Random processes imitation in fatigue studies

Irina Gadolina[1], Alexei Erpalov[2], and Nelly Dinyaeva[3]

•

[1]IMASH RAS, Moscow, Russia, gadolina@mail.ru
[2] South Ural State University, Chelyabinsk, Russia, aerpalov@ya.ru
[3] Moscow Aviation Institute (National Research University), Moscow, Russia, n.dinyaeva@yandex.ru

Abstract

Modelling random processes traditionally supposes working with the spectral density. Although many engineering problems require the knowledge of spectral density, the specific character of fatigue damage accumulation dictates the different approach – namely, the consideration of the distribution of random values of the local extremes, which is responsible for fatigue damage accumulation. There is a need in developing the methods of random loading imitation in the experimental and numerical study of fatigue. According to the up-to-date situation in science in fatigue, both opposing approaches should be considered - the time domain and the frequency domain. The proposed method, which consists of two stages, meets that requirement. The performed case study based on laboratory fatigue testing confirms its applicability.

Keywords: metal fatigue, random loading, imitation,

1. Introduction

The quality of machines and equipment depends on the advanced quality management. The repair plan is also important [1]. The reliability of the industrial production should be guaranteed by reliable testing and design methods for estimation durability. Objective hazards that threaten the performance and durability of machine parts are the processes of degradation of their elements. It is necessary to consider the fatigue process caused by natural factors, namely, alternating loading. Therefore, the engineers need a tool for loads assessing, taking into account their random nature.

In fatigue studies under random loading, two main approaches are widely used [2]. They correspond to the time domain and the frequency domain. Both have their own areas of applications. Investigations in the frequency domain [3] are important while considering problems with studying the impacts of the particular frequencies (resonant effect), modal analysis. They are also important while treating the enormous data storage while using the method of Critical Plane Approach [4]. On the other hand, the well-proved fatigue accumulation problems mostly based on the information about cycles, their extremum values, their order of appearance. It is worth mentioning that the problem of the fatigue crack propagation almost ultimately based on information about extremums (not the frequencies) [5]. Applicable to automobile parts, the study [6belec] consider the random loading strictly in a time domain. A discussion goes on [7,8], which of the two methods is correct: the time-domain methods (mostly, Rainflow [9 endo]) or the frequency domain [3]. It goes out that choosing the particular domain for investigation dictates some additional requirements [10]. Registration of the random loading process is different [11]. It follows from the fact that treating in the time domain supposes the higher precision of peaks registration.

RT&A, No 2 (68)
Volume 17, June 2022

It is important to have a tool to estimate the spectral density of the processes, which are given only by a sequence of extrema. The example of such task is in paper [10]. The aim of this paper to provide such opportunity for the researchers.

2. Methods

This problem partly was first studied in [12]. During testing and numerical modelling in fatigue studies, the investigators should introduce the randomness in one or another way because it exists in service. The conditions of exploitation vary. Some factors are hard to control, etc. The proposed method consists of two steps, namely:

2.1. Random sequence of extremums generation

As mentioned in the Introduction, the most direct path to fatigue estimation is the Rainflow [2, 9], which operates with the extremums' sequence, namely local maximums and minimums of the random process. It would not be justified to repeat this sequence without change in fatigue studies because it would not have reflected the randomness in service and takes a lot of time. The good decision was proposed earlier in [13]. It introduced the so-called Markov's matrixes to create the variability.

Later, the method of target Markov's matrixes was developed, intended to reflect the service conditions of a particular object of investigation. The main idea of filling up the matrix is schematically presented in Fig.1. One by one, the half-cycles (ranges) are entered into the Markov's matrix (Fig.1, b). All important information is being presented in this way, namely, the maximum amplitude in realization, distributions of the values of the half-cycles, the number of their repetition during the period of investigation. The information concerning the sequence of the events is being lost during this procedure. The investigation of the impact of the sequence effect on the crack-propagation stage was reported in [14].

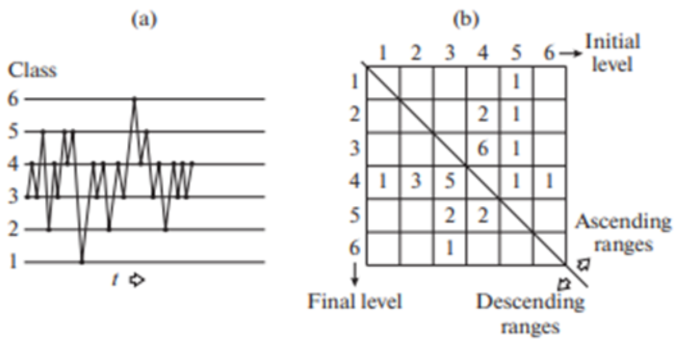


Figure 1: *Filling up the target matrix*

Next, the numerical modelling basing on the filled up earlier matrix is performed. The random number generator is used here. In this way, the engineers get the tool for the experimental or numerical study of fatigue by taking into account the random character of the loading, which is intrinsic to loading in service. As the result of this modelling, the investigators get the sequence of extremums. This sequence is sufficient for estimating the longevity of the objects using the time domain loading approach like Rainflow, [2,9].

2.2. Construction of the continuous function

Unfortunately, it is impossible to estimate the spectral density only the sequence of extremums generated by described above method. The function does not possess the property of smoothness; its firsts derivate is not continuous. It is worth mentioning that there is a need to estimate spectral density [10] of the modelled processes, particularly for the application of spectral methods (frequency domain) for longevity estimation. Although those methods are at some extend doubtful, they are still widespread [3]. As we mentioned above, they also have their own field for application: like des [4].

To overcome this problem, a method for introducing smoothness into the process was developed [12]. The adjacent peaks of the sequence are proposed to connect by half-cosines at the period (0,π). At the next step these half-cosines parts are concatenated.

The equations of half-cosines:

$$x(t) = A \cos (wt+\varphi) \tag{1}$$

where x(t)– is the part of the continuous extrapolating function, which is defined on the domain t=0… π/w, because the period of the cosine function is T=2π/w, s.

For each half-wave (1) starting from the successive extremum *MAX* or *MIN*, the parameters *A, w*, and φ are unique. The stress amplitude *A* [MPa] is defined as half of the range (modulus) of successive extremes:

$$A=mod\,(MAX_i - MIN_{i+1})/2 \;\vee\; A=mod\,(MAX_i - MIN_i)/2 \tag{2}$$

The obtained in this way sequence of the random reading forms the continuous random process with continuous first derivate. The concordances and peculiarities of the modelled processes will be analyzed later in the Case study. The main idea of this modelling – that is the values of the turning points, and their sequence remains unchanged. This point is paramount for fatigue estimation not only on the stage of crack initiation in fatigue but also during the crack propagation stage [5].

3. Case study

Following an engineering problem, the task was initially formulated to investigate the fatigue resistance of the metal specimens under the impact of the random process with the particular spectral density, shown in Fig. 4 [15]:

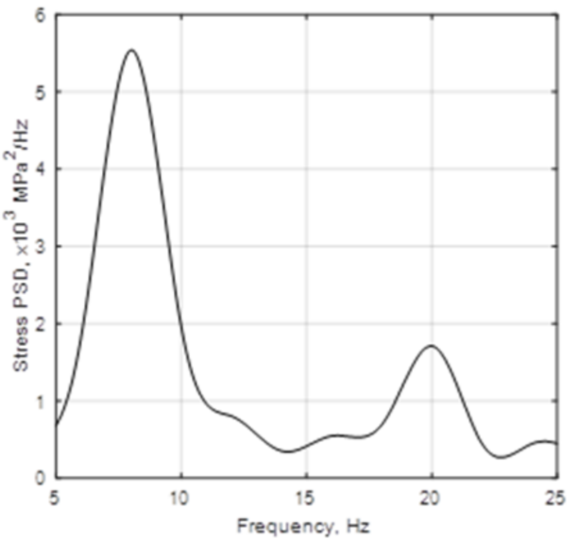


Figure 2: *Target spectral density for testing*

In Fig.3 the testing equipment is shown. Six Al specimens were tested simultaneously.

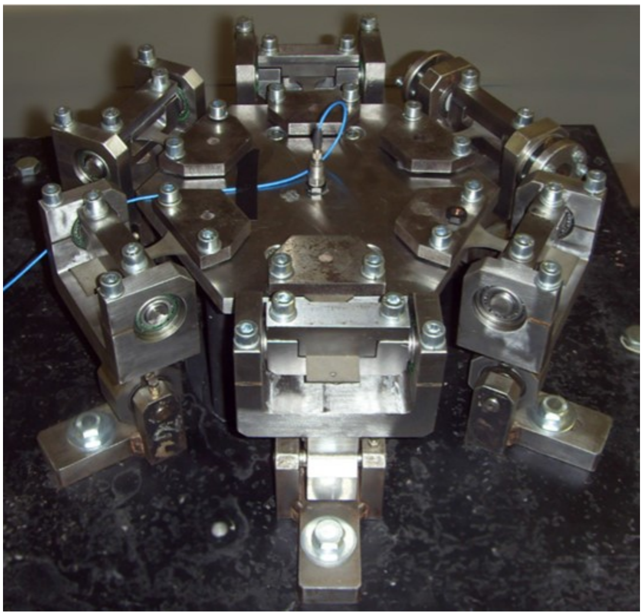


Figure 3. Testing equipment for regular and random cantilever loading

Fatigue experiment [15] was performed on 4 levels of loading to build the so-called Gassner curve (see also [7]). The example of the loading history is shown in Fig.5.

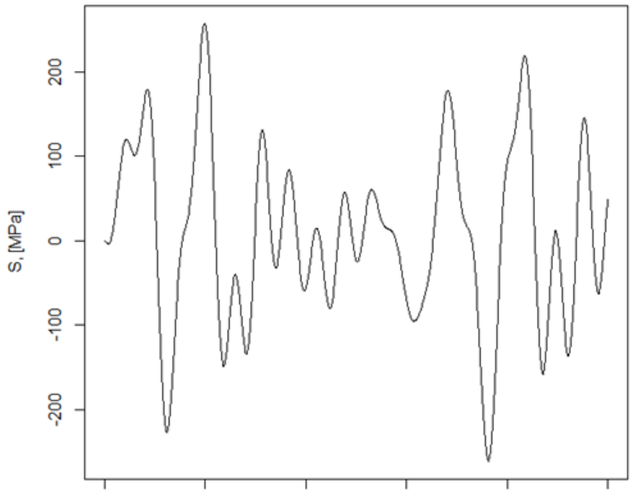


Figure 4. *1-second recording of the stress at the level of root mean square RMS=108 MPa.*

The main characteristics of the imitated random processes were obtained numerically and are shown in Table 1. With the aim of the study, the modelling was based on the laboratory records. For the sake of representativity several modelled trials on the base of laboratory realization were executed. In Table 1 the mean stress of the block is shown, RMS (the root mean square of realization), *I* – is irregularity factor [2] (*I*=*No*/*Ne*, *I*<1; where *No* is the number of crossings of the middle level line and *Ne* is the number of extremums. It is worth mentioning, that value *I* depends on the level number, in other words, on the registration precision. Spectra fullness *V* is defined as follows:

$$V = \sqrt[m]{\frac{1}{n}\sum h_i(\frac{\sigma_{ai}}{\sigma_{amax}})^{\wedge m}} \tag{3}$$

The value of V < 1 and is dimensionless. In formula (3) m is the slope coefficient of the fatigue curve; n is the total number of cycles in the block; h_i is the number of cycles at the i-th step; σ_{ai} is the current value of the stress amplitude; σ_{amax} is the maximum amplitude in the block. As can be seen from the formula, the fullness ratio of the spectrum *V* depends not only on the spectra form, but also on fatigue exponent m.

Table 1: *The main characteristics of random realizations: initial and modelled ones*

	Mean value [MPa]	RMS, [MPa]	*I*	*V*
Experimental realization	0.0244	108	0.67	0.56
1-st modelled realization	-0.0123	107	0.66	0.53
2-nd	-0.0056	99	0.68	0.55
…				
10-th	0.0342	104	0.66	0.57

Unlike the widespread practice of ignoring the time factor during the cycle counting, in this study, following the aim of the investigation, not only the peaks were selected, but also the half-periods of quasi-cycles.

According to the proposed method, each stress range: *r* [MPa] is associated with the following half period: *hT* [s]. The scatterplot of two random variables is shown in Fig. 6. The estimated correlation was cor=0.76. Also, the regression equation by the least square method was estimated:

$$r = -161+10.77\,hT \tag{4}$$

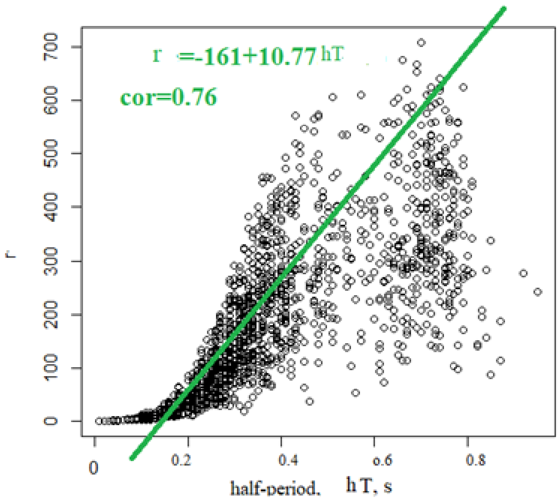


Figure 5. *The Scatterplot of half-periods hT and ranges r (with a regression line).*

The regression equation (3) was used for modelling the cosine curves in extrapolation. Unlike equation (1), in which the frequency is assumed to be the same for all half-cycles, a variable frequency value is used to form a more realistic process. To determine the frequency, the regression equation (3) was used. Equation (4) was used during modelling.

4. Results and discussion

After performing the tests with sufficient repetition of specimens at each loading level, the Gassner curve has been built [7,15]. This curve is analogous to the fatigue curve, but instead of the constant amplitudes, a maximum value in the block is shown. The realization of the random process, modelled on the base of spectral density (Fig. 4), was recorded, and the Rainflow cycle counting procedure [9] was performed.

The main statistical characteristics of a few modelled realizations were shown in Table 1. The spectral density of the modelled realization #1 is slightly differ from target spectral density. Anyway, the first and second frequencies coincide.

A much better coincidence is shown for Rainflow distributions. Several replicas were compared among each other as well with the distribution of initial realization. The good coincidence of parameters is also evident from Table 1.

5. CONCLUSIONS

Machine parts would be deteriorated due to fatigue produced by repeated loading, so the engineers should consider the fatigue impact.

Those impacts are of random nature, so the special imitation method was developed. This method works in time, as well in the frequency domain.

The method was approbated in the Case study based during the laboratory fatigue experiment under random loading.

The better coincidence takes place in the time domain. To improve modelling in a way it works equally well in both domain, addition experiment and researches are needed.

References

1. Tijjani A. Waziri, Ibrahim Yusuf. On age replacement policy of a system involving minimal repair/RT&A, No 4 (59). Volume 15, December 2020
2. V. P. Kogaev, Calculations for Strength for Variables in time Strains (1993)
3. Benasciutti, D., Tovo, R. (2007). Frequency-based fatigue. analysis of non-stationary switching random loads. Fatigue & Fracture of Engineering Materials &. Structures, 30(11):1016–1029, 2007.
4. Luca Susmel, David Taylor. The Theory of Critical Distances to estimate lifetime of notched components subjected to variable amplitude uniaxial fatigue loading/ International Journal of Fatigue 2011. 33(7):900-911. DOI: 10.1016/j.ijfatigue.2011.01.012
5. R. Sunder, Journal of ASTM Intern, 9, JAI103940, (2012)
6. E. Bellec at al. Modelling and identification of fatigue load spectra: Application in the automotive industry/International Journal of Fatigue. Volume 149, August 2021, 106222
7. Gadolina I.V., Makhutov N.A., Erpalov A.V. Varied approaches to loading assessment in fatigue studies/International Journal of Fatigue. 2021. T. 144. C. 106035.
8. J. Martinez, IJF 132 105327 (2020)
9. M. Matsuishi, T. Endo. Fatigue of metals subject to varying stress. Presented to the Japan Society of Mechanical Engineers, Fukuoka, Japan. 1968. Japan Society of Mechanical Engineers, Fukuoka, Japan. (1968)

10. M. Böhm, AIP Conference Proceedings 2029(1):020005. (2018)
11. Gadolina I.V., Lisachenko N.G., Svirskiy Y.A., Dubin D.A. The choice of sampling frequency and optimal method of signals digital processing in problems of a random loading process treating to assess durability/Inorganic Materials. 2020. Т. 56. № 15. С. 1551-1558.
12. Gadolina, I., & Zaynetdinov, R. (2020). Building the Continuous Random Process Out of The Specified Sequence of Turning Points for Fatigue Testing. Reliability: Theory & Applications, 15(1), 33-41.
13. R. Fisher, E. Haibach, Behavior of Steel Under Cyclic Loads, 368–405 (1983)
14. Gadolina I.V., Plotnikov E.V., Bautin A.A. Studying the crack growth rate variability by applying the Willenborg's model to the Markov's simulated trials/Advances in Intelligent Systems and Computing (см. в книгах). 2020. Т. 1127 AISC. С. 175-184.
15. A. Erpalov, L. Shefer. Fatigue-based Classification of Loading Processes. Procedia Engineering 2nd International Conference on Industrial Engineering (ICIE-2016) 150, 144-149 (2016)

A NOVAL APPLICATION OF DUANE PROCESS FOR MODELING TWO GRADED MANPOWER SYSTEM WITH DIRECT RECRUITMENT IN BOTH THE GRADES

Ch. Ganapathi Swamy[1], K. Srinivasa Rao [2]

•

[1]Department of Community Medicine, GSL Medical College, Rajahmundry, Andhra Pradesh, India. Email: ganesh051981@gmail.com
[2]Dept.of statistics, Andhra University, Visakhapatnam, Andhra Pradesh, India. Email: ksraoau@yahoo.co.in

Abstract

Human Resource Management and other companies rely heavily on manpower models. Manpower planning was a prerequisite for effective organization administration. The construction and analysis of two graded manpower models with direct Duane recruiting processes in both graduates is the subject of this paper. Duane's recruitment procedure was capable of identifying time-dependent recruitments. Poisson and non-homogeneous Poisson processes are used in the Duane recruitment process as precise instances for specified parameter values. It is assumed that the organization has two grades and that the recruitment procedure is based on the Duane Process. The processes of leaving and promotion are Poisson processes. The model's transient behavior was investigated by deriving unambiguous expressions for system characteristics such as the mean number of employees in each grade, the mean durational stay of an employee in each grade, and the variance of number of employees in each grade using differential equations. The model's sensitivity analysis of parameters shows that the Duane recruiting process has a substantial impact on system performance indicators. It was also noted that this model incorporates rates of recruitment that are increasing, decreasing, or stable. This model proved helpful in analyzing organizational manpower issues.

Keywords - Two graded Manpower model, Duane process, Time dependent recruitment rate, Sensitivity analysis.

I. Introduction

Planning the organization with the manpower structure in mind was a prerequisite for getting the most out of the resources. Due to their usefulness in creating strategies for Human Resource Development and resource allocation, much work has been reported on manpower models. Seal was a pioneer in the mathematical modelling of labour systems [1]. Silock looked at the observable fact of labour yield, which is related to the study of demography [2]. Bartholomew examined manpower models based on probability distributions of an employee's total service time in the organization [3] [4]. Ugwuowo and Mc Clean, as well as Wang, have examined manpower models and various approaches to their development [5] [6].

The parameters of the manpower model, such as the mean number of employees in each grade, the mean duration of stay an employee in each grade, and the variation of number of employees in each grade, were required for effective analysis and design of manpower systems Srinivasa Rao [7]. Kannan Nilakantan investigated the manpower models staffing rules and their extension to individual outsourcing [8]. Jeeva and Geetha looked at manpower models in a hazy environment [9]. Lalithadevi and Srinivasan used geometric process and shock models to investigate a single graded manpower system [10]. Osagiede and Ekhosuehi used continuous-time Markov chains and sparse stochastic measures to investigate Manpower models [11].

The graded manpower models with poisson processes have been examined by Srinivasa Rao, K. et al., Kondababu et al., and Govinda Rao et al. They implied that the hiring procedure was time-sensitive [12] [13] 14] [15] [16]. Parameswari, K., and Srinivasan [17] used a geometric technique to investigate the reduction in manpower for a two-graded system. Amudha.T and Srinivasan.A, discussed the problem of time to recruitment for a two-graded system, taking into account the loss of personnel in the form of an I.I.D Exponential random variable sequence [18]. Saral, L. et al. established a two-tiered personnel structure and a recruiting policy based on two thresholds [19]. Srividhya, K. et al. investigated manpower loss in a multi-graded organization [20]. Jayanthi et al. (2018) looked at a single graded manpower system and looked at the time to recruitment with a break-even point [21]. Tamas Banyai et al. used Markov chains to study a model for analyzing human resource use [22]. Arokkia Saibe,P et al. investigated two stochastic models based on the assumption that manpower shortages and inter-policy decision delays constitute two distinct sequences of independent and identically distributed random variables with two distinct breakdown thresholds [23]. They assumed that the hiring procedure was time-sensitive.

However, in many real-world circumstances in corporate offices and government agencies, the recruitment process was time-sensitive and did not necessarily follow the Poisson process. As a result, non stationary models must be considered for correct analysis. Srinivasa Rao et al. [24] recently developed two graded manpower models based on non-homogeneous Poisson recruitment. Srinivasa Rao,K et al. [25] have studied on two grade manpower model with Duane recruitment process. They realized that the recruitment rate was linearly proportional to time and that the duration between recruitment was distributed in an exponential manner. However, the recruiting rate in many organizations may increase/decrease/remain steady and time-dependent. The time-dependent non-stationary recruitment process can be fully characterized by the Duane process, which follows a Weibull distribution of inter-recruitment time. Little is reported in the literature on two hierarchical workforce models that use the Duane recruitment process directly at both grades. Therefore, this paper uses the Duane recruitment process of both graduates to develop and analyze a two-step manpower model. Poisson and non-homogeneous processes are two examples of the Duane process. The concept can be applied to a variety of organizations thanks to the recruitment in both grades. The remainder of the paper was laid out as follows:

II. Two graded manpower model with direct recruitment:

Consider a personnel system with two grades, each of which has its own recruitment process. The grade I recruiting process was supposed to be a Duane process, with the mean recruitment rate being a power function of time t and the form $\lambda_1(t) = a_1b_1t^{b_1-1}$. The grade II recruiting process was considered to be a Duane process with a mean recruitment rate of $\lambda_2(t) = a_2b_2t^{b_2-1}$. A Poisson process with parameter is used to promote students from grade I to grade II. Poisson processes with parameters and are used in the grade I and grade II leaving processes, respectively. "Figure 1" depicted a schematic diagram depicting the two-grade manpower concept.

The grade I recruiting process was supposed to be a Duane process, with the mean recruitment rate being a power function of time t and the form $\lambda_1(t) = a_1b_1t^{b_1-1}$. The grade II recruiting process was considered to be a Duane process with a mean recruitment rate of $\lambda_2(t) = a_2b_2t^{b_2-1}$. A Poisson process with parameter is used to promote students from grade I to grade II. Poisson processes with parameters and are used in the grade I and grade II leaving processes, respectively. "Figure 1" depicted a schematic diagram depicting the two-grade manpower concept.

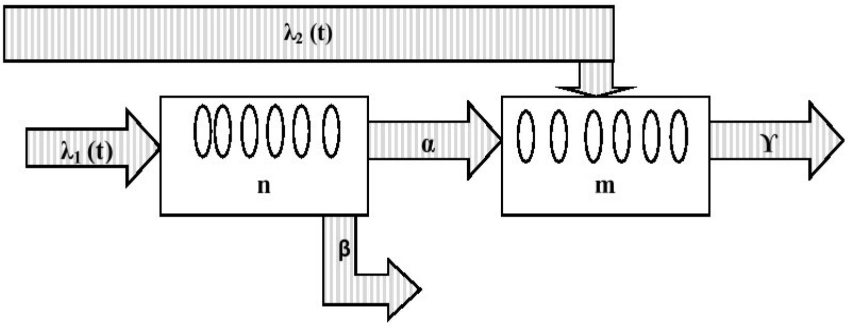


Figure 1: *Two grade manpower model with direct Duane recruitment process.*

Let $P_{n,m}(t)$ denote the probability that there are 'n' employees in grade-1 and 'm' employees in grade-2 at time 't' in the organization. Then the difference-differential equations of the model are

$$\frac{\partial P_{n,m}(t)}{\partial t} = -[\lambda_1(t) + \lambda_2(t) + n\alpha + n\beta + m\gamma]P_{n,m}(t) + [\lambda_1(t)]P_{n-1,m}(t)$$

$$+[\lambda_2(t)]P_{n,m-1}(t) + \alpha(n+1)P_{n+1,m-1}(t) + \beta(n+1)P_{n+1,m}(t) + (m+1)\gamma P_{n,m+1}(t) \quad (1)$$

$$\frac{\partial P_{n,0}(t)}{\partial t} = -[\lambda_1(t) + \lambda_2(t) + n\alpha + n\beta]P_{n,0}(t) + [\lambda_1(t)]P_{n-1,0}(t) + \beta(n+1)P_{n+1,0}(t) + \gamma P_{n,1}(t) \quad (2)$$

$$\frac{\partial P_{0,m}(t)}{\partial t} = -[\lambda_1(t) + \lambda_2(t) + +m\gamma]P_{0,m}(t) + [\lambda_2(t)]P_{0,m-1}(t) + \alpha P_{1,m-1}(t) + P_{1,m}(t) + (m+1)\gamma P_{0,m+1}(t) \quad (3)$$

$$\frac{\partial P_{0,0}(t)}{\partial t} = -[\lambda_1(t) + \lambda_2(t)]P_{0,0}(t) + \beta P_{1,0}(t) + (m+1)\gamma P_{0,1}(t) \quad (4)$$

Let $P(S_1,S_2; t)$ be the joint probability generating function then

$$P(S_1, S_2; t) = \sum_n \sum_m S_1^n S_2^m Pn, m(t) \quad (5)$$

Multiplying the equations (1) to (4) with corresponding $S_1^n S_2^m$ and summing over all n=0, 1, 2,…;

m= 0,1,2,…; we get

$$\frac{\partial P}{\partial t} = [\alpha(S_2 - S_1) + \beta(1 - S_1)]\frac{\partial P}{\partial S_1} + \gamma[1 - S_2]\frac{\partial P}{\partial S_2} - [\lambda_1(t)(1 - S_1) + \lambda_2(t)(1 - S_2)]P$$

$$\frac{\partial P}{\partial t} - [\alpha(S_2 - S_1) + \beta(1 - S_1)]\frac{\partial P}{\partial S_1} - \gamma[1 - S_2]\frac{\partial P}{\partial S_2} = [\lambda_1(t)(S_1 - 1) + \lambda_2(t)(S_2 - 1)]P \quad (6)$$

After simplification, we get

$$\frac{\partial P}{\partial t} + [\alpha(S_1 - S_2) + \beta(S_2 - 1)]\frac{\partial P}{\partial S_1} + \gamma[S_2 - 1]\frac{\partial P}{\partial S_2} = [\lambda_1(t)(S_1 - 1) + \lambda_2(t)(S_2 - 1)]P \quad (7)$$

Solving the equation (7) by Lagrangian's method, the auxiliary equations are

$$\frac{\partial t}{1} = \frac{\partial S_1}{\alpha(S_1 - S_2) + \beta(S_2 - 1)} = \frac{\partial S_2}{\gamma(S_2 - 1)} = \frac{\partial P}{[\lambda_1(t)(S_1 - 1) + \lambda_2(t)(S_2 - 1)]P} \quad (8)$$

With the initial conditions that there are N employees in grade-1 and M employees in grade-2 in the organization at time t=0. i.e., $P_{N,M}(0) = 1$ and $P_{0,0}(t) = 0$ for t > 0.

To solve the equation (8) the functional forms of $\lambda_1(t)$ and $\lambda_2(t)$ are required.

Since the recruitment processes follow Duane processes we have the mean recruitment rates of grade-1and grade-2 in the system are $\lambda_1(t) = a_1 b_1 t^{b1-1}$ and $\lambda_1(t) = a_2 b_2 t^{b2-1}$ respectively, where $\lambda>0$, a_1, b1, a_2 and b_2 are constants
Solve the equations in (8) we get

$$A = (S_2 - 1)e^{-\gamma t}$$

$$B = e^{-(\alpha+\beta)t}\left[(S_1 - 1) + \frac{\alpha(S_2-1)}{\gamma-(\alpha+\beta)}\right]$$

$$C = S_1^N S_2^M \left\{\frac{-a_1 b_1 e^{-(\alpha+\beta)t}(s_1-1)}{[\alpha+\beta]} - \frac{a_1 b_1\ \alpha e^{-(\alpha+\beta)t}(s_2-1)}{[\gamma-(\alpha+\beta)][\alpha+\beta]} + \frac{a_1 b_1\ \alpha e^{-\gamma t}(s_2-1)}{[\gamma-(\alpha+\beta)][\gamma]} - \frac{a_2 b_2 e^{-\gamma t}(s_2-1)}{\gamma}\right\}$$

where $S_1^N = \left[1-(1-S_1)e^{-(\alpha+\beta)t} - \frac{\alpha(1-S_2)}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]^N$, $S_2^M = [1-(1-S_2)e^{-\gamma t}]^M$.

A,B and C are arbitrary constants. (9)

The joint probability generating function of the number of employee in grade-1 and in grade-2 is

$$\begin{aligned} P(S_1, S_2;\ t) = C.\mathrm{Exp}\ &\left\{a_1 b_1 (S_1-1)\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv\right]\right. \\ &+\frac{a_1 b_1\, \alpha\,(S_2-1)}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\right] \\ &\left.+a_2 b_2 (S_2-1)\left[e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_2-1)}dv\right]\right\} \end{aligned} \tag{10}$$

Substituting the value of 'C' from equation (9) in the equation (10), the joint probability generating function of the number of employees in the grade-1 and grade-2 are obtained as

$$\begin{aligned} P(S_1, S_2;\ t) = &\left\{\exp\left\{a_1 b_1 (S_1-1)\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + \right.\right. \\ &\frac{a_1 b_1\, \alpha\,(S_2-1)}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\right] + \frac{a_1 b_1\, \alpha\,(S_2-1)}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + \\ &\left.\left. a_2 b_2 (S_2-1)\left[e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_2-1)}dv - \frac{e^{-\gamma t}}{\gamma}\right]\right\}[XY]\right\} \end{aligned}$$

Where $X = S_1^N = \left[1-(1-S_1)e^{-(\alpha+\beta)t} - \frac{\alpha(1-S_2)}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]^N$,

$$Y = S_2^M = [1-(1-S_2)e^{-\gamma t}]^M,\ |S_1| < 1, |S_2| < 1 \tag{11}$$

III. Characteristics of the model:

The characteristics of the model are obtained by using the equation (11).Expanding $P(S_1, S_2;\ t)$and collecting the constant terms, we get the probability that there is no employee in the organization as

$P_{0,0}(t) =$

$$\Big\{\exp\Big\{a_1 b_1(-1)\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big] + \frac{a_1 b_1\,\alpha(-1)}{\gamma-(\alpha+\beta)}\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\Big] + \Big[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big] + a_2 b_2(-1)\Big[e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_2-1)}dv - \frac{e^{-\gamma t}}{\gamma}\Big]\Big\}[X_1 Y_1]\Big\}$$

Where, $X_1 = \Big[1 - e^{-(\alpha+\beta)t} - \frac{\alpha}{\gamma-(\alpha+\beta)}\big[e^{-(\alpha+\beta)t} - e^{-\gamma t}\big]\Big]^N$ *and* $Y_1 = [1 - e^{-\gamma t}]^M$ (12)

Taking $S_2 = 1$ in equation (11), we get the probability generating function of the number of employees in grade-1 as

$$P(S_1;t) = \Big[\big[1-(1-S_1)e^{-(\alpha+\beta)t}\big]^N\Big]\Big[\exp\Big[a_1 b_1(S_1-1)\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big]\Big]\Big],$$

$|S_1| < 1$ (13)

Expanding $P(S_1;t)$ and collecting the constant terms, we get the probability that there is no employee in grade-1 as

$$P_0(t) = \Big[\big[1-e^{-(\alpha+\beta)t}\big]^N\Big]\Big[\exp\Big[a_1 b_1(-1)\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big]\Big]\Big] \tag{14}$$

The mean number of employees in grade-1 is

$$L_1(t) = \big[Ne^{-(\alpha+\beta)t}\big] + a_1 b_1\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big] \tag{15}$$

The probability that there is at least one employee in grade-1 is

$U_1(t) = 1 - P_0(t)$

$$U_1(t) = 1 - \Big[\big[1-e^{-(\alpha+\beta)t}\big]^N\Big]\Big[\exp\Big[a_1 b_1(-1)\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big]\Big]\Big] \tag{16}$$

The average duration of stay of an employee in grade-1 is

$W_1(t) = \frac{L_1(t)}{(\alpha+\beta)[1-P_0(t)]}$

$$W_1(t) = \frac{\big[Ne^{-(\alpha+\beta)t}\big] + a_1 b_1\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v} v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big]}{(\alpha+\beta)\Big[1-\Big[\big[1-e^{-(\alpha+\beta)t}\big]^N\Big]\Big[\exp\Big[a_1 b_1(-1)\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v} v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big]\Big]\Big]\Big]} \tag{17}$$

The variance of the number of grade-1 is

$$V_1(t) = \big[Ne^{-(\alpha+\beta)t}\big]\big[1-e^{-(\alpha+\beta)t}\big] + a_1 b_1\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\Big] \tag{18}$$

Coefficient of variation of the number of employees in grade-1 is $cv_1(t) = \frac{\sqrt{V_1(t)}}{L_1(t)}$ where, $V_1(t)$ and $L_1(t)$ are given in equations (18) and (15) respectively (19)

Similarly, taking $S_1 = 1$ in equation (11), then we get the probability generating function of the number of employees in grade-2 as

$$P(S_2;t) = \left\{ Exp\left[\frac{a_1 b_1 \alpha (S_2 - 1)}{\gamma - (\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\right] + \frac{a_1 b_1 \alpha (S_2-1)}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + \frac{a_1 b_1 \alpha (S_2-1)}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right]\right]\right\}\left[\left[1 - \frac{\alpha(1-S_2)}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]^N\right]$$

$$[[1-(1-S_2)e^{-\gamma t}]^M] \qquad \text{where, } |S_2| < 1 \tag{20}$$

Expanding $P(S_2;t)$ and collecting the constant terms, we get the probability that there is no employee in grade-2 as

$$P_0(t) = \left\{ \text{Exp}\left[\frac{a_1 b_1 \alpha (-1)}{\gamma - (\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\right] + \frac{a_1 b_1 \alpha (-1)}{\gamma - (\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + a_2 b_2 (-1)\left[e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_2-1)}dv - \frac{e^{-\gamma t}}{\gamma}\right]\right]\right\}$$

$$\left[1 - \frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]^N [[1-e^{-\gamma t}]^M] \tag{21}$$

The mean number of employees in grade-2 is

$$L_2(t) = Me^{-\gamma t} + N\left[\frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right] + \left[\frac{a_1 b_1 \alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\right] + \frac{a_1 b_1 \alpha}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + a_2 b_2\left[e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_2-1)}dv - \frac{e^{-\gamma t}}{\gamma}\right]\right] \tag{22}$$

The probability that there is at least one employee in grade-2 is

$$U_2(t) = 1 - P_0(t)$$

$$= 1 - \left\{ \text{Exp}\left[\frac{a_1 b_1 \alpha (-1)}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v}\, v^{(b_1-1)}dv - e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_1-1)}dv\right] + \frac{a_1 b_1 \alpha (-1)}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + a_2 b_2 (-1)\left[e^{-\gamma t}\int_0^t e^{\gamma v}\, v^{(b_2-1)}dv - \frac{e^{-\gamma t}}{\gamma}\right]\right]\right\} \left[1 - \frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]^N [[1-e^{-\gamma t}]^M] \tag{23}$$

The average duration of stay of employees in grade-2 is

$$W_2(t) = \frac{\begin{gathered} Me^{-\gamma t} + N\left[\frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right] \\ + \left[\frac{a_1 b_1 \alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v} v^{(b_1-1)} dv - e^{-\gamma t}\int_0^t e^{\gamma v} v^{(b_1-1)} dv\right] \\ + \frac{a_1 b_1 \alpha}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + a_2 b_2\left[e^{-\gamma t}\int_0^t e^{\gamma v} v^{(b_2-1)} dv - \frac{e^{-\gamma t}}{\gamma}\right]\right] \end{gathered}}{\begin{gathered} \gamma\Big\{1 - \Big[\mathrm{Exp}\Big[\frac{a_1 b_1 \alpha (-1)}{\gamma-(\alpha+\beta)}\Big[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v} v^{(b_1-1)} dv \\ - e^{-\gamma t}\int_0^t e^{\gamma v} v^{(b_1-1)} dv\Big] + \frac{a_1 b_1 \alpha(-1)}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] \\ + a_2 b_2 (-1)\left[e^{-\gamma t}\int_0^t e^{\gamma v} v^{(b_2-1)} dv - \frac{e^{-\gamma t}}{\gamma}\right]\Big] \\ \left[\left[1 - \frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]^N\right]\left[\left[1 - e^{-\gamma t}\right]^M\right]\Big\} \end{gathered}} \tag{24}$$

The variance of the number of employees in grade-2 is

$$V_2(t) = Me^{-\gamma t}(1 - e^{-\gamma t})$$

$$+N\left[\frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]\left[\left[1 - \left[\frac{\alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t} - e^{-\gamma t}\right]\right]\right]\right.$$

$$+\frac{a_1 b_1 \alpha}{\gamma-(\alpha+\beta)}\left[e^{-(\alpha+\beta)t}\int_0^t e^{(\alpha+\beta)v} v^{(b_1-1)} dv - e^{-\gamma t}\int_0^t e^{\gamma v} v^{(b_1-1)} dv\right]$$

$$+\frac{a_1 b_1 \alpha}{\gamma-(\alpha+\beta)}\left[\frac{e^{-\gamma t}}{\gamma} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\right] + a_2 b_2\left[e^{-\gamma t}\int_0^t e^{\gamma v} v^{(b_2-1)} dv - \frac{e^{-\gamma t}}{\gamma}\right] \tag{25}$$

Coefficient of variation of the number of employees in grade-2 is

$cv_2(t) = \frac{\sqrt{V_2(t)}}{L_2(t)}$ where, $V_2(t)$ and $L_2(t)$ are given in equations (25) and (22) respectively. (26)

The mean number of employees in the organization is $L = L_1 + L_2$ where, $L_1(t)$ and $L_2(t)$ are given in equations (15) and (22) respectively. (27)

IV. Numerical illustration and results

A numerical illustration was used to explain the model's behaviour in this subdivision. For the recruiting, advancement, and leaving rates of the organization, several values of the parameters were explored. Because the manpower model's performance characteristics were particularly time-sensitive, the transient behaviour of the model was investigated by computing performance measures with the following set of values for the model parameters: t= 1.5,2,2.5,3and 3.5; α=3,4,5,6 and 7; β=4,5,6,7 and 8; γ=5,6,7,8 and 9; a_1=5,10,15,20 and 25; b_1=3,4,5,6 and 7; a_2=5,10,15,20 and 25; b_2=3,4,5,6 and 7; N=1000,1100,1200,1300 and 1400 ; M=600,700,800,900 and 1000.

Performance measures such as the mean number of employees in grades I and II, the mean duration of stay of a grade I employee and in grade II, the variance of the number of employees in grades I and II, and the coefficient of variation of the number of employees in grades I and II were computed and presented in Tables 1 and 2. Figures 2a, 2b, 3a and 3b show the link between parameters and performance measures.

Table 1 demonstrated that the performance indicators in grades I and II were extremely time sensitive. The mean number of employees in grade I and grade 2 in the organization increased from 4.0180 to 24.1946 and 8.1050 to 45.6831, respectively, while time (t) varied from 1.5 to 3.5. When all other factors were held constant, the mean period of stay of an employee in grade I and grade II in the company increased from 0.5845 to 3.4564 and 1.7138 to 9.1367, respectively.

When all other parameters were held constant, the mean number of employees in grade I was not influenced and in grade II it increased from 8.4203 to 8.6416.When all other parameters were held constant, the mean duration of stay of an employee in grade I was not influenced and in grade II it increased from 1.7533 to 1.7848.

When all other parameters were held constant, the mean number of employees in grade I was not influenced and in grade II it increased from 8.4203 to 8.6416.When all others parameters were held constant, the mean duration of stay of an employee in grade I was not influenced and in grade II it increased from 1.7533 to 1.7848.

Table 1: *Values of L_1, L_2, W_1 and W_2 for different values of parameters*.

t	α	β	γ	a_1	b_1	a_2	b_2	N	M	L_1	L_2	W_1	W_2
1.5	3	4	5	5	3	5	3	1000	600	4.0180	8.1050	0.5845	1.7138
2	3	4	5	5	3	5	3	1000	600	7.4352	13.5421	1.0628	2.7769
2.5	3	4	5	5	3	5	3	1000	600	11.9497	22.0582	1.7071	4.4219
3	3	4	5	5	3	5	3	1000	600	17.5364	32.7960	2.5052	6.5599
3.5	3	4	5	5	3	5	3	1000	600	24.1946	45.6831	3.4564	9.1367
1.5	3	4	5	5	3	5	3	1000	600	4.0180	8.1050	0.5845	1.7138
1.5	4	4	5	5	3	5	3	1000	600	3.5804	8.4013	0.4604	1.7508
1.5	5	4	5	5	3	5	3	1000	600	3.2370	8.6505	0.3744	1.7862
1.5	6	4	5	5	3	5	3	1000	600	2.9553	8.8632	0.3118	1.8188
1.5	7	4	5	5	3	5	3	1000	600	2.7189	9.0462	0.2646	1.8483
1.5	7	5	5	5	3	5	3	1000	600	2.5174	8.7440	0.2282	1.8003
1.5	7	6	5	5	3	5	3	1000	600	2.3435	8.4931	0.1994	1.7634
1.5	7	7	5	5	3	5	3	1000	600	2.1921	8.2806	0.1763	1.7350
1.5	7	8	5	5	3	5	3	1000	600	2.0589	8.0977	0.1573	1.7130
1.5	7	8	6	5	3	5	3	1000	600	2.0589	6.5928	0.1573	1.2559
1.5	7	8	7	5	3	5	3	1000	600	2.0589	5.7212	0.1573	0.9933
1.5	7	8	8	5	3	5	3	1000	600	2.0589	5.0994	0.1573	0.8147
1.5	7	8	9	5	3	5	3	1000	600	2.0589	4.6108	0.1573	0.6856
1.5	7	8	9	10	3	5	3	1000	600	4.1178	5.9837	0.2791	0.7104
1.5	7	8	9	15	3	5	3	1000	600	6.1767	7.3565	0.4126	0.8309
1.5	7	8	9	20	3	5	3	1000	600	8.2356	8.7293	0.5492	0.9739
1.5	7	8	9	25	3	5	3	1000	600	10.2944	10.1021	0.6863	1.1236
1.5	7	8	9	25	4	5	3	1000	600	19.7548	15.5521	1.3170	1.7280
1.5	7	8	9	25	5	5	3	1000	600	35.6026	24.0877	2.3735	2.6764
1.5	7	8	9	25	6	5	3	1000	600	61.6965	37.3243	4.1131	4.1471
1.5	7	8	9	25	7	5	3	1000	600	104.0989	57.6922	6.9399	6.4102
1.5	7	8	9	25	7	10	3	1000	600	104.0989	60.9278	6.9399	6.7698
1.5	7	8	9	25	7	15	3	1000	600	104.0989	64.1634	6.9399	7.1293
1.5	7	8	9	25	7	20	3	1000	600	104.0989	67.3990	6.9399	7.4888
1.5	7	8	9	25	7	25	3	1000	600	104.0989	70.6346	6.9399	7.8483
1.5	7	8	9	25	7	25	4	1000	600	104.0989	84.7664	6.9399	9.4185
1.5	7	8	9	25	7	25	5	1000	600	104.0989	107.9303	6.9399	11.9923
1.5	7	8	9	25	7	25	6	1000	600	104.0989	145.3699	6.9399	16.1522
1.5	7	8	9	25	7	25	7	1000	600	104.0989	205.2306	6.9399	22.8034
1.5	3	4	5	5	3	10	3	1000	600	4.0180	8.1050	0.5845	1.7138
1.5	3	4	5	5	3	10	3	1100	600	4.0208	8.1838	0.5849	1.7230
1.5	3	4	5	5	3	10	3	1200	600	4.0235	8.2627	0.5853	1.7327
1.5	3	4	5	5	3	10	3	1300	600	4.0263	8.3415	0.5856	1.7428
1.5	3	4	5	5	3	10	3	1400	600	4.0290	8.4203	0.5860	1.7533
1.5	3	4	5	5	3	10	3	1400	700	4.0290	8.4756	0.5860	1.7610
1.5	3	4	5	5	3	10	3	1400	800	4.0290	8.5309	0.5860	1.7688
1.5	3	4	5	5	3	10	3	1400	900	4.0290	8.5863	0.5860	1.7767
1.5	3	4	5	5	3	10	3	1400	1000	4.0290	8.6416	0.5860	1.7848

The performance metrics in grade I and grade II employees in the organization were extremely sensitive to time, as shown in **Table 2**. When other parameters were held constant, it was discovered that time (t) varies from 1.5 to 3.5, the variance of the number of employees in grade I and grade II increased from 4.0180 to 24.1946 and 8.1042 to 45.6831, respectively, and the coefficient of variation of the number of employees in both grades decreased from 0.4989 to 0.2033 and 0.3512 to 0.1480.

When the promotion rate (α) from grade I to grade II increased from 3 to 7, the variance of the number of employees in grade I decreased from 4.0180 to 2.7189 and increased from 8.1042 to 9.0456, and the coefficient of variation of the number of employees in grade I increased from 0.4989 to 0.6065 and decreased from 0.3512 to 0.3325, when all other parameters remained constant.

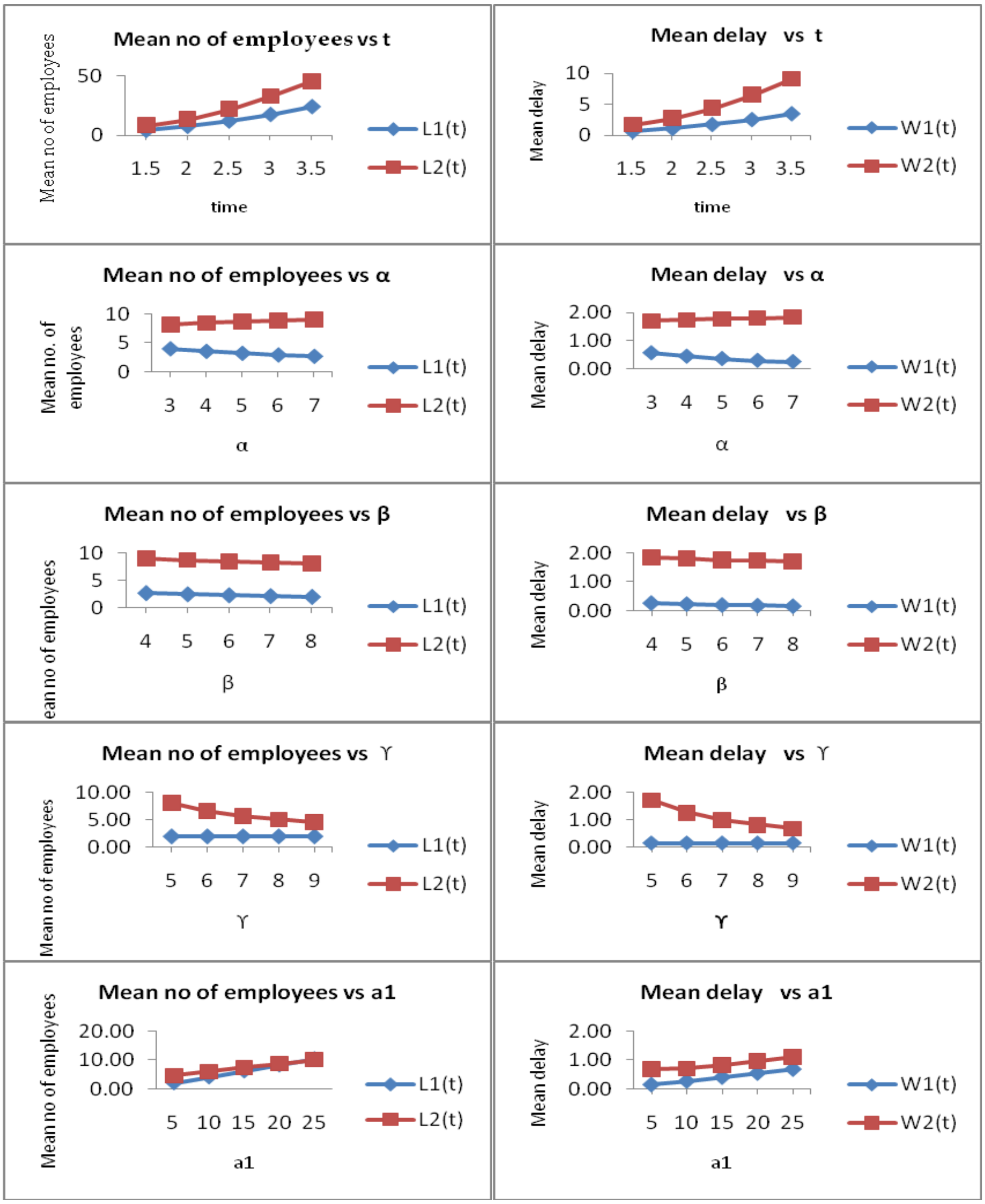


Figure 2a: *Relation between the parameters and performance measures*

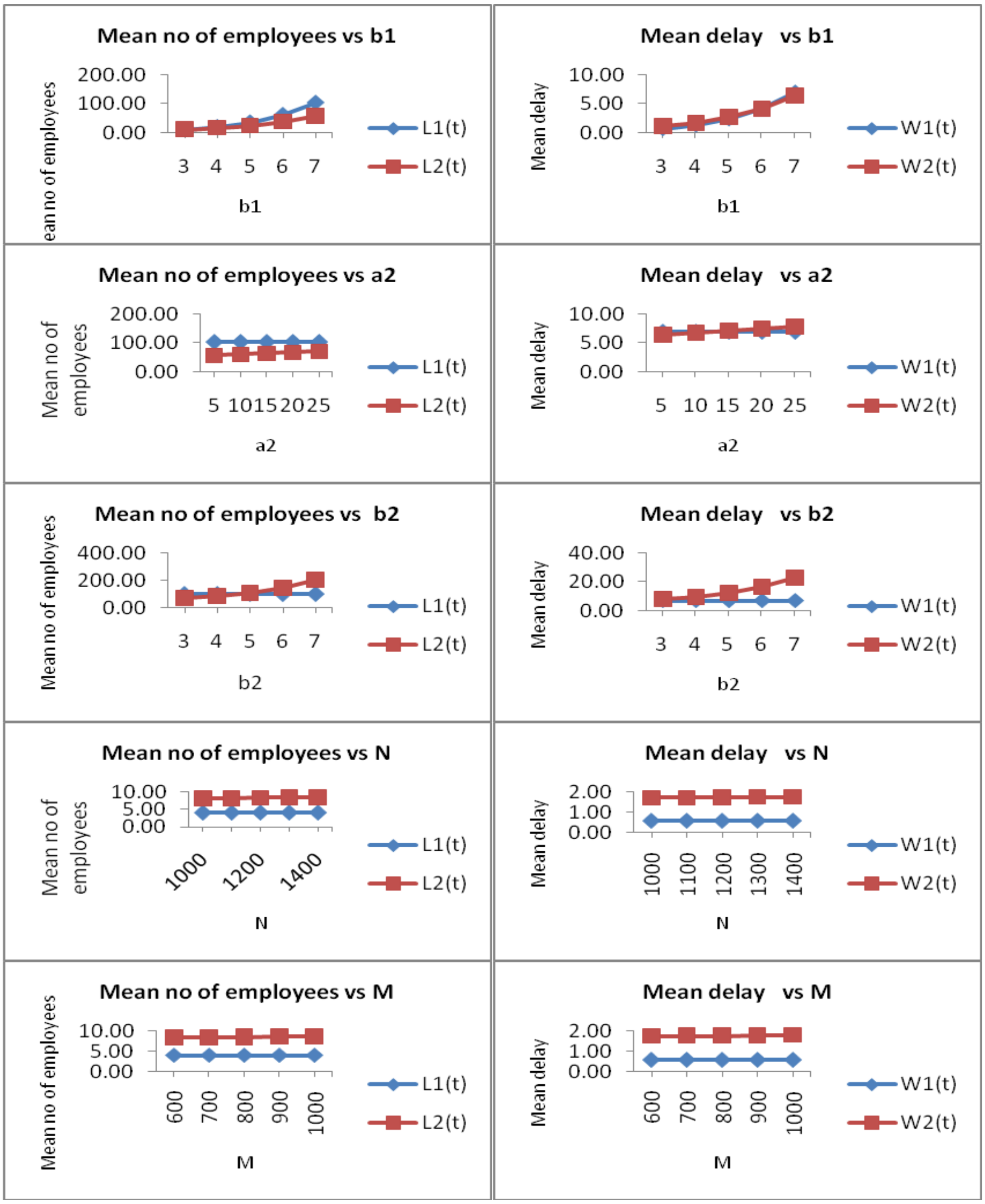


Figure 2b: *Relation between the parameters and performance measures*

When the leaving rate (β) of an employee in grade I increases from 4 to 8, the variance of the number of employees in grade I and grade II decreases from 2.7189 to 2.0589 and 9.0456 to 8.0974, respectively, while the coefficient of variation of the number of employees in both grades I and II increases from 0.6065 to 0.6969 and 0.3325 to 0.3514, respectively, when other parameters remain constant.

When the leaving rate (γ) of an employee in grade II increases from 5 to 9, the variance of the number of employees in grade I is unaffected, while in grade II it decreases from 8.0974 to 4.6108. When other parameters are held constant, the coefficient of variation of the number of employees in grade I is unaffected, while in grade II it increases from 0.3514 to 0.4657. When the recruitment rate parameter (a_1) of employees in grade I was changed from 5 to 25, the variance of the number of employees in grade I and grade II increased from 2.0589 to 10.2944 and 4.6108 to 10.1021 respectively, while the coefficient of variation of the number of employees in both grade I and grade II decreased from 0.6969 to 0.3117 and 0.4657 to 0.3146 when the other parameters remained constant.

Table 2: *Values of V_1,V_2, CV_1 and CV_2 for different values of parameters.*

T	α	β	γ	a_1	b_1	a_2	b_2	N	M	V_1	V_2	CV_1	CV_2
1.5	3	4	5	5	3	5	3	1000	600	4.0180	8.1042	0.4989	0.3512
2	3	4	5	5	3	5	3	1000	600	7.4352	13.5421	0.3667	0.2717
2.5	3	4	5	5	3	5	3	1000	600	11.9497	22.0582	0.2893	0.2129
3	3	4	5	5	3	5	3	1000	600	17.5364	32.7960	0.2388	0.1746
3.5	3	4	5	5	3	5	3	1000	600	24.1946	45.6831	0.2033	0.1480
1.5	3	4	5	5	3	5	3	1000	600	4.0180	8.1042	0.4989	0.3512
1.5	4	4	5	5	3	5	3	1000	600	3.5804	8.4006	0.5285	0.3450
1.5	5	4	5	5	3	5	3	1000	600	3.2370	8.6498	0.5558	0.3400
1.5	6	4	5	5	3	5	3	1000	600	2.9553	8.8626	0.5817	0.3359
1.5	7	4	5	5	3	5	3	1000	600	2.7189	9.0456	0.6065	0.3325
1.5	7	5	5	5	3	5	3	1000	600	2.5174	8.7435	0.6303	0.3382
1.5	7	6	5	5	3	5	3	1000	600	2.3435	8.4927	0.6532	0.3431
1.5	7	7	5	5	3	5	3	1000	600	2.1921	8.2802	0.6754	0.3475
1.5	7	8	5	5	3	5	3	1000	600	2.0589	8.0974	0.6969	0.3514
1.5	7	8	6	5	3	5	3	1000	600	2.0589	6.5928	0.6969	0.3895
1.5	7	8	7	5	3	5	3	1000	600	2.0589	5.7212	0.6969	0.4181
1.5	7	8	8	5	3	5	3	1000	600	2.0589	5.0994	0.6969	0.4428
1.5	7	8	9	5	3	5	3	1000	600	2.0589	4.6108	0.6969	0.4657
1.5	7	8	9	10	3	5	3	1000	600	4.1178	5.9837	0.4928	0.4088
1.5	7	8	9	15	3	5	3	1000	600	6.1767	7.3565	0.4024	0.3687
1.5	7	8	9	20	3	5	3	1000	600	8.2356	8.7293	0.3485	0.3385
1.5	7	8	9	25	3	5	3	1000	600	10.2944	10.1021	0.3117	0.3146
1.5	7	8	9	25	4	5	3	1000	600	19.7548	15.5521	0.2250	0.2536
1.5	7	8	9	25	5	5	3	1000	600	35.6026	24.0877	0.1676	0.2038
1.5	7	8	9	25	6	5	3	1000	600	61.6965	37.3243	0.1273	0.1637
1.5	7	8	9	25	7	5	3	1000	600	104.0989	57.6922	0.0980	0.1317
1.5	7	8	9	25	7	10	3	1000	600	104.0989	60.9278	0.0980	0.1281
1.5	7	8	9	25	7	15	3	1000	600	104.0989	64.1634	0.0980	0.1248
1.5	7	8	9	25	7	20	3	1000	600	104.0989	67.3990	0.0980	0.1218
1.5	7	8	9	25	7	25	3	1000	600	104.0989	70.6346	0.0980	0.1190
1.5	7	8	9	25	7	25	4	1000	600	104.0989	84.7664	0.0980	0.1086
1.5	7	8	9	25	7	25	5	1000	600	104.0989	107.9303	0.0980	0.0963
1.5	7	8	9	25	7	25	6	1000	600	104.0989	145.3699	0.0980	0.0829
1.5	7	8	9	25	7	25	7	1000	600	104.0989	205.2306	0.0980	0.0698
1.5	3	4	5	5	3	10	3	1000	600	4.0180	8.1042	0.4989	0.3512
1.5	3	4	5	5	3	10	3	1100	600	4.0208	8.1830	0.4987	0.3495
1.5	3	4	5	5	3	10	3	1200	600	4.0235	8.2617	0.4985	0.3479
1.5	3	4	5	5	3	10	3	1300	600	4.0263	8.3405	0.4984	0.3462
1.5	3	4	5	5	3	10	3	1400	600	4.0290	8.4193	0.4982	0.3446
1.5	3	4	5	5	3	10	3	1400	700	4.0290	8.4746	0.4982	0.3435
1.5	3	4	5	5	3	10	3	1400	800	4.0290	8.5298	0.4982	0.3424
1.5	3	4	5	5	3	10	3	1400	900	4.0290	8.5851	0.4982	0.3412
1.5	3	4	5	5	3	10	3	1400	1000	4.0290	8.6404	0.4982	0.3402

When the recruitment rate parameter (b_1) in grade I is changed from 3 to 7, the variance of the number of employees in grade I and grade II increases from 10.2944 to 104.0989 and 10.1021 to

57.6922, respectively, while the coefficient of variation of the number of employees in both grades I and II decreases from 0.3117 to 0.0980 and 0.3146 to 0.1317, respectively, when the other parameters remain constant.

When the recruitment rate parameter (a2) in grade II changes from 5 to 25, the variance of the number of employees in grade I does not change and in grade II it increases from 57.6922 to 70.6346, while the coefficient of variation of the number of employees in grade I does not change and in grade II it decreases from 0.1317 to 0.1190.

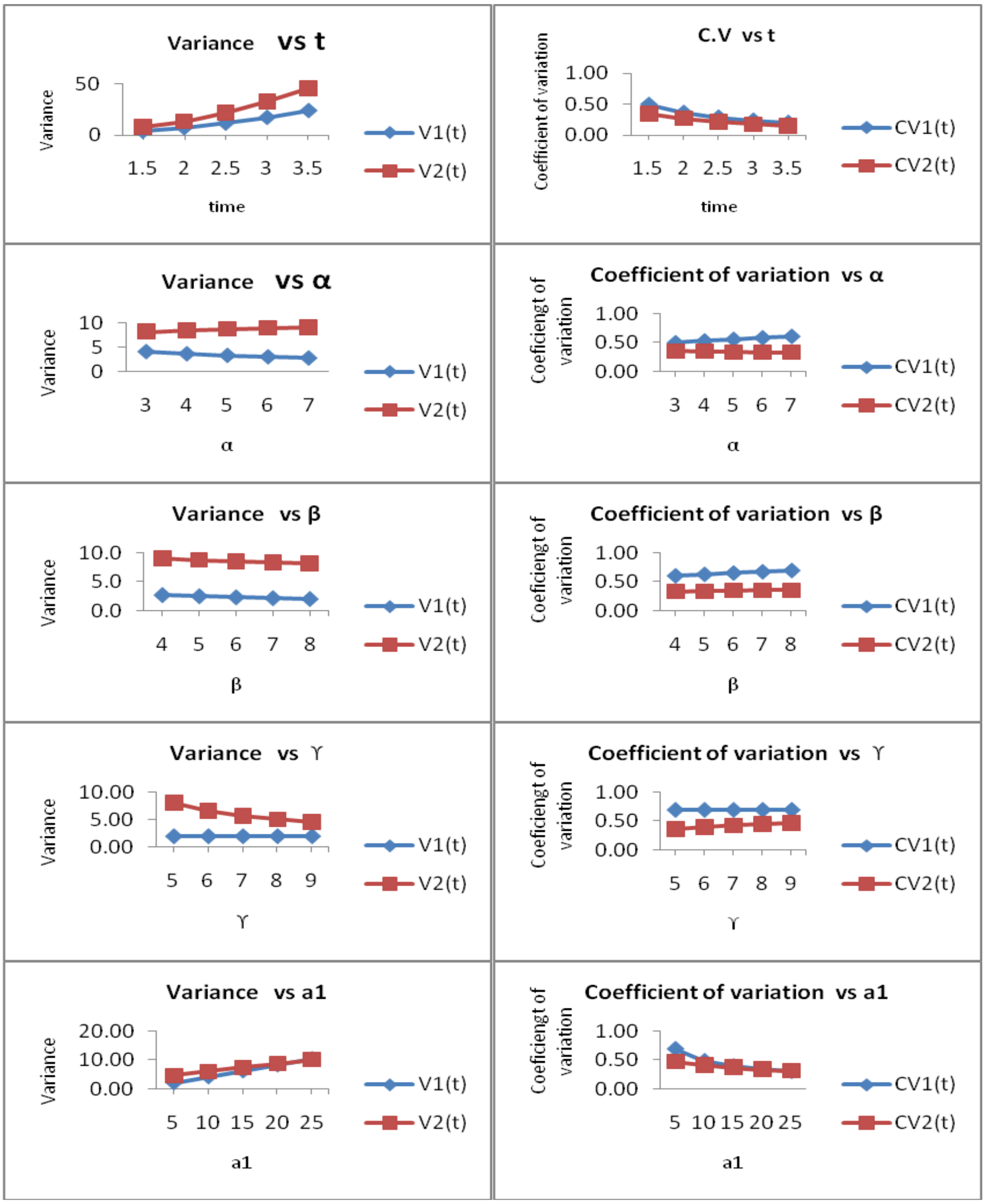


Figure 3a: *Relation between the parameters and performance measures.*

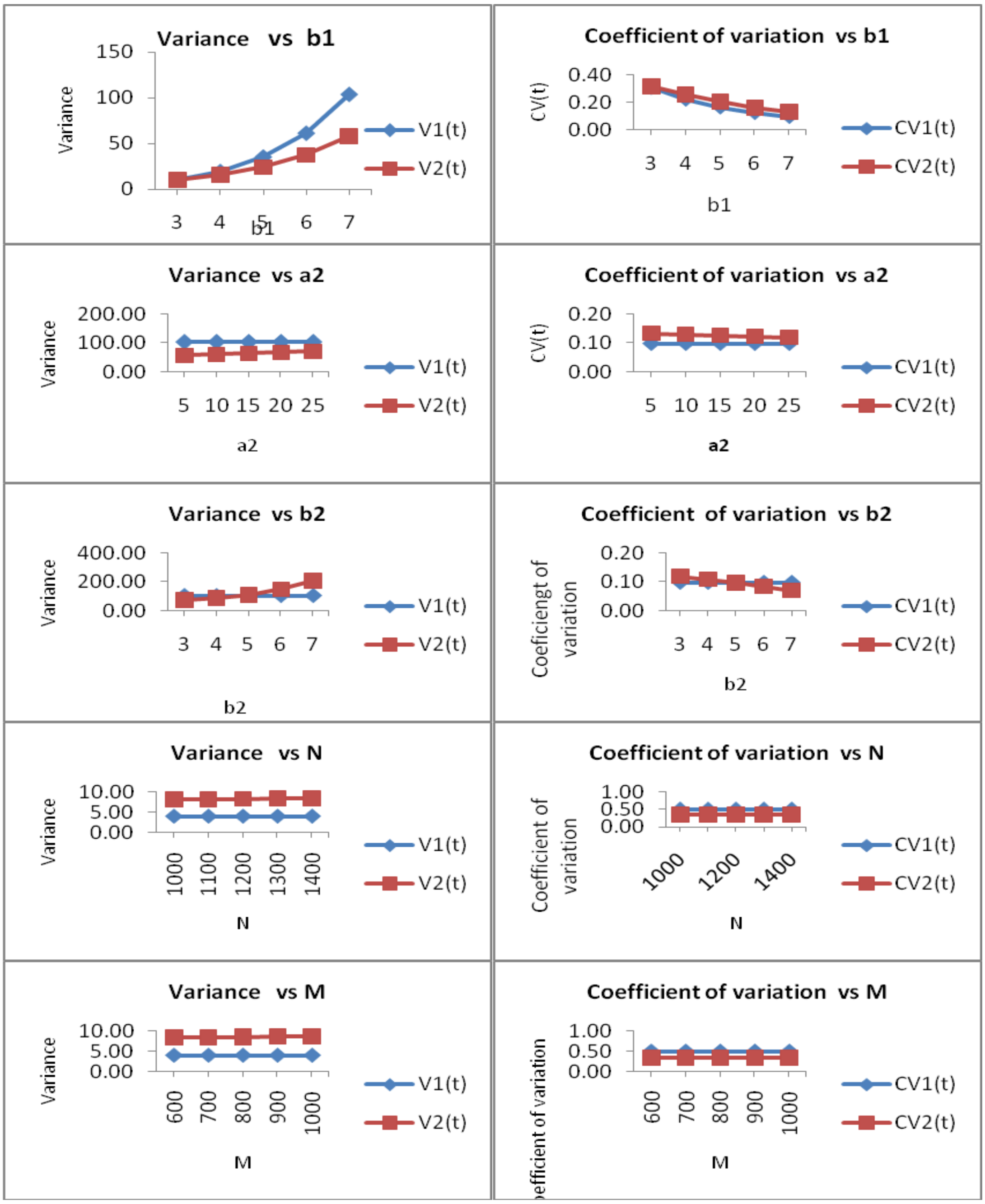


Figure 3b: *Relation between the parameters and performance measures.*

When the recruitment rate parameter (b2) of employees in grade II varies from 3 to 7, the variance of the number of employees in grade I is unaffected, and in grade II it increases from 70.6346 to 205.2306, while the coefficient of variation of the number of employees in grade I is unaffected, and in grade II it decreases from 0.1190 to 0.0698.

When other parameters were held constant, the initial number of employees in grade I (N) varied from 1000 to 1400, the variance of the number of employees in grade I and grade II increased from 4.0180 to 4.0290 and 8.1042 to 8.4193, respectively, and the coefficient of variation of the number of employees in grade I and grade II decreased from 0.4989 to 0.4982 and 0.3512 to 0.3446.

When other parameters were fixed, the variance of the number of employees in grade I was not influenced and in grade II it was increasing from 8.4193 to 8.6404, coefficient of variation of the number of employees in grade I was not influenced and in grade II it was decreasing from 0.3446 to 0.3402, when the primary number of employees in grade II (M) varies from 100 to 500.

V. Sensitivity analysis of the model

The model was sensitivity tested with respect to the value of time (t), recruitment rates $\lambda 1(t)$ and $\lambda 2(t)$, promotion rate parameter (α), and leaving parameters(β) and (γ) of both grade I and grade II, as well as all other parameters combined on the mean number of employees in grade I and grade II, mean duration of stay of an employee in grade I and grade II, and variance of the number of employees in grade I and grade II.

For different value of $t,\alpha,\beta,\gamma,a_1,b_1,a_2$ and b_2 the mean number of employees in grade I and in grade II, mean duration of stay of an employee in grade I and in grade II, the variance of the number of employees in grade I and in grade II were computed and presented in Table 3a and 3b with variation of -15%,-10%,-5%,0%,5%,10% and 15% of the model parameters.

Time had a significant impact on the performance measurements (t). The mean number of employees, mean duration of stay of an employee, and variation of the number of employees in grade I and grade II increased when t increased from -15 % to 15%.

The mean number of employees, mean duration of stay of an employee, and variation of the number of employees in grade I decreased and in grade II it was increasing as the promotion rate parameter (α) increased from -15 % to 15%. The mean number of employees, mean duration of stay of an employee, and variation of the number of employees in grade I and grade II decreased as the leaving rate parameter (β) in grad-1 increased from -15 % to 15%.When the leaving rate parameter (γ) in grad-2 is increased from -15 % to 15%, the mean number of employees, mean duration of stay of

Table 3a:*The values of $L_1(t)$,$L_2(t)$,$W_1(t)$,$W_2(t)$,$V1(t)$ and $V_2(t)$ for different Values of t, α, β, γ, a_1, b_1, a_2 and b_2*

Para -meters	Performance Measure	-15%	-10%	-5%	0%	+5%	+10%	+15%
t=2	L1	5.2497	5.9333	6.6624	7.4356	8.2524	9.1123	10.0152
	L2	16.7277	18.7380	20.9764	23.4109	26.0219	28.7973	31.7300
	W1	0.7539	0.8499	0.9530	1.0629	1.1792	1.3019	1.4308
	W2	3.5231	3.9175	4.3407	4.7972	5.2901	5.8205	6.3878
	VI	5.2497	5.9333	6.6624	7.4356	8.2524	9.1123	10.0152
	V2	16.7276	18.7379	20.9763	23.4109	26.0219	28.7973	31.7300
α=3	L1	7.8716	7.7206	7.5754	7.4356	7.3010	7.1711	7.0457
	L2	23.0525	23.1764	23.2957	23.4109	23.5220	23.6292	23.7329
	W1	1.2022	1.1528	1.1065	1.0629	1.0218	0.9831	0.9466
	W2	4.7743	4.7801	4.7879	4.7972	4.8075	4.8187	4.8305
	VI	7.8716	7.7206	7.5754	7.4356	7.3010	7.1711	7.0457
	V2	23.0525	23.1764	23.2957	23.4109	23.5220	23.6292	23.7329
β=4	L1	8.0287	7.8206	7.6232	7.4356	7.2571	7.0870	6.9247
	L2	23.7293	23.6162	23.5103	23.4109	23.3172	23.2286	23.1447
	W1	1.2549	1.1854	1.1216	1.0629	1.0086	0.9585	0.9120
	W2	4.8301	4.8173	4.8064	4.7972	4.7895	4.7833	4.7784
	VI	8.0287	7.8206	7.6232	7.4356	7.2571	7.0870	6.9247
	V2	23.7293	23.6161	23.5103	23.4109	23.3171	23.2286	23.1446

Table 3b: *The values of $L_1(t)$,$L_2(t)$,$W_1(t)$,$W_2(t)$,$V1(t)$ and $V_2(t)$ for different Values of t, α, β, γ, a_1, b_1, a_2 and b_2*

Para -meters	Performance Measure	-15%	-10%	-5%	0%	+5%	+10%	+15%
γ=5	L1	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	L2	26.8975	25.5983	24.4462	23.4109	22.4709	21.6107	20.8184
	W1	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629
	W2	6.3982	5.7752	5.2484	4.7972	4.4064	4.0650	3.7647
	VI	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	V2	26.8974	25.5983	24.4461	23.4109	22.4709	21.6107	20.8184
a1=5	L1	6.3205	6.6922	7.0639	7.4356	7.8074	8.1791	8.5508
	L2	22.8696	23.0500	23.2305	23.4109	23.5913	23.7717	23.9521
	W1	0.9046	0.9572	1.0100	1.0629	1.1158	1.1688	1.2218
	W2	4.7704	4.7742	4.7834	4.7972	4.8146	4.8351	4.8582
	VI	6.3205	6.6922	7.0639	7.4356	7.8074	8.1791	8.5508
	V2	22.8696	23.0500	23.2305	23.4109	23.5913	23.7717	23.9521
b1=3	L1	4.7683	5.5453	6.4298	7.4356	8.5783	9.8752	11.3459
	L2	22.2169	22.5701	22.9665	23.4109	23.9089	24.4666	25.0909
	W1	0.6870	0.7953	0.9200	1.0629	1.2257	1.4108	1.6209
	W2	4.6951	4.7717	4.7796	4.7972	4.8525	4.9345	5.0407
	VI	4.7683	5.5453	6.4298	7.4356	8.5783	9.8752	11.3459
	V2	22.2169	22.5701	22.9664	23.4109	23.9089	24.4666	25.0909
a2=10	L1	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	L2	20.4589	21.4429	22.4269	23.4109	24.3948	25.3788	26.3628
	W1	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629
	W2	4.1923	4.3939	4.5955	4.7972	4.9988	5.2004	5.4021
	VI	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	V2	20.4589	21.4429	22.4269	23.4109	24.3948	25.3788	26.3628
b2=3	L1	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	L2	16.4841	18.5091	20.8066	23.4109	26.3602	29.6977	33.4714
	W1	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629
	W2	3.3778	3.7927	4.2635	4.7972	5.4015	6.0854	6.8587
	VI	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	V2	16.4841	18.5091	20.8066	23.4109	26.3602	29.6977	33.4714
N=1500	L1	4.0256	4.0276	4.0297	4.0318	4.0338	4.0359	4.0380
	L2	13.4547	13.5138	13.5729	13.6321	13.6912	13.7503	13.8094
	W1	0.5855	0.5858	0.5861	0.5864	0.5866	0.5869	0.5872
	W2	2.8212	2.8257	2.8307	2.8360	2.8418	2.8479	2.8543
	VI	4.0256	4.0276	4.0297	4.0318	4.0338	4.0359	4.0380
	V2	13.4537	13.5128	13.5719	13.6310	13.6901	13.7491	13.8082
M=500	L1	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	L2	23.4075	23.4086	23.4097	23.4109	23.4120	23.4131	23.4143
	W1	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629	1.0629
	W2	4.7969	4.7970	4.7971	4.7972	4.7973	4.7974	4.7975
	VI	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356	7.4356
	V2	23.4075	23.4086	23.4097	23.4109	23.4120	23.4131	23.4143

The mean number of employees, mean duration of stay of an employee, and variation of the number of employees in grade I and grade II increased when the recruitment rate parameter (a1) in grade I increased from -15 % to 15%. The mean number of employees, mean duration of stay of an employee, and variation of the number of employees in grade I and grade II increased when the recruitment rate parameter (b1) in grade I increased from -15 % to 15%.

When the recruitment rate parameter (a2) of employees in grade II goes from -15 % to 15%, the mean number of employees, the mean duration of stay of an employee, and the variance of the number of employees in grade I are unaffected, while they increase in grade II. When the recruitment rate parameter (b2) of employees in grade II is increased from -15 % to 15%, the mean number of employees, mean duration of stay of an employee, and variance of the number of employees in grade I are unaffected, but they increase in grade II.

V1.Comparative Studies of the models:

In this part, a comparison of the generated model with a model with homogeneous Poisson recruitment was shown. Table 4 shows the performance measures of both models for various values of t=1.6, 1.7, 1.8, 1.9, and 2.

Table 4: *Comparative study of models with Homogeneous and Non-Homogeneous Recruitments.*

t	Characteristics Measured	Non-Homogeneous recruitment	Homogeneous recruitment	Deference	Percentage of Variation
t=1.6	L1	2.1021	0.7348	1.3673	65.0445
	L2	4.7680	2.3187	2.4493	51.3695
	W1	0.3421	0.2017	0.1404	41.0406
	W2	1.1085	0.6329	0.4756	42.9048
	VI	2.1021	0.7348	1.3673	65.0445
	V2	4.7680	2.3187	2.4493	51.3695
t=1.7	L1	2.2347	0.7245	1.5102	67.5795
	L2	4.7067	1.9718	2.7349	58.1065
	W1	0.3575	0.2008	0.1567	43.8322
	W2	1.1500	0.6342	0.5158	44.8522
	VI	2.2347	0.7245	1.5102	67.5795
	V2	4.7067	1.9718	2.7349	58.1065
t=1.8	L1	2.3724	0.7193	1.6531	69.6805
	L2	4.7803	1.7598	3.0205	63.1864
	W1	0.3738	0.2003	0.1735	46.4152
	W2	1.2040	0.6611	0.5429	45.0914
	VI	2.3724	0.7193	1.6531	69.6805
	V2	4.7803	1.7598	3.0205	63.1864
t=1.9	L1	2.5127	0.7168	1.7959	71.4729
	L2	4.9365	1.6303	3.3062	66.9746
	W1	0.3906	0.2001	0.1905	48.7711
	W2	1.2579	0.6972	0.5607	44.5743
	VI	2.5127	0.7168	1.7959	71.4729
	V2	4.9365	1.6303	3.3062	66.9746
t=2.0	L1	2.6543	0.7155	1.9388	73.0437
	L2	5.1432	1.5513	3.5919	69.8378
	W1	0.4079	0.2000	0.2079	50.9684
	W2	1.3082	0.7320	0.5762	44.0453
	VI	2.6543	0.7155	1.9388	73.0437
	V2	5.1432	1.5513	3.5919	69.8378

Table 4 shows that the percentage variation of the performance measures between the two models increased as time progressed. The assumption of the Duane recruitment process was found to have a considerable impact on all of the manpower model's performance measures. Time has a substantial impact on system performance, and the proposed model can more correctly forecast system performance.

VII. Conclusion

The purpose of this work is to build and analyze a two-graded manpower model with direct recruitment in both grades for non-stationary recruitment. The Duane recruitment procedure was capable of identifying recruitments that were time-dependent. The model's characteristics were obtained explicitly to assist HR Managers in adopting optimal operating policies, such as the mean number of employees in each grade, the mean duration of stay of an employee in each grade, the variance of the number of employees in each grade, and the coefficient of variation of an employee in each grade in the organization. The model's sensitivity study revealed that the Duane recruitment process has a considerable impact on system performance indicators. A comparison of the suggested model with Poisson recruitment reveals that the proposed model predicts system properties more accurately. When the recruiting was done in a time-dependent manner, the performance measures could be anticipated more correctly and realistically using the evolving model. This model can also be expanded by factoring in cost considerations while determining the best values for the parameters, which will be done elsewhere.

References

[1] Seal,H,. (1945), “The mathematics of a population composed of k stationary strata each recruited from the stratum below the supported at the lowest level by a uniform number”, Biometrika, Vol. 33, pp. 226 –230.

[2] Silcock, H (1954), “The Phenomenon of the labor turnover”, J.R. Statist. Soc. A.117, No. 9, pp. 429 – 440.

[3] Bartholomew, D. J. (1963)."A multistage renewal process. J.R". Statist. Soc. B. 25, 152-168.

[4] Bartholomew,D. J (1971)."The statistical approach to manpower planning", Statistician, 20(1), 3-26.

[5] Ugwuowo, F.I. And Mc Clean, S.I. (2000), “Modeling heterogeneity in a manpower system”: a review Applied Stochastic Models and Data Analysis, Vol.16, No.2, pp. 99-110.

[6] Wang, J. (2005), “A review of operations research applications in workforce planning and potential modeling of military training”, DSTO Systems Sciences Laboratory, Edinburgh South Australia 5111, pp.1-37.

[7] Srinivasa Rao, K., Srinivasa Rao, V. And Vivekananda Murthy, M (2006), “On two graded manpower planning model”, OPSEARCH, Vol. 43, No. 3, pp. 117-130.

[8] Kannan Nilakantan et al., (2014), “Evaluation of staffing policies in Markov manpower systems and their extension to organizations with outsource personnel”, Journal of the Operational Research Society, Vol. 66, pp. 1324–1340.

[9] Jeeva, M and Geetha, N (2013), “Recruitment Model in Manpower Planning Under Fuzzy Environment”, British Journal of Applied Science & Technology, Vol. 3 ,issue 4,pp. 1380–1390.

[10] Lalitha, A Devi And Srinivasan,A (2014), “A Stochastic Model On The Time To Recruitment For A Single Grade Manpower System With Attrition Generated By A Geometric Process of inter-Decision Times Using Univariate Policy Of Recruitment”, Blue Ocean Research Journals, Vol. 3, No 7 .pp.12-15.

[11] A.Osagiede, A and Ekhosuehi,U.V (2015),"Finding a continuous-time Markov chain via sparse stochastic measuresesn manpower systems", Journal of the Nigerian Mathematical Society, Vol. 34,No. 1, pp. 94-105.

[12] Murthy.M.Vivekananda, Srinivasa Rao, V And Srinivasa Rao. K (2003), "Three graded manpower planning model", Proceeding of APORS, pp. 410-418.

[13] Konda Babu And Srinivasa Rao (2013), "Studies on two graded manpower model with bulk recruitment in both the grades", International Journal of Human Resource Management Research Vol. 1, No. 2, pp. 10-15.

[14] Srinivasa Rao, And Konda Babu (2014), "Grade Manpower Model with Bulk Recruitment in First Grade", International Journal of Human Resource Management Research and Development, 1465-6612(print), Vol.4, No.1, pp.1-37.

[15] Govinda rao and Srinivasa Rao (2013), "Bivariate manpower model for permanent and temporary grades under equilibrium", Indian Journal of Applied Research Vol. 3, No.8, pp. 63 – 66.

[16] Govinda Rao and Srinivasa Rao (2014), "Manpower model with three level recruitment in the initial grade", International Research Journal of Management Science and Technology, Vol.5, No.9, pp. 79-102.

[17] Parameswari, K And Srinivasan. (2016), "Estimation of variance of time to recruitment for a two-grade manpower system with two types of decisions when the wastages form a geometric process", International Journal of Mathematics Trends and Technology, Volume 33, Number 3, pp.161-165.

[18] Amudha,T and Srinivasan,A.-(2016), "Estimation of Variance of Time to Recruitment for a Two Grade Manpower System with Single Source of Depletion Using a Different Probabilistic Analysis", International Journal of Scientific Engineering and Applied Science, Volume-2, wassue-7,pp.135-140.

[19] Saral Et Al, L. (2017) ,"Estimation of mean time to recruitment for a two graded manpower system with two thresholds, different epoch for exits and correlated inter-decisions under correlated wastage", International Educational Scientific Research Journal, Volume : 3 ,Issue : 3, pp.1-6.

[20] Srividhya,K & Sendhamizhselvi,S (2017), "Mean time to recruitment for a multi grade manpower system with single threshold, single source of depletion when wastages form an order statistics",nternational Journal of Current Research and Modern Education, Special issue, pp.30-37.

[21] Jayanthi, L, S & Uma K P, (2018)," Determination of manpower model for a single grade system with two sources of depletion and two components for threshold", EAI Endorsed Transactions on Energy Web and information Technologies, Volume 5, Issue 20,pp.1-10.

[22] Tamás Bányai., Christian Landschützer., And Ágota Bányai., (2018)," Markov-Chain Simulation-Based Analysis of Human Resource Structure: How Staff Deployment and Staffing Affect Sustainable Human Resource Strategy", Sustainability, Volume 10, pp. 1-21.

[23] Arokkia Saibe.,P (2019)," Variance of Time to Recruitment with A Record of N Decision", International Journal of Scientific Research and Reviews, 8(1), 1207-1212.

[24] Srinivasa Rao, Mallikharjuna Rao (2015), "On two graded manpower model with non-homogeneous Poisson recruitments", International Journal of Advanced Computer and Mathematical Sciences. Vol 6, wassue3. Pp40-66.

[25] Srinivasa rao,K and Ganapathi Swamy,ch(2019)" On Two Grade Manpower Model With Duane Recruitment Process", Journal of Applied Science and Computations, Volume VI, Issue, pp 220-235.

Parallel System Analysis with Priority and Inspection Using Semi-Markov Approach

Neetu Dabas
•
Department of Statistics, AIAS,
Amity University, Noida, Uttar Pradesh-201 313
Email: neetudabas2@gmail.com

Reetu Rathee
•
Department of Statistics, AIAS,
Amity University, Noida, Uttar Pradesh-201 313
Email: rrathee@amity.edu

Abstract

In this paper a parallel system has been discussed with the idea of priority to preventive maintenance over replacement. The system has two identical units and facility of inspection is given to the failed unit before repair/replacement. There is a single server who play four-in-one role of inspection, replacement, repair and preventive maintenance and comes immediately when required. Units are failed with constant rate whereas failure time is random. The distribution of time for repair activities is arbitrary and there rates follow exponential distribution. The random variable associate with different rates are stochastically independent. Mathematical expression for several reliability terms like MTSF, availability, busy period analysis for server , expected number of visits by the server and cost benefit are obtained by using semi-markov process and regenerative point technique. Graphs are drawn to find the effect of various parameters on MTSF, Availability and profit.

Keywords: parallel system, priority, preventive maintenance, replacement, inspection

I. Introduction

The world is moving day by day towards the smart technology. Advance development of technology has significantly increased cost and complexity of industrial systems. Thus it has become an essential to operate industrial systems with minimum down time in order to achieve optimized production, increase profit and to avoid the losses. Hence, the need for reliability modeling and analysis of complex industrial systems is inevitable. Reliability analysis of parallel systems has been broadly studied by many researchers because parallel configuration is more reliable then series. Dhillon and Viswanath [6] analyzed a parallel system with the common-cause failure. Sridharan and Kalyani [7] gives common-cause failure analysis of a two non-identical unit parallel system using GERT technique.

A parallel (2n-2) system was investigated by E.Papageorgious and G. Kokolakis [9] where two units start operation simultaneously and any one of them was replaced by one of (n-2) warm standby units upon failure. Reliability of parallel systems was studied based on multiple competing dependent failure process by Sanling and David [10]. This problem was formulated for

the conditioning on shock sets because not all shocks cause degradation. N. Sharma and J.P.Singh Joorel [8] investigate a two unit parallel system with inspection and preparation time for replacement. To improve the reliability PM and Inspection are widely used in modern engineering systems. PM can help to extend the life time of an equipment, increase the productivity and hence decrease unexpected maintenance spending. Inspection aims to find the defect of the system and type of the defect. Goel and Gupta [13] considered a two identical unit parallel system with the concept of PM, inspection and two types of repair. They assumed that the time to failure, commencement to PM and inspection are constant while repair and maintenance times are arbitrarily distributed. M.K.Kakkar and J.Bhatti [12] purposed a two dissimilar parallel unit framework under the presumption that the unit may also fail during the preventive maintenance (PM). Wang and Lin [5] found a methodology to optimize the non-periodic maintenance for a series-parallel system. Shruti [11] analyzed a stochastic model with two units subject to routine inspection, maintenance and replacement. In this research routine inspection is conducting over operative unit and after inspection either the unit is maintained or it failed. Repair and replacement of the unit is based upon guarantee period of the equipment. In many research priority concept is also used to make the system better in performance and hence more profitable. P. Kumar, A. Bharti, and A. Gupta [3] investigated and analyze a two unit parallel system in which priority was given to one unit over other. In this system priority unit was repairable and non-priority unit was non-repairable and preference to repair of priority unit was given over replacement of non-priority unit. R.Rathee and D.Pawar [4] introduced a reliability model of a parallel system in which priority to repair over replacement was given using maximum operation and repair times.A.Kumar and S.C.Malik [1] considered a computer system with two identical units and in each unit H/W and S/W components works together. Priority was given to PM of the unit over S/W replacement under certain assumptions. C.Aggarwal and N.Ahlawat [2] done the profit analysis of a standby system with priority to PM over repair by considering rest of the server between repairs.

In the present study a parallel system is investigate under some assumptions. Priority is given to PM over replacement with inspection. Inspection is done to find the type of failure. Semi-Markovian approach and regenerative-point-technique are used to obtain numerical expressions for various reliability terms such as MTSF, Availability, busy period analysis for server, expected number of visits by the server and cost benefit. Graphical interpretation is done to visualize the effect of several parameters (related to repair activities) on obtained reliability measurable terms.

For practical implication of system one of the example is a parallel compressor rack system. Jim Coats [14] gives the commercial and industrial applications of parallel compressor racks. There are several advantages of this system like low installation cost, capacity control, redundancy and maximum efficiency etc.

Figure 1: *Parallel compressor Rack*

II. Notations for System Model

S_i : $States\ of\ the\ sytem\ (i = 1,2,\ldots.,15)$
λ : Constant failure rate
$a/b/\propto_0$: repair/replacement/PM rate of the system
upm/UPM: Unit is under PM/continuosly under PM
wpm/WPM: Unit is waiting for PM/continue waiting for PM
FUi/FWi: Failed unit under inspection/waiting for inspection
FUI/FWI: Failed unit continuously under inspection/waiting for inspection
FUr/FUrp : Failed unit under repair/replacement
FUR/FURP: Failed unit continuously under repair/replacement
μ_i: $mean\ sojourn\ time\ in\ the\ state\ S_i$
m_{ij}: Contribution to mean sojourn time in state S_i when the system transits directly to state S_j
h(t)/f(t)/r(t)/g(t): pdf of the inspection/repair/replacement/PM time

III. System Description and Assumptions

A parallel system with two identical units is studied under some practical assumptions which are given below:

- Initially both the units are in operative condition
- Failure rate is constant
- Only one server operator is taken to do all repair activities
- Repair is perfect or units restore in initial condition after repair
- Post failure inspection is done to find unit is repairable or replaced by new
- Time taken for repair activities is arbitrary and there rates follow exponential distribution
- The random variable associate with different rates are stochastically independent

Table 1: *Description of the states*

States	Description
S_0	Both the units are in normal mode
S_1	One unit is operative and other is failed under inspection
S_2	Resume for PM

S_3	One is working and other is failed under replacement
S_4	One unit is continuously under inspection from previous state and other is waiting for inspection
S_5	One is working and other is failed under repair
S_6	One unit is continuously under inspection from previous state and other is waiting for PM
S_7	One unit is working and other under PM
S_8	One unit is continuously under replacement from previous state and other is waiting for inspection
S_9	One unit is under repair and other is continuously waiting for inspection from previous state
S_{10}	One unit is under replacement and other is continuously waiting for inspection from previous state
S_{11}	One unit is continuously under repair from previous state and other is waiting for inspection
S_{12}	One unit is continuously under repair from previous state and other is waiting for PM
S_{13}	One unit is under repair and other is continuously waiting for PM from previous state
S_{14}	One unit is under PM and other is continuously waiting for replacement from previous state
S_{15}	One unit is continuously under PM from previous state and other is waiting for inspection

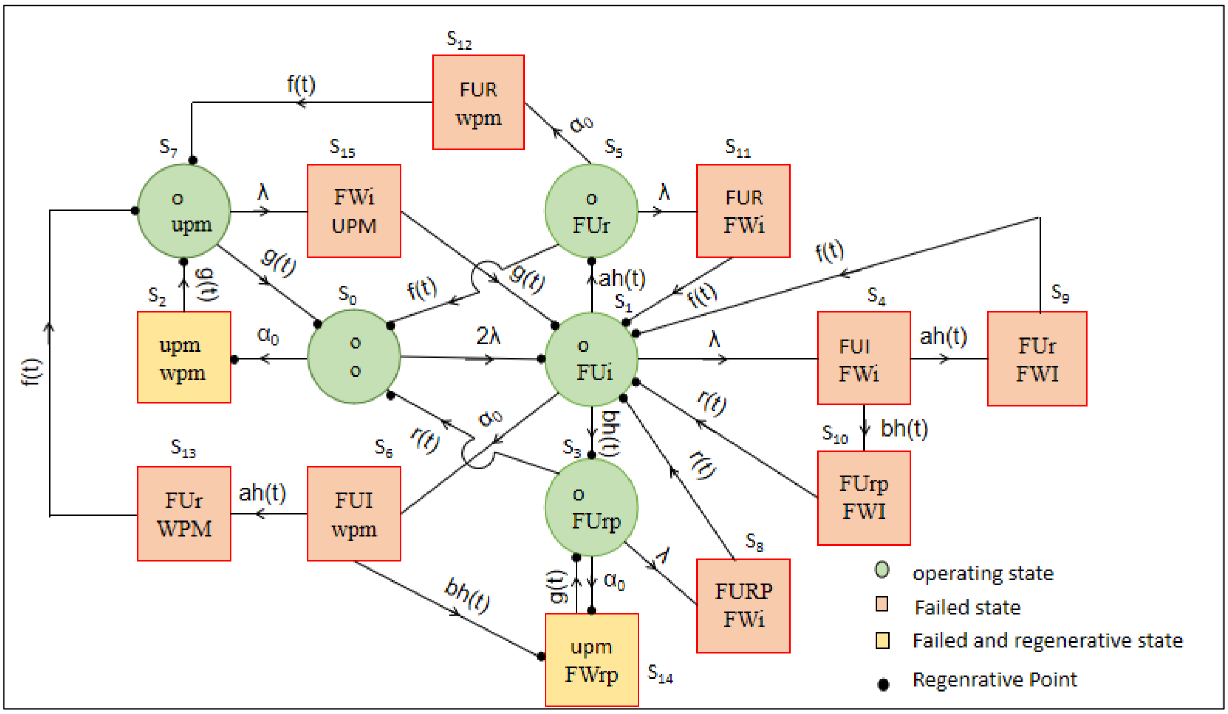


Figure 2: *Transition State Diagram*

IV. Formulation and Stochastic Analysis of the Model

I. Transition Probabilities & Mean Sojourn Times (μ_i)

Steady- state transition probabilities from regenerative state i to state j are given by the formula

$$p_{ij} = Q_{ij}(\infty) = \int_0^\infty q_{ij}(t)dt \quad (1)$$

Here $Q_{ij}(t)$ / $q_{ij}(t)$ are the c.d.f/p.d.f from state i to state j in (0,t) time.

$p_{01} = \frac{2\lambda}{2\lambda+\alpha_0}$, $p_{02} = \frac{\alpha_0}{2\lambda+\alpha_0}$

$p_{13} = bh^*(\lambda+\propto_0)$, $p_{14} = \frac{\lambda}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$, $p_{15} = ah^*(\lambda+\propto_0)$, $p_{16} = \frac{\propto_0}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$

$p_{30} = r^*(\lambda+\propto_0)$, $p_{38} = p_{31.8} = \frac{\lambda}{\lambda+\propto_0}(1-r^*(\lambda+\propto_0))$, $p_{3,14} = \frac{\propto_0}{\lambda+\propto_0}(1-r^*(\lambda+\propto_0))$

$p_{49} = p_{6,13} = a$, $p_{4,10} = p_{6,14} = b$

$p_{50} = f^*(\lambda+\propto_0)$, $p_{5,11} = p_{51.11} = \frac{\lambda}{\lambda+\propto_0}(1-f^*(\lambda+\propto_0))$, $p_{5,12} = p_{57.12} = \frac{\propto_0}{\lambda+\propto_0}(1-f^*(\lambda+\propto_0))$

$p_{70} = g^*(\lambda)$, $p_{7,15} = p_{71.15} = 1-g^*(\lambda)$,

$p_{11.49} = \frac{\lambda a}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$, $p_{11.4,10} = \frac{\lambda b}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$,

$p_{1,14.6} = \frac{\propto_0 b}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$, $p_{17.6,13} = \frac{\propto_0 a}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$,

$p_{27} = p_{81} = p_{91} = p_{10,1} = p_{11,1} = p_{12,7} = p_{13,7} = p_{14,3} = p_{15,1} = 1$

It can verified that

$p_{01} + p_{02} = p_{13} + p_{15} + p_{11.49} + p_{11.4,10} + p_{17.6,13} + p_{1,14.6} = p_{50} + p_{51.11} + p_{57.12} = p_{30} + p_{31.8} + p_{3,14} = p_{49} + p_{4,10} = p_{6,13} + p_{6,14} = p_{70} + p_{71.15} = p_{27} = p_{81} = p_{91} = p_{10,1} = p_{11,1} = p_{12,7} = p_{13,7} = p_{14,3} = p_{15,1} = 1$

And $\mu_i'^s$ are given by the formula

$$\mu_i = E(t) = \int_0^\infty P(T>t)\,dt = \sum_j m_{ij} \quad (2)$$

and

$$m_{ij} = \frac{d[Q_{ij}^{**}(s)]}{ds}|s=0 \quad (3)$$

$\mu_0 = \frac{1}{2\lambda+\alpha_0}$, $\mu_1 = \frac{1}{\lambda+\alpha_0}(1-h^*(\lambda+\alpha_0))$, $\mu_3 = \frac{1}{\lambda+\alpha_0}(1-r^*(\lambda+\alpha_0))$

$\mu_5 = \frac{1}{\lambda+\alpha_0}(1-f^*(\lambda+\alpha_0)$, $\mu_7 = \frac{1}{\lambda}(1-g^*(\lambda))$

$\mu_1' = [\frac{1}{\lambda+\alpha_0} + \frac{\lambda b}{\alpha(\lambda+\alpha_0)} + \frac{1}{\gamma} + \frac{a}{\alpha}](1-h^*(\lambda+\alpha_0))$

$\mu_3' = \frac{(\beta+\lambda)}{\beta(\lambda+\alpha_0)}(1-f^*(\lambda+\alpha_0))$, $\mu_5' = \frac{1}{\propto}$, $\mu_7' = \frac{1}{\theta}$

II. Reliability & Mean Time to System Failure (MTSF)

Let $\Phi_i(t)$ be the cdf of first passing time from the state S_i to the state in which failure occur and we take absorbing state as the failed state. So, the expressions for $\Phi_i(t)$ from which MTSF of discussed system is obtained are given as

$$\Phi_i(t) = \sum_{i,j} Q_{ij}(t) \circledS \Phi_j(t) + \sum_{i,k} Q_{ik}(t) \quad (4)$$

Where j is the operating regenerative state to which the given regenerative state i can transit and k is the failed state to which state i can directly transit.

If we take LST of above relation (4) and solved them for $\Phi_0^{**}(s)$, we have

$$R^*(s) = \frac{1-\Phi^{**}(s)}{s} \quad (5)$$

The system reliability is obtained by taking Inverse Laplace transform of (5) and MTSF is given by the formula

$$\text{MTSF} = \lim_{s\to 0}\frac{1-\Phi^{**}(s)}{s} = \frac{\text{N}}{\text{D}} \quad (6)$$

Where,

$$N = \mu_0 + \mu_1 p_{01} + \mu_3 p_{01} p_{13} + \mu_5 p_{01} p_{15} \text{ and } D = 1 - p_{01} p_{13} p_{30} - p_{01} p_{15} p_{50} \quad (7)$$

III. Analysis of Availability

Let $A_i(t)$ be the probability of system working at time 't' w.r.t the condition that system goes to regenerative state S_i at t = 0. We have the relations for $A_i(t)$ as

$$A_i(t) = M_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © A_j(t) \quad (8)$$

Where j is any successive regenerative state to which the regenerative state i can transit through n transitions.

M_i (t) is the probability that the system in up state S_i up to the time t without visiting to any other regenerative state.

$$M_0(t) = e^{-(2\lambda + \alpha_0)t} \quad (9)$$

$$M_1(t) = e^{-(\lambda+\alpha_0)t}\overline{H(t)} \quad (10)$$

$$M_3(t) = e^{-(\lambda+\alpha_0)t}\overline{R(t)} \quad (11)$$

$$M_5(t) = e^{-(\lambda+\alpha_0)t}\overline{F(t)} \quad (12)$$

$$M_7(t) = e^{-(\lambda t)}\overline{G(t)} \quad (13)$$

Now, if we use LT of (8) and solved it for $A_0^*(s)$.We get the result for steady state availability as

$$A_0(\infty) = \lim_{s\to 0} sA_0^*(s) = \frac{N_1}{D_1} \quad (14)$$

Where

$$N_1 = \mu_0 A + (\mu_1 + \mu_5 p_{15})B + \mu_3 C + \mu_7 D \quad (15)$$

$$D_1 = (\mu_0 + \mu_2 p_{02})A + (\mu_1' + \mu_5' p_{15})B + \mu_3' C + \mu_7' D + \mu_{14} E \quad (16)$$

IV. Busy Period Analysis for Server

Let $B_i^I(t), B_i^R(t), B_i^{Rp}(t), B_i^P(t)$be the probability of busy period of server during inspection, repair, replacement and PM at instant't' with the given condition that the system go to regenerative state Si at t=0. The recursive relations for $B_i^I(t), B_i^R(t), B_i^{Rp}(t), B_i^P(t)$ are as follows:

$$B_i^I(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © B_j^I(t) \quad (17)$$

$$B_i^R(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © \text{B}_j^R(t) \quad (18)$$

$$B_i^{Rp}(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © \text{B}_j^{Rp}(t) \quad (19)$$

$$B_i^P(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © \text{B}_j^P(t) \quad (20)$$

Where j is any successive regenerative state to which the regenerative state i can transit through n transitions.

$W_i(t)$ is the probability of server busyness at state S_i due to repair activities at time t without making any transition to any other regenerative state or returning to the same via one or more non regenerative state.

Here,

$$W_1(t) = e^{-(\lambda+\alpha_0)t}\overline{H(t)} + (\lambda e^{-(\lambda+\alpha_0)t} © 1)\overline{H(t)} + (\alpha_0 e^{-(\lambda+\alpha_0)t} © 1)\overline{H(t)} \quad (21)$$

$$W_5(t) = e^{-(\lambda+\alpha_0)t}\overline{F(t)} + (\lambda e^{-(\lambda+\alpha_0)t} © 1)\overline{F(t)} + (\alpha_0 e^{-(\lambda+\alpha_0)t} © 1)\overline{F(t)} \quad (22)$$

$$W_3(t) = e^{-(\lambda+\alpha_0)t}\overline{R(t)} + (\lambda e^{-(\lambda+\alpha_0)t} © 1)\overline{R(t)} \quad (23)$$

$$W_2(t) = \overline{G(t)} = W_{14}(t), W_7(t) = e^{-(\lambda)t}\overline{G(t)} + (\lambda e^{-(\lambda)t} © 1)\overline{G(t)}$$

Take LT of (17) to (20) and solving for $B_0^{I^*}(s), B_0^{R^*}(s), B_0^{Rp^*}(s), B_0^{P^*}(s)$ .The busy time in inspection, repair, replacement and preventive maintenance for server is given by

$$B_0^I(\infty) = \lim_{s\to 0} sB_0^{I^*}(s) = \frac{N_2}{D_1} \quad (24)$$

$$B_0^R(\infty) = \lim_{s\to 0} sB_0^{R^*}(s) = \frac{N_3}{D_1} \quad (25)$$

$$B_0^{Rp}(\infty) = \lim_{s\to 0} sB_0^{Rp^*}(s) = \frac{N_4}{D_1} \quad (26)$$

$$B_0^P(\infty) = \lim_{s\to 0} sB_0^{P^*}(s) = \frac{N_5}{D_1} \quad (27)$$

Here,

$N_2 = W_1^*(0)B, N_3 = W_5^*(0)p_{15}B, N_4 = W_3^*(0)C$

$N_5 = W_2^*(0)p_{02}A + W_7^*(0)D + W_{14}^*(0)E$ and D1 is mentioned above. (28)

V. Expected Number of Visits by The Server

Consider $I_0(t), R_0(t), Rp_0(t), Pm_0(t)$ as the expected number of visits make by the server for inspection, repair, replacement and PM in (0, t] .We have the following recursive relations for $I_0(t), R_0(t), Rp_0(t), Pm_0(t)$ are

$$I_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + I_j(t)) \quad (29)$$

$$R_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + R_j(t)) \quad (30)$$

$$Rp_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + Rp_j(t)) \quad (31)$$

$$Pm_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + Pm_j(t)) \quad (32)$$

Where j is any successive regenerative state to which the regenerative state i can transit through n transitions and and C = 1 if j is the state where the server does the job afresh, otherwise C = 0.

Take LST of above equations solving for $I_0^{**}(s), R_0^{**}(s), Rp_0^{**}(s), Pm_0^{**}(s)$.The expected number of inspections, repairs, replacements, and preventive maintenance by the server is given by (per unit time)

$$I_0(\infty) = \lim_{s\to 0} sI_0^{**}(s) = \frac{N_6}{D_1} \quad (33)$$

$$R_0(\infty) = \lim_{s\to 0} sR_0^{**}(s) = \frac{N_7}{D_1} \quad (34)$$

$$Rp_0(\infty) = \lim_{s\to 0} sRp_0^{**}(s) = \frac{N_8}{D_1} \quad (35)$$

$$Pm_0(\infty) = \lim_{s\to 0} sPm_0^{**}(s) = \frac{N_9}{D_1} \quad (36)$$

Where,

$N_6 = B, N_7 = (p_{11.49} + p_{15} + p_{17.6,13})B,\ N_8 = (p_{11.4,10} + p_{1,14.6} + p_{13})B,$

$N_7 = p_{02}A + D + E$ and D_1 is already mentioned. (37)

Here A, B, C, D & E are

$$A = (1 - p_{3,14})(p_{15}p_{50} + p_{70}(p_{15}p_{57.12} + p_{17.6,13})) + (p_{13} + p_{1,14.6})p_{30} \quad (38)$$

$$B = (1 - p_{3,14})(1 - p_{02}p_{70}) \quad (39)$$

$$C = (p_{13} + p_{1,14.6})(1 - p_{02}p_{70}) \quad (40)$$

$$D = (1 - p_{3,14})(p_{15} + p_{17.6,13} - p_{01}p_{15}p_{50} - p_{15}p_{51.11}) + p_{02}p_{30}(p_{13} + p_{1,14.6}) \quad (41)$$

$$E = (p_{13}p_{3,14} + p_{1,14.6})(1 - p_{02}p_{70}) \quad (42)$$

VI. Profit Analysis

In steady state the profit function of the system model can be obtained as

$$P = k_0A_0 - k_1B_0^I - k_2B_0^R - k_3B_0^{Rp} - k_4B_0^{Pm} - k_5I_0 - k_6R_0 - k_7Rp_0 - k_8Pm_0 \quad (43)$$

Here,

P = Profit function of system model

k_0 = Revenue per unit up − time of the system
k_1, k_2, k_3, k_4 = Cost per unit time of the server when it is busy in inspection, repair, replacement, preventive maintenance
k_5, k_6, k_7, k_8 = Cost per unit time for inspection, repair, replacement, preventive maintenance

V. Analytical Study of the Model

In the present study concept of priority to PM over replacement with inspection is used. To see the applicability of this situation particular cases are taken for the included parameters like $F(t) = \propto e^{-\propto t}, r(t) = \beta e^{-\beta t}, h(t) = \gamma e^{-\gamma t}, g(t) = \theta e^{-\theta t}$. Results are obtained in form of tables and graphs by taking random values for the given parameters. On the basis of these cases numerical and graphical results are obtained which shows the effect of these parameters on MTSF, Availability and Profit function of the system. From the obtained results we conclude that PM does not effect the MTSF and as the failure rate increases the availability and the profit of the system is decreases. When the repair activities rate is increases then availability and profit is also increases. From the results we obtained that the system is highly profitable if we increase the PM rate at the very first stage when failure rate is very low but as the failure rate increases the system get more profit by enhancing the inspection rate. Tabular and graphical results are given below:

Table 2: *Values of MTSF w.r.t various parameters*

Failure rate	α=2.1,β=2,a=0.6,b=0.4, γ=1.3,α0=3	α= 4.1	β=5	γ= 3	α0= 3.1	a= 0.4, b= 0.6
0.1	0.33274	0.33275	0.33275	0.33279	0.32204	0.33274
0.2	0.33113	0.33119	0.33119	0.33130	0.32056	0.33113
0.3	0.32875	0.32887	0.32886	0.32908	0.31838	0.32875
0.4	0.32577	0.32596	0.32594	0.32627	0.31563	0.32577
0.5	0.32235	0.32260	0.32258	0.32302	0.31247	0.32234
0.6	0.31859	0.31890	0.31887	0.31943	0.30899	0.31858
0.7	0.31458	0.31495	0.31492	0.31558	0.30527	0.31457
0.8	0.31040	0.31082	0.31079	0.31153	0.30138	0.31039
0.9	0.30610	0.30657	0.30653	0.30736	0.29737	0.30609

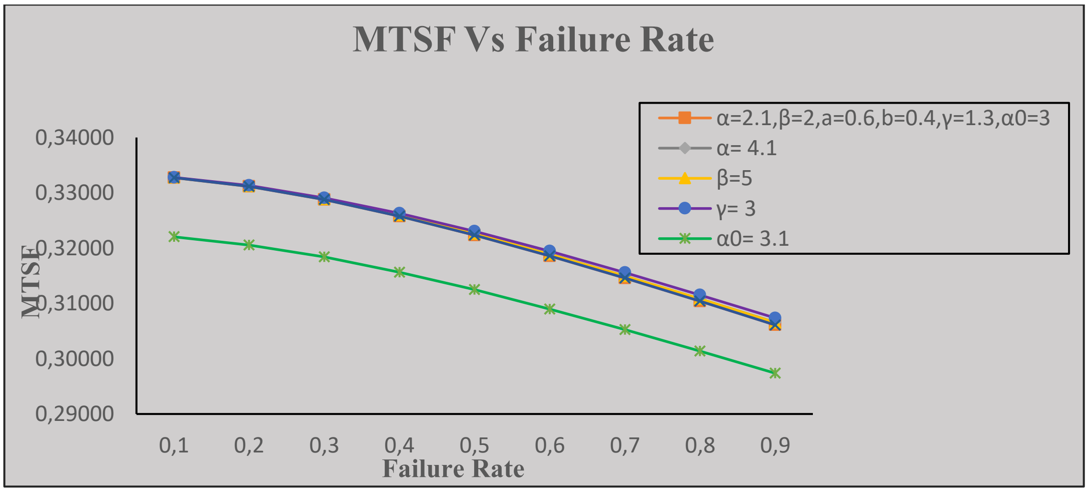


Figure 3: *MTSF VS Failure Rate*

Table 3: *Values of Availability w.r.t various parameters*

Failure Rate	α=2.1,β=2,a=0.6,b=0.4,γ=1.3, α0=3,θ=1.4	α=4.1	β=5	a=0.4,b=0.6	γ=3	α0=3.1	θ=2
0.1	0.53503	0.53911	0.54102	0.52635	0.55017	0.53213	0.56972
0.2	0.45652	0.47175	0.46481	0.45459	0.48943	0.45348	0.48158
0.3	0.42387	0.43952	0.43365	0.41815	0.46238	0.42072	0.45061
0.4	0.39668	0.41256	0.40718	0.38870	0.43924	0.39350	0.42408
0.5	0.37357	0.38955	0.38436	0.36427	0.41912	0.37041	0.40101
0.6	0.35358	0.36960	0.36443	0.34357	0.40138	0.35048	0.38067
0.7	0.33606	0.35206	0.34683	0.32572	0.38556	0.33303	0.36257
0.8	0.32052	0.33647	0.33114	0.31012	0.37131	0.31758	0.34631
0.9	0.30661	0.32248	0.31705	0.29630	0.35836	0.30376	0.33159

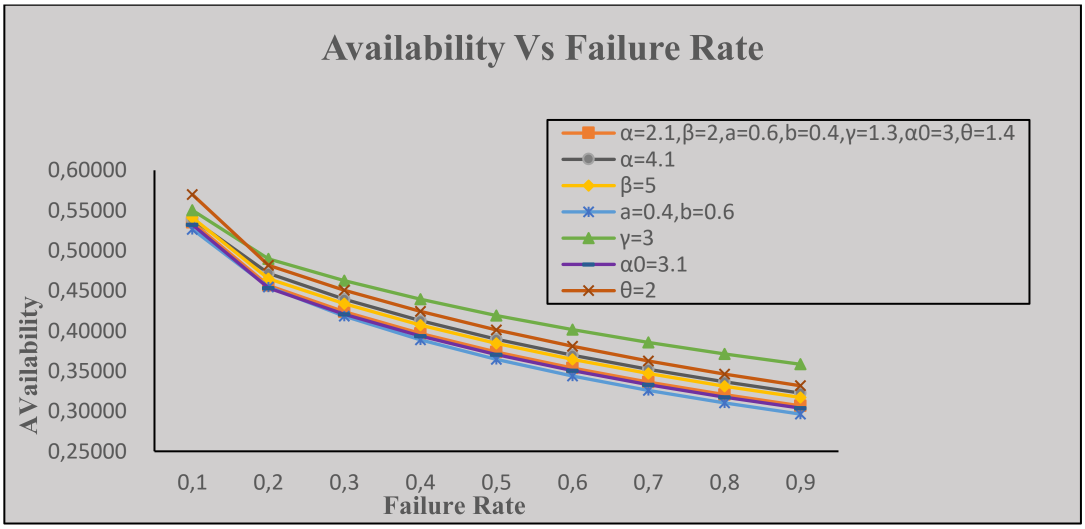


Figure 4: *Availability VS Failure Rate*

Table 4: *Values of Profit w.r.t various parameters*

Failure Rate	α=2.1,β=2,a=0.6,b=0.4,γ=1.3, α0=3,θ=1.4	α=4.1	β=5	a=0.4,b=0.6	γ=3	α0=3.1	θ=2
0.1	7355.86	7416.93	7463.24	7208.33	7569.95	7312.12	7759.37
0.2	6014.32	6239.56	6166.20	5957.81	6474.39	5970.56	6258.77
0.3	5391.47	5618.42	5571.51	5271.16	5911.14	5347.74	5648.91
0.4	4873.10	5098.06	5067.35	4713.77	5427.64	4830.27	5130.70
0.5	4433.01	4653.80	4633.45	4249.77	5006.14	4391.57	4683.55
0.6	4053.32	4268.64	4255.27	3855.69	4634.00	4013.52	4292.79
0.7	3721.35	3930.44	3922.08	3515.48	4301.91	3683.31	3947.67
0.8	3427.88	3630.28	3625.81	3217.79	4002.92	3391.63	3640.12
0.9	3166.00	3361.48	3360.25	2954.37	3731.65	3131.53	3363.95

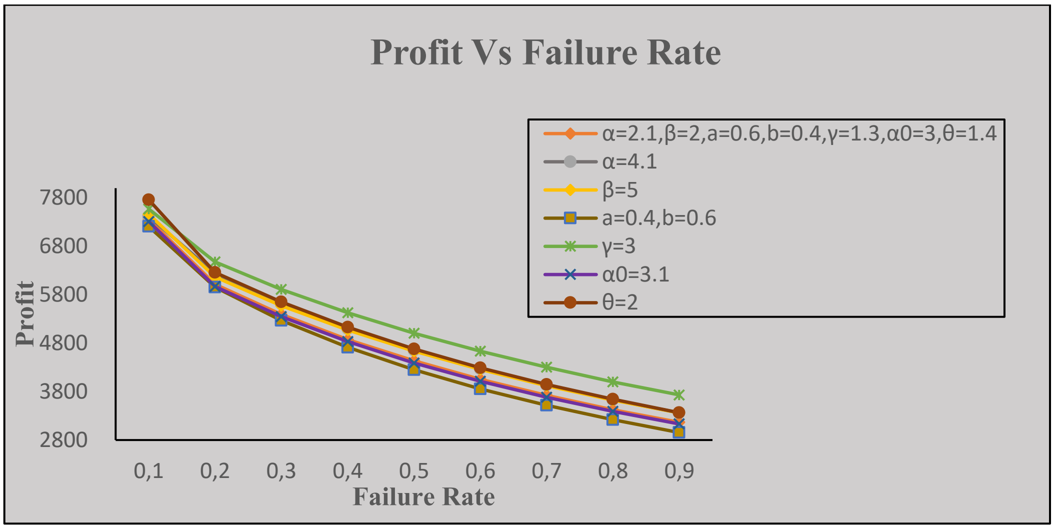


Figure 5: *Profit VS Failure Rate*

VI. Discussion

The performance of the system model for different values of failure and repair rates is shown by the table 2, table 3, and table 4 of MTSF, availability and profit analysis w.r.t. various parameters. Also, the behavior of the MTSF, availability and profit analysis depicted by figure 3, figure 4 and figure 5 respectively.

From the table 2 and figure 3 we conclude that the MTSF is decrease with the increase of the failure rate. If we fix all the parameters and change the value of repair rate α=4.1, replacement rate β=5 and inspection rate γ=3 one by one the MTSF is increase. And the MTSF is decrease if the preventive maintenance is increase by α_0 = 3.1 and the other parameters are fixed.

From the table 3 and figure 4 we analyze the availability of the system. By fixing all the parameters and changing in other parameter one by one we found that availability is increase in we change in α=4.1, replacement rate β=5, inspection rate γ=3 and preventive maintenance rate θ=2. The availability is decline with the increase of preventive maintenance by α_0 = 3.1 and inter change the repair and replacement rate (a=0.4 and b=0.6).

From the table 4 and figure 5 we conclude that profit of the function is decrease by increasing the preventive maintenance by α_0 = 3.1 and inter change the repair and replacement rate (a=0.4 and b=0.6) one by one with the fixed of all other parameters. If we change in α=4.1, replacement rate β=5, inspection rate γ=3 and preventive maintenance rate θ=2 found that profit of the system model is enhance as compared to the fixed values of all parameters.

References

[1] Kumar,A. and Malik,S.C. (2012). Stochastic modelling of a computer system with priority to PM over software replacement subject to maximum operation and repair times. *International Journal of computer applications, vol.43(3)*.

[2] Aggarwal,C., Ahlawat,N. and Malik,S.C. (2021). Profit analysis of a standby repairable system with priority to preventive maintenance and rest of server between repairs. *Journal of Reliability and Statistical Studies, Vol. 14(1), 57-80.*

[3] Kumar,P., Bharti,A. and Gupta,A. (2012). Reliability analysis of a two non-identical unit system with repair and replacements having correlated failure and repairs. J*ournal of Informatics and Mathematical Sciences, Vol.4(3), pp.339-350.*

[4] Rathee,R. and Pawar,D. (2018). Reliability modelling and analysis of a parallel system with priority to repair over replacement subject to maximum operation and repair times. *International Journal of Trend in Scientific Research and Development, vol.2(5)*.

[5] Wang,C.H. and Lin,T.W. (2011). Minimisation of non-periodic preventive maintenance cost in series-parallel systems. *Defence Science Journal, Vol. 61(1), pp. 44-50.*

[6] Dhillon,B.S. and Viswanath,H.C. (1991). Reliability analysis of a non-identical unit parallel system with common-cause failures. *Microelectronics Reliability,Vol. 31(2–3), pp. 429-441*.

[7] Sridharan,V. and Kalyani,T.V. (2002). Stochastic analysis of a non-identical two unit parallel system with common-cause failure using GERT technique. *International Journal of Information and Management Sciences, vol.13(1).*

[8] Sharma,N. and Joorel,J.P (2019). Study of a stochastic model of a two unit system with inspection and replacement under multi failure. *RT&A, Vol.14(3).*

[9] Papageorgious,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, pp. 821-827.*

[10] Song,S., and Coit,D.W. (2013). Reliability analysis for parallel systems with different component shock sets. *Proceedings of the 2013 Industrial and Systems Engineering Research Conference.*

[11] Rani,S.(2015). A stochastic model with routine, inspection, maintenance and replacement. *International Journal of Recent Research Aspects, Vol.2(3), pp.54-58.*

[12] Kakkar,M., Bhatti,J. and Chitkara,A.K. (2016). Reliability analysis of two dissimilar parallel unit repairable system with failure during preventive maintenance. *Management Science Letter, vol.6(4),285-296.*

[13] Goel,L.R., Sharma,G.C. and Gupta,P. (1986). Reliability analysis of a system with preventive maintenance, inspection and two types of repair. *Microelectronics Reliability, vol.26(3), pp.429-433.*

[14] Coats,J. (2016). Refrigerated Food for thought: parallel compressor racks in commercial/industrial applications. *https://www.achrnews.com/articles/132261.*

AN UNIQUE OPTIMAL SOLUTION FOR TYPE – III TRIANGULAR INTUITIONSTIC FUZZY TRANSPORTATION ISSUE

Indira Singuluri

•

Vignan's Institute of Information Technology(A), Duvvada, Visakhapatnam
indira.singuluri@gmail.com

N. Ravishankar

•

Gitam Deemed to be university, GIS, Visakhapatnam
drravi68@gmail.com

CH. Uma Swetha

•

Anil Neerukonda Institute of Science & Technology, Visakhapatnam
umaswethachitta@gmail.com

Abstract

In real-life decisions, usually we happen to suffer through different states of uncertainties. In order to counter these uncertainties, in this paper, the author formulated a transportation problem in which availability, demand and costs are mixed terms of real, triangular intuitionistic fuzzy numbers. In this paper, a simple method for solving type-3 intuitionistic fuzzy transportation is applied. So, the proposed method gives the optimal solution directly. The solution procedure is illustrated with the help of numerical examples.

Keywords: IFN, TIFN, IFTP of type - 3, Optimum Solution.

1. Introduction

The fuzzy set (FS) hypothesis was at first developed by [9] is useful from numerous points of view in various applications in different fields. The idea fuzzy numerical writing computer programs was created by Tanaka et al in 1947 outlining of fuzzy choice of [2]. Idea of Intuitionistic fuzzy sets (IFS's) recommended by [1] are chiefly valuable to manage numerous exemptions, disarray and ambiguities. The IFS's different the force of enrollment (MF) and the power of non-participation (NMF) of a component in the set. Uncertainties' assistance leader to concur the power of satisfaction, force of non-satisfaction and power of vulnerability for transfer and furthermore help to make determinant at stronghold degree of endorsement and non-endorsement for transportation cost (TC) in any transportation issue (TP). Also, without a doubt coming to dynamic issues IFS turned into an extreme technique which is for the most part closable. In like manner, its boss to use IFS diverged from FS to adapt to issues which our own dynamic or unworthiness. In

[6], investigate a general report on TP in fuzzy climate. Along these lines, IFS's are utilized by numerous creators for various ideal issues. [3] presented math activities of IFS's. Various analysts are likewise chipped away at and with IFS's. PSK method for solving mixed and type-4 intuitionistic fuzzy transportation issue (IFTP) was introduced by [8] which limit, requests and cost are availability, demand and costs are mixed terms of real, triangular intuitionistic fuzzy numbers and an algorithmic methodology for tackling IFTP was introduced [4]. By utilizing [8] paper in this article we address mathematical model. [7] Presented another positioning capacity utilizing centroid of centroids of IFN's.

In this article, we will acquaint another transportation strategy with acquire an ideal arrangement in an IFTP. For the new transportation method mathematical model is tackled. Plinth of article is managed: Section 2 quintessence goal, Section 3 gives Ranking capacity, Section 4 arrangements goal of IFTP of type-3 and computational system, region 5 consists Numerical model, at long last conclusion given in region 6.

2. Preliminaries

In this part a couple of essential definitions and math tasks are examined.

Intuitionistic Fuzzy Set (IFS): An IFS $\overset{\sim IFS}{A}$ in X an IFS is described as an object of following design

$$\overset{\sim IFS}{A} = \{\langle x, \mu_{\underset{A}{\sim IFS}}(x), \nu_{\underset{A}{\sim IFS}}(x)\rangle : x \in X\}$$

where the functions $\mu_{\underset{A}{\sim IFS}} : X \to [0, 1]$ and $\nu_{\underset{A}{\sim IFS}}: X \to [0, 1]$ defines degree of Enrollment work and non-participation element $x \in X$, respectively and $0 \leq \mu_{\tilde{A}^{IFS}}(x), \nu_{\tilde{A}^{IFS}}(x) \leq 1$, for every $x \in X$.

Intuitionistic Fuzzy Numbers (IFN's): A subset of IFS, $\overset{\sim IFS}{A} = \{\langle x, \mu_{\underset{A}{\sim IFS}}(x), \nu_{\underset{A}{\sim IFS}}(x)\rangle : x \in X\}$, of real line $\Re$ is called an IFN if the following holds:

(i) $\exists m \in \Re, \mu_{\tilde{A}^{IFS}}(m) = 1 \, and \, \nu_{\tilde{A}^{IFS}}(m) = 0$

(ii) $\mu_{\underset{A}{\sim IFS}} : \Re \to [0, 1]$ is continuous and for every $x \in \Re, 0 \leq \mu_{\tilde{A}^{IFS}}(x), \nu_{\tilde{A}^{IFS}}(x) \leq 1$ holds.

Enrollment work and non-participation capacity of $\widetilde{A}^{IFS}$ is as follows,

$$\mu_{\underset{A}{\sim IFS}}(x) = \begin{cases} f_1(x), & x \in [m - \alpha_1, m) \\ 1, & x = m \\ h_1(x), & x \in (m, m + \beta_1] \\ 0, & otherwise \end{cases} \quad \text{and} \quad \nu_{\underset{A}{\sim IFS}}(x) = \begin{cases} 1, & x \in (-\infty, m - \alpha_2) \\ f_2(x), & x \in [m - \alpha_2, m) \\ 0, & x = m, x \in [m + \beta_2, \infty) \\ h_2(x), & x \in (m, m + \beta_2] \end{cases}$$

Where $f_i(x)$ and $h_i(x)$; $i = 1,2$ are strictly increasing and decreasing functions in $[m - \alpha_i, m)$ and $(m, m - \beta_i]$ respectively. α_i and β_i are left and right spreads of $\mu_{\underset{A}{\sim IFS}}(x)$ and $\nu_{\underset{A}{\sim IFS}}(x)$ respectively.

Triangular Intuitionistic Fuzzy Number (TIFN): A TIFN $\overset{\sim IFN}{A}$ is an IFS in $\Re$ with the following Enrollment function $\mu_{\underset{A}{\sim IFN}}$ and non-participation capacity $\nu_{\underset{A}{\sim IFN}}$ defined by

$$\mu_{\tilde{A}^{IFS}} = \begin{cases} 0, & x < a_1 \\ \frac{x - a_1}{a_2 - a_1}, & a_1 \leq x \leq a_2 \\ 1, & x = a_2 \\ \frac{a_3 - x}{a_3 - a_2}, & a_2 \leq x \leq a_3 \\ 0, & otherwise \end{cases} \quad \text{and} \quad \nu_{\tilde{A}^{IFS}} = \begin{cases} 1, & x < a_1' \\ \frac{a_1' - x}{a_2 - a_1'}, & a_1' \leq x \leq a_2 \\ 0, & x = a_2 \\ \frac{x - a_2}{a_3' - a_2}, & a_2 \leq x \leq a_3' \\ 1, & otherwise \end{cases}$$

Where $a'_1 \leq a_1 \leq a_2 \leq a_3 \leq a'_3$. This TIFN is denoted by $\overset{\sim IFN}{A} = (a_1, a_2, a_3; a'_1, a_2, a'_3)$ in Figure 1.

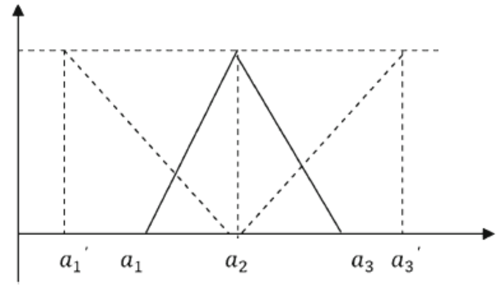


Figure 1: *Participation and non-enrollment elements of TIFN*

Arithmetic operations of TIFN:

For any two TIFN's $\overset{\sim IFN}{A} = (a_1, a_2, a_3; a'_1, a_2, a'_3)$ and $\overset{\sim IFN}{B} = (b_1, b_2, b_3; b'_1, b_2, b'_3)$, arithmetic operations are as follows,

(i) Addition: $\overset{\sim IFN}{A} \oplus \overset{\sim IFN}{B} = (a_1 + b_1, a_2 + b_{2,}, a_3 + b_3; a'_1 + b'_1, a_2 + b_2, a'_3 + b'_3)$

(ii) Subtraction: $\overset{\sim IFN}{A} - \overset{\sim IFN}{B} = (a_1 - b_3, a_2 - b_{2,}, a_3 - b_1; a'_1 - b'_3, a_2 - b_2, a'_3 - b'_1)$

(iii) Multiplication: $\overset{\sim IFN}{A} \otimes \overset{\sim IFN}{B} = (a_1b_1, a_2b_{2,}, a_3b_3; a'_1b'_1, a_2b_2, a'_3b'_3)$

(iv) Scalar multiplication: $k \times \tilde{A}^{IFN} = \begin{cases} (ka_1, ka_2, ka_3; ka'_1, ka_2, ka'_3), k \geq 0 \\ (ka_3, ka_2, ka_1; ka'_3, ka_2, ka'_1), k < 0 \end{cases}$

3. Ranking Function

Ranking function is taken from [7], i.e., the ranking function is defined [7], for Trapezoidal and triangular Intuitionistic fuzzy number as

$$R(\tilde{A}^{IFN}) = \left(\frac{a_1 + b_1 + 2(a_2 + b_3) + 5(a_3 + b_2) + (a_4 + b_4)}{18}\right)\left(\frac{4w_1 + 5w_2}{18}\right)$$

$$R(\tilde{A}^{IFN}) = \left(\frac{(a_1 + b_1) + 14a_2 + (a_4 + b_4)}{18}\right)\left(\frac{4w_1 + 5w_2}{18}\right)$$

Consider $w_1 = w_2 = 1$, we get ranking function is

$$R(\tilde{A}^{IFN}) = \left(\frac{(a_1 + b_1) + 14a_2 + (a_4 + b_4)}{36}\right)$$

Comparison of TIFN's: To differentiate TIFN's and each other, we need to rank them. A capacity like R: F(R) →R, which maps each TIFN's into real line, is called positioning capacity. Here, $F(\Re)$means the arrangement of all TIFN's.

By using ranking function "R", TIFN's can be compared. Let $\overset{\sim IFN}{A} = (a_1, a_2, a_3; a'_1, a_2, a'_3)$ and $\overset{\sim IFN}{B} = (b_1, b_2, b_3; b'_1, b_2, b'_3)$are two TIFN's thenR $\left(\tilde{A}^{IFN}\right) = \frac{a_1+14a_2+a_3+a'_1+a'_3}{36}$and R $\left(\tilde{B}^{IFN}\right) = \frac{b_1+14b_2+b_3+b'_1+b'_3}{36}$then the orders are defined as follows

(i) $\overset{\sim IFN}{A} > \overset{\sim IFN}{B}$ if R $\left(\overset{\sim IFN}{A}\right) > R\left(\overset{\sim IFN}{B}\right)$,

(ii) $\overset{\sim IFN}{A} < \overset{\sim IFN}{B}$ if R $\left(\overset{\sim IFN}{A}\right) < R\left(\overset{\sim IFN}{B}\right)$, and

(iii) $\overset{\sim IFN}{A} = \overset{\sim IFN}{B}$ if R $\left(\overset{\sim IFN}{A}\right) = \mathrm{R}\left(\overset{\sim IFN}{B}\right)$

Ranking function R also holds the following properties:

(i) R $\left(\overset{\sim IFN}{A}\right)$ + R $\left(\overset{\sim IFN}{B}\right)$ = R $\left(\overset{\sim IFN}{A} + \overset{\sim IFN}{B}\right)$, (ii) R $\left(\mathrm{k}\overset{\sim IFN}{A}\right)$ = k R $\left(\overset{\sim IFN}{A}\right)$ ∀k ∈ $\boldsymbol{R}$

4. Mathematical Formulation of Triangular Intuitionistic Fuzzy transportation problem (TIFTP) of Type - III and proposed method

4.1. TIFTP of type - III

Consider Examine a TP with 'm' vendors and 'n' insistent. c_{ij} is value of transiting one module of outcome from i^{th} vendor to j^{th} insistent.

$\tilde{a}_i^{IFN} = (a_1^i, a_2^i, a_3^i; a_1^{i'}, a_2^i, a_3^{i'})$ be IF extent at i^{th} vendor.

$\tilde{b}_j^{IFN} = (b_1^i, b_2^i, b_3^i; b_1^{i'}, b_2^i, b_3^{i'})$ be IF abundant at j^{th} insistent.

$\tilde{x}_{ij}^{IFN} = (x_1^{ij}, x_2^{ij}, x_3^{ij}; x_1^{i'j'}, x_2^{ij}, x_3^{i'j'})$ be IF quantity transformed from i^{th} vendor to j^{th} insistent

Then balanced IFTP of type - III is given by

$$Min\tilde{Z}^{IFN} = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} \times x_{ij}^{IFN}$$

$$s.t.\sum\nolimits_{j=1}^{n} \tilde{x}_{ij}^{IFN} = \tilde{a}_i^{IFN} \quad, i = 1,2,\ldots,m$$

$$\sum\nolimits_{i=1}^{m} \tilde{x}_{ij}^{IFN} = \tilde{b}_j^{IFN} \quad, j = 1,2,\ldots,n$$

$$\tilde{x}_{ij}^{IFN} \geq \tilde{0}; i = 1,2,\ldots,m; j = 1,2,\ldots,n$$

The TP is termed as type – III TIFTP having availability, demand and costs are mixed terms real, fuzzy and TIFN's. To find optimum solution TIFTP of type – III, we are using the following transportation strategy.

4.2. Proposed Transportation strategy (Used in [5])

Stage 1: In the given transportation problem calculate the differences between maximum and minimum cost for each row and column.
Stage 2: Find sum of row difference and column difference and denote row sum by R and column sum by C. Identify Maximum sum of row and column. Select maximum difference in row and column.
Stage 3: Choose cell having most minimal expense in row and column identified in stage 2.
Stage 4: Make a feasible assignment to cell picked in stage 5. Delete fulfilled row/column.
Stage 5: Repeat technique until all the designations has been made.
Stage 6: The Optimum solution and triangular intuitionistic optimum value is attained in step 5, is optimum solution $\{x_{ij}\}$ and triangular intuitionistic fuzzy optimum value is$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} \otimes x_{ij}$.

5. Numerical Example

Example for TIFTP of type – III: In this province, subsist numerical example ([8]) is solved by using above transportation strategy. Consider the 3×3TIFTP of type – III

Table 1: *TIFTP of Type - III*

	D_1	D_2	D_3	$Supply(\tilde{a}_i^{IFN})$
O_1	(8,10,12; 6,10,14)	4	(10,15,20)	(4,6,8; 3,6,9)
O_2	3	(6,12,18)	(4,6,8; 2,6,10)	8
O_3	(4,8,12)	(3,4,5; 1,4,6)	6	(2,5,8)
$Demand\,\tilde{b}_j^{IFN}$	(3,4,5)	(2,6,10; 1,6,11)	9	

The above TIFTP of type – III table can be rewrite as

Table 2: *Modified TIFTP of Type - III*

	D_1	D_2	D_3	$Supply$ $\tilde{a}_i^{IFN}$
O_1	(8,10,12; 6,10,14	(4,4,4; 4,4,4)	(10,15,20; 10,15,20)	(4,6,8; 3,6,9)
O_2	(3,3,3; 3,3,3)	(6,12,18; 6,12,18)	(4,6,8; 2,6,10)	(8,8,8; 8,8,8)
O_3	(4,8,12; 4,8,12	(3,4,5; 1,4,6)	(6,6,6; 6,6,6)	(2,5,8; 2,5,8)
$Demand$ $\tilde{b}_j^{IFN}$	(3,4,5; 3,4,5)	(2,6,10; 1,6,11)	(9,9,9; 9,9,9)	

Here, we have
$(4,6,8;3,6,9)\oplus(8,8,8;8,8,8)\oplus(2,5,8;2,5,8)=$
$(3,4,5;3,4,5)\oplus(2,6,10,1,6,11)\oplus(9,9,9;9,9,9)=(14,19,24;13,19,25).$
Accordingly, problem is balanced. Beyond comparison, foregoing IFN's encounter ranking functional values of $\tilde{a}_i^{IFN}$'s, $\tilde{b}_j^{IFN}$'s and costs as under:

$$R(\tilde{a}_1^{IFN})=3, R(\tilde{a}_2^{IFN})=4, R(\tilde{a}_3^{IFN})=2.5,$$
$$R(\tilde{b}_1^{IFN})=2, R(\tilde{b}_2^{IFN})=3, R(\tilde{b}_3^{IFN})=4.5,$$
$$R(\tilde{c}_{11}^{IFN})=5, R(\tilde{c}_{12}^{IFN})=2, R(\tilde{c}_{13}^{IFN})=8,$$
$$R(\tilde{c}_{21}^{IFN})=2, R(\tilde{c}_{12}^{IFN})=6, R(\tilde{c}_{13}^{IFN})=3,$$
$$R(\tilde{c}_{31}^{IFN})=4, R(\tilde{c}_{32}^{IFN})=2, R(\tilde{c}_{13}^{IFN})=3.$$

Table 3: *Row and Column difference table*

	D_1	D_2	D_3	$Supply$ ($\tilde{a}_i^{IFN}$)	Row $Diff$
O_1	(8,10,12; 6,10,14)	(4,4,4; 4,4,4)	(10,15,20; 10,15,20)	(4,6,8; 3,6,9)	6
O_2	(3,3,3; 3,3,3)	(6,12,18; 6,12,18)	(4,6,8; 2,6,10)	(8,8,8; 8,8,8)	4
O_3	(4,8,12; 4,8,12)	(3,4,5; 1,4,6)	(6,6,6; 6,6,6)	(2,5,8; 2,5,8)	2
$Demand$ ($\tilde{b}_j^{IFN}$)	(3,4,5; 3,4,5)	(2,6,10; 1,6,11)	(9,9,9; 9,9,9)		$R=12$
$Col.diff$	3	4	5	$C=12$	

The problem given in Table 3, transformed in Table 4 by using the Stage 2 and assign first allocation using stage 4 of proposed method.

Table 4: *First Allocation Table*

	D_1	D_2	D_3	Supply $(\tilde{a}_i^{IFN})$	Row diff
O_1	(8,10,12; 6,10,14)	(4,4,4; 4,4,4) $\boxed{(4,6,8;3,6,9)}$	(10,15,20; 10,15,20)	~~(4,6,8; 3,6,9)~~ **0**	**6**
O_2	(3,3,3; 3,3,3)	(6,12,18; 6,12,18)	(4,6,8; 2,6,10)	(8,8,8; 8,8,8)	4
O_3	(4,8,12; 4,8,12)	(3,4,5; 1,4,6	(6,6,6; 6,6,6)	(2,5,8; 2,5,8)	2
Demand $(\tilde{b}_j^{IFN})$	(3,4,5; 3,4,5)	~~(2,6,10; 1,6,11)~~ (-6,0,6; -8,0,8)	(9,9,9; 9,9,9)		$R = 12$
Col. diff	3	4	5	$C = 12$	

Using stage 4 of proposed method remove S_1 from Table 4. New reduced shown in Table 5 again apply the proposed procedure.

Table 5: *New Reduced Table*

	D_1	D_2	D_3	Supply $(\tilde{a}_i^{IFN})$	Row Diff
O_2	(3,3,3; 3,3,3)	(6,12,18; 6,12,18)	(4,6,8; 2,6,10)	(8,8,8; 8,8,8)	4
O_3	(4,8,12; 4,8,12)	(3,4,5; 1,4,6)	(6,6,6; 6,6,6)	(2,5,8; 2,5,8)	2
Demand $(\tilde{b}_i^{IFN})$	(3,4,5; 3,4,5)	(−6,0,6; −8,0,8)	(9,9,9; 9,9,9)		$R = 6$
Col. diff	2	4	0	$C = 6$	

Table 6: *Second Allocation table*

	D_1	D_2	D_3	Supply $(\tilde{a}_i^{IFN})$	Row Diff
O_2	(3,3,3; 3,3,3)	(6,12,18; 6,12,18)	(4,6,8; 2,6,10)	(8,8,8; 8,8,8)	4
O_3	(4,8,12; 4,8,12)	(3,4,5; 1,4,6) $\boxed{(-6,0,6;-8,0,8)}$	(6,6,6; 6,6,6)	~~(2,5,8; 2,5,8)~~ (−4,5,14; −6,5,16)	2
Demand $(\tilde{b}_i^{IFN})$	(3,4,5; 3,4,5)	~~(−6,0,6; −8,0,8)~~ **0**	(9,9,9; 9,9,9)		$R = 6$
Col. diff	2	4	0	$C = 6$	

Again, applying Stage 5 of proposed strategy, all allocations are made as shown in Table 7.

Table 7: *Final allocation table*

	D_1	D_2	D_3	$Supply\ (\tilde{a}_i^{IFN})$
O_1		$(\mathbf{4,6,8};\ \mathbf{3,6,9})$		(4,6,8; 3,6,9)
O_2	$(\mathbf{3,4,5};\ \mathbf{3,4,5})$		$(\mathbf{3,4,5};\ \mathbf{3,4,5})$	(8,8,8; 8,8,8)
O_3		$(\mathbf{-6,0,6};\ \mathbf{-8,0,8})$	$(\mathbf{4,5,6};\ \mathbf{4,5,6})$	(2,5,8; 2,5,8)
$Demand\ (\tilde{b}_j^{IFN})$	(3,4,5; 3,4,5)	(2,6,10; 1,6,11)	(9,9,9; 9,9,9)	

IF optimum solution in terms of TIFN's:

$$\tilde{x}_{12}{}^{IFN} = (4,6,8;3,6,9), \tilde{x}_{21}{}^{IFN} = (3,4,5;3,4,5), \tilde{x}_{23}{}^{IFN} = (3,4,5;3,4,5),$$
$$\tilde{x}_{32}{}^{IFN} = (-6,0,6;-8,0,8), \tilde{x}_{33}{}^{IFN} = (4,5,6;4,5,6).$$

Hence, total TIFTP of type – 3 optimum cost = (41,90,139; 21,90,159).

6. Conclusion

In this article, a new method has evolved which provides the opportunity to find the optimal objective value of the TIFTP of Type 3in terms of mixed intuitionistic fuzzy numbers. Based on current examination, it very well may be presumed that it is much simple to apply proposed strategy when contrasted with existing techniques [7], for tracking down the ideal optimum solution of TIFTP of type – III. Consequently, it is smarter to utilize the proposed technique rather than existing strategies for tackling TIFTP of type – III.

References

[1] Atanassov, K.T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets Syst.*20, pp.87–96.

[2] Bellman, R., Zadeh, L.A. (1970). Decision making in fuzzy environment. *Manag. Sci.* 17(B), 141–164.

[3] Chakraborty, D., Jana, D.K. and Roy, T.K. (2015). Arithmetic operations on generalized intuitionistic fuzzy number and its applications to transportation problem. *Opsearch.*52, 431–471.

[4] Hussain, R. J., Kumar, P. S. (2012). Algorithmic approach for solving intuitionistic fuzzy transportation problem. *Appl Math Sci* 6(80):3981–3989.

[5] Indira Singuluri and Ravishankar, N. (2021). A novel transportation approach to solving type-2 triangular intuitionistic fuzzy transportation problems. *RT&A*. No 4 (65) Volume 16, December.

[6] Ismail Mohideen, S. and Senthil Kumar, A. (2010). A Comparative Study on Transportation Problem in fuzzy environment. *International Journal of Mathematics Research*. Vol.2, no. 1, pp. 151-158.

[7] Pardhasaradhi, B., Madhuri and M. V. & Ravi Shankar, N. (2017) Ordering of Intuitionistic fuzzy numbers using centroid of centroids of Intuitionistic fuzzy numbers, International Journal of Mathematics Trends and Technology, Vol. 52, no. 5, pp.276-285.

[8] Senthil Kumar, P. (2019). PSK Method for Solving Mixed and Type-4 Intuitionistic Fuzzy Solid Transportation Problems. *International Journal of Operations Research and Information Systems.* Volume 10, Issue 2.

[9] Zadeh, L.A. (1965). Fuzzy sets. *Inf. Comput*. 8, pp.338–353.

RT&A, No 2 (68)
Volume 17, March 2022

ZECH DISTRIBUTION: DERIVATION, PROPERTIES AND APPLICATIONS TO REAL LIFE DATA

Sunday Adeyeye[1], Ademola Adewara[2], Emmanuela Akarawak[3], Adeyinka Ogunsanya[4], Alabbasi Jamal[5]

•

[1,3] Department of Mathematics, University of Lagos, Akoka, Nigeria.
[2] Distance Learning Institute, University of Lagos, Akoka, Nigeria.
[4] Department of Statistics, University of Ilorin, Ilorin, Nigeria.
[5] Al – Nahrain University, Baghdad, Iraq.
E- Mail: [1] jadeyeyesunday@yahoo.com, [2] jadewara@unilag.edu.ng,
[3] eakarawak@unilag.edu.ng, [4] ogunsanyaadeyinka@yahoo.com
[5] jamal.alabbasi@nahrainuniv.edu.iq

Abstract

The roles of heavy – tailed distribution in modelling real life events, especially in financial and actuarial sciences, cannot be over – emphasized. In this paper, a new heavy right – tailed, three – parameter continuous distribution with increasing hazard rate called Zech distribution is developed. The proposed model is very suitable for modelling heavy right- tailed data. Zech distribution is the reciprocal of the random variable which follows Gompertz- Inverse – Exponential (GoIE) distribution and it does not involve addition of extra parameter, thereby removing the cumbersomeness in the estimation process posed by other methods involving additional extra parameters, especially where more than three parameters are involved. The statistical properties of the new distribution such as survival function, hazard function, cumulative hazard function, reversed hazard function, quantile function, order statistics, moments, mean, median, variance, skewness, and kurtosis were derived. The Linear representation of the pdf of the newly developed distribution revealed that its probability density function is a weighted exponential distribution. Also, method of maximum likelihood was used in estimating the model's parameters. The simulation results revealed that as the sample sizes increased, the root mean squared errors decreased which showed that the parameters of Zech distribution are stable. The proposed distribution was applied to two real life data sets. The results showed that Zech distribution performs better than Gompertz Inverse Exponential distribution, Weibull Exponential distribution and Gompertz Exponential distribution.

Keywords: Zech distribution, Gompertz Inverse Exponential Distribution, maximum likelihood estimation, simulation studies, moments, linear representation.

1. Introduction

Probability distributions play a crucial role in modelling naturally occurring phenomena. In probability theory and statistics, an inverse distribution is the distribution of the reciprocal of a random variable. To model real life events, there is need for the extension of the classical forms of distributions so as to have a better fit to the real data. Several methods of extending distributions have been proposed in the literature. Among these is 'Inverse Distribution' which does not increase the number of parameter(s) of the parent distribution but provides a better fit. This is a strong motivation for studying inverse distribution as prescribed by the principle of parsimony. Eliwa [12] proposed Inverse Gompertz distribution which was found to out – perform other six competing distributions. The Gompertz Inverse Exponential distribution proposed by Pelumi [5] is good for modelling right – tailed data. Said [13] introduced Extended Inverse Weibull distribution whose density function can be expressed as a linear combination of the Inverse Weibull densities with increasing and decreasing hazard rates. Ogunsanya [11] developed Weibull Inverse Rayleigh distribution which is an extension of a one – parameter Inverse Rayleigh distribution that incorporated a transformation of the Weibull distribution and Log – logistic distribution as quantile functions. El – Gohari A [2] proposed Generalized Gompertz distribution which is a new generalization of the Exponential, Gompertz and Generalized Exponential distributions. The main advantage of this new distribution is that it has increasing or constant or decreasing or bathtub curve failure rate depending upon the shape parameter. It is this property that makes it suitable for survival analysis. Adewara [3] introduced Gompertz Exponential distribution which is an extension of Exponential distribution by using the Gompertz Generalized family of distributions proposed by Morad [4]. To increase the flexibility of Gompertz Exponential distribution, an extra shape parameter was added to it leading to the introduction of Exponentiated Gompertz Exponential distribution by Adewara [8].

The motivation for this study is to derive a distribution which will be more flexible for modelling heavy right – tailed data and to obtain interesting properties of the new model. Therefore, the inverse of '*Gompertz Inverse Exponential distribution*', which will henceforth be called *Zech* distribution is proposed. The adoption of the name '*Zech* distribution' is to avoid the confusion which might arise from using the name: Inverse Gompertz Inverse Exponential Distribution.

Given the cumulative distribution function (cdf) of a random variable Y, the distribution function of a random variable $X = \frac{1}{Y}$ is the reciprocal or inverse of the random variable Y. This implies that the cumulative distribution function $G(x)$ is the inverse function of $F(y)$. This is easier if Y is a continuous random variable and $F(y)$ is strictly on positive supports. The cumulative distribution function of inverse distribution is derived according to the method below:

$$\begin{aligned} G_X(x) &= P(X \leq x) \\ &= P\left(\frac{1}{Y} \leq X\right) \\ &= P\left(x \geq \frac{1}{Y}\right) \\ &= 1 - P\left(x \leq \frac{1}{Y}\right) \\ &= 1 - F\left(\frac{1}{Y}\right) \end{aligned} \tag{1}$$

The cumulative distribution function (cdf) and probability density function (pdf) of Gompertz Inverse Exponential distribution are given in equations (2) and (3) respectively.

$$F(y) = 1 - e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\frac{\theta}{y}}\right]^{-\beta}\right\}} \quad ; y > 0, \alpha > 0, \beta > 0, \theta > 0 \tag{2}$$

$$f(y) = \alpha \frac{\theta}{y^2} e^{-\frac{\theta}{y}} \left[1 - e^{-\frac{\theta}{y}}\right]^{-\beta-1} e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\frac{\theta}{y}}\right]^{-\beta}\right\}} \; ; \; y > 0, \alpha > 0, \beta > 0, \theta > 0 \tag{3}$$

II. Zech Distribution

The Zech distribution is the reciprocal of Gompertz Inverse Exponential distribution. The cumulative distribution function of Zech distribution is stated in the following theorem.

Theorem 1: If a non – negative random variable Y follows the Gompertz inverse Exponential distribution expressed as Y~GIE$(y; \alpha, \theta, \beta)$. Assuming a new random variable $X = \frac{1}{y}$ is defined, then the random variable X follows *Zech* distribution, written as X~Zech$(x; \theta, \alpha, \beta)$ with the cdf in equation (4).

$$G(x) = e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}} \; ; \; x > 0, \alpha > 0, \beta > 0, \theta > 0 \tag{4}$$

Proof:

From equation (1), $G(x) = 1 - F\left(\frac{1}{y}\right)$

$$G(x) = 1 - \left[1 - e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right]$$

$$G(x) = e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}$$

The result of the first derivative, with respect to x, of equation (4) is the probability density function of *Zech* distribution given by (5).

$$g(x) = \alpha\theta e^{-\theta x}\left[1 - e^{-\theta x}\right]^{-\beta-1} e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}; \quad x > 0, \alpha > 0, \beta > 0, \theta > 0 \tag{5}$$

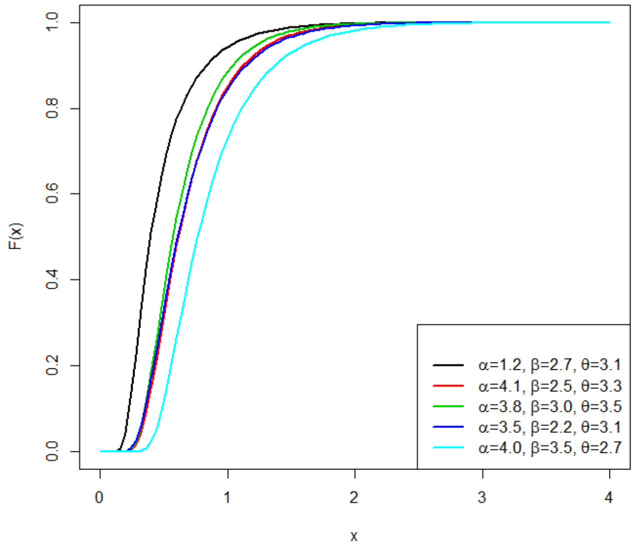


Figure 1: CDF Plots of Zech distribution.

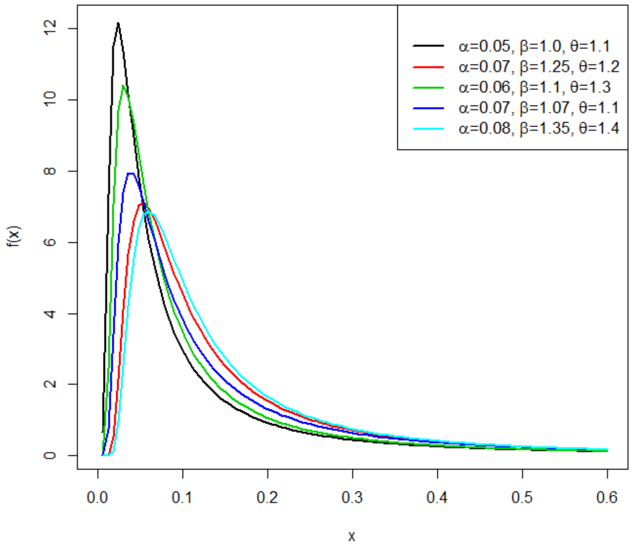


Figure 2: PDF Plots of Zech distribution.

Figures 1 and 2 illustrate some of possible shapes of the cumulative density function and probability density function respectively of Zech distribution.

III. Estimation of Parameters

The method of Maximum likelihood is used to estimate the parameters of *Zech* distribution.
Assuming each of the random samples $x_1, x_2, \ldots, x_n$ follows the pdf of *Zech* distribution, the likelihood function is given by

$$L(x_1, x_2, \ldots, x_n; \alpha, \beta, \theta) = \prod_{i=1}^{n}\left\{\alpha\theta e^{-\theta x_i}\left[1 - e^{-\theta x_i}\right]^{-\beta-1} e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x_i}\right]^{-\beta}\right\}}\right\} \tag{6}$$

Let l denote the log-likelihood function, that is, let

$$l = log\, L(x_1, x_2, \ldots, x_n; \alpha, \beta, \theta) \tag{7}$$

$$l = nlog\alpha + nlog\theta - \theta\sum_{i=1}^{n} x_i - (\beta + 1)\sum_{i=1}^{n}\log\left(1 - e^{-\theta x}\right) + \frac{\alpha}{\beta}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} \tag{8}$$

The solutions of simultaneous equations obtained from $\frac{dl}{d\alpha} = 0$, $\frac{dl}{d\beta} = 0$ *and* $\frac{dl}{d\theta} = 0$ are the maximum likelihood estimates of the parameters α, β *and* θ. Thus,

$$\frac{dl}{d\alpha} = \frac{n}{\alpha} + \frac{1}{\beta}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} \tag{9}$$

$$\frac{dl}{d\beta} = \frac{\alpha}{\beta}\sum_{i=1}^{n}\left\{\left[1 - e^{-\theta x_i}\right]^{-\beta}\ln\left(1 - e^{-\theta x_i}\right)\right\} - \frac{\alpha}{\beta^2}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} - \sum_{i=1}^{n}\ln\left(1 - e^{-\theta x_i}\right)$$
$$- \frac{\alpha}{\beta^2}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} \tag{10}$$

$$\frac{dl}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} x_i - (\beta + 1)\sum_{i=1}^{n}\left(x_i e^{-\theta x_i}\right)\left(1 - e^{-\theta x_i}\right)^{-1} + \alpha\sum_{i=1}^{n} x_i e^{-\theta x_i}\left(1 - e^{-\theta x_i}\right)^{-\beta-1} \tag{11}$$

Equating $\frac{dl}{d\alpha} = 0$, $\frac{dl}{d\beta} = 0$ and $\frac{dl}{d\theta} = 0$, we have

$$\frac{n}{\alpha} + \frac{1}{\beta}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} = 0 \tag{12}$$

$$\frac{\alpha}{\beta}\sum_{i=1}^{n}\left\{\left[1 - e^{-\theta x_i}\right]^{-\beta}\ln\left(1 - e^{-\theta x_i}\right)\right\} - \frac{\alpha}{\beta^2}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} - \sum_{i=1}^{n}\ln\left(1 - e^{-\theta x_i}\right)$$
$$- \frac{\alpha}{\beta^2}\sum_{i=1}^{n}\left\{1 - \left[1 - e^{-\theta x_i}\right]^{-\beta}\right\} = 0 \tag{13}$$

$$\frac{n}{\theta} + \sum_{i=1}^{n} x_i - (\beta + 1)\sum_{i=1}^{n}\left(x_i e^{-\theta x_i}\right)\left(1 - e^{-\theta x_i}\right)^{-1} + \alpha\sum_{i=1}^{n} x_i e^{-\theta x_i}\left(1 - e^{-\theta x_i}\right)^{-\beta-1} = 0 \tag{14}$$

The Maximum likelihood estimate for parameter α can be obtained from (12) in the form below for a given β and θ

$$\hat{\alpha} = \frac{-n\beta}{\sum_{i=1}^{n}\left\{1-\left[1-e^{-\theta x_i}\right]^{-\beta}\right\}} \tag{15}$$

To obtain the MLE of β and θ, equation (15) can be substituted into equations (13) and (14). The resulting system of non – linear equations can be solved numerically.

IV. Linear Representation

Theorem 2: The pdf of *Zech* distribution is a weighted function of an exponential distribution with rate parameter $\theta(1 + i)$.

Proof.

$$g(x) = \alpha\theta e^{-\theta x}\left[1 - e^{-\theta x}\right]^{-\beta-1} e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}$$

Using the exponential expansion

$$e^{x} = \sum_{k=0}^{\infty}\frac{x^k}{k!} \tag{16}$$

$$g(x) = e^{-\theta x}\left[1 - e^{-\theta x}\right]^{-\beta-1}\sum_{k=0}^{\infty}\left\{1 - \left[1 - e^{-\theta x}\right]^{-\beta}\right\}^{k}\frac{\alpha\theta}{k!}\left(\frac{\alpha}{\beta}\right)^{k} \tag{17}$$

From the mixture representation,

$$(1 - z)^{k} = \sum_{j=0}^{\infty}\frac{\Gamma(\mathrm{k} + \mathrm{j})}{\Gamma(k)j!}z^{j}(-1)^{j} \tag{18}$$

$$\left\{1 - \left[1 - e^{-\theta x}\right]^{-\beta}\right\}^{k} = \sum_{j=0}^{\infty}(-1)^{j}\frac{\Gamma(\mathrm{k} + 1)}{\Gamma(k)j!}\left[1 - e^{-\theta x}\right]^{-\beta j} \tag{19}$$

Inserting equation (18) into (16), we have

$$g(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{\Gamma(\mathrm{k} + \mathrm{j})}{\Gamma(k)j!}(-1)^{j}\frac{\alpha\theta}{k!}\left(\frac{\alpha}{\beta}\right)^{k} e^{-\theta x}\left[1 - e^{-\theta x}\right]^{-\beta j}\left[1 - e^{-\theta x}\right]^{-\beta-1} \tag{20}$$

Simplifying, $\left[1 - e^{-\theta x}\right]^{-\beta j}\left[1 - e^{-\theta x}\right]^{-\beta-1} = \left[1 - e^{-\theta x}\right]^{-[\beta(j+1)+1]}$ (21)

Inserting equation (20) into equation (19), we have

$$g(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{\Gamma(\mathrm{k+j})}{\Gamma(k)j!}(-1)^j\frac{\alpha\theta}{k!}\left(\frac{\alpha}{\beta}\right)^k e^{-\theta x}\left[1-e^{-\theta x}\right]^{-[\beta(j+1)+1]} \tag{22}$$

By using mixture representation,

$$(1-z)^{-k} = \sum_{i=0}^{\infty}\frac{\Gamma(\mathrm{k+i})}{\Gamma(k)i!}z^i \tag{23}$$

$$\left[1-e^{-\theta x}\right]^{-[\beta(j+1)+1]} = \sum_{i=0}^{\infty}\frac{\Gamma[(\beta(j+1)+1)+i]}{\Gamma[\beta(j+1)+1]i!}e^{-\theta ix} \tag{24}$$

$$g(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\frac{\Gamma(\mathrm{k+j})}{\Gamma(k)j!}(-1)^j\frac{\alpha\theta}{k!}\left(\frac{\alpha}{\beta}\right)^k\frac{\Gamma[(\beta(j+1)+1)+i]}{\Gamma[\beta(j+1)+1]i!}e^{-\theta ix}e^{-\theta x} \tag{25}$$

$$g(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\frac{\Gamma(\mathrm{k+j})}{\Gamma(k)j!}(-1)^j\frac{\alpha\theta}{k!}\left(\frac{\alpha}{\beta}\right)^k\frac{\Gamma[(\beta(j+1)+1)+i]}{\Gamma[\beta(j+1)+1]i!}e^{-\theta x(1+i)} \tag{26}$$

$$\text{Let}(-1)^j\frac{\Gamma(\mathrm{k+j})}{\Gamma(k)j!}\frac{\Gamma[(\beta(j+1)+1)+i]}{\Gamma[\beta(j+1)+1]i!} = w_{i,j,k} \tag{27}$$

$$g(x) = \frac{\alpha\theta}{k!}\left(\frac{\alpha}{\beta}\right)^k\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}w_{i,j,k}e^{-\theta(1+i)x} \tag{28}$$

$$g(x) = \frac{\theta\alpha^{k+1}}{\beta^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}w_{i,j,k}e^{-\theta(1+i)x} \tag{29}$$

$$g(x) = \frac{\theta(1+i)\alpha^{k+1}}{\beta^k k!\,(1+i)}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}w_{i,j,k}e^{-\theta(1+i)x} \tag{30}$$

$$g(x) = \frac{\alpha^{k+1}}{\beta^k k!\,(1+i)}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}w_{i,j,k}\left[\theta(1+i)e^{-\theta(1+i)x}\right] \tag{31}$$

V. Reliability Properties

The reliability function can be obtained from

$$S(x) = 1 - G(x) \tag{32}$$

Therefore, the survival function of *Zech* distribution is given as

$$S(x) = 1 - e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}} \;\; ; \;\; x > 0, \alpha > 0, \beta > 0, \theta > 0 \tag{33}$$

The hazard function of *Zech* distribution is obtained from

$$h(x) = \frac{g(x)}{S(x)} \tag{34}$$

The hazard function of *Zech* distribution is given by

$$h(x) = \frac{\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}}{1-e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}}; \;\; x > 0, \alpha > 0, \beta > 0, \theta > 0 \tag{35}$$

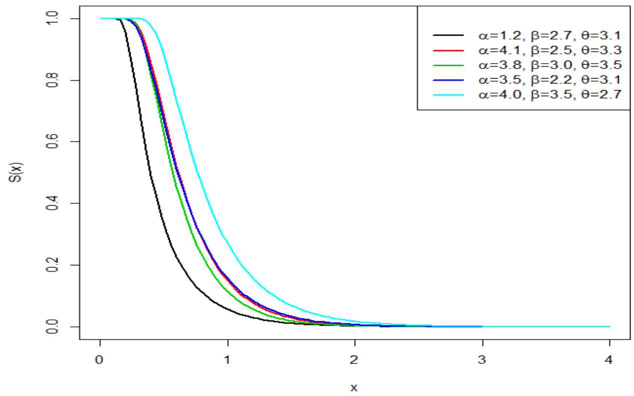


Figure 3: Survival Plots of Zech distribution.

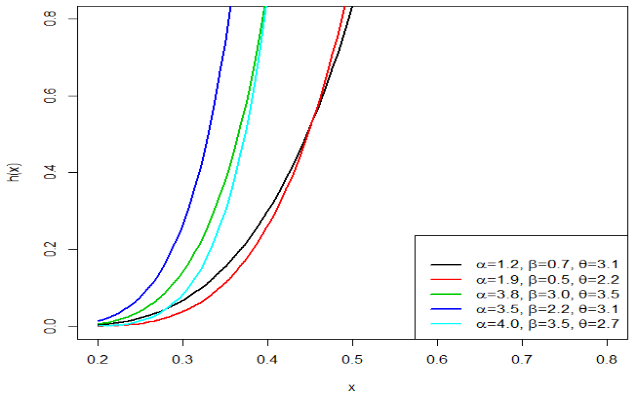


Figure 4: Hazard plots of Zech distribution.

Figures 3 and 4 illustrate some of possible shapes of the Survival function and Hazard function respectively of Zech distribution.

The cumulative hazard function, H(x) of a continuous random variable X from *Zech* distribution is derived from

$$H(x) = -\log(S(x)) \tag{36}$$

Substituting equation (33) into equation (36)

$$H(x)_{Zech} = -\log\left(1 - e^{\frac{\alpha}{\beta}\left\{1-[1-e^{-\theta x}]^{-\beta}\right\}}\right) \tag{37}$$

The reversed hazard function of a random variable x of *Zech* distribution is obtained from

$$r(x) = \frac{g(x)}{G(x)} \tag{38}$$

Therefore,

$$r(x)_{Zech} = \alpha\theta e^{-\theta x}[1 - e^{-\theta x}]^{-\beta-1} \tag{39}$$

VI. Quantile Function and Median of *Zech* distribution.

Quantile function is very important for generating random numbers which can be used for simulation studies. Aside that, it can also be used for finding quantiles i.e. quartiles, octiles, deciles and percentiles of a distribution which are necessary for deriving the measures of skewness and kurtosis.

The quantile function is derived by inverting the cdf

$$Q(u) = G^{-1}(u) \tag{40}$$

$$Q(u) = -\frac{1}{\theta}\left\{\ln\left[1 - \left(1 - \frac{\beta}{\alpha} \ln u\right)^{-\frac{1}{\beta}}\right]\right\} \tag{41}$$

Where $u \sim Uniform\ (0,1)$.

To generate random numbers from Zech distribution, it is sufficient that

$$x = -\frac{1}{\theta}\left\{\ln\left[1 - \left(1 - \frac{\beta}{\alpha} \ln u\right)^{-\frac{1}{\beta}}\right]\right\} \tag{42}$$

The median of *Zech* distribution can be obtained by substituting $u = 0.5$ in equation (41) as follows:

$$Median = -\frac{1}{\theta}\left\{\ln\left[1 - \left(1 - \frac{\beta}{\alpha} \ln 0.5\right)^{-\frac{1}{\beta}}\right]\right\} \tag{43}$$

Other quantiles can also be derived from (41) by substituting appropriate values of $''u''$.

VII. Quantile - based measures of skewness and kurtosis

The measure of Skewness(S) and Kurtosis (K) of Zech distribution using quantile function, defined by Galton [6] and Moors [7] are given by equations (44) and (45) respectively.

$$S = \frac{Q\left(\frac{6}{8}\right)-2Q\left(\frac{4}{8}\right)+Q\left(\frac{2}{8}\right)}{Q\left(\frac{6}{8}\right)-Q\left(\frac{2}{8}\right)} \tag{44}$$

$$K = \frac{Q\left(\frac{7}{8}\right)-Q\left(\frac{5}{8}\right)+Q\left(\frac{3}{8}\right)-Q\left(\frac{1}{8}\right)}{Q\left(\frac{6}{8}\right)-Q\left(\frac{2}{8}\right)} \tag{45}$$

$Q\left(\frac{1}{8}\right), Q\left(\frac{2}{8}\right), Q\left(\frac{3}{8}\right), Q\left(\frac{4}{8}\right), Q\left(\frac{5}{8}\right), Q\left(\frac{6}{8}\right)$, and $Q\left(\frac{7}{8}\right)$ can be obtained by substituting $\frac{1}{8}, \frac{2}{8}, \frac{3}{8}, \frac{4}{8}, \frac{5}{8}, \frac{6}{8}, and\ \frac{7}{8}$ for u respectively in equation (41). Therefore,

$$Q\left(\frac{1}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{1}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{46}$$

$$Q\left(\frac{2}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{47}$$

$$Q\left(\frac{3}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{3}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{48}$$

$$Q\left(\frac{4}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{4}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{49}$$

$$Q\left(\frac{5}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{5}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{50}$$

$$Q\left(\frac{6}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{51}$$

$$Q\left(\frac{7}{8}\right) = -\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{7}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\} \tag{52}$$

The skewness of *Zech* distribution is derived by substituting the values of $Q\left(\frac{6}{8}\right)$, $Q\left(\frac{4}{8}\right)$ and $Q\left(\frac{2}{8}\right)$ into equation (44).

Therefore,
$$S_{Zech} = \frac{\frac{2}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{4}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}}{\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}} \tag{53}$$

Simplifying (53) by factoring out $\left(\frac{1}{\theta}\right)$, the result shows that symmetry or asymmetry of *Zech* distribution is independent of parameter θ.

$$S_{Zech} = \frac{2\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{4}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}}{\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\beta}{\alpha}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}} \tag{54}$$

The kurtosis of *Zech* distribution is derived by substituting the values of $Q\left(\frac{7}{8}\right), Q\left(\frac{5}{8}\right), Q\left(\frac{3}{8}\right), Q\left(\frac{1}{8}\right), Q\left(\frac{1}{8}\right)\ and\ Q\left(\frac{2}{8}\right)$ into equation (45)

$$K_{Zech}=\frac{\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{5}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{7}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{3}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{1}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}}{\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\frac{1}{\theta}\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}} \quad (55)$$

Simplifying equation (55) by factoring out $\left(\frac{1}{\theta}\right)$, the result shows that the peakedness or otherwise of *Zech* distribution is independent of parameter θ.

$$K_{Zech}=\frac{\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{5}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{7}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{3}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{1}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}}{\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{2}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}-\left\{\ln\left[1-\left(1-\frac{\alpha}{\beta}ln\left(\frac{6}{8}\right)\right)^{-\frac{1}{\beta}}\right]\right\}} \quad (56)$$

VIII. Distribution of Order Statistics

Let $x_1, x_2, \dots, x_n$ be a random sample from a cdf and pdf of *Zech* distribution as defined in (4) and (5) respectively. The pdf of jth order statistics of any random variable X is given by:

$$f_{j:n}(x)=\frac{n!}{(j-1)!\,(n-j)!}g(x)G(x)^{j-1}[1-G(x)]^{n-j} \quad (57)$$

From (56), putting the pdf of the jth order statistics of *Zech* distribution,

$$f_{j:n}(x)=\frac{n!}{(j-1)!(n-j)!}\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}.\left\{e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right\}^{j-1}\left\{1-\left(e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right)\right\}^{n-j} \quad (58)$$

Simplifying equation (58),

$$f_{j:n}(x)=\frac{n!}{(j-1)!(n-j)!}\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}.\left\{e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right\}^{j}\left\{1-\left(e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right)\right\}^{n-j}$$
(59)

Therefore, the distribution of minimum and maximum order statistics for the *Zech* distribution is given by $f_{1:n}(x)$ $i.e\ when\ j=1$ and $f_{n:n}(x)$ respectively in equations (60) and (61) respectively.

$$f_{1:n}(x)=\frac{n!}{(j-1)!\,(n-1)!}\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}.e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\left\{1-\left(e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right)\right\}^{n-1}$$

After some simplifications,

$$f_{1:n}(x)=n\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}.e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\left\{1-\left(e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right)\right\}^{n-1} \quad 60)$$

Also,

$$f_{n:n}(x)=\frac{n!}{(n-1)!\,(n-n)!}\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}.\left\{e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right\}^{n}\left\{1-\left(e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right)\right\}^{n-n}$$

After some simplifications,

$$f_{n:n}(x)=n\alpha\theta e^{-\theta x}\left[1-e^{-\theta x}\right]^{-\beta-1}.\left\{e^{\frac{\alpha}{\beta}\left\{1-\left[1-e^{-\theta x}\right]^{-\beta}\right\}}\right\}^{n} \quad (61)$$

IX. Moments of Zech distribution

The moment about the Origin of *Zech* distribution is derived as follows: recall from the linear expansion of the pdf of *Zech* distribution,

$$g(x)=\frac{\alpha^{k+1}}{\beta^k k!(1+i)}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}w_{i,j,k}\left[\theta(1+i)e^{-\theta(1+i)x}\right]$$

The rth moment about the origin of a random variable X is given by

$$E(X^r) = \int x^r f(x)dx \tag{62}$$

$$E(X^r) = \int_0^\infty x^r \frac{\alpha^{k+1}}{\beta^k k!\,(1+i)} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} [\theta(1+i)e^{-\theta(1+i)x}]\, dx \tag{63}$$

$$E(X^r) = \frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \int_0^\infty x^r e^{-\theta(1+i)x}\, dx \tag{64}$$

From gamma expansion,

$$\frac{\Gamma a}{b^a} = \int_0^\infty x^{a-1} e^{-bx}\, dx \tag{65}$$

$a - 1 = r, a = r + 1\,, b = \theta(1+i)$

$$\int_0^\infty x^r e^{-\theta(1+i)x}\, dx = \frac{\Gamma(r+1)}{[\theta(1+i)]^{r+1}} \tag{66}$$

$$E(X^r) = \frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \frac{\Gamma(r+1)}{[\theta(1+i)]^{r+1}} \tag{67}$$

The first moment about the origin represents the mean of *Zech* distribution. This can be done by setting $r = 1$,

$$E(X) = \frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \frac{\Gamma(2)}{[\theta(1+i)]^2} \tag{68}$$

The second moment about the origin of *Zech* distribution is

$$E(X^2) = \frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \frac{\Gamma(3)}{[\theta(1+i)]^3} \tag{69}$$

The variance of *Zech* distribution is obtained from

$$Var(X) = E(X^2) - [E(X)]^2 \tag{70}$$

$$Var(X) = \frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \frac{\Gamma(3)}{[\theta(1+i)]^3} - \left[\frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \frac{\Gamma(2)}{[\theta(1+i)]^2}\right]^2 \tag{71}$$

The Moment about the Mean of Zech distribution is thus derived.

The rth central moment of a random variable X having *Zech* distribution is given by

$$E((x-\mu)^r) = \int_0^\infty (x-\mu)^r f(x)dx \tag{72}$$

$$E((x-\mu)^r) = \int_0^\infty (x-\mu)^r \frac{\alpha^{k+1}}{\beta^k k!\,(1+i)} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \left[\theta(1+i)e^{-\theta(1+i)x}\right] dx \tag{73}$$

$$E((x-\mu)^r) = \frac{\theta\alpha^{k+1}}{\beta^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \int_0^\infty (x-\mu)^r \left[e^{-\theta(1+i)x}\right] dx \tag{74}$$

By setting $y = x - \mu,\ \frac{dy}{dx} = 1, \quad dx = dy, \quad x = y + \mu$

$$E((x-\mu)^r) = \frac{\theta\alpha^{k+1}}{B^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \int_0^\infty y^r\, e^{-\theta(1+i)(y+\mu)} dy \tag{75}$$

$$E(X^r) = \frac{\theta\alpha^{k+1}}{B^k k!} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty w_{i,j,k} \int_0^\infty y^r\, e^{-\theta(1+i)y} . e^{-\theta(1+i)\mu} dy \tag{76}$$

$$E((x-\mu)^r) = \frac{\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\, e^{-\theta(1+i)\mu}\int_0^{\infty} y^r\, e^{-\theta(1+i)y} dy \tag{77}$$

Using the Gamma function expansion,

$$\frac{\Gamma a}{b^a} = \int_0^{\infty} x^{a-1}e^{-bx}\,dx$$

$$a-1 = r,\quad a = r+1,\quad b = \theta(1+i)$$

$$E((x-\mu)^r) = \frac{\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\, e^{-\theta(1+i)\mu}\,\frac{\Gamma_{(r+1)}}{[\theta(1+i)]^{(r+1)}} \tag{78}$$

X. Moment - based measures of Skewness and Kurtosis

The skewness of *Zech* distribution based on central moment is given as

$$\text{Skewness} = \frac{\mu_3^2}{\mu_2^3} \tag{79}$$

Where $\mu_3 = Third\ central\ momen$
$\mu_2 = Second\ central\ moment$

$$\mu_3 = \frac{6\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^4} \tag{80}$$

$$\mu_2 = \frac{2\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^3} \tag{81}$$

$$Skewness = \frac{\left[\frac{6\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^4}\right]^2}{\left[\frac{2\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^3}\right]^3} \tag{82}$$

$$\text{Kurtosis} = \frac{\mu_4}{\mu_2^2} \tag{83}$$

But $\mu_4 = \frac{24\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^5}$ (84)

The Kurtosis, based on central moment of *Zech* distribution is given by

$$Kurtosis = \frac{\frac{24\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^5}}{\left[\frac{2\theta\alpha^{k+1}}{B^k k!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} w_{i,j,k}\,\frac{e^{-\theta(1+i)\mu}}{[\theta(1+i)]^3}\right]^2} \tag{85}$$

XI. Results

I. Simulation Studies

The behavior of the parameters of Zech distribution was investigated through simulation studies using R statistical software. Data were replicated 1000 times. A random sample of sizes 50, 100, 150 and 200 were selected. The parameters were varied as follows: $\alpha = 0.5$, $\theta = 0.5$, and $\beta = 0.5$; and $\alpha = 1$, $\theta = 1$, and $\beta = 1$; and $\alpha = 1.5$, $\theta = 1.5$, and $\beta = 1.5$ respectively. The maximum likelihood estimates of the true parameters, the bias, standard error and Root Mean Square Error were obtained from the simulation. The results are shown in Tables 1, 2 and 3.

Results of Simulation Studies

Table 1: *Simulation study at* $\alpha = 0.5$, $\theta = 0.5$, *and* $\beta = 0.5$

N	Parameters	Means	Bias	Std. Error	RMSE
50	$\alpha = 0.5$	0.0339	0.4661	0.0152	0.0174
	$\theta = 0.5$	0.6382	- 0.1382	0.1931	0.0621
	$\beta = 0.5$	0.3763	0.1237	0.2300	0.0678
100	$\alpha = 0.5$	0.0466	0.4534	0.0170	0.0130
	$\theta = 0.5$	0.5295	- 0.0295	0.1348	0.0367
	$\beta = 0.5$	0.5720	-0.0720	0.1457	0.0382
150	$\alpha = 0.5$	0.0667	0.4333	0.0185	0.0111
	$\theta = 0.5$	0.4228	0.0772	0.0849	0.0238
	$\beta = 0.5$	0.5170	- 0.0170	0.1100	0.0271
200	$\alpha = 0.5$	0.0818	0.4182	0.0221	0.0105
	$\theta = 0.5$	0.4815	0.0185	0.0927	0.0215
	$\beta = 0.5$	0.6211	- 0.1211	0.0987	0.0222

Table 2: *Simulation study at* $\alpha = 1.0$, $\theta = 1.0$, *and* $\beta = 1.0$

N	Parameters	Means	Bias	Std. Error	RMSE
50	$\alpha = 1.0$	0.0627	0.9373	0.0446	0.0299
	$\theta = 1.0$	0.2452	0.7548	0.4291	0.0926
	$\beta = 1.0$	0.1369	0.8631	0.4911	0.0991
100	$\alpha = 1.0$	0.0834	0.9166	0.0363	0.0191
	$\theta = 1.0$	0.8306	0.1694	0.1951	0.0442
	$\beta = 1.0$	0.9365	0.0635	0.2667	0.0516
150	$\alpha = 1.0$	0.1469	0.8531	0.0536	0.0189
	$\theta = 1.0$	1.0107	- 0. 0107	0.1898	0.0356
	$\beta = 1.0$	0.9909	0.0091	0.2452	0.0404
200	$\alpha = 1.0$	0.1609	0.8391	0.0515	0.0160
	$\theta = 1.0$	0.8606	0.1394	0.1484	0.0272
	$\beta = 1.0$	0.9995	0.0005	0.1901	0.0308

Table 3: *Simulation study at* $\alpha = 1.5$, $\theta = 1.5$, *and* $\beta = 1.5$

N	Parameters	Means	Bias	Std. Error	RMSE
50	$\alpha = 1.5$	0.0636	1.4364	0.5580	0.0334
	$\theta = 1.5$	1.5227	- 0.0227	0.5782	0.1075
	$\beta = 1.5$	1.9951	- 0.4951	0.5450	0.1044
100	$\alpha = 1.5$	0.0821	1.4179	0.0517	0.0227
	$\theta = 1.5$	1.2049	0.2951	0.3506	0.0592
	$\beta = 1.5$	1.9077	-0.4074	0.3208	0.0566
150	$\alpha = 1.5$	0.1249	1.3751	0.0621	0.02035
	$\theta = 1.5$	1.0912	0.4088	0.2543	0.04120
	$\beta = 1.5$	0.1713	1.3287	0.2802	0.04320
200	$\alpha = 1.5$	0.2188	1.2812	0.0825	0.02031
	$\theta = 1.5$	1.3248	0.1752	0.2275	0.03370
	$\beta = 1.5$	1.6958	- 0.1958	0.2539	0.03560

The tables 1, 2 and 3, for each of the selected true parameter values show that as the sample sizes increase, the root mean square errors decrease. This implies that the parameters of Zech distribution are stable.

II. Applications to Real Life Data Sets

The performance of Zech distribution, as well as goodness of fit tests when fitted to the real life data sets is hereby compared with other three – parameter distributions such as Gompertz Inverse Exponential (GIE) distribution, Weibull Exponential (WE) distribution and Gompertz Exponential (GE) distribution are provided in tables 6 and 7 respectively.
To select the best among the competing distributions, the following statistical criteria are used: Negative Log-likelihood, Akaike Information criterion (AIC) and Bayesian Information criterion (BIC). The distribution having the least value of the criteria above is adjudged to be the best. Also, the goodness of fit tests like Kolmogorov-Smirnov Statistic (KS) and Anderson-Darling Statistic (ADS) are also computed to select the best fit. The best distribution has the least value of the statistics above.

Data 1: The first data set represents the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli observed and reported by Bjerkedal [9] and used by Adewara [1].
10, 33, 44, 56, 59, 72, 74, 77, 92, 93, 96, 100,100,102, 105, 107, 107, 108, 108, 108, 109, 112, 113, 115,116, 120, 121, 122, 122, 124, 130, 134, 136, 139, 144, 146,153, 159, 160, 163, 163, 168, 171, 172, 176, 183, 195, 196,197, 202, 213, 215, 216, 222, 230, 231, 240, 245, 251, 253,254, 254, 278, 293, 327, 342, 347, 361, 402, 432, 458, 555

Table 4: *Descriptive Statistics for data 1.*

Min	1st Quartile	Median	Mean	3rd Quartile	Max.	Standard Deviation	Skewness	Kurtosis
10.0	108.0	149.5	176.8	224.0	555.0	103.4549	1.341869	4.991056

Table 5: *Performance rating for the fitted models using data 1.*

Distributions	Estimates	-LL	AIC	BIC	KS	ADS
Zech	$\hat{\alpha} = 5.1884384$ $\hat{\beta} = -0.6383047$ $\hat{\theta} = 0.0128380$	424.8790	855.7579	862.5879	0.0835	0.4925
GIE	$\hat{\alpha} = 0.02683774$ $\hat{\beta} = 1.88823061$ $\hat{\theta} = 22.03993914$	427.6661	861.3322	868.1622	0.1095	1.0777
WE	$\hat{\alpha} = 1.157801294$ $\hat{\beta} = 1.340608545$ $\hat{\theta} = 0.002981762$	431.3887	868.7774	875.6074	0.1195	1.9894
GE	$\hat{\alpha} = 0.004089507$ $\hat{\beta} = 1.08290306$ $\hat{\theta} = 0.002751085$	434.3901	874.7802	881.6102	0.1759	2.6168

The distributions tested showed the performances of each. The results revealed that the distribution with the lowest value of –LL, AIC, BIC, KS and ADS is considered to be the best. From Table 5, *Zech* distribution had the least value of 424.8790 for –LL, AIC= 855.7579, BIC = 862.5879, KS = 0.0835 and ADS = 0.4925 hence, it was considered the best fitted distribution among other distributions.

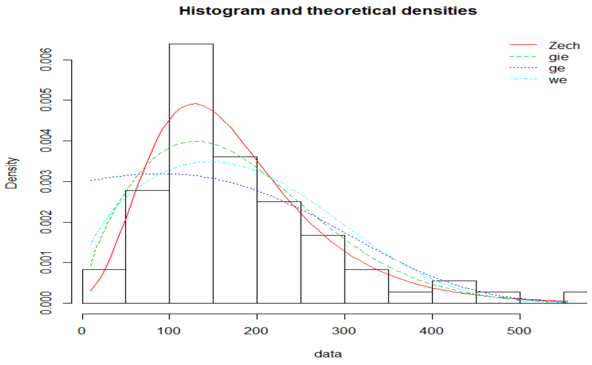


Figure 5: *Histogram and theoretical densities for data 1*

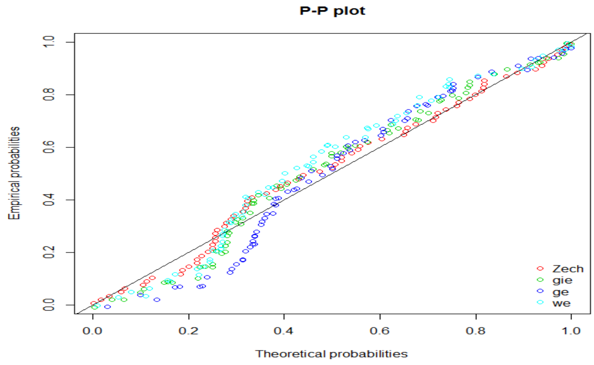


Figure 6: *P – P plot for data 1*

Figures 5 and 6 respectively depict the performance of the new distribution with survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli observed, this is compared to other distributions mentioned in the research.

Data 2: The second dataset represents the gauge of length of 10mm observed by Mohammed [10]

1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020

Table 6: *Descriptive Statistics for data 2.*

Min	1st Quartile	Median	Mean	3rd Quartile	Max.	Standard Deviation	Skewness	Kurtosis
1.901	2.554	2.996	3.059	3.422	5.020	0.6209216	0.6178407	3.286345

Table 7: *Performance rating for the fitted models using data 2.*

Distributions	Estimates	-LL	AIC	BIC	KS	ADS
Zech	$\hat{\alpha} = 238.223663$ $\hat{\beta} = -5.177056$ $\hat{\theta} = 1.971310$	56.5097	119.0194	125.4488	0.0885	0.3571
GIE	$\hat{\alpha} = 330.740054$ $\hat{\beta} = -41.32409$ $\hat{\theta} = 18.50223$	57.28159	120.5632	126.9926	0.0903	0.3846
WE	$\hat{\alpha} = 1.8379671$ $\hat{\beta} = 3.7232674$ $\hat{\theta} = 0.1842052$	63.6584	133.3168	139.7462	0.0986	1.1866
GE	$\hat{\alpha} = 0.944957015$ $\hat{\beta} = 1.480487938$ $\hat{\theta} = 0.008481201$	69.1480	144.2960	150.7254	0.1389	1.9840

The distributions tested showed the performances of each. The results revealed that Zech distribution had the lowest value of –LL, AIC, BIC, KS and ADS and is considered to be the best. From Table 7, *Zech* distribution had the least value of = 56.5097 for –LL, AIC=119.0194, BIC =125.4488, KS = 0.0885 and ADS= 0.3571 hence, it was considered the best fitted distribution among other distributions.

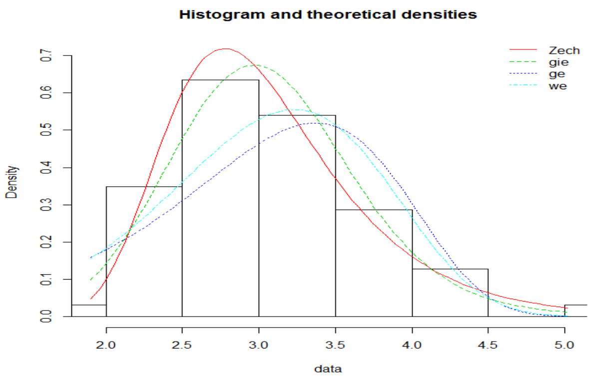


Figure 7: *Histogram and theoretical densities for data 2*

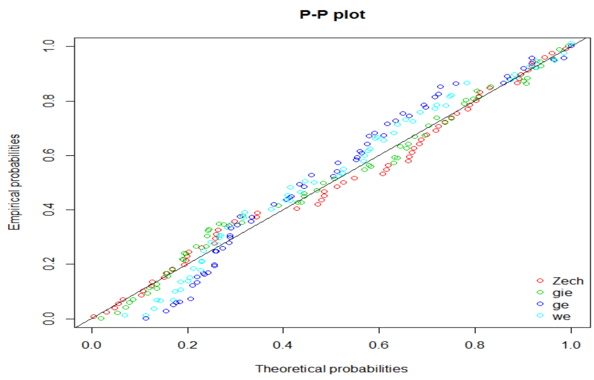


Figure 8: *P – P plot for data 2*

Figures 7 and 8 respectively depict the performance of the new distribution with gauge of length of 10mm data, and compared to other distributions mentioned in the research

XII. Discussion

The histogram and theoretical densities plot shows that *Zech* distribution fits data 1 and 2 best. Also, the probability plot i.e. (the PP plot) which compares the empirical cdf of data sample with specified theoretical cumulative distribution, reveals that *Zech* distribution is closer to the line than the remaining three fitted models. Tables 1, 2 and 3, for each of the selected true parameter values show that as the sample sizes increase, the Root Mean Square errors decrease which shows that the parameters of Zech distribution are stable.

Tables 4 and 6 show that the data is skewed to the right. Interestingly, the shape of the pdf graph of *Zech* distribution is also positively skewed. Also, the Kurtosis values of 4.991056 and 3.2866345 suggest that the data is leptokurtic. Data 1 and 2 have a kurtosis of 1.991056 and 0.2866345 respectively above that of normal distribution which is 3.0

Tables 5 and 7, show that Zech distribution has the lowest value of –LL, AIC, BIC, KS and ADS and is considered to be the best when compared with the competing distributions such as Gompertz Inverse Exponential distribution, Weibull Exponential distribution and Gompertz Exponential distribution.

XIII. Conclusion

In this paper, a new three – parameter continuous distribution named Zech distribution was proposed from Gompertz Inverse Exponential distribution. Its probability density function was plotted and the result revealed a heavy positively skewed distribution which is suitable for modelling heavily right-tailed data. Several statistical and mathematical properties of the new distribution were derived. The results of the simulation studies revealed that the parameters of the new distribution are stabled and as the sample sizes increased, the Root Mean Square (RMS) errors decreased. The applications to two real life data sets showed that Zech distribution has the lowest -LL, AIC, BIC, KS and ADS when compared with other competing distributions such as Gompertz Inverse Exponential distribution, Weibull Exponential distribution and Gompertz Exponential distribution used in this research paper.

References

[1] Adewara, J. A, Adeyeye, J. S and Thron, C. P "Properties and Applications of the Gompertz Distribution". *International Journal of Mathematical Analysis and Optimization: Theory and Applications*. Vol. 2019, No. 1, pp 443 – 454, 2019, http://ijmao.unilag.edu.ng/article/view/346

[2] El – Gohari A, Ahmad A and Adel Naif Al – Otaibi. "The Generalized Gompertz Distribution". *Applied Mathematics Modelling*, 37(2013) 13 – 24, http://dx.doi.org/10.2016/j.apm.2011.05.017

[3] Adewara J .A and Adeyeye J .S. "The Comparative Study of Gompertz Exponential Distribution and Other Three – Parameter Distributions of Exponential Class". *Covenant Journal of Physical & Life Sciences (CJPL).* Vol. 8, No. 1, June 2020. ISSN: p. 2354 – 3574, e. 2354 – 3485, http://journals.covenantuniversity.edu.ng/index.php/cjpl

[4] Morad A. and Indranil G. "The Gompertz – G family of distributions". *Journal of statistical theory and practice.* 11(1), 179 – 207, DOI: 10.1080/15598608.2016.1267668, (20

[5] Pelumi, E .O, Mundher, A. K, Hilary O. and Abiodun O. "The Gompertz Inverse Exponential distribution". *Cogent Mathematics and Statistics*. 5(1), DOI: 10.1080/25742558.2018.1507122

[6] Galton F. "Enquires into Human Faculty and its Development". *Macmillan and Company, London.* 1983.

[7] Moors, J. J. "A quantile alternative for kurtosis". *The Statistician*, 1988, 37, 25 – 32.

[8] Adewara, J. A, Adeyeye, J. S, Mundher, A K and Aako O .L. "Exponentiated Gompertz Exponential Distribution: Derivation, Properties and Applications". *ISTATiSTiK: Journal of the Turkish Statistical Association*. Vol. 13, No. 1, January 2021, pp. 12 – 28, ISSN 1300 – 4077 |21|1|12|28

[9] Bjerkedal, T. (1960), "Acquisition of Resistance in Guinea Pigs Infected with Different Doses of Virulent Tubercle Bacilli". *American Journal of Hygiene,* 72, 130-148. https://doi.org/10.1093/oxfordjournals.aje.a120129

[10] Mohammed T, Mundher, A K, Pelumi, E O and Moudher K A. "A New Version of the Exponentiated Burr -X distribution". *Journal of Physics Conference Series.* March 2021. DOI: 10.1088/1742 – 6596/1818(1):012116

[11] Ogunsanya S .A, Yahya W. B, Adegoke T. M, Iluno C, Aderele O. R. and Ekum M. I. "A New Three – Parameter Weibull Inverse Rayleigh Distribution: Theoretical Development and Applications". *Mathematics and Statistics* 9(3):249 – 272, 2021. DOI: 10.13189/ms.2021.0903

[12] Eliwa M .S, El-Morshedy M. and Mohamed I. "Inverse Gompertz Distribution: Properties and Different Estimation Methods with Application to Complete and Censored Data". *Annals of Data Science*. DOI: 10.1007/s40745-018-0173-0. 11 August, 2018.

[13] Said A, Ahmed Z. A, Elbatal I, Elgarhy M. "The Extended Inverse Weibull Distribution: Properties and Applications". *Hindawi.* Vol. 2020, Article ID 3297693, October (2020)

A New Effective Approach to Solve Fuzzy Transportation Problems

Kaushik A Joshi

•

Department of First Year Engineering
Prof Ram Meghe College of Engineering and Management Badnera-Amravati (MS India)
kaushikanandjoshi@gmail.com

Kirankumar L. Bondar

•

P. G. Department of Mathematics, Govt Vidarbh Institute of Science and Humanities, Amravati.
klbondar_75@rediffmail.com

Pandit U. Chopade

•

Research Supervisor, Department of Mathematics,
D.S.M's Arts Commerce and Science College ,Jintur.
chopadepu@rediffmail.com

Abstract

In this article, we will take a look at FTP, and then present a way to solve many such problems by using the affected method for FN level. Some of the numbers in FTP may be sharp or sharp numbers. In many decision-making problems, numbers are represented in terms of FN. FN can be normal or oblique, triangular or trapezoidal or any other FN LR. So, some FNs do not compare immediately. First, we convert QF such as price, quantity, supply and demand, into accurate quantities by using our system, and then using sophisticated algorithms, we solve and solve the problem. The new system is a configuration, easy to install and can be used for all types of TP, or to increase or decrease the target function. In the end, this process is illustrated by digital examples.

Keywords: Fuzzy Transportation Problem, Fuzzy Numbers, Optimization, Ranking of Fuzzy Numbers

Abbreviations

Optimal Solution	:	OS
Linear Programming Problem	:	LPP
Fuzzy Number	:	FN
Transportation Problem	:	TP
Fuzzy Transportation Problem	:	FTP
Shipping Cost	:	SC
Transportation Model	:	TM
Fuzzy Quantities	:	FQ
Fuzzy Solution	:	FS

Feasible Solution	:	FeS
Transportation Cost	:	TC
Right Hand Side	:	RHS
Transportation Tableau	:	TT

I. Introduction

TP is a unique LPP that manifests itself in many useful applications. In this issue, we determine the best shipping method between source or source and destination. Many terms unrelated to transportation have this system. Suppose *n* from the start will give me a place to use some products. Let b_k be the number of products in the beginning *k*, and c_l be the number of products required in place *l*. Also, we assume that the cost of shipping a unit of product from the beginning k to the destination l is d_{kl}, then we leave y_{kl} to represent the cost from the beginning *k* to point *l*. In the case of SC, it is thought that the corresponding number of upgrades from each start to anywhere to reduce the total SC would be LPP. The TMs have fast and easy delivery and supply chain to reduce costs. When the price range and the number of supply and demand are well known, several algorithms are developed to optimize TP. But, in the real world, there are many cases where the cost factors and the amount of supply and demand are FQ. FTP is the TP of TC, the supply and the required number are FQs. Most of the current systems only provide a clear solution for FTP. [1-3] have devised a system to fix FTP. [4] obtained FS for two steps reducing the FTP value of the donor and the required trapezoidal FNs. [5] has developed a system, i.e. a zero-path path, to find an operating system for FTP where all parameters are trapezoidal FNs.

In many decision-making problems, data is represented in the FN system. In FTP, all parameters is FN. FN can be normal or irregular, triangular or trapezoidal. So, some FNs do not compare immediately. Comparisons between two or more NFs in terms of the numbers are one of the key topics, and how to describe the level of NF is one of the key topics. Introducing several methods for processing FN. Here, we want to use the introductory method for the design of FN, by [6]. Now we want to apply this method to all FTP, where all parameters can be trapezoidal FN, triangular FN, FN LR arbitrary, normal FN or negative FN. This process is very easy to understand and apply. Finally, the operating system of the problem can be accessed as FN or net number.

II. Mathematical Formulation of a FTP

In mathematics, useful work can be said as follows:
Minimize

$$w = \sum_{k=1}^{n} \sum_{l=1}^{n} d_{kl} y_{kl} \tag{2.1}$$

Subject to

$$\begin{aligned} &\sum_{k=1}^{n} y_{kl} = b_k && k = 1, 2, 3..., n \\ &\sum_{k=1}^{n} y_{kl} = c_l && l = 1, 2, 3..., n \\ &y_{kl} \geq 0 && k = 1, 2, 3..., n \quad l = 1, 2, 3..., n \end{aligned} \tag{2.2}$$

where d_{kl} is the TP of a unit from kth source to lth point, and the y_{kl} value must be a positive or negative number, which is transmitted from kth source to lth point. What is clear is that the absolute conditions for the LPP given in (2.1) to find a solution are as follows

$$\sum_{k=1}^{m} b_k = \sum_{l=1}^{m} c_l \tag{2.3}$$

that is, assume that the available sum is equal to the required sum. If this is not true, the source or destination may be added. It should be noted that this problem has FeS if the (2.2) condition is satisfactory. Now the problem is to determine y_{kl}, so that the total TC total.

In mathematics, FTP can be defined as:

$$w = \sum_{k=1}^{n} \sum_{l=1}^{n} d^*_{kl} y_{kl} \tag{2.4}$$

Subject to

$$\begin{aligned} &\sum_{k=1}^{n} y_{kl} = b^*_k && k = 1,2,3\ldots,n \\ &\sum_{k=1}^{n} y_{kl} = c^*_l && l = 1,2,3\ldots,n \\ &y_{kl} \geq 0 \qquad k = 1,2,3\ldots,n && l = 1,2,3\ldots,n \end{aligned} \tag{2.5}$$

of the number of TCs d^*_{kl}, supply b^*_k and demand c^*_l is FQ. An important and absolute condition for non-essential LPP given in (2.4) and (2.5) to find a solution is that

$$\sum_{k=1}^{m} b_k^* \cong \sum_{l=1}^{m} c_l^* \tag{2.6}$$

A large number of systems are provided for FTP. Some of them are based on FN level. Some of the NF-level systems, for example, have limitations, are difficult to compute, or lack understanding, making them ineffective and useful applications, especially in the decision-making process. However, in some of these methods, such as those compared to FN as their centroid point [7-11], the decision maker does not work it each with comparisons between FN. However, there are some ways to compare FNs individually [12-14]. It is not always the case that there is no point in the nature of uncertainty and incorrect title information, but these situations often occur in practice when expressing language words. For this reason, when comparing two FNs, it is natural that the results of the comparison are inaccurate or, at least, parametric, due to its own nature and specifications. This can also be seen in the variability of practitioners in the fuzzy set theory. We clearly see that non-parametric decision makers and non-parametric practitioners perform better than non-parametric practitioners with respect to experimental data [15, 16]. Two factors play an important role in the decision-making process: Contributor's decision-making and decision-making process, ease of total. This essay attempts to provide a system of degrees and compares NF to account as much as possible of the factors mentioned above. The expected mechanism was also discussed in the centroid system [17, 18].

III. Definition of an Arbitrary FN

FN has been described in various forms. We use the next FN definition very accurately. We present the FN B_w strongly by the two prescribed paths *(B (s), B * (s))*, where $0 \leq s \leq w$ and w are the fluctuations between zero and one $(0 \leq w \leq 1)$, with a parametric shape that meets the requirements:

- *B (s)* is a continuous function that does not go down the top left *[0, w]*.

- *B * (s)* is a continuous incremental function left at the top *[0, w]*.
- *B (s) ≤ B * (s), 0 ≤ s ≤ w.*

The net number "*i*" represents only *B (s) = B * (s) = i, 0 ≤ s ≤ w*. By proper definition, the holes FNs *{B (s), B * (s)}*, becoming convex cone F1 are isomorphic and isometric embedded in the Banach hole. If B is FN then cut β of B is *[B *] β = [B (β), B * (β)]*, *0 ≤ β ≤ w*. If *w = 1*, the coefficient described above is called the normal FN.

IV. An Approach for Making Ranking FN

As mentioned earlier, it seems that the parametric method of FN comparison, mainly in the theory of non-parametric determination, is better than non-parametric methods. For example, in a centroid system from a study [19], FN was compared to their Euclidean origin from the beginning. Negative FN is not included in the Cheng centroid core system. Sometime later, however, [20] tried to solve this problem by using the area between the centroid point at the beginning. But their study was also flawless. [21] found that the regional systems of a study lead to some time in unintelligible planning. That study showed a marked eye pattern. But their method is not parametric and is only available in normal FN. It is well known that non-parametric methods compared to FN have some setbacks in practice.

According to the above definition of FN, as $B_w = (B(s), B^*(s)), (0 \le s \le w)$ u_n FN, then the value $N(B_w)$, assigned to B_w for decision. Levels greater than "*β*" calculated as follows:

$$N_\beta(B_w^*) = \frac{1}{2}\int_\beta^w \{B(s) + B^*(s)\}ds \quad \text{where} \quad 0 \le \beta < 1$$

This number will be used as a basis for comparing FN with a resolution level higher than β.

V. Trapezoidal and Triangular FNs

The two major classes of FN, commonly used for practical and easy-to-use purposes, are the "trapezoidal and triangular FN", some ways of bringing the FN and the trapezoidal FN closer and closer, see also [22-24] and therefore there is no concern in this.

I. Triangular FNs

FN, B is a triangular FN called *(σ, n, α)* where *σ*, *n* and *α* is a real number and its functions are given by $u_B(y)$ below,

$$u_B(y) = 0 \qquad \text{for } y \le \sigma$$

$$u_B(y) = \frac{y-\sigma}{n-\sigma} \qquad \text{for } \sigma \le y \le n$$

$$u_B(y) = 1 \qquad \text{for } y = n$$

$$u_B(y) = \frac{\alpha - y}{\alpha - n} \qquad \text{for } n \le y \le \alpha$$

$$u_B(y) = 0 \qquad \text{for } y \ge \alpha$$

According to the description above of the triangular FN, let *B = (B (s), B * (s)), (0 ≤ s ≤ 1)* one FN, then the value *N(B)*, to assign with B calculated as follows:

$$N - Tra(B) = \frac{1}{2}\int_0^1 \{B(s) + B^*(s)\}ds = \frac{1}{4}\left[2n + \sigma + \alpha\right]$$

that is very useful for calculations.

II. Trapezoidal FNs

A FN, *B* is a trapezoidal FN identified by the symbol *(σ, n, m, α)*, where *σ, m, n, α* are real numbers and the membership function $u_B(y)$ is given below.

$$\begin{array}{ll} u_B(y) = 0 & \text{for } y \le \sigma \\ u_B(y) = \dfrac{y-\sigma}{n-\sigma} & \text{for } \sigma \le y \le n \\ u_B(y) = 1 & \text{for } n \le y \le m \\ u_B(y) = \dfrac{m-y}{\alpha-m} & \text{for } n \le y \le \alpha \\ u_B(y) = 0 & \text{for } y \ge \alpha \end{array}$$

According to the above definition of the FN trapezoid, let *B = (B (s), B * (s)), (0 ≤ s ≤ 1)* FN, then the value of *N (B)*, set B, as the following is calculated. :

$$N - Tra(B) = \frac{1}{2}\int_0^1 \{B(s) + B^*(s)\}ds = \frac{1}{4}\left[n + m + \sigma + \alpha\right]$$

that is very useful for calculations.

VI. A Newly Developed Approach for Solving FTP

We are now introducing a new approach to FTP solutions where key contributors, resources and requests are FNs. FNs in each problem can be triangular, trapezoidal, or any FN or mixture. The FTP operating system can be downloaded explicitly or implicitly.

First Step: Calculate the values of N (.) For each fuzzy data, TC d_{kl}, supply b_k, and c_l demand values, which are FQ.

Second Step: By replacing *N (* d_{kl} *), N (* b_k *)* and *N (* c_l *)* which are fragile values with d_{kl}, b_k, and c_l values which are FQ, you select a new fragile TP.

Third Step: Fix a new TP net, through the old system, and get the problematic network. Note that each answer in the laboratory will have a specific *(n-m-1)* point FS. We know that OS y_{kl} must be an integer or an integer, but OS y_{kl} for net TP may be an integer or not an integer, because the RHS of the problem is an FN which is an integer. If you accept a soft solution, stop. The OS is in your hands. If you want the kind of nonsense solution, go to the next step.

Forth Step: Determine where the FS are not missing in the TT. The background is a rooted tree, that is, there must be at least one cell in each row and in each column of the TT. In addition, the foundation must be wood, that is, cells *(n-m-1)* cells will not have a circle. Therefore, there are rows and columns with only one main cell. Starting with these cells, calculate the nonlinear core solution, and continue until you obtain a *(n-m-1)* foundation solution.

VII. Examples

The following example may be useful to clarify the proposed procedure:

Example: Consider the following FTP which is in [25]. All data for this problem are trapezoidal FNs. We want to use our method to solve it, then we will compare the results.

Table 1: *FNs, Demand and Supply*

	1	2	3	4	Supply
1	4,5,6,7	4,6,7,9	12,14,15,17	8,10,11,14	4,9,10,13
2	3,4,5,7	2,3,4,5	8,9,10,11	3,4,5,6	3,4,5,6
3	6,8,9,11	8,11,12,15	15,18,19,22	10,12,13,15	8,13,15,20
Demand	8,10,11,13	4,8,9,13	4,6,7,9	4,5,6,7	

According to the description above of the trapezoidal FN or $A = (A(s), A^*(s))$, $(0 \le s \le 1)$ FN, then the value $N(A)$, is assigned and Calculated as part:

$$N - Tra(A) = \frac{1}{2}\int_0^1 \{A(s) + A^*(s)\}ds = \frac{1}{4}\left[n + m + \sigma + \alpha\right]$$

Thus, we obtain the values of $N(B_{kl})$, $N(b_k)$ and $N(c_l)$ according to the recommended formula:

Table 2: *Trapezoidal FNs*

FNs	Trapezoidal FNs
B11=4,5,6,7	N-Tra(B11)=1/4(4+5+6+7)=5.5
B12=4,6,7,9	N-Tra(B12)=1/4(4+6+7+9)=6.5
B13=12,14,15,17	N-Tra(B13)=1/4(12+14+15+17)=14.5
B14=8,10,11,14	N-Tra(B14)=1/4(8+10+11+14)=10.75
B21=3,4,5,7	N-Tra(B21)=1/4(3+4+5+7)=4.75
B22=2,3,4,5	N-Tra(B22)=1/4(2+3+4+5)=3.5
B23=8,9,10,11	N-Tra(B23)=1/4(8+9+10+11)=9.5
B24=3,4,5,6	N-Tra(B24)=1/4(3+4+5+6)=4.5
B31=6,8,9,11	N-Tra(B31)=1/4(6+8+9+11)=8.5
B32=8,11,12,15	N-Tra(B32)=1/4(8+11+12+15)=11.5
B33=15,18,19,22	N-Tra(B33)=1/4(15+18+19+22)=18.5
B34=10,12,13,15	N-Tra(B34)=1/4(10+12+13+15)=12.5

and fuzzy supplies are given as:

Table 3: *Trapezoidal Fuzzy Supplies*

Supply	Trapezoidal Fuzzy Supply
b1=4,9,10,15	N-Tra(b1)=1/4(4+9+10+15)=9.5
b2=3,4,5,6	N-Tra(b2)=1/4(3+4+5+6)=4.5
b3=8,13,15,20	N-Tra(b3)=1/4(8+13+15+20)=14

and fuzzy demands are given as:

Table 4: *Trapezoidal Fuzzy Demands*

Demand	Trapezoidal Fuzzy Demand
c1=8,10,11,13	N-Tra(c1)=1/4(8+10+11+13)=10.5
c2=4,8,9,13	N-Tra(c2)=1/4(4+8+9+13)=8.5
c3=4,6,7,9	N-Tra(c3)=1/4(4+6+7+9)=6.5
c4=4,5,6,7	N-Tra(c4)=1/4(4+5+6+7)=5.5

total fuzzy supply is given as: T = (9, 20, 24, 35) as well as total fuzzy demand is given as: F = (11, 20, 24, 33), hence:

Table 5: *Trapezoidal Total Fuzzy Demands and Supply*

Demand / Supply	Trapezoidal Fuzzy Demand
T=9,20,24,35	N-Tra(T)=1/4(9+20+24+35)=22
F=11,20,24,33	N-Tra(F)=1/4(11+20+24+33)=22

Since *N-Tra (T) = N-Tra (F)*, the given problem is a valid problem. Now using our system, we convert FTP to pure TP. So, we have these reduced FTP:

Table 6: *Reduced FTP*

	1	2	3	4	Supply
1	6.4	7.4	15.4	11.65	10.4
2	5.65	4.4	10.4	5.4	5.4
3	9.4	12.4	19.4	13.4	15
Demand	11.4	9.4	7.4	6.4	

As shown in Table 6, the defuzzification results of FN obtain values of non-numerical values. Therefore, the existence of a negative value in TP is next to the fact that the solution of net TP is not important. Note that the solution and the useful function are quantitative, because its matrix is uni-modular (26). If we solve the new problem, we will get the following answers:

$$y_{12} = 9.4\ ,\ y_{13} = 5\ ,\ y_{23} = 5.4\ ,\ y_{31} = 11.4\ ,\ y_{33} = 5\ ,\ y_{34} = 6.4\ ,$$

and the total value of the problem is $y_0 = 160$.

Table 7: *Reduced Solution*

	1	2	3	4	Supply
1		9.4	5		14.4
2			5.4		5.4
3	11.4		5	6.4	22.8
Demand	11.4	9.4	15.4	6.4	

Now we can go back to the original problem and get the FS of FTP based on the data in table-7.

Table 8: *Reduced FNs, Demand and Supply*

	1	2	3	4	Supply
1		4,8,9,13	-6,3,5,14		4,9,10,15
2			3,4,5,6		3,4,5,6
3	8,10,11,13		-6,2,6,14	4,5,6,7	8,13,15,20
Demand	8,10,11,13	4,8,9,13	4,6,7,9	4,5,6,7	

where the default OS for FTP provided is:

$y^*_{12} = (4,8,9,13)$; $y^*_{13} = (-6,3,5,14)$; $y^*_{23} = (3,4,5,6)$; $y^*_{31} = (8,10,11,13)$; $y^*_{33} = (-6,2,6,14)$; $y^*_{34} = (4,5,6,7)$

The results are the same as to the previous studies (4, 8). Note that the benefits of the changes are the same, but the benefits of objective work are different. The best CT net value for the problem provided by the Pandian system is 132.17 where it is obtained from our system is 160. That seems clear is that there is no single system to compare FN, and that different approaches can meet different desirable requirements.

VIII. Conclusion

In this article, a simple but effective parametric method has been introduced to configure FTP using the FN interface. This method can be used for all types of FTP, either triangular or trapezoidal FN with normal or negative data. The new system is a configuration, easy to install and can be used for all types of TP, or to increase or decrease the target function.

References

[1] Ebrahimnejad A, Verdegay JL.(2018) A new approach for solving fully intuitionistic fuzzy transportation problems. *Fuzzy optimization and decision making*.17(4):447-74.

[2] Singh SK, Yadav SP. (2016) A new approach for solving intuitionistic fuzzy transportation problem of type-2. *Annals of operations research*.243(1):349-63.

[3] Maity G, Mardanya D, Roy SK, Weber G-W. (2019) A new approach for solving dual-hesitant fuzzy transportation problem with restrictions. Sādhanā.44(4):1-11.

[4] Chakraborty D, Jana DK, Roy TK. (2015) A new approach to solve multi-objective multi-choice multi-item Atanassov's intuitionistic fuzzy transportation problem using chance operator. *Journal of intelligent & fuzzy systems*.28(2):843-65.

[5] Roy SK, Ebrahimnejad A, Verdegay JL, Das S. (2018) New approach for solving intuitionistic fuzzy multi-objective transportation problem. *Sādhanā*.43(1):1-12.

[6] Thamaraiselvi A, Santhi R. (2016) A new approach for optimization of real life transportation problem in neutrosophic environment. *Mathematical Problems in Engineering*.

[7] Prakash KA, Suresh M, Vengataasalam S. (2016) A new approach for ranking of intuitionistic fuzzy numbers using a centroid concept. *Mathematical Sciences*. 10(4):177-84.

[8] Ebrahimnejad A. (2015) An improved approach for solving fuzzy transportation problem with triangular fuzzy numbers. *Journal of intelligent & fuzzy systems*.29(2):963-74.

[9] Ebrahimnejad A. (2016) New method for solving fuzzy transportation problems with LR flat fuzzy numbers. *Information Sciences*.357:108-24.

[10] Soto J, Melin P, Castillo O. (2018) A new approach for time series prediction using ensembles of IT2FNN models with optimization of fuzzy integrators. *International Journal of Fuzzy Systems.* 20(3):701-28.

[11] Singh SK, Yadav SP. (2015) Efficient approach for solving type-1 intuitionistic fuzzy transportation problem. *International Journal of System Assurance Engineering and Management*. 6(3):259-67.

[12] Asan U, Kadaifci C, Bozdag E, Soyer A, Serdarasan S. (2018) A new approach to DEMATEL based on interval-valued hesitant fuzzy sets. *Applied Soft Computing*.66:34-49.

[13] Kalid N, Zaidan A, Zaidan B, Salman OH, Hashim M, Albahri OS, et al. (2018) Based on real time remote health monitoring systems: a new approach for prioritization "large scales data" patients with chronic heart diseases using body sensors and communication technology. *Journal of medical systems*.42(4):1-37.

[14] de Farias Aires RF, Ferreira L. (2019) A new approach to avoid rank reversal cases in the TOPSIS method. *Computers & Industrial Engineering*.132:84-97.

[15] Balaji S, Revathi N. (2016) A new approach for solving set covering problem using jumping particle swarm optimization method. *Natural Computing*.15(3):503-17.

[16] Gani AN, Abbas S. (2017) A new method on solving intuitionistic fuzzy transportation problem. *Annals of Pure and Applied Mathematics*.15(2):2279-0888.

[17] Min W, Fan M, Guo X, Han Q. (2017) A new approach to track multiple vehicles with the combination of robust detection and two classifiers. *IEEE Transactions on Intelligent Transportation Systems.*19(1):174-86.

[18] Cömert SE, YAZGAN HR, Sertvuran I, ŞENGÜL H. (2017) A new approach for solution of vehicle routing problem with hard time window: an application in a supermarket chain. *Sādhanā.* 42(12):2067-80.

[19] Fitiwi DZ, De Cuadra F, Olmos L, Rivier M. (2015) A new approach of clustering operational states for power network expansion planning problems dealing with RES (renewable energy source) generation operational variability and uncertainty. *Energy*.90:1360-76.

[20] Bartczuk Ł, Przybył A, Cpałka K. (2016) A new approach to nonlinear modelling of dynamic systems based on fuzzy rules. *International Journal of Applied Mathematics and Computer Science*.26.

[21] Aien M, Rashidinejad M, Firuz-Abad MF. (2015) Probabilistic optimal power flow in correlated hybrid wind-PV power systems: A review and a new approach. *Renewable and Sustainable Energy Reviews*. 41:1437-46.

[22] Magala M, Sammons A. (2015) A new approach to port choice modelling. *Port Management: Springer*; p. 29-56.

[23] Shen M, Yan S, Zhang G. (2016) A new approach to event-triggered static output feedback control of networked control systems. *ISA transactions*.65:468-74.

[24] Barreira-González P, Gómez-Delgado M, Aguilera-Benavente F. (2015) From raster to vector cellular automata models: A new approach to simulate urban growth with the help of graph theory. *Computers, Environment and Urban Systems*.54:119-31.

[25] Kumar R, Edalatpanah S, Jha S, Singh R. (2019) A Pythagorean fuzzy approach to the transportation problem. *Complex & intelligent systems*.5(2):255-63.

[26] Mahajan S, Gupta S. (2021) On fully intuitionistic fuzzy multiobjective transportation problems using different membership functions. *Annals of operations research*.296(1):211-41.

THE LENGTH BIASED NEW QUASI LINDLEY DISTRIBUTION: STATISTICAL PROPERTIES AND APPLICATION

N. W. Andure (Yawale)[1] and R. B. Ade[2*]

•

[1] Department of Statistics, Government Vidarbha Institute of Science and Humanities, Amravati, Maharashtra, India, neetayawale@gmail.com

[2*]Department of Statistics, Government Vidarbha Institute of Science and Humanities, Amravati, Maharashtra, India, rajeshwarb.sc@gmail.com

Abstract

In this paper, a new distribution namely the length biased new quasi-Lindley distribution is proposed with the different weight function. The different mathematical and statistical properties of the proposed distribution are derived and discussed. The survival function, hazard rate function and mean residual life function for the length biased new quasi Lindley distribution is discussed. Also, concepts like stochastic ordering and entropy for proposed distribution are studied. The parameters of the proposed distribution are estimated by using the method of maximum likelihood estimation. The performance of the newly introduced distribution is studied using a real- life data set.

Keywords: Length Biased Distribution, New Quasi Lindley Distribution, Reliability Analysis, Stochastic ordering, Maximum Likelihood Estimation.

I. Introduction

The concept of weighted distributions was first introduced by [8] to model ascertainment bias, weighted distributions were later formalized in a unifying theory by [16]. Weighted distribution is used in a variety of research fields related to reliability, environment, engineering and biomedicine. The weighted distribution reduces to length-biased distribution when the weight function considers only the length of the units. The concept of length-biased sampling was first introduced by [5] and [24]. [14] studied size-biased sampling and related form-invariant weighted distributions. Refer [15] for a general statistical discussion of weighted distributions. [12] proposed a useful result by giving a relationship between the original random variable X and its length-biased form Y when X is either Inverse Gaussian or Gamma distribution. Several researchers have studied length biased versions of different distributions see, [10], [6], [7], [17], [13] and [19].

Initially Quasi Lindley distribution was proposed by [22]. Later New Quasi Lindley (NQL) distribution was studied by [21] for modelling various data sets with probability density function (p.d.f) as follows

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

The NQL distribution in (1) is a mixture of exponential and gamma distribution [exponential (θ) and gamma (2, θ)].
In the present work, length biased new Quasi Lindley distribution is proposed and discussed in next section.

II. Length Biased New Quasi Lindley Distribution

Suppose X be a non-negative random variable with pdf $f(x)$. Let $w(x)$ be the non-negative weight function, and then the pdf of the weighted random variable X_w is given by:

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))} \quad , \quad x > 0$$

where $w(x)$ be a non - negative weight function and

$$E\big(w(x)\big) = \int w(x)f(x)dx < \infty$$

When $w(x) = x^c$,the resulting distribution is termed as weighted distribution. When $w(x) = x$ the resultant is known as size or length biased distribution. In this paper, the length biased version of new quasi-Lindley distribution is proposed. The weight function used is as follows

$$w(x) = \frac{nx}{\theta} \tag{2}$$

According to [1], let X is a non-negative random variable with pdf $f(x)$. Let $w(x)$ be the non-negative weight function then the pdf $f_l(x)$ for a length biased distribution of X is given by:

$$f_l(x) = \frac{w_*(x)w(x)f(x)}{E(w_*(x)w(x))} \quad , \quad x > 0 \tag{3}$$

Assuming the $E\big(w(x)w_*(x)\big) = \int w(x)w_*(x)f(x)dx < \infty$
Provided that $w_*(x) = x$
Using equation (1) and (2) in (3), the pdf of length biased new quasi-Lindley (LBNQL) distribution is

$$f_l(x;\theta,\alpha) = \frac{\frac{nx}{\theta}\frac{x\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x}}{E\big(w(x)w_*(x)\big)}$$

where

$$E\big(w(x)w_*(x)\big) = \int_0^\infty \frac{nx}{\theta}\frac{x\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x}dx = \frac{2n(\theta^2+3\alpha)}{\theta^3(\theta^2+\alpha)}$$

$$f_l(x;\theta,\alpha) = \frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x} \quad ;x > 0, \theta > 0, \alpha > 0 \tag{4}$$

and the cumulative distribution function (CDF) of LBNQL distribution is obtained as

$$F(x) = 1 - Pr(X > x) = 1 - \int_x^\infty f_l(t;\theta,\alpha)dt$$

$$F(x) = 1 - \frac{\theta^4}{2(\theta^2+3\alpha)}\int_x^\infty t^2(\theta+\alpha t)e^{-\theta t}dt$$

$$F(x) = 1 - \frac{\theta^4}{2(\theta^2+3\alpha)}\left(\theta\int_x^\infty t^2\, e^{-\theta t}dt + \alpha\int_x^\infty t^3\, e^{-\theta t}dt\right)$$

after the simplification, the CDF of the LBNQL distribution is

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

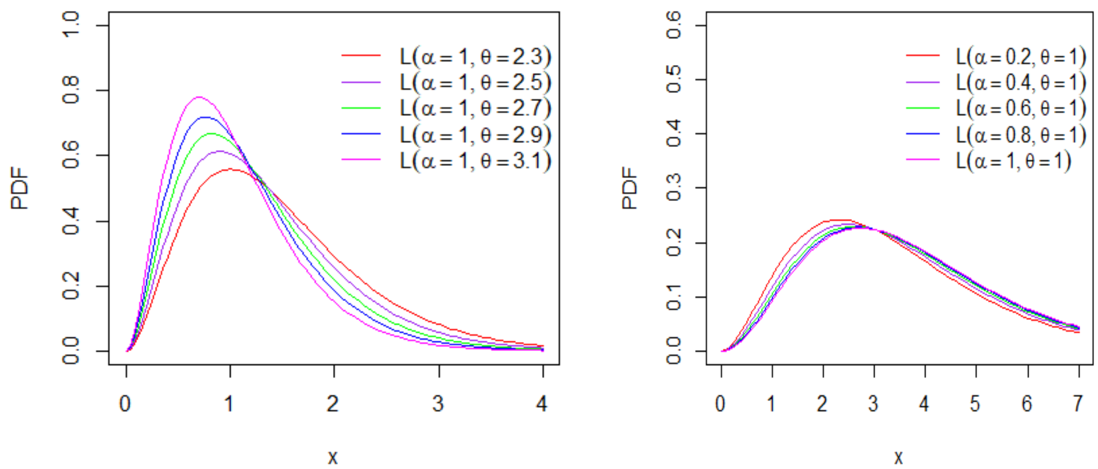


Figure 1: *PDF of LBNQL distribution for different values of θ and α.*

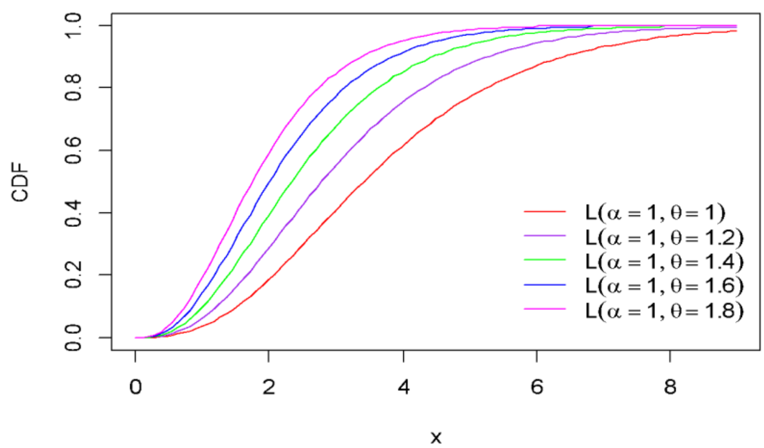


Figure 2: *CDF of LBNQL distribution for different values of θ and α.*

III. Reliability Analysis

In this section, the reliability function or survival function, hazard rate function and mean residual life for the LBNQL distribution is discussed.

I. Survival Function of LBNQL Distribution

The survival function or the reliability function of Length biased new quasi-Lindley distribution (LBNQLD) is defined as

$$S(x) = 1 - F(x) \tag{6}$$

Substituting from equation (5) in (6),

$$S(x) = \left(1 + \frac{\alpha\theta^3 x^3 + x(6\alpha\theta + 2\theta^3) + (3\alpha\theta^2 + \theta^4)x^2}{2(\theta^2 + 3\alpha)}\right) e^{-\theta x} \tag{7}$$

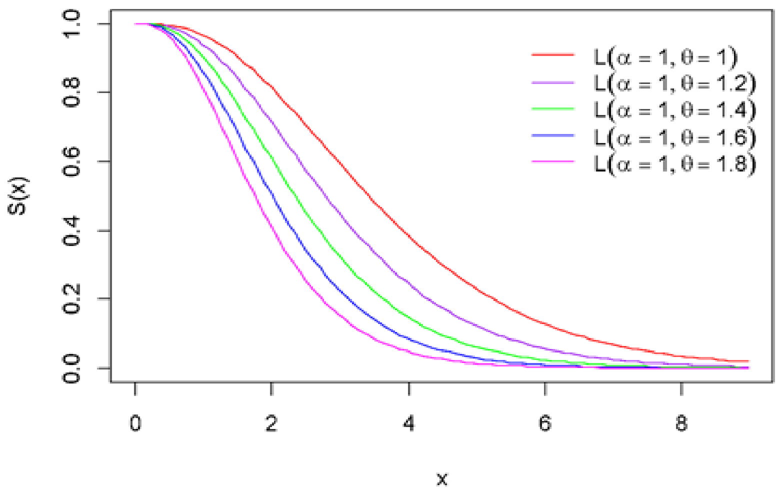


Figure 3: *survival function of LBNQL distribution for different values of θ and α.*

II. Hazard Rate Function of LBNQL Distribution

The basic tool for studying the aging and reliability characteristics of the system is the hazard rate (HR). The hazard function is also known as the hazard rate. Thus, the hazard rate function of the LBNQL distribution is given by

$$h(x) = \frac{f_l(x;\theta,\alpha)}{S(x)} \tag{8}$$

Substitute the value of (4) and (7) in (8),

$$h(x) = \frac{x^2\theta^4(\theta+\alpha x)}{2(\theta^2+3\alpha)+\alpha\theta^3x^3+(6\alpha\theta+2\theta^3)x+(3\alpha\theta^2+\theta^4)x^2}$$

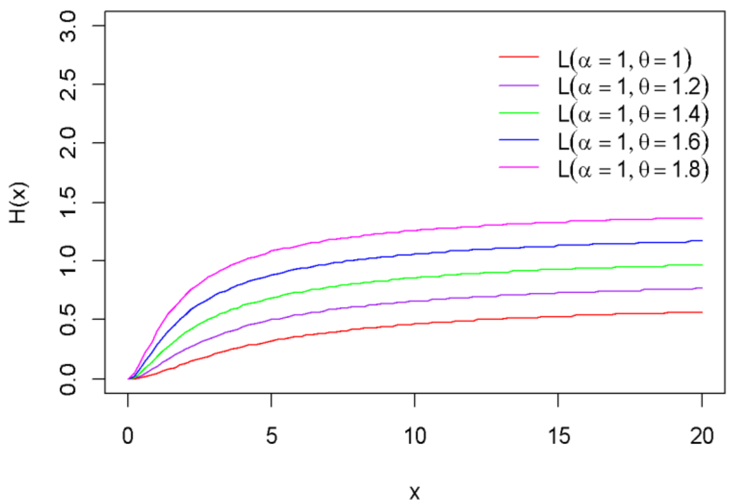


Figure.4: *Hazard rate function of LBNQL distribution for different values of θ and α.*

Figure (4) shows the behavior of hazard function. For different choices of α and θ it shows increasing failure rate.

III. Mean Residual Life Function of LBNQL Distribution

The mean residual life function is defined as

$$m(x) = E[X > x] = \frac{1}{1 - F(x)} \int_x^\infty [1 - F(t)]\, dt$$

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

$$\times \int_x^\infty \left(1 + \frac{\alpha\theta^3 t^3 + t(6\alpha\theta + 2\theta^3) + (3\alpha\theta^2 + \theta^4)t^2}{2(\theta^2 + 3\alpha)}\right) e^{-\theta t} dt$$

After simplifying, we get

$$m(x) = \frac{24\alpha + 6\theta^2 + \alpha\theta^3 x^3 + x(18\alpha\theta + 4\theta^3) + (6\alpha\theta^2 + \theta^4)x^2}{\theta(6\alpha + 2\theta^2 + \alpha\theta^3 x^3 + (3\alpha\theta^2 + \theta^4)x^2 + x(6\alpha\theta + 2\theta^3))}$$

It can be easily verified that $m(0) = \frac{3(\theta^2+4\alpha)}{\theta(\theta^2+3\alpha)} = \mu_1'$

IV. Moments and Associated Measures

Let X denotes the random variable of LBNQL distribution with parameters α and θ then the r^{th} order moment of LBNQL distribution can be defined as

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_l(x;\theta,\alpha)\, dx$$

$$= \int_0^\infty x^{r+2} \frac{\theta^4}{2(\theta^2+3\alpha)} (\theta + \alpha x) e^{-\theta x} dx$$

$$= \frac{\theta^4}{2(\theta^2+3\alpha)} \int_0^\infty x^{r+2} (\theta + \alpha x) e^{-\theta x} dx$$

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

$$\mu_r' = \frac{\theta^4}{2(\theta^2+3\alpha)} \left(\frac{\Gamma(r+3)}{\theta^{r+2}} + \frac{\alpha\Gamma(r+4)}{\theta^{r+4}}\right) \tag{9}$$

Putting $r = 1$ in (9), the mean of LBNQL distribution is given by

$$\mu_1' = E(X) = \frac{3(\theta^2 + 4\alpha)}{\theta(\theta^2 + 3\alpha)}$$

and putting $r = 2,3,4$ in (9), the second, third and fourth raw moments are

$$\mu_2' = E(X^2) = \frac{12(\theta^2+5\alpha)}{\theta^2(\theta^2+3\alpha)}, \quad \mu_3' = E(X^3) = \frac{60(\theta^2+6\alpha)}{\theta^3(\theta^2+3\alpha)}, \quad \mu_4' = E(X^4) = \frac{360(\theta^2+7\alpha)}{\theta^4(\theta^2+3\alpha)}$$

Therefore,

$$Variance = \sigma^2 = \frac{3(12\alpha^2+8\alpha\theta^2+\theta^4)}{\theta^2(\theta^2+3\alpha)^2}$$

$$Standard\ Deviation = \sigma = \frac{\sqrt{3(12\alpha^2 + 8\alpha\theta^2 + \theta^4)}}{\theta(\theta^2 + 3\alpha)}$$

$$Coefficient\ of\ Variation(c.v) = \frac{\sigma}{\mu} = \frac{\sqrt{3(12\alpha^2 + 8\alpha\theta^2 + \theta^4)}}{3(\theta^2 + 4\alpha)}$$

$$Coefficient\ of\ Dispersion(\gamma) = \frac{\sigma^2}{\mu} = \frac{(12\alpha^2 + 8\alpha\theta^2 + \theta^4)}{\theta(\theta^2 + 3\alpha)(\theta^2 + 4\alpha)}$$

I. Moment Generating Function and Characteristic Function

Let the random variable X follows the LBNQL distribution. By definition of moment generating function of X and using equation (4), we get

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx}\ f_l(x;\theta,\alpha)dx$$

$$= \int_0^\infty (1 + (tx) + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots) f_l(x;\theta,\alpha)dx$$

$$= \int_0^\infty \sum_{r=0}^\infty \frac{(tx)^r}{r!} f_l(x;\theta,\alpha)dx$$

$$= \sum_{r=0}^\infty \frac{(t)^r}{r!} \int_0^\infty x^r\ f_l(x;\theta,\alpha)dx$$

$$= \sum_{r=0}^\infty \frac{(t)^r}{r!} E(X^r) \tag{10}$$

Substituting from equation (9) in (10),

$$M_X(t) = \sum_{r=0}^\infty \frac{(t)^r}{r!} \left\{ \frac{\theta^4}{2(\theta^2 + 3\alpha)} \left(\frac{\Gamma(r+3)}{\theta^{r+2}} + \frac{\alpha\Gamma(r+4)}{\theta^{r+4}} \right) \right\}$$

Similarly, the characteristic function of LBNQL distribution is obtained as

$$\phi_X(t) = M_X(it)$$

$$\Rightarrow \phi_x(t) = M_X(it) = \sum_{r=0}^\infty \frac{(it)^r}{r!} \left\{ \frac{\theta^4}{2(\theta^2 + 3\alpha)} \left(\frac{\Gamma(r+3)}{\theta^{r+2}} + \frac{\alpha\Gamma(r+4)}{\theta^{r+4}} \right) \right\}$$

V. Entropy

The concept of entropy is important in a variety of topics such as probability and mathematics, physics, communication theory and economics. The entropy of random variable X is a measure of the variability of uncertainty.

I. R´enyi Entropy

R´enyi entropy [18] is important in nature and mathematics as an indicator of diversity. R´enyi entropy is also important in quantum data, where it can be used as a catch measure. R´enyi entropy is provided by

$$Re(\delta) = \frac{1}{1-\delta} log \left(\int_0^\infty f_l{}^\delta(x;\theta,\alpha)\, dx \right)$$

where $\delta > 0$ and $\delta \neq 1$

$$Re(\delta) = \frac{1}{1-\delta} log \left(\int_0^\infty \left(\frac{\theta^4}{2(\theta^2 + 3\alpha)} x^2(\theta + \alpha x) e^{-\theta x} \right)^\delta dx \right)$$

$$Re(\delta) = \frac{1}{1-\delta} log \left(\left(\frac{\theta^5}{2(\theta^2 + 3\alpha)} \right)^\delta \int_0^\infty x^{2\delta} \left(1 + \frac{\alpha}{\theta} x\right)^\delta e^{-\theta\delta x} dx \right)$$

putting $\left(1 + \frac{\alpha}{\theta} x\right)^\delta = \sum_{j=0}^\infty \binom{\delta}{j} \left(\frac{\alpha}{\theta} x\right)^j$

$$Re(\delta) = \frac{1}{1-\delta} log \left\{ \left(\frac{\theta^5}{2(\theta^2+3\alpha)} \right)^\delta \sum_{j=0}^\infty \binom{\delta}{j} \left(\frac{\alpha}{\theta}\right)^j \int_0^\infty (x)^{2\delta+j+1-1}\, e^{-\theta\delta x} dx \right\}$$

$$Re(\delta)=\frac{1}{1-\delta}log\left\{\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^{\delta}\sum_{j=0}^{\infty}\binom{\delta}{j}\left(\frac{\alpha}{\theta}\right)^{j}\frac{\Gamma(2\delta+j+1)}{(\theta\delta)^{2\delta+j+1}}\right\}$$

II. Tsallis Entropy

A generalization of Boltzmann-Gibbs (B.G) statistical properties initiated by Tsallis has received great attention. This generalization of BG statistic was proposed firstly by introducing the mathematical expression of Tsallis entropy by [23] for a continuous random variable and is defined as follows

$$S_\lambda=\frac{1}{1-\lambda}\left(1-\int_0^\infty f_l^{\lambda}(x;\theta,\alpha)dx\right)$$

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

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

putting $\left(1+\frac{\alpha}{\theta}x\right)^{\lambda}=\sum_{j=0}^{\infty}\binom{\lambda}{j}\left(\frac{\alpha}{\theta}x\right)^{j}$

$$S_\lambda=\frac{1}{1-\lambda}\left\{1-\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^{\lambda}\sum_{j=0}^{\infty}\binom{\lambda}{j}\left(\frac{\alpha}{\theta}\right)^{j}\int_0^\infty(x)^{2\lambda+j}\ e^{-\theta\lambda x}dx\right\}$$

$$S_\lambda=\frac{1}{1-\lambda}\left\{1-\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^{\delta}\sum_{j=0}^{\infty}\binom{\lambda}{j}\left(\frac{\alpha}{\theta}\right)^{j}\frac{\Gamma(2\lambda+j+1)}{(\theta\lambda)^{2\lambda+j+1}}\right\}$$

VI. Order Statistic of LBNQL Distribution

Order statistic have a central role in statistical theory. Suppose $X_{(1)},X_{(2)},\dots\dots\dots\dots,X_{(n)}$ be the continuous ascending order statistic. The probability density function of the j^{th} order statistic $X_{(j)}$ for $1\le j\le n$ is

$$f_{X_{(j)}}(x)=\frac{n!}{(j-1)!(n-j)!}[F(x)]^{j-1}\,[1-F(x)]^{n-j}f_1(x) \tag{11}$$

Substitute the value of pdf and cdf of LBNQL distribution in (11), we get

$$f_{X_{(j)}}(x)=\frac{n!}{(j-1)!(n-j)!}\left[1-\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{j-1}$$
$$\times\left[\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{n-j}\times\frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x} \tag{12}$$

Put $j=1$ in equation (12), the probability density function of first order statistics of LBNQL distribution.

$$f_{X_{(1)}}(x)=n\left[\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{n-1}\times\frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x}$$

Put $j=n$ in equation (12), the probability density function of n^{th} order statistics of LBNQLD.

$$f_{X_{(n)}}(x)=n\left[1-\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{n-1}$$

$$\times \frac{\theta^4}{2(\theta^2+3\alpha)} x^2(\theta+\alpha x)e^{-\theta x}$$

VII. Stochastic ordering

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

1) Stochastic order $(X \leq_{st} Y)$ if $F_X(x) \geq F_Y(x)$ for all x
2) Hazard rate order $(X \leq_{hr} Y)$ if $h_X(x) \geq h_Y(x)$ for all x
3) Mean residual life order $(X \leq_{mrl} Y)$ if if $m_X(x) \leq m_Y(x)$ for all x
4) Likelihood ratio order $(X \leq_{lr} Y)$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x

The following important interrelations due to [20] are well-known for establishing stochastic ordering of distributions

$$\begin{array}{c} X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y \\ \Downarrow \\ X \leq_{st} Y \end{array}$$

The LBNQL distribution is ordered with respect to the strongest 'likelihood ratio ordering' as shown in the following theorem:

Theorem 7.1: Let $X \sim LBNQL(\theta_1, \alpha_1)$ and $Y \sim LBNQL(\theta_2, \alpha_2)$.If $\theta_1 = \theta_2$ and $\alpha_1 \leq \alpha_2$ 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: From the pdf of LBNQL distribution (4), we have

$$\frac{f_X(x)}{f_Y(x)} = \frac{{\theta_1}^4({\theta_2}^2+3\alpha_2)}{{\theta_2}^4({\theta_1}^2+3\alpha_1)}\left(\frac{\theta_1+\alpha_1 x}{\theta_2+\alpha_2 x}\right)e^{-(\theta_1-\theta_2)x} \qquad ;x>0$$

Now

$$\log\frac{f_X(x)}{f_Y(x)} = \log\left[\frac{{\theta_1}^4({\theta_2}^2+3\alpha_2)}{{\theta_2}^4({\theta_1}^2+3\alpha_1)}\right] + \log(\theta_1+\alpha_1 x) - \log(\theta_2+\alpha_2 x) - (\theta_1-\theta_2)x$$

This gives

$$\frac{d}{dx}\left(\log\frac{f_X(x)}{f_Y(x)}\right) = \frac{\alpha_1}{(\theta_1+\alpha_1 x)} - \frac{\alpha_2}{(\theta_2+\alpha_2 x)} - (\theta_1-\theta_2)$$
$$= \frac{\alpha_1\theta_2 - \alpha_2\theta_1}{(\theta_1+\alpha_1 x)(\theta_2+\alpha_2 x)} - (\theta_1-\theta_2)$$

Case I: If $\theta_1 = \theta_2$ and $\alpha_1 \leq \alpha_2$, then $\frac{d}{dx}\left(\log\frac{f_X(x)}{f_Y(x)}\right) < 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$**.**

Case II: If $\alpha_1 = \alpha_2$ and $\theta_1 \geq \theta_2$, then $\frac{d}{dx}\left(\log\frac{f_X(x)}{f_Y(x)}\right) < 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Thus, LBNQL distribution follows the strongest likelihood ratio ordering**.**

VIII. Bonferroni and Lorenz curve

The most important inequality curves are called Bonferroni and Lorenz curve, which have some application in applied science such as economics, reliability, demography and medicine. Bonferroni and Lorenz curves are proposed by [3]. The Bonferroni and Lorentz curves for the LBNQL distribution is obtained as

$$B(p) = \frac{1}{p\mu}\int_0^q x\, f_l(x,\theta,\alpha)dx$$

and

$$L(p) = pB(p) = \frac{1}{\mu}\int_0^q x\, f_l(x;\theta,\alpha)dx$$

where $E(x) = \mu = \frac{3(\theta^2+4\alpha)}{\theta(\theta^2+3\alpha)}$ and $q = F^{-1}(p)$

$$\therefore\ B(p) = \frac{\theta(\theta^2+3\alpha)}{p3(\theta^2+4\alpha)}\int_0^q x^3 \frac{\theta^4}{2(\theta^2+3\alpha)}(\theta+\alpha x)e^{-\theta x}dx$$

$$B(p) = \frac{\theta^5}{p6(\theta^2+4\alpha)}\int_0^q x^3\,(\theta+\alpha x)e^{-\theta x}dx$$

after simplification,

$$B(p) = \frac{\theta^2\gamma(4,\theta q)+\alpha\gamma(5,\theta q)}{p6(\theta^2+4\alpha)}$$

and

$$L(p) = pB(p) = \frac{1}{6(\theta^2+4\alpha)}(\theta^2\gamma(4,\theta q)+\alpha\gamma(5,\theta q))$$

IX. Maximum Likelihood Estimation

The method of maximum likelihood is the most frequently used method of parameter estimation given in [4]. The maximum likelihood method of estimation has been adopted to estimate the unknown parameter α and θ of the LBNQL distribution. Consider the random sample of size n from the LBNQL distribution, the likelihood function is given by

$$L(x;\alpha,\theta) = \left(\frac{\theta^4}{2(\theta^2+3\alpha)}\right)^n \prod_{i=1}^n x_i^{\,2}\,(\theta+\alpha x_i)e^{-\theta\sum_{i=1}^n x_i}$$

The log likelihood function is given by

$$\log l = 4n\log\theta - n\log\big(2(\theta^2+3\alpha)\big) + 2\sum_{i=1}^n \log x_i + \sum_{i=1}^n \log(\theta+\alpha x_i) - \theta\sum_{i=1}^n x_i \tag{13}$$

Now maximize the above log-likelihood function given in equation (13) to get maximum likelihood estimate of unknown parameters of length biased new quasi-Lindley distribution. For this purpose, take the first derivative of the above log-likelihood equation with respect to parameters α and θ and equate to zero respectively.

$$\Rightarrow \frac{4n}{\theta} - \frac{2n\theta}{(\theta^2+3\alpha)} + \sum_{i=1}^n \frac{1}{\theta+\alpha x_i} - \sum_{i=1}^n x_i = 0 \tag{14}$$

$$\Rightarrow \frac{-3n}{(\theta^2+3\alpha)} + \sum_{i=1}^n \frac{x_i}{\theta+\alpha x_i} = 0 \tag{15}$$

Equations (14) and (15) are nonlinear equation. The exact solution of above equation is not possible numerically. Above nonlinear equations are solved with the help of R Software.

Using the large sample property of MLE, $\hat{\lambda}$ can be treated as being approximately normal with mean λ and variance covariance matrix equal to the inverse of the expected information matrix, i.e.,

$$\sqrt{n}(\hat{\lambda} - \lambda) \to N(0, I^{-1}(\lambda))$$

$I(\lambda)$ is the information matrix then its inverse matrix is $I^{-1}(\lambda)$. The $I(\hat{\lambda})$ variance-covariance matrix is essentially equal to the inverse of the expected information matrix $I^{-1}(\hat{\lambda})$, the observed information matrix is given by

$$I(\lambda) = -\frac{1}{n}\begin{bmatrix} E\left(\frac{\partial^2 \log l}{\partial \theta^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \theta \partial \alpha}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \alpha \partial \theta}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha^2}\right) \end{bmatrix}$$

where $I(\lambda)$ is Fisher's Information Matrix.

$$\frac{\partial^2 \log l}{\partial \theta^2} = \frac{-4n}{\theta^2} + \frac{2n(\theta^2 - 3\alpha)}{(\theta^2 + 3\alpha)^2} - \sum_{i=1}^{n} \frac{1}{(\theta + \alpha x_i)^2}$$

$$\frac{\partial^2 \log l}{\partial \alpha^2} = \frac{9n}{(\theta^2 + 3\alpha)^2} - \sum_{i=1}^{n} \left(\frac{x_i^2}{(\theta + \alpha x_i)^2}\right)$$

$$\frac{\partial^2 \log l}{\partial \theta \partial \alpha} = \frac{\partial^2 \log l}{\partial \alpha \partial \theta} = \frac{6n\theta}{(\theta^2 + 3\alpha)^2} - \sum_{i=1}^{n} \left(\frac{x_i}{(\theta + \alpha x_i)^2}\right)$$

Since λ being unknown, $I^{-1}(\lambda)$ is estimated by using $I^{-1}(\hat{\lambda})$ and we obtain the asymptotic confidence intervals for α and θ. Hence the approximate $100(1 - \psi)\%$ confidence interval for α and θ are respectively given by

$$\hat{\alpha} \pm z_{\frac{\psi}{2}} \sqrt{I_{\alpha\alpha}{}^{-1}(\hat{\lambda})}, \qquad \hat{\theta} \pm z_{\frac{\psi}{2}} \sqrt{I_{\theta\theta}{}^{-1}(\hat{\lambda})}$$

Where $z_{\frac{\psi}{2}}$ is the ψ^{th} percentile of the standard distribution.

X. Application

In this section, one real life data set is analyzed for the purpose of illustration to show the usefulness and flexibility of the LBNQL distribution. The LBNQL model is compared with other distributions, such as, New Quasi Lindley (NQL) distribution, [21], length biased weighted New Quasi Lindley (LBWNQL) distribution [9]. The ML Estimates of the unknown parameters are determined for the LBNQL distribution and two other models along with goodness of fit test.

Data set I: Following data depicts the fatigue life of some aluminum's coupons cut in specific manner (see,[2]). The dataset (after subtracting 65) is given below:

5, 25, 31, 32, 34, 35, 38, 39, 39, 40, 42, 43, 43, 43, 44, 44, 47, 47, 48, 49, 49, 49 ,51, 54, 55, 55 ,56 ,56, 56, 58, 59, 59, 59, 59, 59, 63, 63, 64 ,64, 55, 65, 65, 65, 66, 66, 66, 66, 67, 67, 67, 68, 69, 69, 69 ,69, 71, 71, 72, 73, 73, 73, 74, 74, 76, 76, 77, 77, 77, 77, 77 ,77, 79, 79, 80, 81, 83, 83, 84, 86, 86, 87, 90, 91, 92, 92, 92 ,92, 93, 93, 94, 97, 98, 98, 99, 101, 101, 103, 105, 109, 139, 147

The data set is modeled by LBNQL distribution and compared with the New Quasi Lindley, length biased weighted New Quasi Lindley distribution. Table 1 describes estimated unknown parameters, -log likelihood (-LL), the values of the AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and K-S Statistics calculated for above data using LBNQL, LBWNQL, and NQL distributions.

Table 1. *Estimate and goodness of fit measures under considered distribution based on data set.*

Model	Estimated Parameter		-2LL	AIC	BIC	K-S	P-Value
	$\hat{\alpha}$	$\hat{\theta}$					
LBNQLD	**74.01668**	**0.05801**	**945.7779**	**949.7779**	**955.0082**	**0.14788**	**0.2403**
LBWNQLD	425.13398	0.04350	962.6170	966.6170	971.8473	0.18094	0.1436
NQL	168.18633	0.02900	992.7213	996.7213	1001.9516	0.23557	0.05719

From table 1 it can be seen that the value of the statistics -2LL, AIC, BIC values of LBNQL distribution are comparatively smaller than the other distributions on a data set. Therefore, the result shows that LBNQL distribution provides a significantly better fit than the other models.

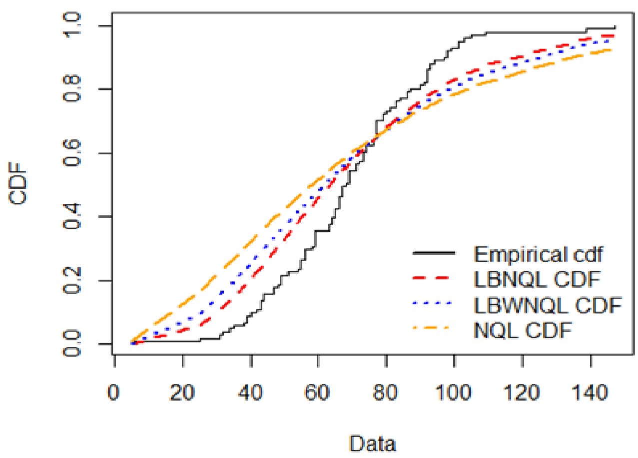


Figure 5: *Empirical CDF and fitted CDF plot of data set.*

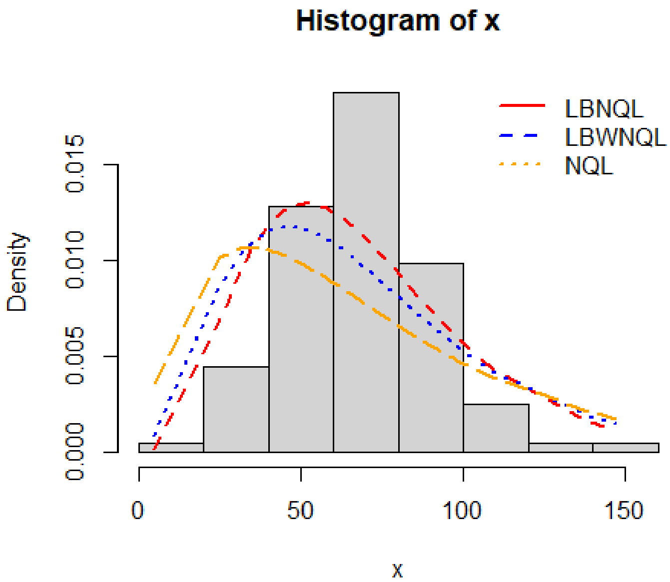


Figure 6: *fitted PDF plot of data set.*

XI. Conclusion

In the present study, a Length biased NQL distribution is proposed. Some statistical properties along with Reneyi entropy, and Tsallis entropy, Bonferroni and Lorenz curves have been discussed. Various reliability properties such as hazard rate function, mean residual life function, stochastic orderings have been obtained. It is proved that LBNQL distribution follows the strongest likelihood ratio ordering. For different choices of the parameters α and θ increasing failure rate is observed. The parameters of the proposed distribution are obtained by using the maximum likelihood estimation technique. Finally, the new proposed distribution is tested by applying to a real-life data set and compared with new quasi Lindley distribution and length biased weighted new quasi Lindley distribution. It is observed from the table 1 that LBNQL distribution gives better fit over both distributions on a data set.

References

[1] Al-kadim, K. A. and Hussain, N. A. (2014). New Proposed Length-Biased Weighted Exponential and Rayleigh Distribution with Application. Mathematical Theory and Modelling.4(7),137-152

[2] Birnbaum, Z. W. and Saunders, S. C. (1969). Estimation for a Family of Life Distributions with Applications to Fatigue. Journal of Applied Probability, 6(2), 328–347.

[3] Bonferroni, CE. Elementi di statisca gnerarale, Seeber, Firenze (1930).

[4] Casela, G. and Berger, R. L. Statistical inference. Pacific Grove, Calif: Brooks/Cole Pub. (1990).

[5] Cox, D. R. 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, (1969).

[6] Das, K. K. and Roy, T. D. (2011). Applicability of length biased generalized Rayleigh distribution. Advances in Applied Science Research, 2, 320–327

[7] Das, K. K. and Roy, T. D. (2011a). On some length-biased weighted Weibull distribution. Advances in Applied Science Research, 2, 465–475.

[8] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, Ann. Eugenics, 6, 13-25.

[9] Ganaie, R. A. and Rajgopalan, V. (2021). Length biased Weighted New Quasi Lindley Distribution: Statistical Properties and Applications. Pakistan Journal of Statistics and Operation Research. 17(1), 123-136.

[10] Gove, J. (2003). Moment and maximum likelihood estimators for Weibull distributions under length- biased and area-biased sampling. Environmental and Ecological Statistics. 10. 455-467.

[11] Jones, O., Maillardet, R. and Robinson, A. Introduction to scientific Programming and Simulation Using R, New York: Taylor and Francis Group, (2009).

[12] Khattree, R. (1989). Characterization of inverse-Gaussian and gamma distributions through their length-biased distributions. IEEE Transactions on Reliability, 38, 610-611.

[13] Nanuwong, N. and Bodhisuwan, W. (2014). Length Biased Beta-Pareto Distribution and Its Structural Properties with Application. Journal of Mathematics and Statistics, 10(1), 49-57.

[14] Patil, G. P. and Ord, J. K. (1976). On size biased sampling and related form invariant weighted distributions. Sankhya, Series B 38, 48-61

[15] Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size biased sampling with applications to wildlife populations and human families. Biometrics 34, 179-189.

[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] 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. Modern Applied Statistical Methods, 11, 225–260.

[18] Renyi, A. (1961). On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Contributions to the Theory of Statistics, Berkeley, California: University of California Press, (1): 547–561

[19] Seenoi, P., Supapa, K. T. and Bodhisuwan, W. (2014). The length biased exponentiated inverted Weibull Distribution. International Journal of Pure and Applied Mathematics, 92, 191–206.

[20] Shaked, M. and Shanthikumar, J. G. Stochastic Orders and Their Applications, Academic Press, New York. (1994).

[21] Shanker, R. and Ghebretsadik, A. H. (2013). A New Quasi Lindley distribution. International Journal of Statistics and Systems, 8(2), 143-156.

[22] Shanker, R. and Mishra, A. (2013). A quasi Lindley distribution. African Journal of Mathematics and Computer Science Research. Vol. 6(4), 64-71.

[23] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibb's statistics. *Journal of statistical physics*, 52(1-2), 479-487.

[24] Zelen, M. (1974). Problems in Cell Kinetics and Early Detection of Disease. Reliability and Biometry, SIAM, Philadelphia, 701-726.

POWER WEIGHTED AKASH DISTRIBUTION WITH PROPERTIES AND APPLICATIONS

Rama Shanker[1] and Kamlesh Kumar Shukla[2*]

•

[1]Department of Statistics, Assam University, Silchar, India
Shankerrama2009@gmail.com
[2] Department of Mathematics, School of Sciences, Noida International University, Gautam Budh Nagar, India, kkshukla22@gmail.com
*Corresponding Author

Abstract

In In this paper power weighted Akash distribution (PWAD) which includes weighted Akash distribution (WAD), power Akash distribution (PAD) and Akash distribution as particular cases has been proposed and investigated. Its moments, hazard rate function and mean residual life function have been discussed. Method of maximum likelihood estimation has been discussed for estimating the parameters of the distribution. Applications of the proposed distribution to two real lifetime datasets have been presented and compared with other one parameter, two-parameter and three-parameter well-known lifetime distributions.

Keywords: Akash distribution, Weighted Akash distribution, Power Akash distribution, Hazard rate function, stochastic ordering, Maximum Likelihood estimation, Applications.

I. Introduction

Shanker and Shukla [1] proposed a two-parameter weighted Akash distribution (WAD) having parameters θ and α and defined by its probability density function (pdf) and cumulative distribution function (cdf)

$$f_1(y;\theta,\alpha)=\frac{\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)}\frac{y^{\alpha-1}}{\Gamma(\alpha)}(1+y^2)e^{-\theta y};y>0,\theta>0,\alpha>0 \tag{1.1}$$

$$F_1(y;\theta,\alpha)=1-\frac{[\theta^2+\alpha(\alpha+1)]\Gamma(\alpha,\theta y)+(\theta y)^\alpha(\theta y+\alpha+1)e^{-\theta y}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)} \tag{1.2}$$

where $\Gamma(\alpha)$ and $\Gamma(\alpha,z)$ are the complete gamma function and the upper incomplete gamma function defined as

$$\Gamma(\alpha)=\int_0^\infty e^{-t}t^{\alpha-1}dt\,;\alpha>0 \tag{1.3}$$

$$\Gamma(\alpha,z)=\int_z^\infty e^{-y}t^{\alpha-1}dt;\alpha>0,z\geq 0 \tag{1.4}$$

Its structural properties including moments, hazard rate function, mean residual life function, estimation of parameters and applications for modeling survival time data has been discussed by Shanker and Shukla [1]. Shanker and Shukla [2] discussed various moments based properties including coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion of weighted Akash distribution and its applications to model lifetime data from biomedical sciences and engineering.

Shanker and Shukla [2] proposed a power Akash distribution (PAD) having parameters θ and α and defined by its pdf and cdf

$$f_2(y;\theta,\beta)=\frac{\beta\theta^3}{(\theta^2+2)}\left(1+y^{2\beta}\right)y^{\beta-1}e^{-\theta y^\beta};y>0,\theta>0,\beta>0 \tag{1.5}$$

$$F_2(y;\theta,\beta)=1-\left[1+\frac{\theta y^\beta\left(\theta y^\beta+2\right)}{\theta^2+2}\right]e^{-\theta y^\beta};y>0,\theta>0,\beta>0 \tag{1.6}$$

Note that the PAD is a convex combination of Weibull (α,θ) and a generalized gamma $(2,\alpha,\theta)$distribution with mixing proportion$\frac{\theta^2}{\theta^2+2}$. Shanker and Shukla [1] has discussed the properties of PAD including the shapes of the density, hazard rate functions, moments, skewness and kurtosis measures, estimation of parameters using maximum likelihood estimation and application to model a real lifetime data from engineering. Recall that WAD and PAD reduces to Akash distribution at $\alpha=1$, and $\beta=1$ respectively. The Akash distribution proposed by Shanker [3] is defined by its pdf and cdf

$$f_3(y;\theta)=\frac{\theta^3}{\theta^2+2}(1+y^2)e^{-\theta y};y>0,\theta>0 \tag{1.7}$$

$$F_3(y;\theta)=1-\left[1+\frac{\theta y(\theta y+2)}{\theta^2+2}\right]e^{-\theta y};y>0,\theta>0 \tag{1.8}$$

Shanker [3] has discussed its various statistical and mathematical properties including shapes of the density. Moments and moments based measures, hazard rate function, mean residual life function, stochastic ordering, mean deviations, order statistics, Bonferroni and Lorenz curves, Renyi entropy measure, stress-strength reliability, estimation of parameter using both the method of moment and the maximum likelihood estimation and application to model lifetime data from engineering and biomedical sciences.

In the present paper, a three - parameter power weighted Akash distribution which includes Akash distribution, WAD, and PAD as particular cases, has been proposed and discussed. Its raw moments have been given. The survival function and the hazard rate function of the distribution have been derived and their shapes have been discussed for varying values of the parameters. The estimation of its parameters has been discussed using maximum likelihood method. Finally, the goodness of fit and the applications of the distribution have been explained through two real lifetime datasets and the fit has been compared with other one parameter, two-parameter and three-parameter lifetime distributions.

II. Power weighted Akash distribution

Assuming the power transformation $X = Y^{\frac{1}{\beta}}$ in (1.1), the pdf of the random variable Xcan be obtained as

$$f_4(x;\theta,\alpha,\beta) = \frac{\beta\theta^{\alpha+2}}{\theta^2+\alpha^2+\alpha}\frac{x^{\beta\alpha-1}}{\Gamma(\alpha)}\left(1+x^{2\beta}\right)e^{-\theta x^{\beta}};x > 0,\theta > 0,\alpha > 0,\beta > 0 \tag{2.1}$$

We would call the distribution in (2.1) as the power weighted Akash distribution (PWAD). It can be easily verified that the WAD(θ,α) in (1.1), PAD(θ,β) in (1.5) and Akash(θ)in (1.7) are the special cases of PWAD for $(\beta = 1)$,$(\alpha = 1)$and $(\alpha = \beta = 1)$, respectively.

It can be easily verified that PWAD is a convex combination of generalized gamma distribution (GGD) having parameters (θ,α,β)proposed by Stacy (1962) and GGD having parameters $(\theta,\alpha+2,\beta)$.

We have

$$f_4(x;\theta,\alpha,\beta) = pg_1(x;\theta,\alpha,\beta) + (1-p)g_2(x;\theta,\alpha+2,\beta),$$

where

$$p = \frac{\theta^2}{\theta^2+\alpha^2+\alpha}$$
$$g_1(x;\theta,\alpha,\beta) = \frac{\beta\theta^{\alpha}}{\Gamma(\alpha)}x^{\beta\alpha-1}e^{-\theta x^{\beta}};x > 0,\theta > 0,\alpha > 0,\beta > 0$$
$$g_2(x;\theta,\alpha,\beta) = \frac{\beta\theta^{\alpha+2}}{\Gamma(\alpha+2)}x^{\beta(\alpha+2)-1}e^{-\theta x^{\beta}};x > 0,\theta > 0,\alpha > 0,\beta > 0.$$

Graphs of density function of PWAD for varying values of parameters θ,αandβ have been drawn and presented in figure 1. It is clear that the natures of PWAD are decreasing, positively skewed, negatively skewed, platykurtic, mesokurtic and leptokurtic for varying values of parameters and hence it can be applied to model lifetime datasets of various natures. It is observed that pdf is increasing for increased value of θ and its pdf is increasing vastly as increased value of θ,αandβrespectively. However, role of α on the shape of the graph more as compared to other parameters.

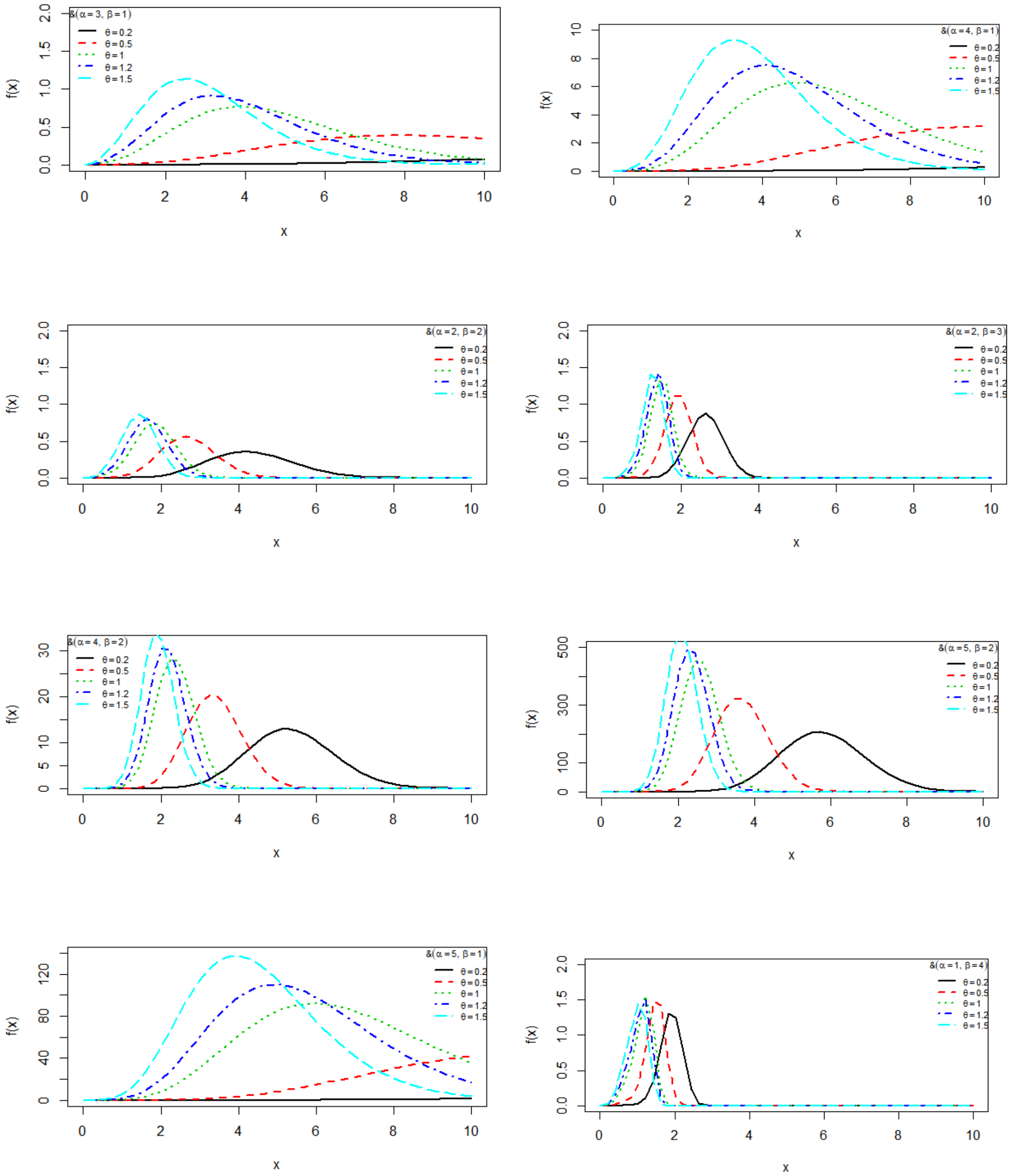


Figure. 2: Graphs of the probability density function of PWAD for varying values of parametersθ, αandβ

III. Reliability Measures

Survival function and Cumulative distribution Function

The survival function $S(x;\theta,\alpha,\beta)$of PWAD can be obtained as

$$S(x;\theta,\alpha,\beta)=P(X>x)=\int_x^\infty f_4(t;\theta,\alpha,\beta)\,dt$$
$$=\frac{\beta\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\int_x^\infty t^{\beta\alpha-1}\left(1+t^{2\beta}\right)e^{-\theta t^\beta}dt$$
$$=\frac{\beta\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\int_x^\infty e^{-\theta t^\beta}t^{\beta\alpha-1}\,dt+\int_x^\infty e^{-\theta t^\beta}t^{\beta\alpha+2\beta-1}\,dt\right]$$

Assuming $u=t^\beta$, which gives $t=(u)^{\frac{1}{\beta}}$and $dt=\frac{1}{\beta}(u)^{\frac{1-\beta}{\beta}}du$, we get

$$S(x;\theta,\alpha,\beta)=\frac{\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\int_{x^\beta}^\infty e^{-\theta u}u^{\alpha-1}\,du+\int_{x^\beta}^\infty e^{-\theta u}u^{\alpha+1}\,du\right]$$
$$=\frac{\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\frac{\Gamma(\alpha,\theta x^\beta)}{\theta^\alpha}+\frac{e^{-\theta x^\beta}(\theta x^\beta+\alpha+1)\left(\theta x^\beta\right)^\alpha+\alpha(\alpha+1)\Gamma(\alpha,\theta x^\beta)}{\theta^{\alpha+2}}\right]$$
$$=\frac{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+\left(\theta x^\beta\right)^\alpha(\theta x^\beta+\alpha+1)e^{-\theta x^\beta}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)},$$

where $\Gamma\left(\alpha,\theta x^\beta\right)$ is the upper incomplete gamma function defined as

$$\Gamma\left(\alpha,\theta x^\beta\right)=\int_{\theta x^\beta}^\infty y^{\alpha-1}e^{-y}dy;\alpha>0,\theta x^\beta>0.$$

It can be easily verified that at $(\beta=1)$,$(\alpha=1)$and $(\alpha=\beta=1)$ the survival function of PWAD reduce to the survival function of WAD, PAD and Akash distribution.

Thus the cdf of PWAD can be given by

$$F_4(x;\theta,\alpha,\beta)=1-S(x;\theta,\alpha,\beta)=1-\frac{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+\left(\theta x^\beta\right)^\alpha(\theta x^\beta+\alpha+1)e^{-\theta x^\beta}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}$$

The natures of the cdf of PWAD for varying values of parameters θ,αandβ are shown in figure 2. From the figure 2, It is observed that distribution function is slightly increasing as increased value of θ.

Hazard Rate Function

The hazard rate function, $h(x;\theta,\alpha,\beta)$, of PWAD can be given by

$$h(x;\theta,\alpha,\beta)=\frac{f_4(x;\theta,\alpha,\beta)}{S(x;\theta,\alpha,\beta)}=\frac{\beta\theta^{\alpha+2}x^{\beta\alpha-1}(1+x^{2\beta})e^{-\theta x^\beta}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+\left(\theta x^\beta\right)^\alpha(\theta x^\beta+\alpha+1)e^{-\theta x^\beta}}.$$

Graphs of $h(x;\theta,\alpha,\beta)$for varying values of parameters θ,αandβ are shown in figure 3. The graphs of $h(x;\theta,\alpha,\beta)$shows that it takes different shapes for varying values of parameters θ,αandβ and it is observed that hazard rate is increasing as increased value of θ,αandβrespectively.

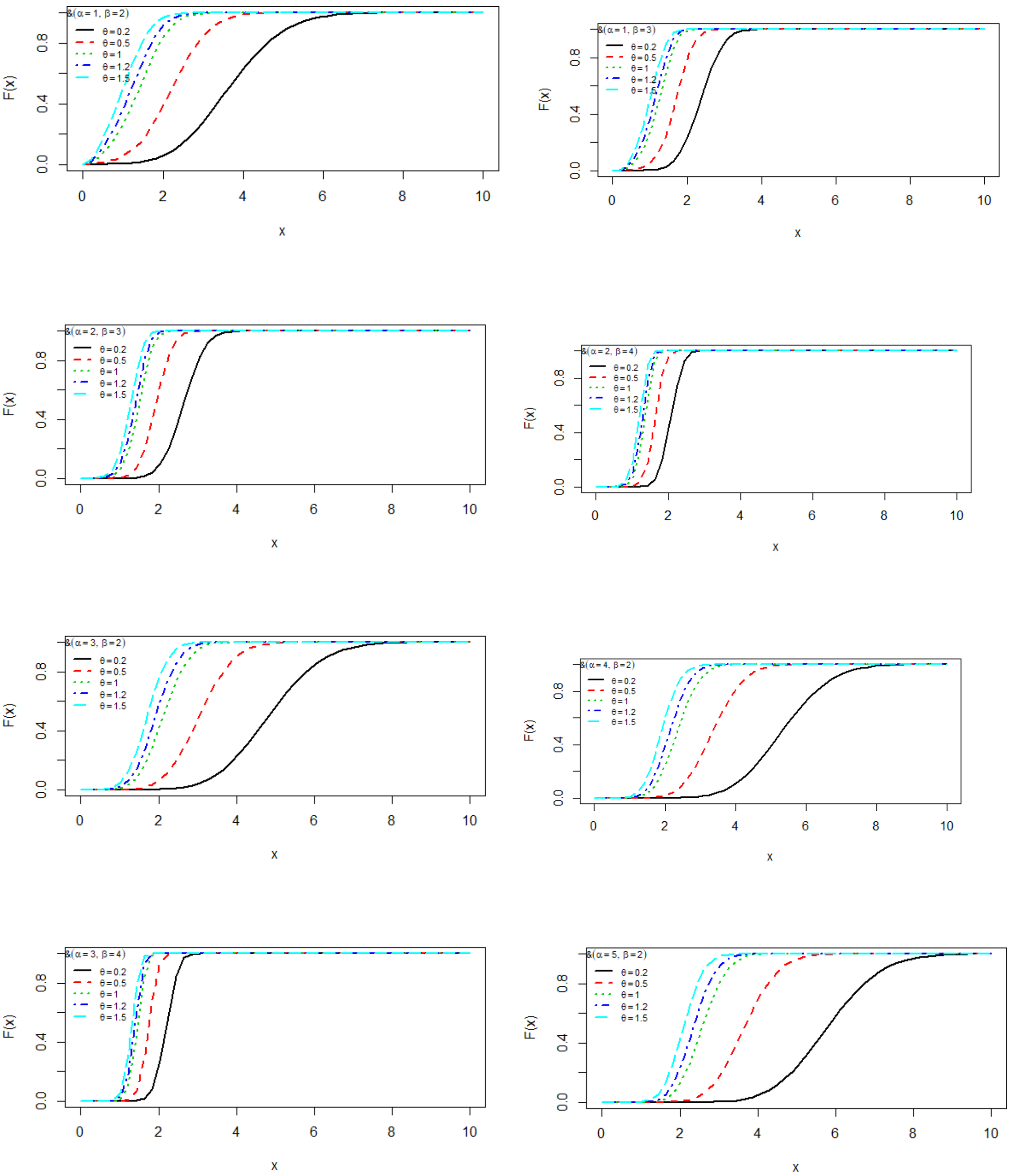


Figure 2: Graphs of the cdf of PWAD for varying values of parameters θ, α and β

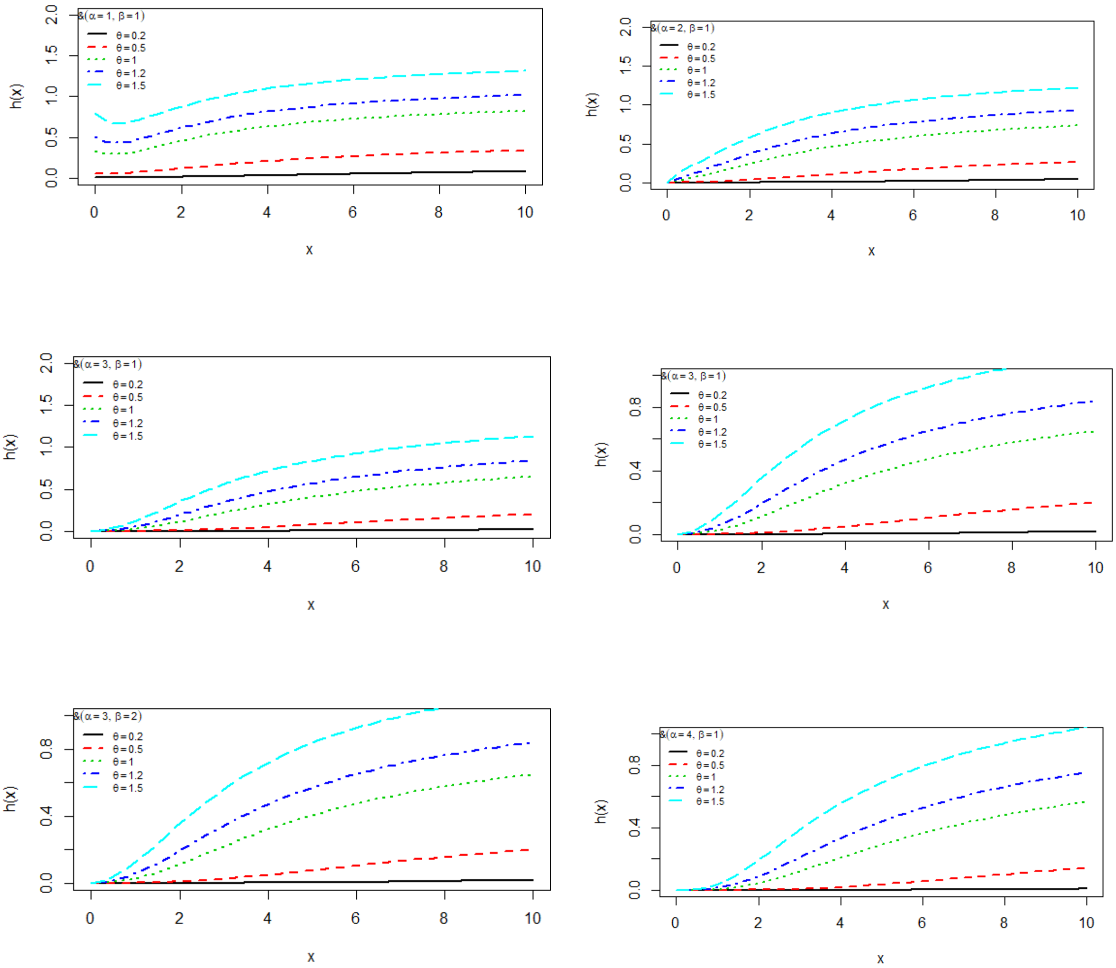


Figure 3: *Graphs of hazard rate function for varying values of parameters* θ, α and β.

Mean Residual Life Function

The mean residual life function, $m(x) = m(x;\theta,\alpha,\beta)$, of PWAD can be obtained as

$$m(x) = m(x;\theta,\alpha,\beta) = \frac{1}{S(x;\theta,\alpha,\beta)}\int_x^\infty t f_4(t;\theta,\alpha,\beta)\,dt - x$$
$$= \frac{\beta\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+\left(\theta x^\beta\right)^\alpha\left(\theta x^\beta+\alpha+1\right)e^{-\theta x^\beta}}\int_x^\infty t^{\beta\alpha}\left(1+t^{2\beta}\right)e^{-\theta t^\beta}\,dt - x$$
$$= \frac{\beta\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+(\theta x^\beta)^\alpha(\theta x^\beta+\alpha+1)e^{-\theta x^\beta}}\left[\int_x^\infty e^{-\theta t^\beta}t^{\beta\alpha}\,dt + \int_x^\infty e^{-\theta t^\beta}t^{\beta\alpha+2\beta}\,dt\right] - x$$

Assuming $u = t^\beta$, which gives $t = (u)^{\frac{1}{\beta}}$and $dt = \frac{1}{\beta}(u)^{\frac{1-\beta}{\beta}}du$, we get

$$m(x) = m(x;\theta,\alpha,\beta) = \frac{\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+(\theta x^\beta)^\alpha(\theta x^\beta+\alpha+1)e^{-\theta x^\beta}}$$
$$\times\left[\int_{x^\beta}^\infty e^{-\theta u}u^{\alpha+\frac{1}{\beta}-1}\,du + \int_{x^\beta}^\infty e^{-\theta u}u^{\alpha+2+\frac{1}{\beta}-1}\,du\right] - x$$
$$= \frac{\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+\left(\theta x^\beta\right)^\alpha\left(\theta x^\beta+\alpha+1\right)e^{-\theta x^\beta}}$$
$$\times\left[\frac{\Gamma\left(\alpha+\frac{1}{\beta},\theta x^\beta\right)}{\theta^{\alpha+\frac{1}{\beta}}} + \frac{\Gamma\left(\alpha+2+\frac{1}{\beta},\theta x^\beta\right)}{\theta^{\alpha+2+\frac{1}{\beta}}}\right] - x$$
$$= \frac{\theta^2\Gamma\left(\alpha+\frac{1}{\beta},\theta x^\beta\right)+\Gamma\left(\alpha+2+\frac{1}{\beta},\theta x^\beta\right)}{\theta^{\frac{1}{\beta}}\left[(\theta^2+\alpha^2+\alpha)\Gamma(\alpha,\theta x^\beta)+\left(\theta x^\beta\right)^\alpha\left(\theta x^\beta+\alpha+1\right)e^{-\theta x^\beta}\right]} - x.$$

The behaviors of $m(x)$of PWAD for varying values of its parameters $\theta,\alpha$and$\beta$are shown in figure 4. It is observed from the figure 4 that overall mean residual value is decreasing as increased value of θ whereas other parameters are kept as constant; however mean residual is very much affected with value ofβ.

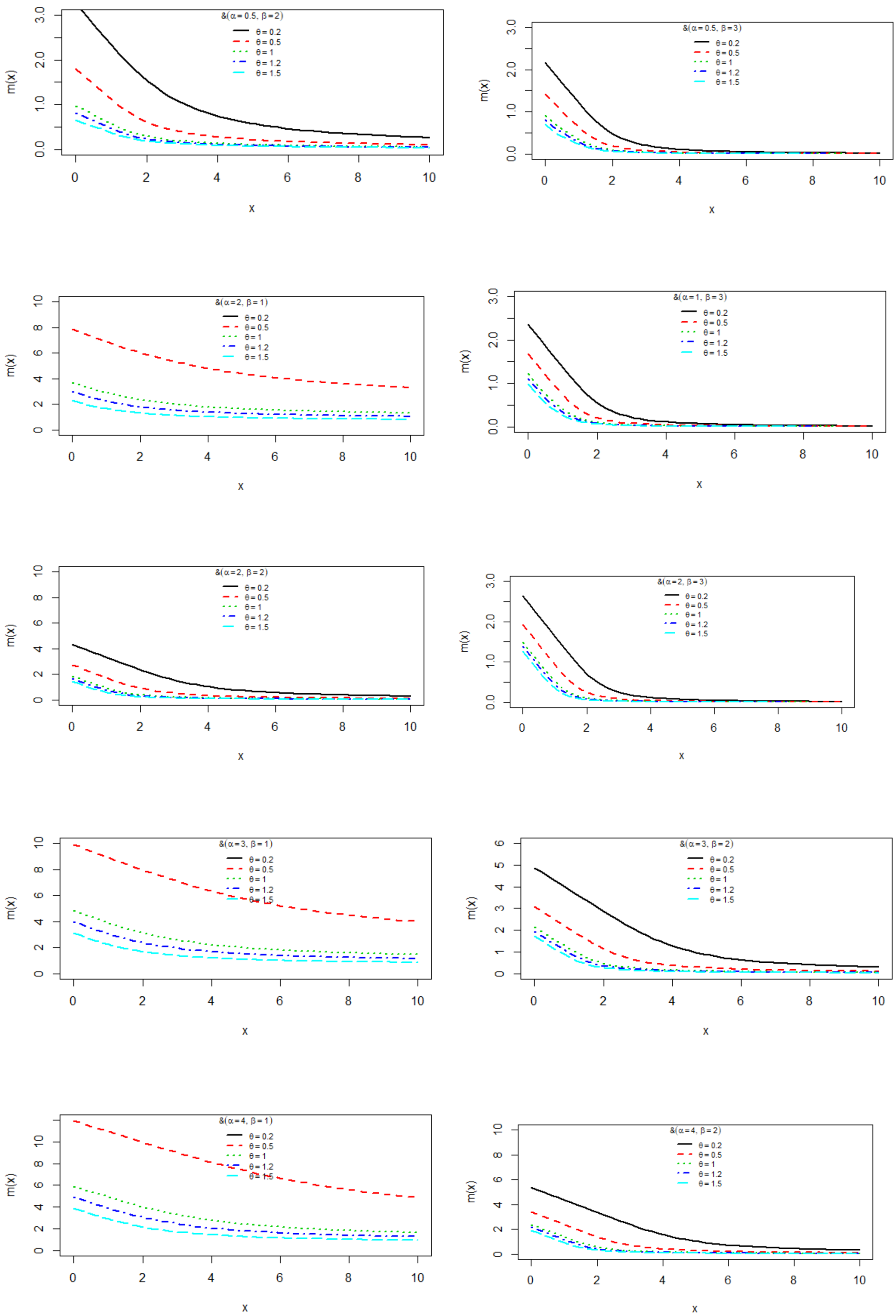


Figure 4: *Graphs of mean residual life function for varying values of parameters* θ, α *and* β.

IV. Moments

The rth moment about origin, $\mu_r{}'$ of PWAD (2.1) can be obtained as

$$\mu_r{}' = E(X^r) = \frac{\beta\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\int_0^\infty x^{\beta\alpha+r-1}\left(1+x^{2\beta}\right)e^{-\theta x^\beta}dx$$

$$= \frac{\beta\theta^{\alpha+2}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\int_0^\infty e^{-\theta x^\beta}x^{\beta\alpha+r-1}\,dx + \int_0^\infty e^{-\theta x^\beta}x^{\beta\alpha+2\beta+r-1}\,dx\right]$$

Assuming $u=\theta x^\beta$, which gives $x=\left(\frac{u}{\theta}\right)^{\frac{1}{\beta}}$ and $dx=\frac{1}{\theta\beta}\left(\frac{u}{\theta}\right)^{\frac{1-\beta}{\beta}}du$, we get

$$\mu_r{}' = \frac{\theta^{\alpha+1}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\int_0^\infty e^{-u}\left(\frac{u}{\theta}\right)^{\frac{\beta\alpha+r}{\beta}-1}du + \int_0^\infty e^{-u}\left(\frac{u}{\theta}\right)^{\frac{\beta\alpha+2\beta+r}{\beta}-1}du\right]$$

$$= \frac{\theta^{\alpha+1}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\frac{1}{\theta^{\alpha+\frac{r}{\beta}-1}}\int_0^\infty e^{-u}u^{\alpha+\frac{r}{\beta}-1}du + \frac{1}{\theta^{\alpha+2+\frac{r}{\beta}-1}}\int_0^\infty e^{-u}u^{\alpha+2+\frac{r}{\beta}-1}du\right]$$

$$= \frac{\theta^{\alpha+1}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\frac{\Gamma\left(\alpha+\frac{r}{\beta}\right)}{\theta^{\alpha+\frac{r}{\beta}-1}} + \frac{\Gamma\left(\alpha+2+\frac{r}{\beta}\right)}{\theta^{\alpha+2+\frac{r}{\beta}-1}}\right]$$

$$= \frac{\theta^{\alpha+1}}{(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}\left[\frac{\theta^2\Gamma\left(\alpha+\frac{r}{\beta}\right)+\Gamma\left(\alpha+2+\frac{r}{\beta}\right)}{\theta^{\alpha+1+\frac{r}{\beta}}}\right]$$

$$= \frac{\theta^2\Gamma\left(\alpha+\frac{r}{\beta}\right)+\Gamma\left(\alpha+2+\frac{r}{\beta}\right)}{\theta^{\frac{r}{\beta}}(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)};\ r=1,2,3,\ldots \tag{4.1}$$

Thus the first four moments about origin of PWAD can be given by

$$\mu_1{}' = \frac{\theta^2\Gamma\left(\alpha+\frac{1}{\beta}\right)+\Gamma\left(\alpha+2+\frac{1}{\beta}\right)}{\theta^{\frac{1}{\beta}}(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}$$

$$\mu_2{}' = \frac{\theta^2\Gamma\left(\alpha+\frac{2}{\beta}\right)+\Gamma\left(\alpha+2+\frac{2}{\beta}\right)}{\theta^{\frac{2}{\beta}}(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}$$

$$\mu_3{}' = \frac{\theta^2\Gamma\left(\alpha+\frac{3}{\beta}\right)+\Gamma\left(\alpha+2+\frac{3}{\beta}\right)}{\theta^{\frac{3}{\beta}}(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}$$

$$\mu_4{}' = \frac{\theta^2\Gamma\left(\alpha+\frac{4}{\beta}\right)+\Gamma\left(\alpha+2+\frac{4}{\beta}\right)}{\theta^{\frac{4}{\beta}}(\theta^2+\alpha^2+\alpha)\Gamma(\alpha)}.$$

Using the relationship between moments about origin and central moments, central moments can be obtained. Since the expressions for central moments are complicated, central moments are not being given.

V. Maximum Likelihood Estimation

Suppose $(x_1,x_2,x_3,\ldots,x_n)$ be a random sample of size n from PWAD (2.1). The natural log likelihood function is thus obtained as

$ln\,L=\sum_{i=1}^n ln\,f_4\,(x_i;\theta,\alpha,\beta)$ $= n[ln\,\beta+(\alpha+2)\,ln\,\theta-ln(\theta^2+\alpha^2+\alpha)-ln\,\Gamma\,(\alpha)]+(\beta\alpha-1)\sum_{i=1}^n ln(x_i)+\sum_{i=1}^n ln\left(1+x_i{}^{2\beta}\right)-\theta\sum_{i=1}^n x_i{}^\beta$ The maximum likelihood estimates (MLEs) of parameters (θ,α,β) of PWAD are the solution of the following nonlinear log likelihood equations

$$\frac{\partial\,ln\,L}{\partial\theta}=\frac{n(\alpha+2)}{\theta}-\frac{2n\theta}{\theta^2+\alpha^2+\alpha}-\sum_{i=1}^n x_i{}^\beta=0$$

$$\frac{\partial\,ln\,L}{\partial\alpha}=n\,ln\,\theta-\frac{n(2\alpha+1)}{\theta^2+\alpha^2+\alpha}-n\psi(\alpha)+\beta\sum_{i=1}^n ln(x_i)=0$$

$$\frac{\partial\,ln\,L}{\partial\beta}=\frac{n}{\beta}+\alpha\sum_{i=1}^n ln\,x_i+\sum_{i=1}^n\frac{2x_i{}^{2\beta}\,ln(x_i)}{1+x_i{}^{2\beta}}-\theta\sum_{i=1}^n x_i{}^\beta\,ln(x_i)=0$$

where $\bar{x}$ is the sample mean and $\psi(\alpha) = \frac{d}{d\alpha} ln\,\Gamma(\alpha)$ is the digamma function. These three natural log- likelihood equations do not seem to be solved directly because they cannot be expressed in closed forms. However, the MLE's of parameters (θ, α, β) can be obtained directly by solving the log likelihood equation using Newton-Raphson iteration method available in R –Software till sufficiently close estimates of parameters are obtained.

VI. Applications

In this section, the applications and goodness of fit of the PWAD have been discussed for two real lifetime datasets. The fit is compared with one parameter lifetime distributions including exponential distribution, Lindley distribution proposed by Lindley [4] and studied by Ghitany et al [5], Akash distribution; two-parameter lifetime distributions including Weibull distribution introduced by Weibull [6], Gamma distribution, Generalized exponential distribution (GED) introduced by Gupta and Kundu [7], Power Lindley distribution (PLD) proposed by Ghitany et al [8], Shukla distribution (SD) proposed by Shukla and Shanker [9],Weighted Lindley distribution (WLD) introduced by Ghitany et al [10] and PAD and WAD and three-parameter lifetime distributions including generalized gamma distribution (GGD) introduced by Stacy[11] and generalized Lindley distribution (GLD) suggested by Zakerzadeh and Dolati [12]. Note that Shanker et al [13] and Shanker[14] have detail discussion on WLD and GLD regarding some important properties and applications for various lifetime data from engineering and biomedical sciences. The first dataset is the data reported by Efron[15] represents the survival times of a group of patients suffering from Head and Neck cancer disease and treated using a combination of radiotherapy and chemotherapy (RT+CT). The second dataset is the data which represents the tensile strength, measured in GPa, of 69 carbon fibers tested under tension at gauge lengths of 20mm and are available in Bader and Priest [16].

Table1 *The data set 1 reported by Efron [15] represent the survival times of a group of patients suffering from Head and Neck cancer disease and treated using a combination of radiotherapy and chemotherapy (RT+CT).*

12.20	23.56	23.74	25.87	31.98	37	41.35	47.38	55.46	58.36
63.47	68.46	78.26	74.47	81.43	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	1776						

Table2 *The following data set 2 represent the tensile strength, measured in GPa, of 69 carbon fibers tested under tension at gauge lengths of 20mm, Bader and Priest [16]*

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	2.140	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	2.570	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	2.880	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 goodness of fit of these distributions for the two datasets, values of $-2\ln L$, AIC (Akaike information criterion), K-S Statistic (Kolmogorov-Smirnov Statistic) and p-value for two datasets have been computed The formulae for AIC and K-S Statistics are as follows: $AIC = -2\ln L + 2k$, and $K - S = \sup_x |F_n(x) - F_0(x)|$, where k being the number of parameters involved in the respective distributions, n is the sample size and $F_n(x)$is the empirical distribution function. The best distribution corresponds to the lower values of$-2\ln L$, AIC and K-S statistic.

Note that the estimates of parameters of the considered distributions are based on maximum likelihood estimates. In this paper, the initial values of the parameters for ML estimates of PWAD have been selected as$\theta = 1.5$,$\alpha = 0.5$ and$\beta = 1.5$for both dataset. In general, it has been observed that the initial values of the parameters can be taken as any positive real numbers, preferably from 0.5 to 5, for any dataset.

The pdf of the fitted distributions are presented in table 3. The ML estimates of parameters of the considered distributions for datasets 1 and 2 are presented in tables 4 and 5. The goodness of fit by K-S statistics for datasets 1 and 2 with considered distributions are presented in tables 6 and 7. The variance-covariance matrix of the parameters (θ, α, β) of PWAD for datasets 1 and 2 are presented in tables 8 and 9. It is obvious from the goodness of fit of the proposed distribution that in tables 4 and 5 it gives better fit than all considered distributions and competes well with GGD. Therefore, PWAD can be considered an important three-parameter lifetime distribution alternative to GGD and other lifetime distributions.

Table 3*: pdf of the fitted distributions*

Distributions	Pdf
Weibull	$f(x;\theta,\alpha) = \theta\alpha x^{\alpha-1}e^{-\theta x^{\alpha}}; x > 0, \theta > 0, \alpha > 0$
Gamma	$f(x;\theta,\alpha) = \frac{\theta^{\alpha}}{\Gamma(\alpha)} e^{-\theta x}x^{\alpha-1}; x > 0, \theta > 0, \alpha > 0$
PLD	$f(x;\theta,\alpha) = \frac{\alpha\theta^2}{(\theta+1)} x^{\alpha-1}(1 + x^{\alpha})e^{-\theta x^{\alpha}}; x > 0, \theta > 0, \alpha > 0$
WLD	$f(x;\theta,\alpha) = \frac{\theta^{\alpha+1}}{\theta+\alpha}\frac{x^{\alpha-1}}{\Gamma(\alpha)}(1 + x)e^{-\theta x}; x > 0, \theta > 0, \alpha > 0$
GED	$f(x;\theta,\alpha) = \theta\alpha\left(1 - e^{-\theta x}\right)^{\alpha-1}e^{-\theta x}; x > 0, \theta > 0, \alpha > 0$
SD	$f(x;\theta,\alpha) = \frac{\theta^{\alpha+1}}{\theta^{\alpha} + \Gamma(\alpha+1)}(1 + x^{\alpha})e^{-\theta x}; x > 0, \theta > 0, \alpha \geq 0$
GGD	$f(x;\theta,\alpha,\beta) = \frac{\beta\theta^{\alpha}}{\Gamma(\alpha)} x^{\beta\alpha-1}e^{-\theta x^{\beta}}; x > 0, \theta > 0, \alpha > 0, \beta > 0$
GLD	$f(x;\theta,\alpha,\beta) = \frac{\theta^{\alpha+1}}{\theta+\beta}\frac{x^{\alpha-1}}{\Gamma(\alpha+1)}(\alpha + \beta x)e^{-\theta x}$
Lindley	$f(x;\theta) = \frac{\theta^2}{\theta+1}(1 + x)e^{-\theta x}; x > 0, \theta > 0$

Table 4: *Summary of the ML estimates of parameters for dataset 1*

Model	ML Estimates		
	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$
PWAD	11.8734	27.3026	**0.1804**
GLD	0.00473	0.05243	**5.07505**
GGD	11.25540	27.72340	**0.18220**
SD	0.00458	0.02380	
WAD	0.0090	0.0165	……..
PAD	0.16751	0.55764	……..
WLD	0.00531	0.21191	……
PLD	0.05301	0.68893	…….
GED	0.00482	1.09367	……
Gamma	0.00489	1.08501	……
Weibull	0.00710	0.92327	…….
Akash	0.01344	…….	……..
Lindley	0.00892	…….	……..
Exponential	**0.00447**	……	…….

Table 5: *Summary of the ML estimates of parameters of dataset 2*

Model	ML Estimates		
	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$
PWAD	0.2918	1.7049	**2.7229**
GLD	9.39076	22.71981	**4.77105**
GGD	0.30440	3.58610	**2.64830**
SD	5.9922	17.1611	
WAD	9.7584	22.2327	….
PAD	0.16964	3.06033	…..
WLD	9.62655	22.89383	…..
PLD	0.0500	3.8680	……
GED	2.03307	87.28471	……
Gamma	9.53843	23.38184	…..
Weibull	0.00558	5.33523	……
Akash	0.96472	……	…….
Lindley	0.65450	…….	…….
Exponential	**0.40794**	…….	…….

Table 6: *Summary of Goodness of fit by K-S Statistic for dataset 1*

Model	$-2\ln L$	AIC	K-S	p-value
PWAD	555.67	561.67	0.081	**0.921**
GLD	564.09	570.09	0.150	**0.248**
GGD	555.64	561.64	0.079	**0.921**
SD	564.00	568.00	0.147	**0.267**
WAD	580.32	584.32	0.219	**0.023**
PAD	559.10	563.10	0.108	**0.635**
WLD	565.91	569.91	0.161	**0.181**
PLD	560.78	564.78	0.118	**0.529**
GED	563.93	567.93	0.145	**0.280**
Gamma	564.10	568.10	0.149	**0.249**
Weibull	563.71	567.71	0.298	**0.005**
Akash	609.92	611.92	0.279	**0.001**
Lindley	579.16	581.16	0.219	**0.025**
Exponential	**564.01**	**566.01**	**0.145**	**0.282**

Table 6 represents the goodness of fit by K-S Statistic for data set-1, It is observed that AIC and P-value from K-S test were found almost minimum and maximum in comparison to all other included distributions respectively. Therefore, it may be concluded that PWAD is better fits than other included distributions except GGD (Generalized Gamma distribution). Hence, PWAD can be considered an important lifetime distribution for modeling lifetime data.

Table 7: *Summary of Goodness of fit by K-S Statistic for dataset 2*

Model	$-2\ln L$	AIC	K-S	p-value
PWAD	97.93	103.93	0.037	**0.999**
GLD	101.96	107.96	0.056	**0.979**
GGD	100.58	106.58	0.044	**0.999**
SD	184.35	188.35	0.290	**0.000**
WAD	99.95	103.95	0.057	**0.976**
PAD	98.02	102.02	0.038	**0.999**
WLD	100.04	104.04	0.058	**0.974**
PLD	98.12	102.12	0.044	**0.998**
GED	109.24	113.24	0.095	**0.558**
Gamma	100.07	104.07	0.058	**0.973**
Weibull	99.31	103.31	0.060	**0.964**
Akash	224.27	226.27	0.362	**0.000**
Lindley	238.38	240.38	0.401	**0.000**
Exponential	**261.73**	**263.73**	**0.448**	**0.000**

Table 7 represents the goodness of fit by K-S Statistic for data set-2, It is observed that AIC and P-value from K-S test were found almost minimum and maximum in comparison to almost all other included distributions respectively expect PAD and PLD. Therefore, it may be concluded that PWAD is a better fit than other included distributions except PAD (Power Akash distribution) and

PLD (Power Lindley distribution). Further, PWAD competing well with the considered distributions and hence can be an important distribution for lifetime data.

Table 8: *Variance-covariance matrix of the parameters θ, $\alpha$and β of PWAD for dataset* **1**

$$\begin{array}{c} \\ \hat{\theta} \\ \hat{\alpha} \\ \hat{\beta} \end{array} \begin{array}{c} \hat{\theta}\hat{\alpha}\hat{\beta} \\ \begin{bmatrix} 3044.5355 & 4357.7480 & -4.9546 \\ 4357.7480 & 6246.2648 & -7.0633 \\ -4.9546 & -7.0633 & 0.0081 \end{bmatrix} \end{array}$$

Table 9: *Variance-covariance matrix of the parameters θ, $\alpha$and β of PWAD for dataset 2*

$$\hat{\theta}\hat{\alpha}\hat{\beta} \qquad \begin{array}{c} \hat{\theta} \\ \hat{\alpha} \\ \hat{\beta} \end{array} \begin{bmatrix} -12.5447 & -27.6412 & 1.7898 \\ -27.6412 & -60.2546 & 3.9955 \\ 1.7898 & 3.9955 & -0.2499 \end{bmatrix}$$

VII. Conclusions

In the present paper a three-parameter power weighted Akash distribution (PWAD) ,of which two-parameter weighted Akash distribution (WAD), two-parameter power Akash distribution (PAD) and one parameter Akash distribution are particular cases, has been introduced and studied. Its moments, hazard rate function, mean residual life function and stochastic ordering have been discussed. Maximum likelihood estimation has been discussed for estimating the parameters of the distribution. The applications of the proposed distribution have been discussed through two real lifetime datasets. The goodness of fit test of the proposed distribution is a better model for lifetime data than the other well-known one parameter, two-parameter and three-parameter lifetime distributions.

ACKNOWLEDGEMENTS: Authors are grateful to the editor in chief and the anonymous reviewers for their constructive comments to improve the quality of paper.

References

[1] Shanker, R. and Shukla, K.K. (2016): Weighted Akash distribution and Its Application to model lifetime data, International Journal of Statistics, 39 (2), 1138 – 1147.

[2] Shanker, R. and Shukla, K.K. (2017): Power Akash distribution and its Application, *Journal of Applied Quantitative Methods*, 12(3), 1 – 10.

[3] Shanker, R. (2015): Akash Distribution and Its Applications, *International Journal of Probability and Statistics,* 4 (3), 65 – 75.

[4] Lindley, D.V. (1958): Fiducial distributions and Bayes' Theorem, *Journal of the Royal Statistical Society*, Series B, 20, 102 – 107.

[5] Ghitany, M.E. , Atieh, B. and Nadarajah, S. (2008): Lindley distribution and its Application, *Mathematics Computing and Simulation,* 78, 493 – 506.

[6] Weibull, W. (1951): A statistical distribution of wide applicability, *Journal of Applied Mathematics*, 18, 293 – 297.

[7] Gupta, R.D. and Kundu, D. (1999): Generalized Exponential Distribution, *Austalian & New Zealand Journal of Statistics*, 41(2), 173 – 188.

[8] Ghitany, M.E., Al-Mutairi, D.K., Balakrishnan, N., and Al-Enezi, L.J. (2013): Power Lindley distribution and Associated Inference, *Computational Statistics and Data Analysis*, 64,20– 33.

[9] Shukla, K. K. and Shanker, R. (2019): Shukla distribution and its application, *Reliability: Theory and Applications* 54(3), 46-55.

[10] Ghitany ,M.E., Alqallaf ,F., Al-Mutairi, D.K., Husain, H.A. (2011): A two-parameter weighted Lindley distribution and its applications to survival data, *Mathematics and Computers in simulation*, 81, 1190-1201.

[11] Stacy, E.W. (1962): A generalization of the gamma distribution, *Annals of Mathematical Statistical*, 33, 1187 – 1192.

[12] Zakerzadeh, H. and Dolati, A. (2009): Generalized Lindley distribution, *Journal of Mathematical extension*, 3 (2), 13 – 25.

[13] Shanker, R**.,** Shukla, K.K. and Hagos, F. (2016): On Weighted Lindley distribution and Its Applications to Model Lifetime Data**,** *Jacobs Journal of Biostatistics,*1(1), 1- 9.

[14] Shanker, R. (2016): On Generalized Lindley Distribution and its Applications to model lifetime data from biomedical science and engineering, *Insights in Biomedicine,* 1(2), 1 – 6.

[15] Efron, B. (1988): Logistic regression, survival analysis and the Kaplan-Meier curve, *Journal of the American Statistical Association*, 83, 414-425.

[16] Bader,M.G., 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.

COMPREHENSIVE OPTIMIZATION OF SIGNOMIALLY COMBINED-NUMERAL NONLINEARITY CODING PROBLEMS WITH FREE VARIABLES QUANTITY

Dr. K.Srinivasa Rao

•

GITAM (Deemed to be University)
skolli2@gitam.edu

Dr.U.V.Adinarayana Rao

•

GITAM (Deemed to be University)
auppu@gitam.edu

Abstract

Combined-numeral nonlinearity coding problem (CNNLCP) troubles concerning usual restrictions and empirical roles and constant then numeral variable quantity frequently appear in a production project, substance method business, and organization. Even though several optimize techniques need to be established for CNNLCP troubles, these techniques can hold signal relationships together with a particular variable quantity. Thus, this analysis intends a different approach used to explain a signal CNNLCP trouble and set free variable amount towards achieving an internationally optimum explanation. The signal CNNLCP trouble is initially converted into an individual with one certain variable quantity. However, the changed trouble is redeveloped as a curving combined-numeral system as the Convexness of the approaches and piecewise linearization systems. A comprehensive optimal signal CNNLCP trouble can ultimately be realized inside the acceptable inaccuracy. Algebraic models are also introduced to establish the effectiveness of the recommended approach.

Keywords: Comprehensive Optimize, Combined-numeral nonlinearity coding, Set free variable quantity, Convexness.

I. Introduction

Combined-numeral nonlinearity coding (CNNLCP) troubles concerning together constant and distinct variable quantity rise in several claims of a production project, substance method [12,14,26], for instance, combination and project of partings [1–4], no intricate isothermal apparatus webs [20], stage symmetry [28] and frame-conversation webs [29]. Biegler and Grossmann [7] demonstrate optimized procedures that have been affected in development techniques planning. They revealed that pattern and creation troubles had been controlled by nonlinearity coding and CNNLCP types. Floudas et al.. [13] indicated the investigation activity into comprehensive Optimize for 1998–2003, together with the determinist universal optimize improvements into CNNLCPs and connected products. Along with the expanding dependence on demonstrating optimized troubles in functional troubles, several hypothetical and algorithmic influences of CNNLCP have been planned. Though these troubles regularly consist of noncurved roles, the standard local optimize techniques cannot be dealt with to ensure comprehensive optimality. Used to discuss the nonconvexities in CNNLCP troubles, the established procedures can be separated into binary attitudes.

The stochasticity processes consist of arbitrary factors in their pursuit and are dependent on an arithmetical dispute to demonstrate their merging. For example, Salcedo et al. [32] proposed an increased arbitrary examination process for explaining nonlinearity optimize troubles. Hussain and Al-Sultan [19] planned a fusion system for noncurved function minimization by applying the natural method to create examine instructions. Yiu et al. [39] established a Combined slope attitude shown on a virtual hardening process and slope-established system to explain multidimensional noncurved uninterrupted optimize troubles. The experimental method is a variation of stochasticity techniques, for example, the restriction analyses system [16]. The collection of all aspirant explanations that can be produced in each repetition must vary on the present repetition position and be changed by eliminating a subgroup of contestant explanations known as tabu. The meaning of which contestant results are tabu goes upon the changes that have got be there created among current repetition positions. While the tabu analyses are more efficient than virtual galvanizing, these stochasticity systems stated above cannot ensure discovering the universal optimum. Hence, the worth of the explanation is not confirmed. Likewise, the likelihood of finding the universal description reduces when the difficulty volume strengthens.

Determinist procedures in a typical analysis of optimizing methods [7,17,18], several determinist approaches for curved CNNLCP troubles have been evaluated. The processes contain area and constrained (CAC) [9,22,33], widespread binges decay(WBD) [15], outward estimate (OE) [10,11,31], continued reducing plane technique (CRPT)) [37],and simplified disjunctive coding (SDC) [21]. The CAC system can only get the universal explanation when each subproblem can be explained worldwide optimality. The WBD system, the OE system, and the CRPT system cannot explain CNNLCP troubles with noncurved restrictions since the troubles cannot develop a distinctive optimum in the resolution method. Lee and Grossmann [21] planned a resolution system for the SDC simulations that parallel distinct/permanent optimization troubles involving disconnections and nonlinearity inequities and reasoning proposals. The empirical roles and the restrictions in the GDP trouble are expected to be curved and constrained. Maranas and Floudas [25] required a process to produce curved estimators for universal geometric coding challenges via the hollow words' index conversion and straight dryness. Adjiman et al. [1,2] projected two worldwide optimize techniques, SMIN-α BB and GMIN-α BB, for noncurved CNNLCP established on the model of separate off-and-constrained and trust on Optimize or period-created changing-required updates to improve productivity. Even though one likely method to avoid noncurved ties in CNNLCP shows is a reformulation, for example, applying the index revolution to deal with the simplified symmetrical coding troubles.

In which a signal phrase $x_1^{\alpha}\, x_2^{\beta}$ is assigned into an index term $e^{\alpha ln x_1 + \beta ln x_2}$ [12,14,26], the index alteration procedure can only be utilized to precisely certain variable quantity and is thus incapable of trading with noncurved CNNLCP difficulties with set free variable amount. Pörn et al. [30] announced separate Convexness approaches for converting noncurved CNNLCP troubles into curved issues and explaining them by a CNNLCP solver. They recommended an easy conversion, x $+\tau = e^{x}$, to deal with a free distinct varying. Introducing the converted effect into the earliest indication conditions will bring different signal periods, growing computational complications.

II. Transformation of free variables

The mathematical formulation of a signomial CNNLCP problem with free variables considered in this study is expressed as follows:

$$\text{Minimize } f(x,y)$$
$$\text{Subjects to } g_i(x,y) \le 0, \quad i = 1,2,3,\ldots\ldots I \tag{1}$$

$$x = (x_1, x_{2,} \dots \dots . x_p, x_{p+1} \dots \dots, x_n), x_i \ \le x_i \le x^i \tag{2}$$
$$y = (y_1, y_{2,} \dots \dots . y, y_{q+1} \dots \dots, y_m)\, , y_i \ \le y_i \le y^i \tag{3}$$

where$x_i \epsilon R^+$ for $1 \le i \le p$, x_i are constrained set free variable quantity for $1 + p \le i \le n$, y_j are +ve number /distinct variable quantity for $1 \le j \le q, y_j$ are constrained number /distinct variable quantity for $q + 1 \le j \le m, f(x, y)$ and $g_i(x, y)$ are Combined-numeral signal roles, x_i, and x^i . are more diminutive and more significant boundaries of the permanent variable quantity x_i, and y_j and y^j are more minor and more significant boundaries of the distinct variable quantity y_j, respectively.

Let $$x_i = x_i^+ - x_i^-, x_i^+ x_i^- \ge 0, for\ i = p + 1, \dots \dots n, \tag{4}$$
$$y_j = y_j^+ - y_j^-, y_j^+ y_j^- \ge 0, for\ j = q + 1, \dots \dots m, \tag{5}$$
And nonlinearity relationships $x_i^{\alpha_i}\ and\ y_j^{\beta_j}$ are expressed as

$$x_i^{\alpha_i} = \ (x_i^+)^{\alpha_i} + (-1)^{\alpha_i}(x_i^-)^{\alpha_i}, \alpha_i \in Z, for\ i = p + 1, \dots \dots n, \tag{6}$$
$$y_j^{\alpha_j} = \ \left(y_j^+\right)^{\beta_j} + (-1)^{\beta_j}\left(y_j^-\right)^{\beta_j}, \beta_j \in Z, for\ j = q + 1, \dots \dots m, \tag{7}$$

If $x_i^+ > 0$ and $x_i^- = 0\, , then\ x_i\ is$ positive. Otherwise, if $x_i^- > 0\ and\ x_i^+ = 0, then\ x_i\ is\ positive$.
Remark 1

Let, $x_i = x_i^+ - x_i^-, x_i^+ x_i^- \ge 0, and\ \ x_i^+\ and\ x_i^-$ the resulting inequities.
$$x_i^+ \le x_i^- \theta_i$$
$$x_i^- \le x_{-i}(\theta_i - 1). where \theta_i \in [0,1].$$

I. Preposition

Let $x_i^- \in R, 0 \le x_i^+ \le x_i^-, \lambda_i \in [0,1], \epsilon_0 \le \bar{x}_i^+ \le x_i^-, \epsilon_0 > 0,$

Then, $$x_i^+ = \bar{x}_i^+, \lambda_i \Rightarrow \begin{cases} (i) 0 \le x_i^+ \le x_i^-, \lambda_i \\ (ii)\ x_i^-(\lambda_i - 1) + \bar{x}_i^+ \le x_i^+ \le \bar{x}_i^+. \end{cases}$$

Proof :

If $x_i^+ = 0, \Rightarrow$ (i) is initiated, then $\lambda_i = 0$, hence $\bar{x}_i^+ \ \lambda_i = 0$ then $x_i^+ = \bar{x}_i^+ \ \lambda_i$
If $x_i^+ > 0 \Rightarrow (ii)$ is initiated, then $\lambda_i = 1$, hence $x_i^+ = \bar{x}_i^+$ then $x_i^+ = \bar{x}_i^+ \ \lambda_i$
If $\bar{x}_i^+ \ \lambda_i = 0 \Rightarrow \lambda_i = 0\ and\ \ (i)$ is initiated , hence$x_i^+ = 0$ then $x_i^+ = \bar{x}_i^+ \ \lambda_i$
If$\bar{x}_i^+ \ \lambda_i > 0, \Rightarrow \lambda_i = 1\ and\ \ (ii)$ is initiated hence $x_i^+ = \bar{x}_i^+$ and $x_i^+ = \bar{x}_i^+ \ \lambda_i$
$x_i^+ = \bar{x}_i^+ \ \lambda_i$is determined .

Now denote $z^+ and\ \tilde{z}^+$ as below :

$z^+ = x_1^{\alpha_1} \dots \dots \dots . . x_p^{\alpha_p} \left(x_{p+1}^+\right)^{\alpha_{p+1}} \dots \dots . . (x_n^+)^{\alpha_n}$ and
$\tilde{z}^+ = x_1^{\alpha_1} \dots \dots \dots . . x_p^{\alpha_p} \left(\tilde{x}_{p+1}^+\right)^{\alpha_{p+1}} \dots \dots . . (\tilde{x}_n)^{\alpha_n}$, where $\tilde{x}_i^+$ are positive variables .

From Proposition 1,

$z^+ = x_1^{\alpha_1} \dots \dots \dots . . x_p^{\alpha_p} \left(\tilde{x}_{p+1}^+ \lambda_{p+1}\right)^{\alpha_{p+1}} \dots \dots . . (\tilde{x}_n^+ \lambda_n)^{\alpha_n}$ and it is clear that
$$z^+ = \tilde{z}^+ \lambda_{p+1} \dots \dots \dots . \lambda_n, \quad \lambda_i \in [0,1] \tag{8}$$

Remark 2

Let, $\lambda\, , \lambda_i \in [0,1]\ for\ i = p + 1, \dots \dots \dots \dots n. then:$
$$\lambda \ = \lambda_{p+1} \lambda_{p+2} \dots \dots \lambda_n \ \Rightarrow \begin{cases} (i) \lambda \le \lambda_i for\ i = p + 1, \dots n, \\ (ii)\ \ \lambda \ge \sum_{i=p+1}^n \lambda_i - n + p + 1 \end{cases}$$

By discussing Remark 2, Eq (8) becomes

$$z^{+} = \tilde{z}^{+}\lambda, \quad \lambda \in [0,1]. \tag{9}$$

From Proposition 1, Eqs is equivalent to the following two linear inequalities.

$(i) 0 \leq z^{+} \leq \bar{z}\, \lambda,$

$(ii)\ \tilde{z}^{+} + \bar{z}(\lambda - 1) \leq z^{+} \leq \tilde{z}^{+}.$

$\lambda \in \{0,1\}, \bar{z}$ is an upper bound of z^{+}

III. Classification of curved relationships and curved relaxation approaches

Convexness policies used for signal periods are essential techniques for worldwide optimization efforts. Sun et al. [34] planned a Convexness technique used for international optimize efforts with unmodulated roles in various restricted situations. Wu et al.[38] established a more than typical Convexness, then coalification conversion was used to explain a standard worldwide optimize challenge and specific unmodulated estates. CNNLCP problem can be redeveloped with several Convexness techniques interested in a new curved Combined-numeral program resolvable to achieve an almost universal optimum. Björk et al. [8]planned a worldwide optimized system created on Convexness signal conditions. They examined that the correct selection of revolution for Convexness noncurved signal conditions strongly impacts the effectiveness of the optimized method. Tsai et al. [36] also recommended Convexness systems for the signal conditions with trio variable quantity. This analysis introduces general Convexness systems and laws to convert a CNNLCP problem into a curved Combined-numeral system.

I Proposition

Let, $f(x) = c\prod_{i=1}^{n} x_i^{\alpha_i}, x = (x_1, x_2, \dots \dots \dots x_n)$, $c, x_i, \alpha_i \in R$, for all i, is curved if $c \leq 0, x_i \geq 0, \alpha_i \geq 0$(for $i = 1,2, \dots \dots \dots n$)$1 - \sum_i^n \alpha_i \geq 0$

Proof :

Let $H_i(x)$ be the most crucial trivial of a Hessian matrix $H(x) of\ f(x)$. The determinant of $H_i(x)$ $\det H_i(x) = (-1)^i \left(\prod_{j\in J_i}^{i} c\alpha_j\, x_j^{i\alpha_j - 2}\right)\left(\prod_{j\in J_i, J_i \neq \emptyset}^{n} x_J^{i\alpha_J}\right)\left(1 - \sum_{J\in J_i} \alpha_j\right)$.

Since, $\det H_i(x) \geq 0$ when $c \leq 0, x_i \geq 0, \alpha_i \geq 0$ for all i and $1 - \sum_i^n \alpha_i \geq 0, H_i(x)\ i = 1,2, \dots \dots . . n$.

Corollary **:**

1Let$f(x) = c\prod_{i=1}^{n} x_i^{\alpha_i}, x = (x_1, x_2, \dots \dots \dots x_n)$, $c, x_i, \alpha_i R$, for all i, is curved if $c \leq 0, x_i \geq 0, \alpha_i \geq 0$(for $i = 1,2, \dots \dots \dots n$)

II Preposition

A nonlinearity relationship

$s = x_1^{\alpha_1} x_2^{\alpha_2} \dots \dots . . x_n^{\alpha_n}$, where $x_1, x_2, \dots \dots \dots x_n > 0, \alpha_i < 0 (for\ i = 1,2,3 \dots \dots k), and \alpha_i \geq 0 (i = k + 1, k + 2, \dots \dots \dots \dots . . n)$,

$(i)\ \ s = \prod_{i=1}^{k} x_i^{\alpha_i} \prod_{i=k+1}^{n} z_i^{-\alpha_i},$

$(ii)\ \ z_i + L(-x_i^{-1}) \leq 0\ for\ \ i = k + 1, k + 2, \dots \dots \dots \dots . . n$

$(ii)\ \ x_i^{-1} - z_i \leq 0 + L(-x_i^{-1}) \leq 0\ for\ i = k + 1, k + 2, \dots \dots \dots \dots . . n$

Proof :

$$L(-x_i^{-1}) = -x_i^{-1}, z_i = x_i^{-1}\ for\ i = k + 1, k + 2, \dots . n,$$

after (ii) and (iii) since $, z_i > 0\ and - \alpha_i \leq 0\ for\ i = k + 1, k + 2, \dots \dots n, s$ is then a curved period describing to corollary1.

the piecewise straight function *L(f (x))* for approaching the hollow role *f (x)* [27,35]. Splitting should be achieved to close the gap since a large enough reduction can be close to the earliest nonlinearity problem in any to define beforehand precision. Splitting programs for typical SOS Class 2 cases can be noticed, for example, in [5,6].

$\left|\frac{f(x)-L(f(x))}{f(x)}\right|$ is utilized to assess the inaccuracy in the straight calculation. Assume $f\ (x)$ is an empirical task and x^* results from the converted system. In that case, the linearity makes non involve delicacy till $\left|\frac{f(x^*)-L(fx^*)}{f(x^*)}\right| \leq \varepsilon_2$, where ε_2 is the optimum acceptance. If $g(x) < 0$ is a restriction and x^* is the result, then x^* is achievable if $\left|\frac{f(x^*)-L(fx^*)}{f(x^*)}\right| \leq \varepsilon_1$ and $L(g(x^*)) < \varepsilon_1$ where ε_1 is the feasibility acceptance.

III Proposition

Let, $s = -{x_1}^{\alpha_1}{x_2}^{\alpha_2} \ldots\ldots\ldots {x_n}^{\alpha_n}$ where $x_1, x_2 \ldots\ldots\ldots\ldots x_n > 0, 0 \leq a_1 \leq a_2 \leq \cdots \leq a_{k,}$, $0 \geq a_{k+1} + a_{k+2} \geq \cdots \geq a_n$ and $\sum_{i=1}^{r} \alpha_i < 1$ for some most extensive numeral *r , such that r ≤ k,*

$$S = -\prod_{i=1}^{r} x_i^{\alpha_i} \prod_{i=r+1}^{n} z_i^{\beta} \quad , \beta = \frac{1-\sum_{i=1}^{r}\alpha_i}{n-r} \ ,$$

- $z_i + L\left(-x_i^{\frac{\alpha_i}{\beta}}\right) \leq 0$
- $-x_i^{\frac{\alpha_i}{\beta}} - z_i \leq 0$

Where $L\left(-x_i^{\frac{\alpha_i}{\beta}}\right)$ is piecewise linearization function of a hollow period $-x_i^{\frac{\alpha_i}{\beta}}$

Proof :

$$\left(-x_i^{\frac{\alpha_i}{\beta}}\right) = -x_i^{\frac{\alpha_i}{\beta}} \ , \ z_i = x_i^{\frac{\alpha_i}{\beta}} for\ i = r+1, r+2, \ldots\ldots\ldots n,$$

Since $\alpha_i > 0\ \ for\ i = 1,2, \ldots\ldots\ldots r. z_i > 0\ \ for\ i = k+1, k+2, \ldots\ldots\ldots\ldots . n, \beta > 0\ \ and$ $\sum_{i=1}^{r} \alpha_i + (n-r)\beta = 1. s$

Remark 3
Let $f(x) = x^{\alpha}$ for $x > 0$ is curved at what time $\alpha \leq 0$ or $\alpha \geq 1$. $f(x)$ is hollow at what time $0 \leq \alpha \leq 1$.

Remark 4
Let $y \in \{d_1, d_2 \ldots\ldots d_m\}\ \ d_{j+1} > d_j > 0 for\ \ j = 1,2,3, \ldots\ldots m-1$
$y^{\alpha} = \sum_{j=1}^{m} {d_j}^{\alpha} u_j\ \ , where\ \sum_{j=1}^{m} u_j = 1\ u_j \in \{0,1\}$

Remark 5
Let $s = u\,f(x)$ wherever $f(x)$ is a straight serve is equal to the resulting straight variations:

- $\overline{f(x)}(u-1) + f(x) \leq s \leq \overline{f(x)}(u-1) + f(x).$
- $-\overline{f(x)}u \leq s \leq \overline{f(x)}u,$

Where $u \in \{0,1\}, s$ is an unobstructed in symbol flexible, and $\overline{f(x)}$ is the greater duty-bound of $f(x)$,

IV. Examples

I. Example

Minimize $\quad {x_1}^2{x_2}^{-2}x_3 - 2x_2^{0.7}x_3^{0.2} + x_4x_5^{-2} - 2x_1 - 4x_3$
Subject to $\quad x_1 + 6x_2 - x_3 - 5x_4 \leq 2, \qquad (10)$

$x_3^{1.5}x_4 + 0.5x_2 + 3x_1 \le -10$ (11)

$-x_1 - 0.5x_4 + x_5 \le 6,$ (12)
$-7 \le x_1 \le 5,\ \ 1 \le x_2 \le 10,\ \ 1 \le x_3 \le 5,\ \ 2 \le x_4 \le 8,\ \ 2 \le x_5 \le 9,$
$x_1, x_2, x_4, x_5 \in R,\ \ x_3 \in Z.$

By Remark 1,
$x_1 = x_1^{+} - x_1^{-}\ ,\ x_1^{+}, x_1^{-} \ge 0,$

Minimize $(x_1^{+})^2 x_2^{-2} x_3 + (x_1^{-})^2 x_2^{-2} x_3 - 2x_2^{0.7} x_3^{0.2} + x_4 x_5^{-2} - 2x_1 - 4x_3$

Subject to $x_1 = x_1^{+} - x_1^{-},$ (13)
$x_1^{+} \le 5\theta_1,$ (14)
$x_1^{-} \le 7(\theta_1 - 1)\,,$ (15)
$x_1^{+} - x_1^{-} + 6x_2 - x_3 - 5x_4 \le 2$ (16)

$x_3^{1.5}x_4 + 0.5x_2 + 3x_1^{+} - 3x_1^{-} \le -10$ (17)

$-x_1^{+} + x_1^{-} - 0.5x_4 + x_5 \le 6,$ (18)
$0 \le x_1^{+} \le 5,\ \ 0 \le x_1^{-} \le 7, 1 \le\ x_2 \le 10, 1 \le x_3 \le 5, 2 \le x_4 \le 8$
$, 2 \le x_5 \le 9, \theta_1 \in \{0,1\}\,, x_2, x_4, x_5 \in R, x_3 \in Z$

Now we familiarize two severely +ve variables $\tilde{x}_1^{+}, \tilde{x}_1^{-}$ as follows:

$0 \le x_1^{+} \le 5\lambda_1,$ (19)
$\tilde{x}_1^{+} + 5(\lambda_1 - 1) \ \le\ \tilde{x}_1^{+} \ \le\ \tilde{x}_1^{+}$ (20)
$0 \le x_1^{-} \le 7\lambda_2$ (21)
$\tilde{x}_1^{-} + 7(\lambda_2 - 1) \le \tilde{x}_1^{-} \le \tilde{x}_1^{-}$ (22)

For computer implementation,

$\tilde{x}_1^{+}, \tilde{x}_1^{-} \ge \varepsilon_0$ where $\varepsilon_0 = 10^{-7}$ is a zero acceptance. The signomial times $z_1^{+} = (x_1^{+})^2 x_2^{-2} x_3$ and $z_1^{-} = (x_1^{-})^2 x_2^{-2} x_3$ $\tilde{z}_1^{+} = (\tilde{x}_1^{+})^2 x_2^{-2} x_3$ and $\tilde{z}_1^{-} = (\tilde{x}_1^{-})^2 x_2^{-2} x_3$, respectively, where $0 \le z_1^{+} \le \bar{z}\lambda_1,\ \ \tilde{z}_1^{+} + \bar{z}(\lambda_1 - 1) \le z_1^{+} \le \tilde{z}_1^{+}, 0 \le z_1^{-} \le \bar{z}\lambda_2,\ \ \tilde{z}_1^{-} + \bar{z}(\lambda_2 - 1) \le z_1^{-} \le \tilde{z}_1^{-}$

. From Proposition 2, the nonlinearity term
$-2x_2^{0.7}.x_3^{0.2}$ is curved.

The noncurved relationships $x_3^{1.5}x_4$ and $x_4 x_5^{-2}$ can be changed into curved relations and $z_4^{-1} x_5^{-2}$, respectively, anywhere $z_3 = x_3^{-1}$ and $z_4 = x_4^{-1}$
According to Remark 4,

$z_3 = x_3^{-1}$ can be linearized as $z_3 = u_1 + \frac{1}{2}u_2 + \frac{1}{3}u_3 + \frac{1}{4}u_4 + \frac{1}{5}u_5$ where $x_3 = u_1 + 2u_2 + 3u_3 + 4u_4 + 5u_5$.

- The noncurved relationships $(\tilde{x}_1^{+})^2 x_2^{-2} x_3$ and $(\tilde{x}_1^{-})^2 x_2^{-2} x_3$ can be transferred into curved relationships $e^{2y_1^{+} - 2y_2 + y_3}$ and $e^{2y_1^{-} - 2y_2 + y_3}$, respectively, where $y_1^{+} = ln\tilde{x}_1^{-}$, $y_1^{-} = ln\tilde{x}_1^{-}$, $y_2 = \ln x_2$ and $y_3 = \ln x_3$.

Minimize $z_1^{+} + z_1^{-} - 2x_2^{0.7}.x_3^{0.2} + z_4^{-1} x_5^{-2} - 2x_1 - 4x_3$
subject to $x_1 = x_1^{+} - x_1^{-}$
$(x_1^{+} - x_1^{-} + 6x_2 - x_3 - 5x_4 \le 2) - (\tilde{x}_1^{-} + 7(\lambda_2 - 1) \le x_1^{-} \le \tilde{x}_1^{-})$
$z_3^{-1.5} x_4^{-1} + 0.5x_2 + 3x_1^{+} - 3x_1^{-} \le -10,$
$y_1^{+} = L(ln\tilde{x}_1^{+}),\ \ y_1^{-} = L(ln\tilde{x}_1^{-}),\ \ y_2 = L\ln x_2,$
$y_3 == u_1 \ln 1 + u_2 \ln 2 + u_3 \ln 3 + u_4 \ln 4 + u_2 \ln 5,$
$0 \le z_1^{+} \le \bar{z}\lambda_1, e^{2y_1^{-} - 2y_2 + y_3} + \bar{z}(\lambda_1 - 1) \le z_1^{+} \le L\left(e^{2y_1^{-} - 2y_2 + y_3}\right),$
$0 \le z_1^{+} \le \bar{z}\lambda_2, e^{2y_1^{-} - 2y_2 + y_3} + \bar{z}(\lambda_2 - 1) \le z_1^{-} \le L\left(e^{2y_1^{-} - 2y_2 + y_3}\right),$
$x_3 = u_1 + 2u_2 + 3u_3 + 4u_4 + 5u_5,$

$$z_3 = u_1 + \frac{1}{2}u_2 + \frac{1}{3}u_3 + \frac{1}{4}u_4 + \frac{1}{5}u_5$$
$$u_1 + u_2 + u_3 + u_4 + u_5 = 1$$
$$x_4^{-1} - z_4 \leq 0, z_4 + L(-x_4^{-1}) \leq 0,$$
$$\epsilon_0 \leq \tilde{x}_1^+ \leq 5, \epsilon_0 \leq \tilde{x}_1^- \leq 7, 1 \leq x_2 \leq 10, \; 1 \leq x_3 \leq 5, 2 \leq x_4 \leq 8, 2 \leq x_5 \leq 9,$$
$$\theta_1, \lambda_1, \lambda_2 \in \{0,1\}, \; u_1, u_1, u_1, u_1, u_1 \in \{0,1\}, \tilde{x}_1^+, \tilde{x}_1^-, x_2, x_4, x_5 \in R, x_3 \in Z$$

The optimality acceptance and probability acceptance are within the prespecified error of 0.001. The universally optimal explanation found is $(x_1, x_2, x_3, x_4, x_5) = (-5.353, 4.548, 1, 3.787, 2.541)$ along with the empirical cost is 2.803.

Table 1:

Number of variables in the reformulated model of each example				
	Quantity of originals variable quantity	Quantity of further constant variable quantity	Quantity of further second variable quantity	Quantity of originals variable quantity eradicated from construction
Example 1	5	206	206	1
Example 2	2	4	4	1
Example 3	4	201	201	1

II. Example

Minimize $\quad x_1^{0.5}x_2 + 3lnx_1,$

Subject $\quad -x_1 + x_2 \leq 5,$
$x_1^{0.5} - x_2 \leq 6,$
$x_1 \in \{0.1, 0.5, 0.7, 1.2\}, -6 \leq x_2 \leq 4.$

The nonlinearity relationships $x_1^{0.5}x_2, 3lnx_1,$ and $x_1^{0.5}$ are noncurved roles. By Remarks 4 and 5,

Minimize $\quad 0.1^{0.5}s_1 + 0.5^{0.5}s_2 + 0.7^{0.5}s_3 + 1.2^{0.5}s_4 + 3(u_1ln0.1 + u_2ln0.5 + u_3ln0.7 + u_4ln1.2)$
Subject $\quad -0.1u_1 - 0.5u_2 - 0.7u_3 - 1.2u_4 + x_2 \leq 5,$
$u_1 + u_2 + u_3 + u_4 = 1,$
$0.1^{0.5}u_1 + 0.5^{0.5}u_2 + 0.7^{0.5}u_3 + 1.2^{0.5}u_4 - x_2 \leq 6,$
$-6u_i \leq s_i \leq 6u_i, 6(u_i - 1) + x_2 \leq s_i \leq 6(1 - u_i) + x_2, i = 1,2,3,4, s_1, s_2, s_3, s_4$
are unrestricted in sign variables, $u_1, u_2, u_3, u_4 \in \{0,1\}, -6 \leq x_2 \leq 4.$

The converted sequencer can be resolved by LINGO [24] to find the universally optimum explanation $(x_1, x_2) = (0.2, -5.753)$ and the empirical charge −8.705 contained by the optimality acceptance0.001 as the probability acceptance 0.001.

III. Example

Minimize $\quad x_1x_4^3 - x_3 - 0.5x_1^2 x_2^4$
Subject $\quad x_1x_4^{1.5} - x_2 - x_2^{0.5} x_3^{0.4} \leq 4,$
$-x_1 - 2x_2 + x_3 \leq -2,$
$0 \leq x_1 \leq 6, \; 1 \leq x_2 \leq 10, \; 1 \leq x_3 \leq 6, 20 \leq x_4 \leq 30, x_1, x_2, x_3, x_4 \in R$

The nonlinearity relationships $x_1x_4^3$, $x_1^2 x_2$ four and $x_1x_4^{1.5}$ where x_1 has zero-value smaller constrained Table 1 listing the quantity of variable quantity applied in the converted standard of Instance 3. Though the intended system needs the accumulation of variable another amount, variable dual amount, and restrictions, it can prevent imprecision by requiring a little $\epsilon > 0$ smaller bound for x_1. Explaining this system by the planned technique with LINGO [24], the universally optimal explanation achieved is smaller, particularly for $x1$. Demonstrating this system by the intended design and LINGO [24], the universally optimum description achieved is$(x_{1,}\ x_{2,}x_3, x_{4,}) =$ $(0, 4, 6, 20)$and the empirical rate is −6. Still, explaining this system by really requiring $x_1 \geq 0.001$, the universally optimal explanation achieved is$(x_{1,}\ x_{2,}x_3, x_{4,}) =\ (0.001, 10, 6, 20)$, and the empirical rate is 1.995.

V. Conclusions

This analysis intends an optimized technique to discuss a signal CNNLCP difficulty and set the free variable quantity to achieve a comprehensive optimum. The free variable quantity practical trading techniques change over the variable amount and translate the analytical association among the variable quantity in a result period into a set of straight inequities, appropriately combined into the CNNLCP types. Several valuable instructions to essentially Convexness more than universal signal conditions in CNNLCP systems are also produced. Numerical illustrations are demonstrated to provide for the impacts of the recommended system.

References

[1] Adjiman, C.S., Androulakis, I.P., and Floudas, C.A. (1997). Global Optimize of CNNLCP troubles in process synthesis and design. Compute. Chem. Eng. 21, S445–S450.

[2] Adjiman, C.S., Androulakis, I.P., and Floudas, C.A. (2000). Global Optimize of Combined-numeral nonlinearity troubles. AIChE J. 46(9), 1769–1797.

[3] Adjiman, C.S., Androulakis, I.P., Maranas, C.D.,and Floudas, C.A.(1996). A global optimize method α BB for process design. Compute. Chem. Eng. 20, S419–S424.

[4] Aggarwal, A., and Floudas, C.A. (1990). Synthesis of general separation sequences-no sharp separations. Compute. Chem. Eng. 14(6), 631–653.

[5] Beale, E.M.L., Forrest, and J.J.H. (1976). Global optimize using unique ordered sets. Math. Prog. 10, 52–69.

[6] Beale, E.L.M., Tomlin, and J.A. (1970). Exceptional facilities in a general mathematical coding system for noncurved troubles using ordered sets of variables. In: Lawrence, J. (ed.) Proceedings of the Fifth International Conference on Operations Reexamine, pp. 447–454.

[7] Biegler, L.T., Grossmann, and I.E. (2004). Retrospective on Optimization. Compute. Chem. Eng. 28, 1169–1192.

[8] Björk, K.M., Lindberg, P.O., and Wester Lund.T.(2003). Some convexification in global Optimize of troubles containing signal relationships. Compute. Chem. Eng. 27, 669–679.

[9] Borchers, B., and Mitchell, J.E. (1994). An improved branch and bound algorithm for Combined-numeral nonlinearity programs. Compute. Opera. Res. 21(4), 359–367.

[10] Duran, M., and Grossmann, I.E. (1986). An outer-approximation algorithm for a class of Combined-numeral nonlinearity programs. Math. Prog**.** 36**,** 307–339.

[11] Fletcher, R., and Leyffer,. S. (1994). Explaining Combined-numeral nonlinearity programs by outer approximation. Math. Rog. **66**, 327–349.

[12] Floudas, C.A. Determinist Global Optimization: Theory, Techniques, and Application. Kluwer Academic Publishers, Boston 2000.

[13] Floudas, C.A., Akrotirianakis, I.G., Caratzoulas, S., Meyer, C.A., and Kallrath, J. (2005). Global Optimize in the 21st century: advances and challenges. Compute. Chem. Eng. 29, 1185–1202.

[14] Floudas, C.A., and Pardalos, P.M. State of the Art in Global Optimization: Computational Techniques and Applications. Kluwer Academic Publishers, Boston 1996.

[15] Geoffrion, A.M. (1972). Generalized benders decomposition. J. Opt. Theory Appl. 10, 237–260.

[16] Glover, F., Laguna, M.: Tabu Examine. Kluwer Academic Publishers, Boston 1997.

[17] Grossmann, I.E. (2002). Review of nonlinearity Combined-numeral and disjunctive coding techniques. Opt. Eng. 3, 227–252.
[18] Grossmann, I.E., and Biegler, L.T. (2004). Part II. A future perspective on Optimization. Compute. Chem. Eng. 28, 1193–1218.
[19] Hussain, M.F., and Al-Sultan, K.S. (1997). A hybrid genetic algorithm for noncurved function minimization. J. Global Opt. 11, 313–324.
[20] Kokossis, A.C., and Floudas, C.A. (1994). Optimize of complex reactor networks-II: no isothermal operation. Hem. Eng. Sci. 49(7), 1037–1051.
[21] Lee, S., Grossmann, I.E. (2000). New algorithms for nonlinearity generalized disjunctive coding. Compute. Chem. Eng. 24, 2125–2142.
[22] Leyffer, S. (2001). Integrating sap and branch-and-bound for Combined numeral nonlinearity coding. CompStat. Opt. Appl. 18, 295–309.
[23] Li, H.L., and Tsai, J.F. (2005). Treating free variables in generalized geometric global optimize programs. J. Global Opt. 33, 1–13.
[24] LINGO Release 9.0. LINDO System Inc., Chicago 2004.
[25] Maranas, C.D., and Floudas, C.A. (1995). Finding all solutions of nonlinearity-constrained systems of equations. J. Global Opt. 7(2), 143–182.
[26] Maranas, C.D., and Floudas, C.A. (1997). Global optimize generalized geometric coding. Compute. Chem. Eng. 21(4), 351–369.
[27] Martin, A., Moller, M., and Moritz, S. (2006). Combined-numeral models for the stationary case of gas network optimization. Math. Prog. 105, 563–582.
[28] McDonald, C.M., and Floudas, C.A. (1995). Global optimize for the phase and chemical equilibrium problem: application to the NRTL equation. Compute. Chem. Eng. 19(11), 1111–1141.
[29] Papalexandri, K.P., Pistikopoulos, E.N., and Floudas, C.A. (1994). Mass-exchange networks for waste minimization– a simultaneous approach. Chem. Eng. Res. Design 72(A3), 279–294.
[30] Pörn, R., Harjunkoski, I., and Westerlund, T. (1999). convexification of different classes of noncurved CNNLCP troubles. Compute. Chem. Eng. 23, 439–448.
[31] Quesada, I., and Grossmann, I.E. (1992). An LP/NLP-based branch and bound algorithm for curved CNNLCP optimize troubles. Compute. Chem. Eng. 16, 937–947.
[32] Salcedo, R.L., Goncalves, M.J., and Feyo de Azevedo, S. (1990). An improved arbitrary - examine algorithm for nonlinearity Optimization. Compute. Chem. Eng. 14, 1111–1126.
[33] Stubbs, R.A., and Mehrotra, S. (1999). A branch-and-cut method for 0–1 Combined curved coding. Math. Rog. 86, 515–532.
[34] Sun, X.L., McKinnon, and K.I.M., Li, D. (2001). A convexification method for global optimizes troubles with applications to reliability optimization. J. Global Opt. 21, 185–199.
[35] Tomlin, J.A. 1981). A suggested extension of unique ordered sets to non-separable noncurved coding. Troubles. In: Hansen, P. (ed.) Studies on Graphs and Discrete Coding, pp. 359–370.
[36] Tsai, J.F., Li, H.L., and Hu, N.Z. (2002). Global Optimize for signal discrete coding troubles in engineering design. Eng. Opt. 34, 613–622.
[37] Westerlund, T., and Pettersson, F. (1995). An extended cutting plane method for explaining curved CNNLCP troubles. Compute. Chem. Eng. 19, S131–S136.
[38] Wu, Z.Y., Bai, F.S., and Zhang, L.S. (2005). convexification and non-ratification for a general class of global optimize troubles. J. Global Opt. 31, 45–60.
[39] Yiu, K.F.C., Liu, Y., and Teo, K.L. (2004) A hybrid descent method for global Optimization. J. Global Opt. 28,229–23.

RT&A No 2 (68)
Volume 17, June 2022

SELECTION OF SINGLE SAMPLING PLANS BY VARIABLES BASED ON GENERALIZED BETA DISTRIBUTION

R. Vijayaraghavan[a] and A. Pavithra[b]

•

Department of Statistics, Bharathiar University,
Coimbatore 641 046, Tamil Nadu, INDIA
[a]vijaystatbu@gmail.com, [b]pavistat95@gmail.com

Abstract

Statistical quality control (SQC) has wider applications in industries and production engineering. Product control, one of the two major categories of SQC, consists in procedures by which decisions are made on the disposition of one or more lots of finished items or materials produced by manufacturing industries. Sampling inspection by variables in product is the methodology that is employed for deciding about the disposition of a lot of individual units based on the observed measurements on a quality characteristic of randomly sampled units from the lot submitted for inspection. These procedures are defined under the assumption that the quality characteristic is measurable on a continuous scale and the functional form of the probability distribution must be known. Inspection procedures which have been developed based on the implicit assumption that the quality characteristic is distributed as normal with the related properties are found in the literature of sampling inspection procedures. The assumption of normality may not be realized often in practice and it becomes inevitable to investigate the properties of variable sampling plans based on non-normal distributions. In this paper a single sampling plan by variables is formulated and evaluated under the assumption that the quality characteristic is distributed according to a generalized beta distribution of first kind. Procedures are developed for determining the parameters of the proposed plan for specified requirements in terms of producer's and consumer's protection.

Key Words: Consumer's Quality Level, Generalized Beta Distribution, Normal Distribution, Operating Characteristic Function, Single Sampling Plan, Producer's Quality Level.

1. Introduction

Sampling inspection is an activity for taking decisions on one or more lots of finished products which have been submitted for inspection. The decision of either acceptance or rejection of the lots is usually taken by adopting suitable sampling inspection procedures, called sampling plans. Sampling plans are generally categorized into two types, namely, lot-by-lot sampling by attributes and lot-by-lot sampling by variables. In lot-by-lot inspection by attributes, one or more samples of items are drawn from a given lot of manufactured items; each item in the sample(s) is classified as conforming or nonconforming; and the decision of acceptance or rejection of the lot is made based on a specific rule. In lot-by-lot inspection by variables, one or more samples of items are drawn from a given lot; the measurement of a quality characteristic in each sampled item is recorded; and the decision of acceptance or rejection of the lot is made as a function of such measurements. The theory of inspection by variables is applicable when the quality characteristic of sampled items is measurable on a continuous scale and the functional form of the probability distribution is assumed to be known. A variables sampling is advantageous in the sense that it generates more information from each item inspected, requires small sample and provides same protection when compared to attributes sampling. See, [1] and [2].

On the basis of the implicit assumption that the quality characteristic is distributed according to normal with mean μ and standard deviation σ, the concept of variables sampling inspection has been studied by many researchers. Some of the early works on variables sampling inspection are seen in [3], [4], [5], [6] and [7]. Studies relating to sampling plans when the assumption of normality of the quality characteristic fails or the functional form of the underlying distribution deviates from normal or the form of the distribution is not known are also found in the literature of acceptance sampling. [8] – [23] are few references which deal with variables inspection using non-normal distributions.

The problem of designing single sampling plans by variables, when the quality characteristic, *X*, follows a normal distribution with mean μ and standard deviation σ, has been addressed in the past. See, [24]. In the industrial situations, quite often, the assumption of normality may not be valid or the quality characteristic may be distributed according to non-normal distributions. In such cases, the selection of variable sampling plans becomes complicated. However, the literature of acceptance sampling provides procedures for the designing of variables plans when the quality characteristic follows a probability distribution other than normal. A detailed survey on various works related to variable sampling plans with emphasis on non-normality is given in [11]. A computer-aided procedure has been developed in [25] for the identification of the appropriate distribution in designing sampling inspection plans by variables when the quality characteristics are defined by compositional proportions.

A generalized probability density function, termed as double bounded probability density function has been derived in [26]. It is also called a generalized beta distribution of first kind, in which the random variable *X* is defined within the range (0, 1). Practical applications of variables sampling plans using a generalized beta distribution can be visualized for bulk product inspection where the quality characteristics are the compositional proportions, such as proportion of binary mixtures of pharmaceutical powder, percentage of protein in milk powder, fatty acid composition of serum lipid fractions, etc. Sampling inspection plans for compositional fractions based on the beta distribution and the procedure for designing the plans to control the proportion nonconforming levels are discussed in [27].

In this paper, a study on single sampling plans by variables is formulated under the assumption that the quality characteristic is assumed to have a generalized beta distribution which would be appropriate in situations where the quality characteristics are compositional fractions. A procedure for determining the parameters of the proposed plan for specified requirements in terms of producer's and consumer's protection is also developed.

2. Single Sampling Inspection Plans by Variables

A single sampling inspection plan by variables is defined under the following assumptions:

(a) The quality characteristic, denoted by X, is measurable on a continuous scale and has a known form of probability distribution, represented by $F_X(x;\mu,\sigma)$, which is the distribution function of X with mean μ and variance σ^2.

(b) Each individual unit in a submitted lot has a one-sided specification, say, lower specification, L or upper specification, U. If, for a unit, $X > U$ (or $X < L$), the unit is classified as a non-conforming unit.

The operating procedure of a variable sampling plan is as follows:

Step 1: Draw a random sample of n units from a lot and observe the measurements, $x_1, x_2, \ldots, x_n$ of the quality characteristic, X.

Step 2: When σ is known, accept the lot if $\bar{x} + k\sigma \le U$ (or $\bar{x} - k\sigma \ge L$); otherwise, reject the lot, where $\bar{x}$ is the sample mean.

When σ is unknown, accept the lot, if $\bar{x} + ks \le U$ (or $\bar{x} - ks \ge L$); otherwise, reject the lot. Here, $s^2 = \dfrac{1}{n-1}\sum_{i:1}^{n}(x_i - \bar{x})^2$ is an unbiased estimate of σ^2.

Thus, a single sampling plan by variables is designated by two parameters, namely, the sample size, n, and the acceptability constant, k. When these parameters are known, the plan could be implemented. The explicit expressions for n and k can be derived by specifying two points on the operating characteristic curve of the plan, namely, $(p_1, 1-\alpha)$ and (p_2, β), where p_1 and p_2 are termed as producer's quality level (*PQL*) and the consumer's quality level (*CQL*), associated with the producer's risk, α and the consumer's risk, β, respectively. A sampling plan by variables is termed as a known σ or unknown σ plan according as σ is known or unknown.

3. Operating Characteristic Function

An important measure of performance of a variables sampling plan is its operating characteristic function, which is a function of the proportion, *p*, of non-conforming units, called incoming lot quality, and provides the probability, $P_a(p)$, of acceptance of a lot. The plot of $P_a(p)$ against *p* results in a curve, called operating characteristic (OC) curve. For a given upper specification limit, *U*, when σ is known, p and $P_a(p)$ are defined by

$$p = P(X > U|\mu) \tag{1}$$

and $P_a(p) = P(\bar{x} + k\sigma \le U|\mu)$. (2)

PQL and *CQL*, using (1), are defined by

$$PQL = p_1 = P(X > U|\mu_1) \tag{3}$$

and $$CQL = p_2 = P(X > U|\mu_2), \tag{4}$$

where μ_1 and μ_2 are the means of the underlying distribution which results in *PQL* and *CQL*, respectively.

Assume that the random variable, *X*, is modeled by a two-parameter generalized beta distribution. The probability density function and the cumulative distribution function of the generalized beta distribution, according to [26], are respectively given by

$$f(x;a,b) = abx^{a+1}(1-x^a)^{b-1},\ 0 < x < 1 \tag{5}$$

and $$F(x) = 1-(1-x^a)^b, \tag{6}$$

where a > 0 and b > 0 are the shape parameters.

The moments about origin of the distribution are defined by

$$m_r = \frac{b\Gamma\left(1+\frac{r}{a}\right)\Gamma(b)}{\Gamma\left(1+b+\frac{r}{a}\right)}, \; r = 1, 2, 3, 4, \cdots,$$

from which the measures such as mean, variance, skewness and kurtosis can be derived as $\mu = m_1$, $\sigma^2 = \mu_2$, $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ and $\beta_2 = \frac{\mu_4}{\mu_2^2}$, where $\mu_2 = m_2 - m_1^2$, $\mu_3 = m_3 - 3m_2m_1^2 + 2m_1^3$, and $\mu_3 = m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4$.

From (1), (3) and (4), the lot quality levels, *p*, *PQL* and *CL* using standardized beta distribution are defined, respectively, by

$$p = P(T > K_p^*),$$

$$PQL = p_1 = P(T > K_{p_1}^*) \tag{7}$$

and $$CQL = p_2 = P(T > K_{p_2}^*), \tag{8}$$

where $T = \frac{X-\mu}{\sigma}$, $K_p^* = \frac{U-\mu \mid p}{\sigma}$, $K_{p_1}^* = \frac{U-\mu \mid p_1}{\sigma}$ and $K_{p_2}^* = \frac{U-\mu \mid p_2}{\sigma}$.

The producer's risk, α, and the consumer's risk, β, corresponding to *AQL* and *LQL* are, respectively, defined from (2) as

$$\alpha = P(rejecting\, the\, lot \mid \mu = \mu_1) = P(\bar{x} + k\sigma > U \big| \mu = \mu_1) \tag{9}$$

and $$1-\beta = P(rejecting\, the\, lot \mid \mu = \mu_2) = P(\bar{x} + k\sigma > U \big| \mu = \mu_2). \tag{10}$$

When σ is unknown, the estimate *s* is used in the decision criterion, and hence in the evaluation of α and β.

4. Designing Single Sampling Plans by Variables

In the industrial practice, the unknown standard deviation variables plans are more realistic than the known standard deviation variables plans. If the distribution is non-normal, the designing of unknown σ plans is rather complicated. Such problems introducing an expansion factor in terms of measures of skewness and kurtosis are addressed in [12], which also provides a methodology for determining the parameters of sampling plans by variables under the conditions of non-normal populations using the expansion factor. The procedures for the selection of unknown standard deviation sampling plans are provided in [23] giving protection to the producer and consumer under the assumption that the quality characteristics under study follow a Pareto distribution when the measures of skewness and /or kurtosis are specified.

4.1. Case of Unknown Sigma

The methodology proposed in [12] using the expansion factor will, now, be discussed for an unknown sigma plan by variables under the assumption of generalized beta distribution for the quality characteristic, *X*.

In the case of unknown sigma plan, the determination of n and k is usually based on the sampling distribution of $\bar{x}+ks$ or $\bar{x}-ks$. It is known that under the assumption of normal distribution, $\bar{x}$ and s are independent and distributed as normal. Therefore, $\bar{x}+ks$ and $\bar{x}-ks$ are normally distributed. Using these properties, formulae for finding the values of n and k can be obtained. The asymptotic distributions of $\bar{x}+ks$ and $\bar{x}-ks$ are shown to be normal having the means $\mu_y = \mu + k\sigma$ and $\mu_y = \mu - k\sigma$, respectively, and the common variance given by

$$\sigma_Y^2 = \frac{\sigma^2}{n}\left[1+\frac{k^2}{4}(\beta_2-1)\pm k\beta_1\right], \tag{11}$$

where β_1 and β_2 represent the measures of skewness and kurtosis of the underlying distribution. Having defined $Z_p^* = \frac{U-\mu \mid p}{\sigma}$ and acceptance probability function for the case of unknown standard deviation as $P_a(p) = P_r[\bar{x}+k_U s \le U \mid p] = P_r(Y \le U \mid p)$, from [12], the expressions for Z_α, Z_β, Z_{PQL}^* and Z_{CQL}^* corresponding to *α*, *β*, *PQL* and *CQL*, respectively, are as given below:

$$Z_\alpha = \frac{U-(\mu \mid p_1 + k_U\sigma)}{\sigma\sqrt{\frac{e_U}{n_U}}} \tag{12}$$

$$-Z_\beta = \frac{U-(\mu \mid p_2 + k_U\sigma)}{\sigma\sqrt{\frac{e_U}{n_U}}} \tag{13}$$

$$Z_{PQL}^* = k_U + K_\alpha\sqrt{\frac{e_U}{n_U}} \tag{14}$$

$$Z_{CQL}^* = k_U - K_\beta\sqrt{\frac{e_U}{n_U}}\,. \tag{15}$$

Here, $e_U = 1+\frac{k^2}{4}(\beta_2-1)+k\beta_1$ is the expansion factor, which can be used to obtain the known standard deviation plans. When the requirements are specified in terms of the points $(PQL, 1-\alpha)$ and (CQL, β) on the OC curve such that $P_a(PQL) = 1-\alpha$ and $P_a(CQL) = \beta$, the expressions for the plan parameters *n* and *k*, derived from (14) and (15), are as given below:

$$n_U = e_U\left[\frac{Z_\alpha + Z_\beta}{Z_{PQL}^* - Z_{LQL}^*}\right]^2 \tag{16}$$

and
$$k_U = \frac{Z_\alpha Z_{CQL}^* + Z_\beta Z_{PQL}^*}{Z_\alpha + Z_\beta}. \tag{17}$$

In a similar way, when the lower specification limit, *L*, is specified, the expressions for n and k can be derived.

4.2. Numerical Illustration

Suppose that a set of measurements yields $\beta_1 = 0.0377$ and $\beta_2 = 2.0147$ It is desired to determine a variables sampling plan giving protection to the producer and the consumer in terms of $(PQL = 0.01, \alpha = 0.05)$ and $(CQL = 0.06, \beta = 0.10)$. For the given requirements, the values of *a* and b are found as a = 0.65 and b = 0.80. Associated with these values are the mean and standard deviation given by M = 0.5581 and S = 0.2545, respectively. Corresponding to PQL and CQL, the values of Z^*_{PQL} and Z^*_{CQL} are determined from (7) and (8) as 1.6717 and 1.4619, respectively. The normal deviates Z_α and Z_β are obtained as 1.645 and 1.282 by satisfying (9) and (10) for the specified sets of values of α and β. Substituting these values in (16) and (17), the parameters of the desired plan are determined as $n = 195$ and $k = 1.554$. The value of e_U is obtained as 1.67119. Thus, the parameters of a known standard deviation plan, are computed as $n'_U = \frac{n_U}{e_U} = 117$ and $k'_U = k_U = 1.554$.

4.3. Case of Known Sigma

The method of designing known sigma variables sampling plan under the assumption of Burr distribution utilizing the measures skewness and kurtosis is proposed in [15]. A similar procedure is developed here for the known sigma plan by variables when the underlying distribution is a two-parameter generalized beta distribution.

Let M and S be the mean and standard deviation of the two-parameter generalized beta distribution. Then, PQL and CQL are defined by

$$PQL = 1 - F(x_{PQL}; M, S) = (1 - x^a_{PQL})^b, a > 0, b > 0 \tag{18}$$

and $$CQL = 1 - F(x_{CQL}; M, S) = (1 - x^a_{CQL})^b, a > 0, b > 0. \tag{19}$$

where $$\frac{U - \mu|PQL}{\sigma} = \frac{x_{PQL} - M}{S} = Z^*_{PQL} \tag{20}$$

and $$\frac{U - \mu|CQL}{\sigma} = \frac{x_{CQL} - M}{S} = Z^*_{CQL} \tag{21}$$

with Z^*_{PQL} and Z^*_{CQL} being the standardized values of x corresponding to PQL and CQL, respectively.

Assuming that the distribution of $\bar{x}$ is normal, α and β are defined as area under normal curve and are expressed by

$$\alpha = \int_{k_\alpha}^{\infty} \varphi(t)\,dt \tag{22}$$

and $$1 - \beta = \int_{k_{1-\beta}}^{\infty} \varphi(t)\,dt, \tag{23}$$

where $$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}, \tag{24}$$

$$Z_\alpha = \sqrt{n}\frac{K - \mu_1}{\sigma} \tag{25}$$

and $$Z_{1-\beta} = \sqrt{n}\frac{K - \mu_2}{\sigma}. \tag{26}$$

From equations (22) to (26), the expressions for n and k are, respectively, obtained as

$$n_U = \left(\frac{Z_\alpha + Z_\beta}{Z^*_{PQL} - Z^*_{CQL}}\right)^2 = \left(\frac{S[Z_\alpha + Z_\beta]}{x_{PQL} - x_{CQL}}\right)^2 \tag{27}$$

and $$K = U - \frac{\sigma}{S}\left[\frac{Z_\alpha x_{CQL} + Z_\beta x_{PQL}}{Z_\alpha + Z_\beta} - M\right]. \tag{28}$$

If the acceptance criterion is written as $\bar{x} + k\sigma \le U$, according to [15], the expression for k is given by

$$k_U = \left[\frac{Z_\alpha x_{CQL} + Z_\beta x_{PQL}}{S(Z_\alpha + Z_\beta)} - \frac{M}{S}\right]. \qquad (29)$$

5. Determination of n and k of a Variables Sampling Plan

The parameters of a sampling plan by variables can be derived from the generalized beta distribution when the third and fourth moments of the distribution of measurements are known or specified. It is known that the measures of skewness and kurtosis, specified by β_1 and β_2, of a generalized beta distribution are functions of the shape parameters *a* and *b*. Thus, for a specified values of β_1 and β_2, the values of a and *b* can be determined. In order to determine the required sampling plan by variables, the following procedure is followed:

Step 1: Specify β_1 and β_2.

Step 2: Specify the desired protection in terms of $(p_1, 1-\alpha)$ and (p_2, β).

Step 3: Choose the value a and *b* from Table 2 corresponding to the specified values of β_1 and β_2.

Step 4: For specified p_1 and p_2, determine x_{PQL} and x_{CQL} from $F_X(x)$, which is the cumulative distribution function of the generalized beta distribution, satisfying the equations (18) and (19), and obtain Z^*_{PQL} and Z^*_{CQL} from equations (20) and (21).

Step 5: For specified α and β, determine the normal deviates Z_α and Z_β, satisfying the equations (22) and (23).

Step 6: Determine the parameters n_σ and k_σ of the plan as n_U and k_U using equation (27) and (29).

Based on the procedure described, the parameters, n_σ and k_σ, of the sampling plans by variables for a wide range of values of *PQL* and *CQL* are obtained and given in Table 3 for various combination of values of *a* and *b*. The parameters provided in the table yield the maximum producer's risk of 5% and the maximum consumer's risk of 10%. To facilitate the computation of Z^*_{PQL} and Z^*_{CQL}, the mean, *M*, and standard deviation, *S*, are obtained for sets of values of *a* and *b* and provided in Table 1.

5.1. Numerical Illustration

It is desired to have a single sampling plan by variables when the set of measurements drawn from a generalized beta distribution has the measure of skewness and kurtosis specified as $\beta_1 = 0.0654$ and $\beta_2 = 2.1384$. Suppose that the desired protection against an upper specification limit is specified in terms of $(PQL = 0.01, \alpha = 0.05)$ and $(CQL = 0.06, \beta = 0.10)$.

Table 2 yields $a = 0.750, b = 0.50, M = 0.4156$ and $S = 0.2363$ associated with $\beta_1 = 0.0654$ and $\beta_2 = 2.1384$. The values of x_{PQL} and x_{CQL} are determined from (18) and (19) as 0.924 and 0.8458 for the specified $PQL = 0.01$ and $CQL = 0.04$. The standardized deviates Z^*_{PQL} and Z^*_{CQL} are obtained as 2.1512 and 1.8206, respectively, from (20) and (21). The values of Z_α and Z_β are determined as 1.645 and 1.282. On substitution of these values in (27) and (29), the parameters of the desired plan are determined as $n_\sigma = n_U = 78,$ and $k_\sigma = 1.9654$. Table 3, when entered with the specified values of the quality levels, can be used to choose the parameters of the required plan corresponding to *a* = 0.75 and *b* = 0.50.

Table 1: Mean, M, and Standard Deviation, S of Generalized Beta Distribution

a	*b*								
	0.200	0.250	0.300	0.350	0.400	0.450	0.500	0.550	
0.500	0.3694	0.4063	0.4382	0.4661	0.4909	0.5132	0.5333	0.5517	M
	0.1738	0.1868	0.1969	0.205	0.2116	0.2168	0.2211	0.2246	S
0.550	0.3387	0.3759	0.4082	0.4367	0.4622	0.4852	0.506	0.5251	M
	0.1727	0.1871	0.1985	0.2076	0.2151	0.2212	0.2262	0.2303	S
0.600	0.3111	0.3484	0.381	0.4099	0.4359	0.4594	0.4808	0.5004	M
	0.1707	0.1863	0.1989	0.2091	0.2175	0.2244	0.2301	0.2349	S
0.650	0.2862	0.3234	0.3561	0.3853	0.4116	0.4356	0.4574	0.4775	M
	0.168	0.1848	0.1985	0.2097	0.2189	0.2266	0.233	0.2384	S
0.700	0.2638	0.3007	0.3334	0.3627	0.3893	0.4135	0.4357	0.4562	M
	0.1648	0.1827	0.1973	0.2094	0.2195	0.2279	0.235	0.241	S
0.750	0.2435	0.28	0.3126	0.3419	0.3687	0.3931	0.4156	0.4364	M
	0.1612	0.1801	0.1957	0.2086	0.2195	0.2286	0.2363	0.2429	S
0.800	0.2251	0.2611	0.2934	0.3228	0.3495	0.3741	0.3968	0.4178	M
	0.1573	0.1771	0.1936	0.2073	0.2189	0.2287	0.2371	0.2442	S
0.850	0.2084	0.2438	0.2758	0.305	0.3318	0.3565	0.3793	0.4005	M
	0.1533	0.1739	0.1911	0.2056	0.2179	0.2284	0.2373	0.245	S
0.900	0.1932	0.2279	0.2596	0.2886	0.3153	0.34	0.363	0.3843	M
	0.1492	0.1705	0.1884	0.2036	0.2165	0.2276	0.2371	0.2453	S
0.950	0.1793	0.2134	0.2446	0.2734	0.3	0.3247	0.3477	0.3691	M
	0.145	0.1669	0.1855	0.2013	0.2149	0.2265	0.2366	0.2452	S
1.000	0.1667	0.2	0.2308	0.2593	0.2857	0.3103	0.3333	0.3548	M
	0.1409	0.1633	0.1824	0.1988	0.213	0.2251	0.2357	0.2449	S
1.500	0.0852	0.1108	0.136	0.1604	0.1841	0.2068	0.2286	0.2494	M
	0.1029	0.1276	0.15	0.1701	0.1881	0.2042	0.2185	0.2313	S
2.000	0.0476	0.0667	0.0865	0.1068	0.127	0.147	0.1667	0.1859	M
	0.0753	0.0992	0.1219	0.1432	0.1628	0.1808	0.1972	0.2122	S
2.500	0.0284	0.0426	0.0583	0.0749	0.092	0.1095	0.127	0.1444	M
	0.0562	0.0782	0.1001	0.1212	0.1412	0.16	0.1775	0.1936	S
3.000	0.0179	0.0286	0.041	0.0547	0.0693	0.0844	0.1	0.1157	M
	0.043	0.0628	0.0833	0.1037	0.1235	0.1425	0.1604	0.1772	S

Table 1 (Continued)

a	*b*							
	0.600	0.650	0.700	0.750	0.800	0.850	0.900	
0.500	0.5686	0.5841	0.5985	0.6119	0.6243	0.6359	0.6468	M
	0.2274	0.2296	0.2314	0.2328	0.2339	0.2347	0.2352	S
0.550	0.5426	0.5588	0.5738	0.5877	0.6007	0.6129	0.6243	M
	0.2337	0.2365	0.2387	0.2405	0.2419	0.243	0.2439	S
0.600	0.5185	0.5352	0.5507	0.5652	0.5787	0.5914	0.6032	M
	0.2388	0.2421	0.2448	0.247	0.2488	0.2502	0.2513	S
0.650	0.496	0.5132	0.5292	0.5441	0.5581	0.5712	0.5835	M
	0.2429	0.2466	0.2498	0.2524	0.2545	0.2563	0.2577	S
0.700	0.4751	0.4927	0.5091	0.5244	0.5388	0.5523	0.565	M
	0.246	0.2503	0.2539	0.2569	0.2594	0.2615	0.2633	S
0.750	0.4556	0.4736	0.4903	0.506	0.5207	0.5345	0.5475	M
	0.2485	0.2532	0.2572	0.2607	0.2635	0.266	0.268	S
0.800	0.4374	0.4556	0.4726	0.4886	0.5036	0.5177	0.531	M
	0.2503	0.2555	0.2599	0.2637	0.267	0.2697	0.2721	S
0.850	0.4203	0.4387	0.456	0.4723	0.4875	0.5019	0.5155	M
	0.2515	0.2572	0.2621	0.2663	0.2698	0.2729	0.2755	S
0.900	0.4043	0.4229	0.4404	0.4569	0.4723	0.487	0.5008	M
	0.2524	0.2584	0.2637	0.2683	0.2722	0.2756	0.2785	S
0.950	0.3892	0.408	0.4257	0.4423	0.458	0.4728	0.4869	M
	0.2528	0.2593	0.2649	0.2698	0.2741	0.2778	0.281	S
1.000	0.375	0.3939	0.4118	0.4286	0.4444	0.4595	0.4737	M
	0.2528	0.2598	0.2658	0.2711	0.2756	0.2796	0.2831	S
1.500	0.2693	0.2884	0.3066	0.3239	0.3405	0.3564	0.3716	M
	0.2428	0.253	0.2622	0.2704	0.2777	0.2842	0.2901	S
2.000	0.2045	0.2227	0.2402	0.2571	0.2735	0.2893	0.3045	M
	0.2258	0.2382	0.2494	0.2596	0.2689	0.2773	0.2849	S
2.500	0.1616	0.1785	0.1951	0.2113	0.227	0.2423	0.2572	M
	0.2086	0.2223	0.2349	0.2465	0.2572	0.2669	0.2759	S
3.000	0.1315	0.1472	0.1627	0.178	0.1931	0.2078	0.2222	M
	0.1928	0.2074	0.221	0.2336	0.2452	0.2559	0.2658	S

Table 2: Measures, β_1 and β_2, of Skewness and Kurtosis Generalized Beta Distribution

a	*b*								
	0.200	0.250	0.300	0.350	0.400	0.450	0.500	0.550	
0.500	0.0676	0.0342	0.0133	0.0025	0.0001	0.0049	0.0157	0.0316	β_1
	2.4664	2.3723	2.3018	2.2505	2.2149	2.192	2.18	2.1771	β_2
0.550	0.1318	0.0804	0.0441	0.02	0.0059	0.0003	0.0017	0.0091	β_1
	2.5366	2.4147	2.3212	2.2503	2.1977	2.1601	2.135	2.1204	β_2
0.600	0.2138	0.1431	0.0904	0.0522	0.026	0.0096	0.0015	0.0003	β_1
	2.6341	2.4822	2.3641	2.2727	2.2026	2.1499	2.1115	2.085	β_2
0.650	0.3122	0.2205	0.1503	0.0974	0.0585	0.031	0.0131	0.0032	β_1
	2.7563	2.572	2.4276	2.3145	2.2262	2.1579	2.1059	2.0675	β_2
0.700	0.426	0.3114	0.2226	0.154	0.1018	0.0628	0.0348	0.016	β_1
	2.9013	2.6818	2.5094	2.3734	2.266	2.1815	2.1156	2.0651	β_2
0.750	0.5542	0.415	0.306	0.2209	0.1547	0.1037	0.0654	0.0374	β_1
	3.0679	2.8101	2.6076	2.4473	2.3199	2.2186	2.1384	2.0755	β_2
0.800	0.6965	0.5304	0.4	0.2972	0.2162	0.1528	0.1037	0.0664	β_1
	3.255	2.9558	2.721	2.5349	2.3864	2.2677	2.1726	2.0969	β_2
0.850	0.8523	0.6572	0.5036	0.3821	0.2857	0.2092	0.149	0.102	β_1
	3.4621	3.1178	2.8484	2.6349	2.4643	2.3273	2.2169	2.1282	β_2
0.900	1.0216	0.7949	0.6166	0.4751	0.3624	0.2723	0.2005	0.1436	β_1
	3.6887	3.2956	2.9889	2.7464	2.5525	2.3964	2.2702	2.1679	β_2
0.950	1.2042	0.9432	0.7384	0.5758	0.4458	0.3415	0.2577	0.1905	β_1
	3.9347	3.4886	3.142	2.8686	2.6502	2.4741	2.3314	2.2154	β_2
1.000	1.4	1.102	0.8687	0.6836	0.5355	0.4163	0.32	0.2422	β_1
	4.2	3.6964	3.3072	3.0009	2.7566	2.5598	2.4	2.2696	β_2
1.500	4.1297	3.2597	2.6111	2.1141	1.7247	1.4143	1.1633	0.9582	β_1
	7.9786	6.5824	5.5687	4.8073	4.2201	3.7575	3.3868	3.0856	β_2
2.000	8.531	6.5435	5.1512	4.1346	3.3675	2.7735	2.3038	1.9259	β_1
	14.2288	11.0692	8.9317	7.412	6.2892	5.4341	4.7668	4.2355	β_2
2.500	15.1716	11.1897	8.5727	6.7539	5.4343	4.4441	3.6804	3.0784	β_1
	23.8586	17.5151	13.5161	10.8234	8.9175	7.515	6.4505	5.622	β_2
3.000	24.8095	17.5104	13.0029	10.0171	7.9316	6.4136	5.2717	4.3895	β_1
	38.094	26.3968	19.5076	15.1041	12.1123	9.9818	8.4073	7.2085	β_2

Table 2 (Continued)

a	*b*							
	0.600	0.650	0.700	0.750	0.800	0.850	0.900	
0.500	0.052	0.0762	0.1037	0.1342	0.1672	0.2026	0.24	β_1
	2.182	2.1934	2.2107	2.2329	2.2595	2.2899	2.3237	β_2
0.550	0.0216	0.0386	0.0594	0.0835	0.1106	0.1403	0.1723	β_1
	2.1145	2.1162	2.1244	2.1382	2.1569	2.1798	2.2065	β_2
0.600	0.005	0.0148	0.029	0.0469	0.0682	0.0925	0.1193	β_1
	2.0686	2.0606	2.06	2.0656	2.0766	2.0924	2.1124	β_2
0.650	0	0.0026	0.0102	0.0221	0.0377	0.0566	0.0784	β_1
	2.0404	2.0229	2.0135	2.0111	2.0147	2.0236	2.0371	β_2
0.700	0.0048	0.0002	0.0013	0.0071	0.0171	0.0309	0.0478	β_1
	2.0272	2.0001	1.982	1.9717	1.9681	1.9702	1.9775	β_2
0.750	0.0182	0.0063	0.0007	0.0005	0.005	0.0136	0.0258	β_1
	2.0267	1.9899	1.9632	1.945	1.9342	1.9298	1.9309	β_2
0.800	0.0389	0.0196	0.0074	0.0012	0.0002	0.0037	0.0111	β_1
	2.0371	1.9906	1.9552	1.9292	1.9113	1.9003	1.8954	β_2
0.850	0.066	0.0393	0.0204	0.0081	0.0016	0.0001	0.0028	β_1
	2.0571	2.0006	1.9564	1.9226	1.8976	1.8802	1.8694	β_2
0.900	0.0989	0.0646	0.0389	0.0205	0.0085	0.0019	0	β_1
	2.0853	2.0188	1.9657	1.924	1.892	1.8682	1.8515	β_2
0.950	0.137	0.0948	0.0622	0.0377	0.0201	0.0085	0.002	β_1
	2.1209	2.0441	1.9821	1.9325	1.8934	1.8632	1.8407	β_2
1.000	0.1796	0.1295	0.0899	0.0592	0.036	0.0193	0.0082	β_1
	2.163	2.0758	2.0048	1.9471	1.9008	1.8642	1.836	β_2
1.500	0.7889	0.6482	0.5307	0.4322	0.3495	0.28	0.2217	β_1
	2.8381	2.6328	2.4614	2.3174	2.1959	2.0932	2.0062	β_2
2.000	1.6175	1.3629	1.1506	0.9721	0.8211	0.6927	0.5829	β_1
	3.8057	3.4531	3.1605	2.9155	2.7086	2.5329	2.3827	β_2
2.500	2.5949	2.2007	1.8751	1.6033	1.3743	1.1798	1.0136	β_1
	4.9637	4.4316	3.9953	3.633	3.3292	3.0721	2.8529	β_2
3.000	3.6928	3.1325	2.6749	2.2963	1.9795	1.7119	1.484	β_1
	6.273	5.528	4.9247	4.4289	4.0164	3.6697	3.3756	β_2

Table 3: Sample Size, n, and Acceptability Constant, k, of Single Sampling Plans by Variables Based on Generalized Beta Distribution Having a Maximum Producer's Risk of 5 Percent (α = 0.05) and a Maximum Consumer's Risk of 10 Percent (β = 0.10)

PQL	*CQL*	*b*	a								
			0.500	0.550	0.600	0.650	0.700	0.750	0.800	0.850	
0.005	0.01	0.200	244	209	180	155	135	118	104	91	*n*
			2.4205	2.5076	2.5939	2.6792	2.7637	2.8473	2.9298	3.0114	*k*
0.005	0.01	0.250	67	58	50	43	38	34	30	26	*n*
			2.2105	2.2836	2.3556	2.4264	2.4962	2.5648	2.6324	2.6988	*k*
0.005	0.02	0.300	89	77	67	59	51	45	40	36	*n*
			2.1311	2.2	2.2679	2.3346	2.4004	2.4651	2.5289	2.5916	*k*
0.006	0.03	0.350	70	60	53	47	41	36	33	29	*n*
			1.9813	2.041	2.0996	2.157	2.2134	2.2686	2.3228	2.376	*k*
0.007	0.04	0.400	63	55	49	43	38	34	30	27	*n*
			1.863	1.916	1.9678	2.0185	2.068	2.1165	2.1638	2.2102	*k*
0.008	0.05	0.450	61	54	48	42	38	34	30	27	*n*
			1.7649	1.8127	1.8593	1.9047	1.949	1.9922	2.0344	2.0756	*k*
0.01	0.04	0.500	142	125	111	98	88	78	70	64	*n*
			1.7406	1.7878	1.8338	1.8787	1.9226	1.9654	2.0073	2.0483	*k*
0.01	0.07	0.550	142	125	111	98	88	78	70	64	*n*
			1.7406	1.7878	1.8338	1.8787	1.9226	1.9654	2.0073	2.0483	*k*
0.02	0.09	0.600	75	67	60	54	48	44	40	36	*n*
			1.4919	1.5258	1.5584	1.5899	1.6203	1.6497	1.678	1.7054	*k*
0.02	0.03	0.650	2922	2585	2298	2052	1840	1656	1496	1356	*n*
			1.6059	1.6474	1.6878	1.7272	1.7657	1.8033	1.84	1.876	*k*
0.03	0.04	0.700	4520	4014	3582	3211	2889	2610	2366	2152	*n*
			1.5208	1.5578	1.5938	1.6288	1.6629	1.696	1.7283	1.7598	*k*
0.03	0.05	0.750	4520	4014	3582	3211	2889	2610	2366	2152	*n*
			1.5208	1.5578	1.5938	1.6288	1.6629	1.696	1.7283	1.7598	*k*
0.04	0.06	0.800	2002	1787	1603	1444	1305	1185	1079	986	*n*
			1.4037	1.4354	1.466	1.4957	1.5244	1.5523	1.5793	1.6056	*k*
0.04	0.08	0.850	634	567	510	460	417	379	346	317	*n*
			1.3458	1.3749	1.4029	1.4301	1.4563	1.4816	1.5061	1.5299	*k*
0.05	0.10	0.900	501	450	406	367	334	304	279	256	*n*
			1.2823	1.3085	1.3337	1.3579	1.3813	1.4038	1.4255	1.4464	*k*

Table 3 (Continued)

PQL	*CQL*	*b*	a							
			0.900	0.950	1.000	1.500	2.000	2.500	3.000	
0.005	0.01	0.200	81	72	64	24	12	7	4	*n*
			3.092	3.1716	3.2501	3.9699	4.5555	4.993	5.2813	*k*
0.005	0.01	0.250	23	21	19	8	4	2	1	*n*
			2.7642	2.8284	2.8915	3.4603	3.9127	4.2494	4.4768	*K*
0.005	0.02	0.300	32	29	26	10	5	3	2	*N*
			2.6534	2.7141	2.7739	3.3179	3.765	4.1177	4.3816	*K*
0.006	0.03	0.350	26	24	21	9	5	3	2	*n*
			2.4281	2.4792	2.5294	2.9786	3.3394	3.62	3.8287	*k*
0.007	0.04	0.400	25	22	20	9	5	3	2	*n*
			2.2555	2.2999	2.3433	2.7277	3.0319	3.2667	3.4418	*k*
0.008	0.05	0.450	25	22	20	9	5	3	2	*n*
			2.1157	2.155	2.1933	2.5298	2.793	2.9954	3.1469	*k*
0.01	0.04	0.500	58	52	48	21	12	7	5	*n*
			2.0883	2.1275	2.1657	2.5055	2.7778	2.9938	3.162	*k*
0.01	0.07	0.550	58	52	48	21	12	7	5	*n*
			2.0883	2.1275	2.1657	2.5055	2.7778	2.9938	3.162	*k*
0.02	0.09	0.600	33	30	28	14	8	5	4	*n*
			1.7319	1.7575	1.7822	1.9885	2.1343	2.2343	2.2987	*k*
0.02	0.03	0.650	1232	1124	1028	477	261	160	106	*n*
			1.9111	1.9455	1.9791	2.2801	2.5277	2.7323	2.9013	*k*
0.03	0.04	0.700	1963	1796	1648	788	443	278	189	*n*
			1.7904	1.8203	1.8494	2.1056	2.3092	2.4715	2.6002	*k*
0.03	0.05	0.750	1963	1796	1648	788	443	278	189	*n*
			1.7904	1.8203	1.8494	2.1056	2.3092	2.4715	2.6002	*k*
0.04	0.06	0.800	903	830	765	379	220	142	98	*n*
			1.631	1.6558	1.6798	1.8871	2.0463	2.169	2.2627	*k*
0.04	0.08	0.850	291	268	247	125	74	48	34	*n*
			1.5529	1.5751	1.5967	1.7808	1.919	2.023	2.1004	*k*
0.05	0.10	0.900	235	217	201	104	62	41	30	*n*
			1.4666	1.4861	1.5049	1.6621	1.7752	1.856	1.9121	*k*

6. Conclusion

The literature in statistical quality control provides various sampling inspection procedures which been developed based on the assumption that the quality characteristic under study follows a normal distribution. While such procedures are widely used in the industries, the departure from the assumption of normality or the violation of distributional assumptions are the major concern for the industrial practitioners as the decision that is made on the lot disposition in such situations would be inappropriate. Focusing on this vital aspect, in this paper, procedures for designing single sampling plans by variables are devised under the assumption that the quality characteristic is distributed according to a generalized beta distribution of first kind. The procedures and tables presented are appropriate for bulk inspection procedures where the quality characteristics are defined by compositional proportions.

7. Acknowledgments

The authors are grateful to Bharathiar University, Coimbatore for providing necessary facilities to carry out this research work. The second author is indebted to the Department of Science and Technology, India for awarding the DST-INSPIRE Fellowship under which the present research has been carried out.

References

[1] Bowker, A. H. and Goode, H. P. (1952), Sampling Inspection by Variables. McGraw-Hill, New York, Inc.
[2] Montgomery, D. C. (2004), Introduction to Statistical Quality Control, John Wiley & Sons, Inc., New York, USA.
[3] Lieberman, G. J. and Resnikoff, G. J. (1955), 'Sampling Plans for Inspection by Variables', *Journal of the American Statistical Association*, 50, pp. 457 - 516.
[4] Schilling, E. G. (1982), Acceptance Sampling in Quality Control, Marcel Dekker, New York, USA..
[5] Owen, D. B. (1966), 'One-Sided Variables Sampling Plans', *Industrial Quality Control*, 22, pp. 450 - 456.
[6] Owen, D. B. (1967), 'Variables Sampling Plans Based on the Normal Distribution', *Technometrics*, 9, pp. 417 - 423.
[7] Hamaker, H. C. (1979), 'Acceptance Sampling for Percent Defective by Variables and by Attributes', *Journal of Quality Technology*, 11, pp. 139 - 148.
[8] Duncan, A. J. (1986), Quality Control and Industrial Statistics, Richard D. Irwin, Homewood, Illinois, USA.
[9] Srivastava, A. B. L. (1961), 'Variables Sampling Inspection for Non-Normal Samples', *Journal of Science and Engineering Research*, 5, pp. 145 - 152.
[10] Das, N. G. and Mitra, S. K. (1964), 'The Effect of Non-Normality on Plans for Sampling Inspection by Variables', *Sankhya - A*, 26, pp.169 - 176.
[11] Owen, D. B. (1969), 'Summary of Recent Work on Variables Acceptance Sampling with Emphasis on Non-Normality', *Technometrics*, 11, pp. 631 - 637.
[12] Takagi, K. (1972), 'On Designing Unknown-Sigma Sampling Plans based on a Wide Class of Non-Normal Distributions', *Technometrics*, 14, pp. 669 - 678.
[13] Guenther, W. C. (1972), 'Variables Sampling Plans for the Poisson and the Binomial', *Statistica Neerlandica*, 26, pp. 113 - 120.
[14] Guenther, W. C. (1985), 'LQL Like Plans for Sampling by Variables', *Journal of Quality Technology*, 17, pp.155 - 157.
[15] Zimmer, W. J. and Burr, I. W. (1963), 'Variables Sampling Plans based on Non-Normal Populations', *Industrial Quality Control*, 21, pp. 18 - 26.
[16] Aminzadeh, M. S. (1996), 'Inverse-Gaussian Acceptance Sampling Plans by Variables', *Communications in Statistics – Theory and Methods*, 25, pp. 923 - 935.
[17] Suresh, R. P., Ramanathan, T. V. (1997), 'Acceptance Sampling Plans by Variables for a Class of Symmetric Distributions', *Communications in Statistics – Simulation and Computation*, 26, pp.1379 - 1391.
[18] Vijayaraghavan, R., and Geetha, S. (2009a), 'Evaluation of Single Sampling Plans by Variables using Single Parameter Gamma Distribution', National Conference on Quality Improvement Concepts and their Implementation in Higher Educational Institutions, December 11 – 12, Amrita University, Coimbatore, INDIA.

[19] Vijayaraghavan, R., and Geetha, S. (2009b), 'Procedure for Selection of Single Sampling Plans Variables Based on Gamma Distribution', UGC National Level staff Seminar on Applied Statistics, December 24, Sri Sarada College for Women, Salem, INDIA.

[20] Vijayaraghavan, R., and Geetha, S. (2010), 'Procedure for Selection of Single Sampling Plans Variables Based on Laplace Distribution', *Recent Trends in Statistical Research,* Publication Division, Mononmaniam Sundaranar University, pp. 209-218.

[21] Geetha, S., and Vijayaraghavan, R. (2011a), 'Evaluation of Single Sampling Plan by Variables Based on Rayleigh Distribution', National Conference on Statistics for Twenty First Century (NCSTC) & Annual Conference of Kerala Statistical Association, March 17 - 19, 2011, University of Kerala, Trivandrum, INDIA.

[22] Geetha, S., and Vijayaraghavan, R. (2011b), 'Selection of Single Sampling Plans by Variables Based on Logistic Distribution', UGC sponsored National Conference on Recent Trends in Statistical and Computer Applications, March 23 - 24, 2011, Mononmaniam Sundaranar University. Tirunelveli, INDIA.

[23] Geetha, S., and Vijayaraghavan, R. (2013). 'A Procedure for the Selection of Single Sampling Plans by Variables Based on Pareto Distribution', Journal of Quality and Reliability Engineering, Article ID 808741.

[24] Schilling, E. G. and Neubauer, D. V. (2009), *Acceptance Sampling in Quality Control*, Chapman and Hall, New York, 2009.

[25] Wu, H., and Govindaraju, K. (1984), 'Computer-aided Variables Sampling Inspection Plans for Compositional Proportions and Measurement Error Adjustment', *Computers and Industrial Engineering,* 72, pp. 239 - 246.

[26] Kumaraswamy (1980), 'A Generalized Probability Density Function for Double-bounded Random Processes, *Journal of Hydrology*, 46, 79 – 88.

[27] Govindaraju, K., and Kissling, R. C. (2015), 'Sampling Plans for Beta Distributed Compositional Fractions', *Chemometrics and Intelligent Laboratory Systems*, 151, 103 – 107.

RELIABILITY ANALYSIS OF REVERSE OSMOSIS FILTRATION SYSTEM USING COPULA

Anas Sani Maihulla
•
Department of Mathematics, Sokoto State University, Sokoto- Nigeria
anas.maihulla@ssu.edu.ng

Ibrahim Yusuf
•
Department of Mathematical sciences, Bayero University, Kano- Nigeria
iyusuf@buk.edu.ng

Abstract

In this study, the reliability metrics used to assess the strength of a three-subsystem reverse osmosis filtering system. The subsystems include sand filter, carbonated filter, and precision filter. Each subsystem is composed of active components that can operate in series parallel. The system of partial differential equations was built using the mnemonic rule and analytically solved. Other reliability variables that were investigated for determining system strength included availability, reliability, mean time to failure (MTTF), profit analysis, and sensitivity analysis. The Maple software was used to obtain numerical solutions. In addition, a graphical representation of the numerical results was provided to demonstrate the behaviors of reliability characteristics with regard to time and failure rate. The study could assist water treatment firms and their repairers in overcoming some of the challenges faced by repairers of specialized manufacturing and industrial systems working in harsh settings or contaminated environments unfit for human consumption.

Keywords: mnemonic, profit, sensitivity, reliability, repair, system, copula

I. Introduction

The removal of pollutants from drinking water caused by humans is a modern-day technical challenge. You'll learn about the detectable contamination of drinking water caused by anthropogenic (human-made) contaminants that is still present after a quick review of the treatment steps that municipal water goes through before it reaches your tap. Prescription medications, herbicides, and hormones have all been discovered in the drinking water systems of our respective countries. The engineering design method could be used to find solutions to a real-world problem (contaminated water) that could be dangerous to people's health. Water is perhaps the most important nutrient in our diets. In reality, a normal adult needs to drink around 2 liters (8 glasses) of water per day to replace the water lost through the epidermis, respiratory system, and urine. The Reliability, Availability, Maintainability, and Dependability Analysis of a Complex Reverse Osmosis Machine System in Water Purification was studied by Maihulla A. S. et al. [1] Al-Ghouti, M.A et al. [2] evaluated the recent developments and applications for municipal solid waste bottom and fly ashes. The Gumbel-Hougaard family copula was used to forecast the reliability and performance of a small serial solar photovoltaic system for rural use was carried out by Maihulla A.S. and Yusuf I. [3]. Evaluation of Ozone Pretreatment on Reverse Osmosis Flux Parameters for Surface Water Treatment, Ozone by Shalana L. et al. [4]. The study pertaining the Availability and reliability analysis of integrated reverse osmosis was carried out by safder U. et al. [5]. Calixto, Eduardo. [6] Discussed about Reliability, Availability, and Maintainability (RAM Analysis). The RAM analysis and availability optimization of thermal power plant water circulation system using PSO was carried out by Hanumant P. et al. [7]. The study that tackled the Reliability Assessment for Hybrid Systems of Advanced Treatment Units of Industrial Wastewater Reuse Using Combined Event Tree and Fuzzy Fault Tree Analyses. Was conducted by Farzad P. et al. [8]. Hajeeh, M. [11] & Chaudhuri, D. [9] conducted a research titled reliability and availability assessment of reverse osmosis Desalination. The feasibility and reliability of the Life Cycle Assessment for desalination, is a study carried out by Zhou, J. et al. [10]. Revas P. et al. [12] study the Real-Time Implementation of an Expert Model Predictive Controller in a Pilot-Scale Reverse Osmosis Plant for Brackish and Seawater Desalination. Goyal et al. [13] defined the most vulnerable aspect of serial processes such as evaporation systems in the sugar industry and water treatment plants. Using STP, this research deconstructs the efficiency indices of the power generation system. Simple probability theory concepts and the Markovian birth-death process were used to investigate the power system. As a Markov process progresses from one stage to the next, it becomes more complex.

M.F. Idrees [14] also works with the Performance Analysis and Treatment Technologies of a Reverse Osmosis Plan. C. Li, S. Besarati, and colleagues [15] A few years ago, I did research on reverse osmosis desalination using a low temperature supercritical organic Rankine cycle. In order to meet environmental and economic standards, any sophisticated reverse osmosis plant must have automation and reliability. S. Srivastava. [16]. S. Sadri [17] Created a computational model based on diffusion and convection transport mechanisms, as well as the concentration polarization concept, to predict RO membrane performance using a variety of feed water concentrations, feed flow rates, feed water pressures, membrane specifications, and feed water properties. Y. Li et al. [18] conducted a study of the concepts and categorization of membrane distillation, with an emphasis on the variables influencing it and ways to improving its efficacy. Experiments were conducted by E. O. Ezugbe et al. [19] Pure water and NaCl solutions ranging from 15 g/L to 300 g/L, as well as two distinct fiber types and architectures, were used. Vacuum membrane distillation (VMD) is a saltwater desalination technique. The pure water permeability and global heat transfer coefficient of the two systems were compared. The effects of hydrodynamics on global heat and mass transport coefficients are discussed. R. Tundis et al. [20] The chemical profile, antioxidant and anti-obesity properties of concentrated fractions obtained from micro-filtered OMW treated by direct contact membrane distillation were studied (DCMD). Using ultrahigh performance liquid chromatography, several phenols chosen as phytochemical markers were measured (UHPLC). To treat the pollutants found in olive mill wastewater (total organic carbon (TOC), dissolved organic carbon (DOC), total phosphorus (TP), total nitrogen (TN), and total polyphenols), a sequential Direct Contact Membrane Distillation (DCMD) and a Reverse Osmosis (RO) hybrid membrane system were used. The study was conducted by D. Teresa [21]. The influence of permeate flow and pressure on pollutant parameter removals was also studied. One of the biggest challenges that humanity must address in the twenty-first century is the scarcity of freshwater. P. Biniaz et al. [22] explored how an ecologically acceptable, cost-effective, and energy-efficient membrane distillation process might be developed (MD) process can reduce pollution caused by industrial and domestic wastes. Garud R. M. et al. [23] conducted a Short Review on Process and Applications of Reverse Osmosis. The Gumbel–Hougaard family copula was used to model the reliability and performance of a solar photovoltaic system was analyzed by Maihulla A. S. et al. [24]. Y.G. Lee et al. [25] created a model with five input factors (feed temperature, feed total dissolved solids (TDS), trans-membrane pressure (TMP), feed flow rate, and time) and two output parameters (permeate TDS and flow rate) to estimate the performance of a saltwater reverse osmosis (SWRO) desalination plant It was then used to simulate feed water temperature. D. Wirth [26] compared two hollow fiber module designs (inside/out and outside/in).

Our motivation for investigating the reverse osmosis filtration system derives from a severe problem that the water purification businesses are experiencing as a result of purification filter failure. And the slow progress in technological advancement in water purification as a result, as well as its relevance in the lives of people all over the world. Industry is working hard to keep up with the growing complexity of filtration systems. According to the paper's findings, Gumbel-Haugaard Family Copula analysis was performed to examine the filtration system's strength, efficiency, and performance improvement. Users will be able to serve the expense of medical care owing to un-pure water if the strength, efficiency, and performance of the filtration system are assessed. Protect yourself from aquatic pollution. The research is broken into five sections, one of which being the present introduction. Filtration modeling is discussed. Later in the second portion Section 4 contains a discussion and explanation of the results from the third section, which involves an analytical analysis of the system. Section 5 is devoted to the conclusion and ramifications of the findings.

II. Methods

I. Filtration

One of the most common methods for removing these materials is gravity filtration. Water containing solid impurities (e.g., precipitates after water softening) is passed through a porous material, usually sand and gravel layers, in this procedure. The force of gravity pushes the water through the medium. The gaps between the sand and gravel grains allow small water molecules to flow through. Precipitation-derived solids, on the other hand, become trapped in the pores and thus remain in the porous medium. The solid contaminants have been removed from the water that passes through the bottom of the filter.

II. Description of the model

A model with three series-parallel subsystems, A, B, and C, is shown in Figure 1. Two identical units work as 1-out-of-2 in subsystem A (sand filter), three identical units work as 1-out-of-3 in subsystem B (carbonated filter), and two parallel units work as 2-out-of-2 in subsystem C (precision filter). The two types of system failures are partial and complete failures. When a unit in a subsystem fails but the system continues to function, it is called partial failure, whereas total failure occurs when all of the subsystems fail. Copula is used to repair a system that has completely failed. In the system, there are eleven states: eight that are operational and three that are total failure states (see Figure2). The states are described briefly in Table 1.

III. Results

I. Formulation and Solution of Mathematical Model

The following set of difference-differential equations is related to the aforementioned mathematical model based on consideration likelihood and argument continuity.

$$\left(\frac{\partial}{\partial x}+\alpha_1+\alpha_2+2\alpha_3\right)P_0(t)=\int_0^\infty \beta_1 P_1(x,t)dx+\int_0^\infty \beta_2 P_2(y,t)dy+\int_0^\infty \psi(y)P_6(y,t)dy+\int_0^\infty \varphi(x)P_8(x,t)dx+\int_0^\infty \varphi(y)P_9(y,t)dy+\int_0^\infty \varphi(z)P_{10}(z,t)dz \quad (1)$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\alpha_1+\alpha_2+\beta_1\right)P_1(x,t)=0 \quad (2)$$

$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\alpha_1+\alpha_2+\ \beta_2)\ P_2(y,t)=0$ (3)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\alpha_1+\alpha_2+\ \beta_2)\ P_3(y,t)=0$ (4)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\alpha_2+\ \beta_2)\ P_4(y,t)=0$ (5)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\alpha_2+\ \beta_2)\ P_5(y,t)=0$ (6)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\alpha_1+\ \beta_1)\ P_6(x,t)=0$ (7)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\alpha_1+\ \beta_1)\ P_7(x,t)=0$ (8)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}+\varphi(x))\ P_8(x,t)=0$ (9)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial y}+\varphi(y))\ P_9(y,t)=0$ (10)
$(\frac{\partial}{\partial t}+\frac{\partial}{\partial z}+\varphi(z))\ P_{10}(x,t)=0$ (11)

Boundary conditions
$P_1(0,t)=\alpha_1P_0(t)$ (12)
$P_2(0,t)=\alpha_2P_0(t)$ (13)
$P_3(0,t)=\alpha_2^2P_0(t)$ (14)
$P_4(0,t)=\alpha_1\alpha_2P_0(t)$ (15)
$P_5(0,t)=\alpha_1\alpha_2^2P_0(t)$ (16)
$P_6(0,t)=\alpha_1\alpha_2P_0(t)$ (17)
$P_7(0,t)=\alpha_1\alpha_2^2P_0(t)$ (18)
$P_8(0,t)=\alpha_1^2(1+\alpha_1+\alpha_2^2)P_0(t)$ (19)
$P_9(0,t)=\alpha_2^3(1+\alpha_1)P_0(t)$ (20)
$P_{10}(0,t)=2\alpha_3P_0(t)$ (21)

Initial condition $P_0(t)$ = 1 and other transition probability at t=0 are zero
Laplace transformation of (1) – (21)

$$(s+\alpha_1+\alpha_2+2\alpha_3)\bar{P}_0(s)=\int_0^\infty\beta_1\bar{P}_1(x,s)dx+\int_0^\infty\beta_2\bar{P}_2(y,s)dy+\int_0^\infty\psi(y)\bar{P}_6(y,s)dy+\int_0^\infty\varphi(x)\bar{P}_8(x,s)dx+\int_0^\infty\varphi(y)\bar{P}_9(y,s)dy+\int_0^\infty\varphi(z)\bar{P}_{10}(z,s)dz \quad (22)$$

$(s+\frac{\partial}{\partial x}+\alpha_1+\alpha_2+\ \beta_1)\ \bar{P}_1(x,s)=0$ (23)
$(s+\frac{\partial}{\partial y}+\alpha_1+\alpha_2+\ \beta_2)\ \bar{P}_2(y,s)=0$ (24)
$(s+\frac{\partial}{\partial y}+\alpha_1+\alpha_2+\ \beta_2)\ \bar{P}_3(y,s)=0$ (25)
$(s+\frac{\partial}{\partial y}+\alpha_2+\ \beta_2)\ \bar{P}_4(y,s)=0$ (26)
$(s+\frac{\partial}{\partial y}+\alpha_2+\ \beta_2)\ \bar{P}_5(y,s)=0$ (27)
$(s+\frac{\partial}{\partial x}+\alpha_1+\ \beta_1)\ \bar{P}_6(x,s)=0$ (28)
$(s+\frac{\partial}{\partial x}+\alpha_1+\ \beta_1)\ \bar{P}_7(x,s)=0$ (29)
$(s+\frac{\partial}{\partial x}+\varphi(x))\ \bar{P}_8(x,s)=0$ (30)
$(s+\frac{\partial}{\partial y}+\varphi(y))\ \bar{P}_9(y,s)=0$ (31)
$(s+\frac{\partial}{\partial z}+\varphi(z))\ \bar{P}_{10}(z,s)=0$ (32)

Laplace of the Boundary conditions
$\bar{P}_1(0,s)=\alpha_1\bar{P}_0(s)$ (33)
$\bar{P}_2(0,s)=\alpha_2\bar{P}_0(s)$ (34)
$\bar{P}_3(0,s)=\alpha_2^2\bar{P}_0(s)$ (35)
$\bar{P}_4(0,s)=\alpha_1\alpha_2\bar{P}_0(s)$ (36)
$\bar{P}_5(0,s)=\alpha_1\alpha_2^2\bar{P}_0(s)$ (37)
$\bar{P}_6(0,s)=\alpha_1\alpha_2\bar{P}_0(s)$ (38)
$\bar{P}_7(0,s)=\alpha_1\alpha_2^2P_0(s)$ (39)
$\bar{P}_8(0,s)=\alpha_1^2(1+\alpha_1+\alpha_2^2)\bar{P}_0(s)$ (40)

$$\bar{P}_9(0,s) = \alpha_2^3(1+\alpha_1)\,\bar{P}_0(s) \tag{41}$$

$$\bar{P}_{10}(0,s) = 2\alpha_3\bar{P}_0(s) \tag{42}$$

Solving equation (22) to (32) with the help of boundary condition (33) to (42) and applying the below shifting properties of Laplace:

$$\int_0^\infty [e^{-sx}\,.\,e^{-\int_0^x f(x)dx}]dx = L\left\{\frac{1-\bar{S}_f(x)}{s}\right\} = \frac{1-\bar{S}_f(x)}{s} \tag{43}$$

$$\int_0^\infty [e^{-sx}\,.\,f(x)\,e^{-\int_0^x f(x)dx}]dx = L\{\bar{S}_f(x)\} = \bar{S}_f(s) \tag{44}$$

And the identity; $\bar{P}_1(s) = \int_0^\infty \bar{P}_1(x,s)dx$ (45)

$$\bar{P}_1(s) = \bar{P}_1(0,s)\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_3}\right\} \tag{46}$$

$$\bar{P}_2(s) = \bar{P}_2(0,s)\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_2}\right\} \tag{47}$$

$$\bar{P}_3(s) = \bar{P}_3(0,s)\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_2}\right\} \tag{48}$$

$$\bar{P}_4(s) = \bar{P}_4(0,s)\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_2)}{S+\alpha_2}\right\} \tag{49}$$

$$\bar{P}_5(s) = \bar{P}_5(0,s)\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1)}{S+\alpha_1}\right\} \tag{50}$$

$$\bar{P}_6(s) = \bar{P}_6(0,s)\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1)}{S+\alpha_1}\right\} \tag{51}$$

$$\bar{P}_7(s) = \bar{P}_7(0,s)\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1)}{S+\alpha_1}\right\} \tag{52}$$

$$\bar{P}_8(s) = \bar{P}_8(0,s)\left\{\frac{1-\bar{S}_{\varphi(x)}(S)}{S}\right\} \tag{53}$$

$$\bar{P}_9(s) = \bar{P}_9(0,s)\left\{\frac{1-\bar{S}_{\varphi(y)}(S)}{S}\right\} \tag{54}$$

$$\bar{P}_{10}(s) = \bar{P}_{10}(0,s)\left\{\frac{1-\bar{S}_{\varphi(z)}(S)}{S}\right\} \tag{55}$$

Substituting the Laplace of the Boundary conditions i.e (33) to (42) into (43) to (55)

$$\bar{P}_1(s) = \alpha_1\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_3}\right\}\bar{P}_0(s) \tag{56}$$

$$\bar{P}_2(s) = \alpha_2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_2}\right\}\bar{P}_0(s) \tag{57}$$

$$\bar{P}_3(s) = \alpha_2^2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_2}\right\}\bar{P}_0(s) \tag{58}$$

$$\bar{P}_4(s) = \alpha_1\alpha_2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_2)}{S+\alpha_2}\right\}\bar{P}_0(s) \tag{59}$$

$$\bar{P}_5(s) = \alpha_1\alpha_2^2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1)}{S+\alpha_1}\right\}\bar{P}_0(s) \tag{60}$$

$$\bar{P}_6(s) = \alpha_1\alpha_2\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1)}{S+\alpha_1}\right\}\bar{P}_0(s) \tag{61}$$

$$\bar{P}_7(s) = \alpha_1\alpha_2^2\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1)}{S+\alpha_1}\right\}\bar{P}_0(s) \tag{62}$$

$$\bar{P}_8(s) = \alpha_1^2(1+\alpha_1+\alpha_2^2)\left\{\frac{1-\bar{S}_{\varphi(x)}(S)}{S}\right\}\bar{P}_0(s) \tag{63}$$

$$\bar{P}_9(s) = \alpha_2^3(1+\alpha_1)\left\{\frac{1-\bar{S}_{\varphi(y)}(S)}{S}\right\}\bar{P}_0(s) \tag{64}$$

$$\bar{P}_{10}(s) = 2\alpha_3\left\{\frac{1-\bar{S}_{\varphi(z)}(S)}{S}\right\}\bar{P}_0(s) \tag{65}$$

$$\begin{aligned}(s+\alpha_1+\alpha_2+2\alpha_3)\bar{P}_0(s) \quad &= 1+\beta_1\bar{S}_{\beta_1}(S+\alpha_1+\alpha_2)\bar{P}_0(s)+\beta_2\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)\bar{P}_0(s)+[\varphi(x)\bar{S}_{\varphi(x)}(S)\\ &+\varphi(y)\bar{S}_{\varphi(y)}(S)+\varphi(z)\bar{S}_{\varphi(z)}(S)]\,\bar{P}_0(s)\end{aligned} \tag{66}$$

$$s+\alpha_1+\alpha_2+2\alpha_3-[\beta_1\bar{S}_{\beta_1}(S+\alpha_1+\alpha_2)+\beta_2\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)+[\varphi(x)\bar{S}_{\varphi(x)}(S)+\varphi(y)\bar{S}_{\varphi(y)}(S)+\varphi(z)\bar{S}_{\varphi(z)}(S)]]\bar{P}_0(s) = 1 \tag{67}$$

$$D(s) = s+\alpha_1+\alpha_2+2\alpha_3-[\beta_1\bar{S}_{\beta_1}(S+\alpha_1+\alpha_2)+\beta_2\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)+[\varphi(x)\bar{S}_{\varphi(x)}(S)+\varphi(y)\bar{S}_{\varphi(y)}(S)+\varphi(z)\bar{S}_{\varphi(z)}(S)]] \tag{68}$$

Since $D(s)\times\ \bar{P}_0(s)=1$ (69)

$$\bar{P}_{up}(S) = \bar{P}_0(S)+\bar{P}_1(S)+\bar{P}_3(S)+\bar{P}_4(S)+\bar{P}_5(S)+\bar{P}_6(S)+\bar{P}_7(S) \tag{70}$$

$$\bar{P}_{up}(S) = [1+\alpha_1\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_3}\right\}+\alpha_2^2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1+\alpha_2)}{S+\alpha_1+\alpha_2}\right\}+\alpha_1\alpha_2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_2)}{S+\alpha_2}\right\}+\alpha_1\alpha_2^2\left\{\frac{1-\bar{S}_{\beta_2}(S+\alpha_1)}{S+\alpha_1}\right\}+$$

$$\alpha_1\alpha_2\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1)}{S+\alpha_1}\right\}+\alpha_1\alpha_2^2\left\{\frac{1-\bar{S}_{\beta_1}(S+\alpha_1)}{S+\alpha_1}\right\}]\bar{P}_0(s) \quad (71)$$

II. Availability Analysis

Setting all repairs to 1. i.e. $\varphi(x)=\varphi(y)=\varphi(z)=\beta_1=\beta_2=\beta_3=1$ (72)

$\bar{S}_{\phi}(S)=\frac{2.7183}{S+2.7183}$, $\frac{1-\bar{S}_{\phi}(S)}{S}=\frac{1}{S+\phi}$

Taking the values of different parameters as $\alpha_1=0.01, \alpha_1=0.02, \alpha_3=0.03, \alpha_4=0.04$ in (48) Then taking the inverse Laplace transform, we can obtain, the expression for availability as:

$$D(s) = S + 0.12- [\frac{0.03}{S+1.04}+\frac{0.03}{S+1.03}+0.06\,.(\frac{2.7183}{S+2.7183})\,] \quad (73)$$

$$\bar{P}_{up}(S)=[1+\frac{0.03}{S+1.04}+\frac{0.03}{S+1.03}+\frac{0.0018}{S+1.03}] \quad (74)$$

Taking $S_{\alpha_0}(s)=\bar{S}_{\exp[x^\theta+\{\log\varphi(x)\}^\theta]^{1/\theta}}(s)=\dfrac{\exp[x^\theta+\{\log\varphi(x)\}^\theta]^{1/\theta}}{s+\exp[x^\theta+\{\log\varphi(x)\}^\theta]^{1/\theta}}$, $\bar{P}_{\phi}(s)=\frac{\phi}{s+\phi}$ *but* ϕ =1 and $\alpha_1=0.001$, $\alpha_2=0.002$, $\alpha_3=0.003$

And repair rates $\varphi(x)=\varphi(y)=\varphi(z)=\beta_1=\beta_2=\beta_3=1$ in equation (69), and applying the inverse Laplace transform to (69), the expression for system availability is

$$\bar{P}_{up}(t)=\{-0.00004904770570\,e^{-1.011000000\,t}+0.001603099809\,e^{-2.722649460\,t}+(-0.006078301855 - 0.0014178109921)\,e^{(-1.014501644-0.0009938111316\,I)\,t}+(-0.006078301855+0.0014178109921)\,e^{(-1.014501644+0.0009938111316\,I)\,t}+1.010602552\,e^{-0.02864725083\,t}\} \quad (75)$$

Taking t = 0, 10,…,100, availability of the system is obtained and presented in Table 1 below:

Table 1: *Availability variance with respect to time*

Time in days	Availability
0	1.000000
10	0.997679
20	0.997531
30	0.997383
40	0.997235
50	0.997086
60	0.996939
70	0.996791
80	0.996643
90	0.996495
100	0.996347

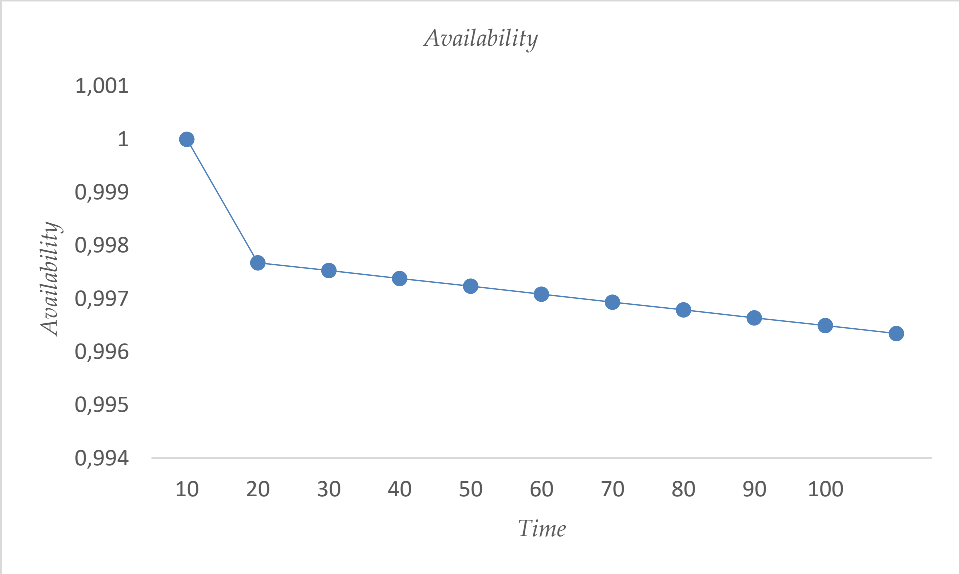


Figure 1: *Variation of availability with time*

III. Reliability Analysis

Letting all repair rates, $\varphi(x) = \varphi(y) = \varphi(z) = \beta_1 = \beta_2 = \beta_3 = 0$ in equation (70), and taking the values of failure rates and employing inverse Laplace transformation, the expression is reliability relation.

$$R(t) = \{ 0.002909090909\, e^{-0.011100000000\, t} + 0.2352941176\, e^{-0.016000000000\, t} + 0.3617967914\, e^{-0.033000000000\, t} + 0.4000000000\, e^{-0.013000000000\, t} \} \quad (76)$$

Using t = 0, 10...100 as time units in equation (72), reliability is determined and shown in Table 2 below.

Table 2: *Variation in Reliability as a Function of Time*

Time in Days	Reliability
0	1.000000
10	0.941981
20	0.887595
30	0.836601
40	0.788773
50	0.743903
60	0.701796
70	0.662272
80	0.625161
90	0.590308
100	0.557565

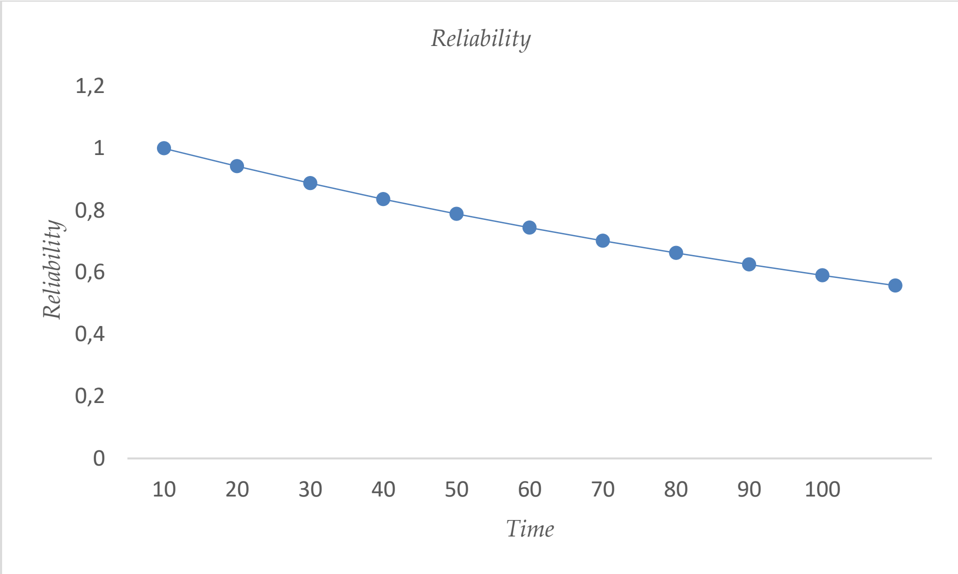


Figure 2. *Variation of reliability with time*

IV. Cost Analysis

The expression for the expected profit incurred in $[0,t)$

$$E_p(t) = K_1 \int_0^t P_{up}(t)dt - K_2 t \tag{77}$$

Taking fixed values of parameters of equation (62), the subsequent equation (69) follows;

$$E_p(t) \quad = \quad E_p(t) = k_1 \begin{Bmatrix} -0.018340e^{-2.87072t} + 0.008479e^{-1.20122t} \\ +0.000094e^{-1.15618t} - 598.072776e^{-0.00016t} \\ +0.000183e^{-1.13000t} + 0.000237e^{-1.12000t} \\ +5985.0821 \end{Bmatrix} - k_2(t) \tag{78}$$

Table 3: *Expected profit as a function of time*

Time in days	$E_p(t)$					
0	0	0	0	0	0	0
10	3.97831	4.97831	5.97831	6.97831	7.97831	8.97831
20	7.95435	9.95435	11.95435	13.95435	15.95435	17.95435
30	11.92892	14.92892	17.92892	20.92892	23.92892	26.92892
40	15.90200	19.90200	23.90200	27.90200	31.90200	35.90200
50	19.87361	24.87361	29.87361	34.87361	39.87361	44.87361
60	23.84373	29.84373	35.84373	41.84373	47.84373	53.84373
70	27.81238	34.81238	41.81238	48.81238	55.81238	62.81238
80	31.77954	39.77954	47.77954	55.77954	63.77954	71.77954
90	35.74523	44.74523	53.74523	62.74523	71.74523	80.74523
100	39.70944	49.70944	59.70944	69.70944	79.70944	89.70944

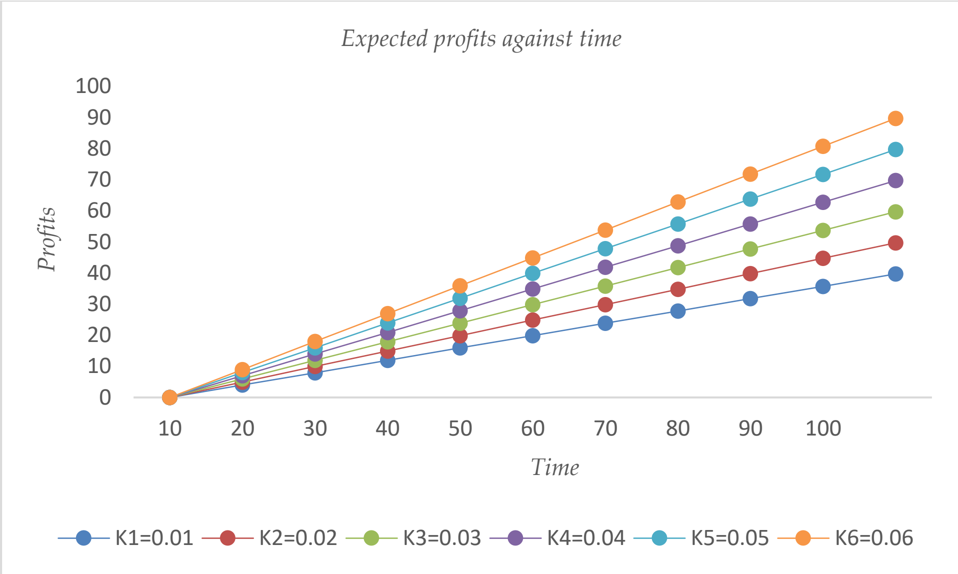


Figure 3: *Variation of profit with time*

V. Formulation and Analysis Mean Time to Failure

Setting repairs to zero in equation (70), the expression for MTTF is defined as follows:
Fixing $\alpha_1 = 0.001$, $\alpha_2 = 0.002$, $\alpha_3 = 0.003$ varying, failure rate in equation (66), MTTF is computed with respect to failure rate as presented in Table 3 below.

Table 4: *Variation of MTTF with failure rates* α_k

Failure rate	MTTF		
	(a)	(b)	(c)
0.001	222.8152	250.2918	401.6008
0.002	200.5337	222.8153	286.5720
0.003	182.3034	200.9009	222.8153
0.004	167.1114	183.0318	182.2731
0.005	154.2567	168.1965	154.2157
0.006	143.2384	155.6951	133.6447
0.007	133.6892	145.0273	117.9162
0.008	125.3336	135.8265	105.5002
0.009	117.9610	127.8176	95.44993

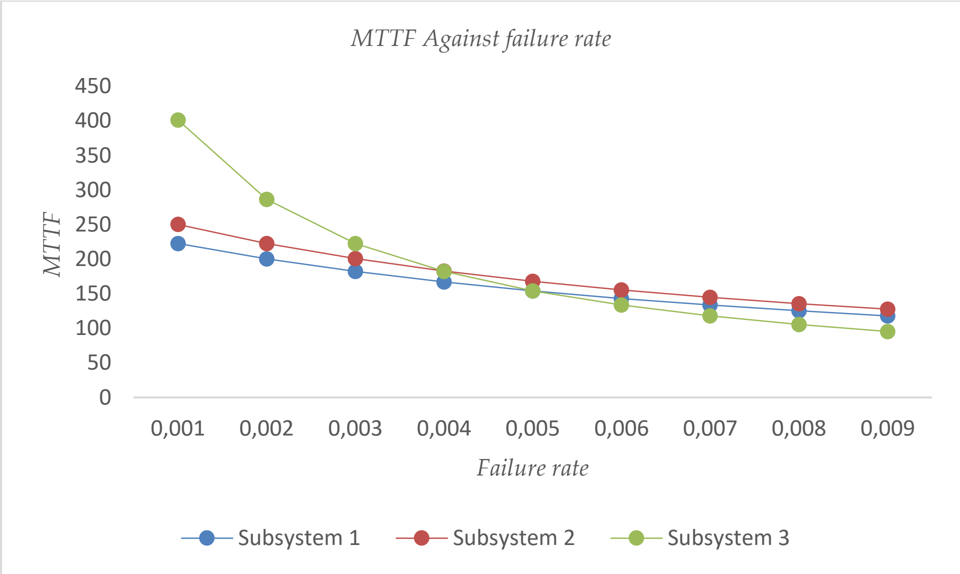


Figure 4: *MTTF as function of Failure rate*

VI. Sensitivity analysis corresponding to MTTF

The partial differentiation of MTTF with respect to the failure rate of the system can be used to investigate the system's sensitivity in MTTF. Using the set of parameters as $\alpha_1 = 0.001$, $\alpha_2 = 0.002$ $\alpha_3 = 0.003$, The MTTF sensitivity can be calculated in the partial differentiation of MTTF, as shown in the Table below and associated graphs in the Figure.

Table 5. *MTTF sensitivity as function of failure rate*

Failure rate	$\frac{\partial(MTTF)}{\alpha_1}$	$\frac{\partial(MTTF)}{\alpha_2}$	$\frac{\partial(MTTF)}{\alpha_3}$
0.001	-24757.25	-30952.89	-100000.61
0.002	-20053.37	-24386.44	-82020.57
0.003	-16573.03	-19689.49	-49563.88
0.004	-13925.95	-16214.28	-33163.29
0.005	-11865.90	-13571.10	-23737.80
0.006	-10231.31	-11514.08	-17826.70
0.007	-8912.61	-9881.903	-13877.30
0.008	-7833.35	-8565.144	-11108.57
0.009	-6938.88	-7487.473	-9092.820

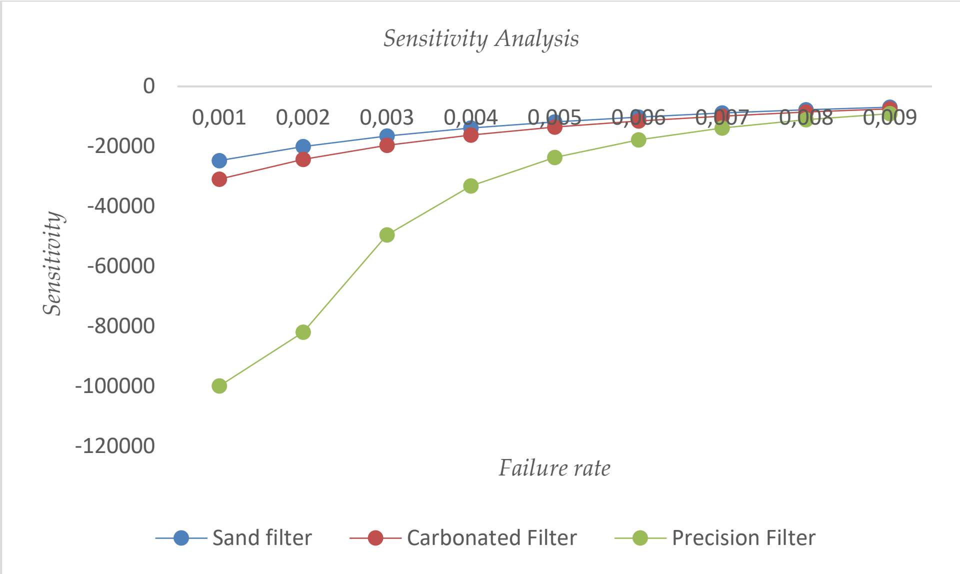


Figure 5: *Sensitivity analysis corresponding to mean time to failure (MTTF)*

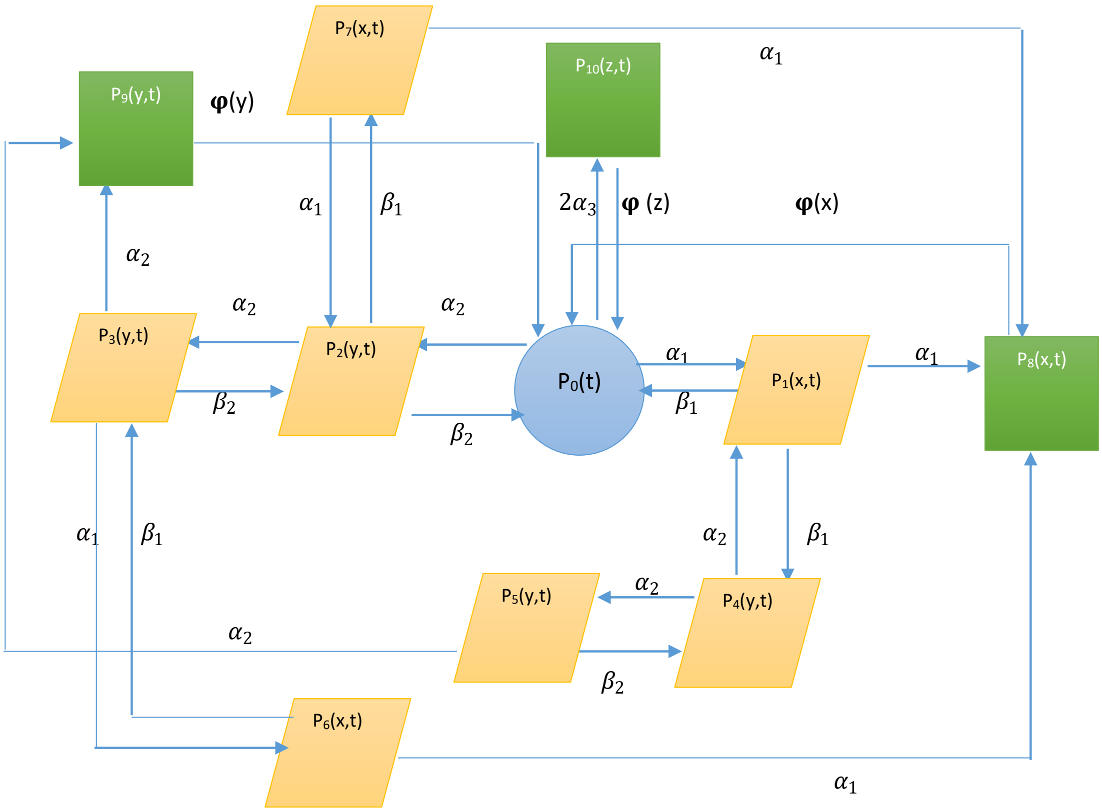
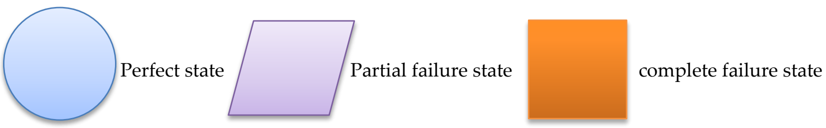


Figure 6: *Filtration system's transition diagram*



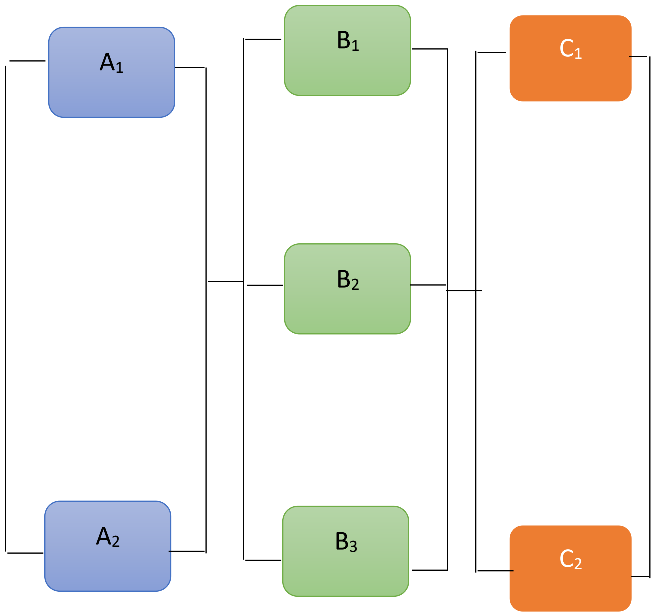


Figure 7: *Reliability block diagram of the filtration system*

VII. Description of the system

α_n: Failure rate for the subsystems. Where n=1,2,3.
β_q: Repair rate for the subsystems with incomplete failure Where q =1,2,3.
φ(k) : Repair rate for the subsystems with complete failure. Where k= x, y, z.
P_0: Denote initial state where the system is working perfectly.
P_1: Denote state with an incomplete failure in subsystem 1 due to failure of first unit and copula repair is busy in repairing the failed unit.
P_2: Denote state with a complete failure in subsystem 2 due to failure of first unit and repair machine is busy in repairing the failed unit.
P_3: Denote state with a incomplete failure in subsystem 2
P_4: Denote state with an incomplete failure state due to failure of first unit in subsystem 1 and one unit from subsystem 2.
P_5: Denote state with an incomplete failure state due to failure of first unit in subsystem 1, first and second units from subsystem 2.
P_6: Denote state with an incomplete failure due to the failure of first and second units from subsystem 2, and one unit from subsystem 1.
P_7: Denote incomplete state of system due to failure of first unit from each of subsystems one and two
P_8: Denote complete failure state in subsystem 1.
P_9: Denote complete failure state in subsystem 2.
P_{10}: Denote complete failure state in subsystem 3.

IV. Discussion

I. Interpretation of the Result and Conclusion

Table1 and Figure 1 provide information how availability and time changes when failure rates are fixed at different values. When failure rates are fixed at lower values α_1 = 0.001, α_2= 0.002 and α_3 = 0.003. Table 1 shows the results, and Figure 1 shows the corresponding result, the system's availability decreases over time. And the decreases from the first point are greater than those from the second. This is due to the units in subsystem 1 having less redundancy. Figure 2 shows that the system is more reliable than it is available. However, this is due to a copula repair for a component that had completely failed. And the system will be assumed to be operational during the repair. The system's availability reduces with time and eventually stabilizes at zero over a sufficiently extended length of time. As a result, the graphical representation of the model shows that one may confidently describe the future behavior of a complex system at any time for any given set of parametric parameters. Figure 2 and table 2 indicates that incorporating copula considerably decreases system reliability with time. The graphical representation of the model shows that one may confidently forecast the future behavior of a complex system at any time for any given set of parametric variables. When the values of availability and reliability in Tables 1 and 2 are compared, it is clear that when repair is offered, the system performs far better than replacement.

Tables 4 and figure 4 yield the mean-time-to-failure (MTTF) of the system with respect to variation in the failure rates,α_1, α_2, and α_3 respectively when all other parameters are held constant Shown by color graphs (green, Pink, and ash) respectively. The system is also examined using the Gumbel-Hougaard family copula. According to the findings, incorporating copula considerably enhances system reliability.

Cost analysis of the system is done in the analytic part of the paper Figure 3 and table 3 shows the variation of cost-profit with variation in the values of parameters. Table 5 and figure 5 displayed the result of sensitivity analysis corresponding to mean time to failure (MTTF) with respect to the failure rates α_k from the table, MTTF sensitivity decreases with increase in α_k.

For the analysis of profit the following are used:

(a) Fixing $K_2 = 0.01$ varying time t from 0 to 100
(b) Fixing $K_2 = 0.02$ varying time t from 0 to 100
(c) Fixing $K_2 = 0.03$ varying time t from 0 to 100
(d) Fixing $K_2 = 0.04$ varying time t from 0 to 100
(e) Fixing $K_2 = 0.05$ varying time t from 0 to 100

Table 4 displayed the result of expected profit $E_p(t)$ with respect to K_2. From the table, it is evident that expected profit increases as K_2 decreases.

II. Conclusion

Due to a lack of data on the filtration system, the current paper developed a reliability modeling approach to investigate the filtration system's overall strength, efficiency, and performance. The reliability, availability, MTTF, and profit function of this paper can all be evaluated. We present a new filtration system model that includes three subsystems: a sand filter, an activated carbon filter, and a precision filter in this study.

The findings of the study suggest that reliability modeling can be used to investigate the strength, efficiency, and performance enhancement of a reverse osmosis (RO) filtering system. The system's strength, efficiency, and performance improvement are determined at this point. This study could be enhanced to incorporate a system with numerous subsystems and several repair machines to reduce repair facility congestion and handle problems using supplemental variable techniques. Among other things, the current effort will benefit water manufacturing and industrial uses that are hazardous to humans.

References

[1] Maihulla, A. S., Yusuf, I., and Bala, S. I. (2021). Performance Evaluation of a Complex Reverse Osmosis Machine System in Water Purification using Reliability, Availability, Maintainability and Dependability Analysis. *Reliability: Theory & Applications*, *16*(3), 115-131.

[2] Al-Ghouti, M.A., Khan, M., Nasser, M.S.; Al-Saad, K.; Heng, O.E. (2021) Recent advances and applications of municipal solid wastes bottom and fly ashes: Insights into sustainable management and conservation of resources. Environ. Technol. Innov. 21:101-267.

[3] Maihulla, A.S., Yusuf, I. (2021). Reliability and performance prediction of a small serial solar photovoltaic system for rural consumption using the Gumbel-Hougaard family copula. *Life Cycle Reliab SafEng* https://doi.org/10.1007/s41872-021-00176-x 10: 347–354.

[4] Shalana L. Brown, Kathleen M. Leonard and Sherri L. Messimer (2008) Evaluation of Ozone Pretreatment on Flux Parameters of Reverse Osmosis for Surface Water Treatment, Ozone: Science&Engineering, DOI: 10.1080/01919510701864031. 30:2, 152 -164.

[5] Safder, Usman & Ifaei, Pouya & Nam, Kijeon & Rashidi, Zhowun & Yoo, ChangKyoo. (2018). Availability and reliability analysis of integrated reverse osmosis - Forward osmosis desalination network. *DESALINATION AND WATER TREATMENT*. 10.5004/dwt.2018.22147. 109: 1-7.

[6] Calixto, Eduardo. (2016). Reliability, Availability, and Maintainability (RAM Analysis). 10.1016/B978-0-12-805427-7.00004-X.

[7] Hanumant P. Jagtap, Anand K. Bewoor, Ravinder Kumar, Mohammad Hossein Ahmadi d, Mamdouh El Haj Assad, Mohsen Sharifpur (2021) RAM analysis and availability optimization of thermal power plant water circulation system using PSO. Energy reports https://doi.org/10.1016/j.egyr.2020.12.025 . 7: 1133-1153

[8] Farzad Piadeh , Mohsen Ahmadi , Kourosh Behzadian. (2018) Reliability Assessment for Hybrid Systems of Advanced Treatment Units of Industrial Wastewater Reuse Using Combined Event Tree and Fuzzy Fault Tree Analyses. *Journal of Cleaner Production* 10.1016/j.jclepro.2018.08.052

[9] Hajeeh, M. and Chaudhuri, D. (2000). Reliability and availability assessment of reverse osmosis. Desalination. 10.1016/S0011-9164(00)00086-2. 130: 185-192.

[10] Zhou, Jin & Chang, Victor & Fane, A.G. (2014). Life Cycle Assessment for desalination: A review on methodology feasibility and reliability. Water research. 10.1016/j.watres.2014.05.017. 61:210-223.

[11] Mohammed A. Hajeeh. (2019) Failure Analysis of Reverse Osmosis Plants. *International Journal of Environmental Science,* 4:102-106.

[12] Rivas-Perez, Raul, Javier Sotomayor-Moriano, Gustavo Pérez-Zuñiga, and Mario E. Soto-Angles (2019). "Real-Time Implementation of an Expert Model Predictive Controller in a Pilot-Scale Reverse Osmosis Plant for Brackish and Seawater Desalination" *Applied Sciences* 9, https://doi.org/10.3390/app9142932 14: 2932.

[13] Goyal, D., Kumar, A., Saini, M. *et al.* Reliability, maintainability and sensitivity analysis of physical processing unit of sewage treatment plant. *SN Appl. Sci.* 1, 1507 (2019). https://doi.org/10.1007/s42452-019-1544-7

[14] M.F. Idrees (2020) "Performance Analysis and Treatment Technologies of Reverse Osmosis Plant" – A Case Study, *Case Studies in Chemical and Environmental Engineering*, https://doi.org/10.1016/j.cscee.2020.100007.

[15] Li, Chennan & Besarati, Saeb & Goswami, Yogi & Stefanakos, Elias & Chen, Huijuan, 2013. Reverse osmosis desalination driven by low temperature supercritical organic rankine cycle, Applied Energy, Elsevier, vol:102(C), 1071-1080.

[16] S. Srivastava, S. Vaddadi, P. Kumar, and S. Sadistap (2018). Design and development of reverse osmosis (RO) plant status monitoring system for early fault prediction and predictive maintenance. journal of Applied Water Science *Applied Water Science* vol. 8, no. 159 pp 1-10.

[17] S. Sadri, R. H. Khoshkhoo and M. Ameri (2016). Multi objective optimization of reverse osmosis desalination plant with energy approach Journal of Mechanical Science and Technology vol. 30, pp 4807–4814.

[18] Y. Li and K. Tian (2009) Application of vacuum membrane distillation in water treatment, *Journal of Sustainable Development*, vol. 2, no. 3, pp. 183–186.

[19] Obotey Ezugbe, E., & Rathilal, S. (2020). Membrane Technologies in Wastewater Treatment: A Review. *Membranes*, https://doi.org/10.3390/membranes10050089 vol. no. *10* (5), 89

[20] Tundis, R., Conidi, C., Loizzo, M. R., Sicari, V., Romeo, R., & Cassano, A. (2021). Concentration of Bioactive Phenolic Compounds in Olive Mill Wastewater by Direct Contact Membrane Distillation. *Molecules (Basel, Switzerland)*, https://doi.org/10.3390/molecules26061808 *26*(6), 1808.

[21] D. Teresa Sponza (2021). Treatment of Olive Mill Effluent with Sequential Direct Contact Membrane Distillation (DCMD)/Reverse Osmosis (RO) Hybrid Process and Recoveries of Some Economical Merits. Advanced Journal of Physics Research and Applications. Vol. 1(1): 001-010.

[22] Biniaz, P., Torabi Ardekani, N., Makarem, M., & Rahimpour, M. (2019). Water and Wastewater Treatment Systems by Novel Integrated Membrane Distillation (MD). Chem Engineering, https://doi.org/10.3390/chemengineering3010008 3(1), 8.

[23] Garud R. M., Kore S. V., Kore V. S., and Kulkarni G. S. (2011) A Short Review on Process and Applications of Reverse Osmosis. www.environmentaljournal.org Volume 1, Issue 3: 233-238,

[24] Maihulla, A.S., Yusuf, I. and Salihu Isa, M. (2021), "Reliability modeling and performance evaluation of solar photovoltaic system using Gumbel–Hougaard family copula", International Journal of Quality & Reliability Management, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/IJQRM-03-2021-0071

[25] Y. G. Lee, Y. S. Lee, J. J. Jeon, S. Lee, D. R. Yang, (2009). Artificial neural network model for optimizing operation of a seawater reverse osmosis desalination plant. Volume 147, Issues 1–3, 10, Pages 180-189.

[26] D. Wirth, C. Cabassu (2002). Water desalination using membrane distillation: comparison between inside/out and outside/in permeation. ScienceDirect Volume 147, Issues 1–3, 10, Pages 139-145.

RT&A No. 2 (68)
Volume 17, June 2022

SENSITIVITY ANALYSIS OF A UREA FERTILIZER PLANT

Deepika Garg[1], Arun Kumar[4*]

•

[1,4]School of Engineering and Sciences, G D Goenka University,
Gurgaon, Haryana, India
deepika.garg@gdgoenka.ac.in, rnkumar535@gmail.com

Vimal Kumar Joshi[2]

•

[2]School of Basic and Applied Sciences, Galgotias University,
Greater Noida, Uttar Pradesh, India
joshiambition2008@gmail.com

Nahid Fatima[3]

•

[3]Humanities & Science, Prince Sultan University Riyadh, Saudi Arabia
drnahidfatima@gmail.com
*Corresponding author: rnkumar535@gmail.com

Abstract

Purpose – This paper presents a sensitivity analysis of a urea fertilizer manufacturing system comprising several sub-systems of differing nature. Design/methodology/approach–A mathematical model is developed for the consistent general repair and disappointment rates for every subsystem. The framework is analyzed by utilizing regenerative point graphical technique; as a result, some recommendations are made for the optimized output. A state transition diagram of the system is developed to find mean time to busy period server, system failure and system availability. Findings – The present study suggests an approach to improve the system performance. The analysis and results outlined in this paper are useful to system managers, training supervisor, engineers and reliability analysts in the manufacturing industry. Originality/ value – The manufacturing system of Urea fertilizer consists of a complex structure with the high risk of machine failure. Machine/ Production failure leads to high risks of economic & environmental loss and worker's safety. To address this challenge effectively, sensitivity analysis of the urea fertilizer plant is discussed for minimizing the risk of machine failure.

Keywords: Reliability, Availability, Server of Busy Period, RPGT

I. Introduction

The plants of urea fertilizer consist of a large number of sub-systems which are inter-connected in series/parallel or both. It is needed for various sub-systems to be remaining perpetually in the up state for the efficient working, But, in reality, they are subject to random failures and replacement take place. The processing of the sub-system depends upon the operating conditions and the repair policies, as a result, its failure are difficult to predict. For the most preferable level of system availability, behavioural analysis is a best mechanism to economize operational parameters.

The analysis of accessibility parameters like reliability, availability, maintainability etc. of different mechanical system can help in improving the quality of synthesis and increase the production. To ensure the system performance, it is necessary to utilize various strategies throughout its service life. A number of researchers [Garg et al. [10], Ram and Manglik [17], kumar et al. [9], Lin [12], Liu and Xie [13], Ni et al. [15]] analyzed the accessibility parameters of different mechanical systems utilizing various strategies. Kumar et al. [11] considered a single-unit system to study the concept of preventive maintenance for all associated variables. Mishra et al. [14] used the Markov approach to discuss the optimal availability of break drum manufacturing system. Kumar and Singh [10] performed the reliability analysis of a complex system which consists of two repairable subsystems connected in series. Kumar et al. [8] discussed the behavior analysis of a bread making system considering five distinct sub-systems consist of mixer, oven, tunnels, divider and proofer useful to the management utilizing RPGT under steady-state. Hua et al. [5] developed a mathematical modeling using the state merging method to analyses a rearranged Markov model to assess the reliability of the phased-mission system (PMS). Gao et al. [3] considered planar slider crank mechanism for two clearance joints to study the reliability sensitivity analysis and optimization design using the Monte Carlo method. Tahir et al. [18] demonstrated a model by incorporating thermal storage, heat pump and demand responses and showed that warm capacity and demand response improve the part of variable manageable force sources. Jindal et al. [6] analyzed the reliability of the plant comprises of one programmed screw-press bio-coal briquetting machine. The behavioral analysis of a washing unit in paper industry for system parameters was discussed by Kumar et al. [7] using the RPGT. Rajbala et al. [16] applied Markov birth-death process for the analysis of the EGR Air Exhaust Pipe (EAEP) manufacturing plant. Agrawal et al. [1] studied the profit analysis of a Water Treatment RO Plant is agreed out by utilizing the RPGT. Dahiya et al. [2] studied the Optimization Using Heuristic Algorithm in Pharmaceutical industry. In this paper, keeping in view the purpose of analyzing real existing industrial system model, a urea fertilizer system is considered.

In fact, Urea fertilizer manufacturing system is a complex type repairable engineering system involving high risk of machine/production failure. Machine/Production failure leads to high risks of economic & environmental loss and worker's safety. That's why sensitivity analysis of the same plant is discussed in the present research. The problem is solved using RPGT to analyze the system parameters. The results describing the system behavior is discussed qualitatively through graphs and tables.

II. Problem Description and Assumptions

I. System Description

The urea fertilizer manufacturing system comprise of nine subsystems connected in series named as Ammonia Making Section (A), Medium Pressure Section (B), Low Pressure Section (C), Pre-vacuum Section (D), Vacuum Section (E), Periling Section (F) and high pressure (P1), medium pressure (P2), low pressure units (P3) as shown in Figure 1.

The performance of the system is best when all units are good but it fails to work when any of the nine sub-systems fail.

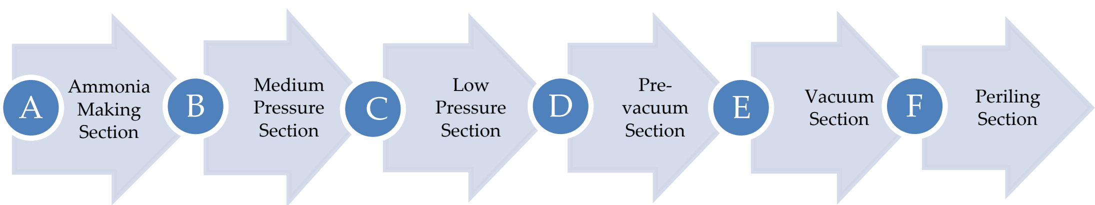


Figure 1: *Urea Fertilizer Making System Network*

II. Notations

$p_n(t)$ $(0 \le n \le 29)$	: Probability that the systems is in state S_n.at time t.
α_i $(1 \le i \le 6)$	: Subsystem's failure rates.
α_7, α_8, α_9	: Failure rate of pressure unit P_1, P_2 and P_3 respectively.
α_0	: Constant failure rate of entire system from any of its operative state.
β_i $(1 \le i \le 6)$	: Subsystem's repair rates.
H	: Repair rate of system failed due to pressure unit P_3
C	: Repair rate of system failed due to common cause failed.
a, b, c, d, e, and f	: Subsystem A, B, C, D, E, and F failed.
S_0	: Initial operative state of the system
S_{21}	: System's failed state due to the failure of pressure unit P_3.
S_2	: System's failed state due to the common cause failure.

III. Assumptions

- The single repair facility is available.
- Medium and low pressure can be obtained from high pressure unit by scientific logic.
- When system fails then only the pressure units will be repair one.

IV. State transition diagram

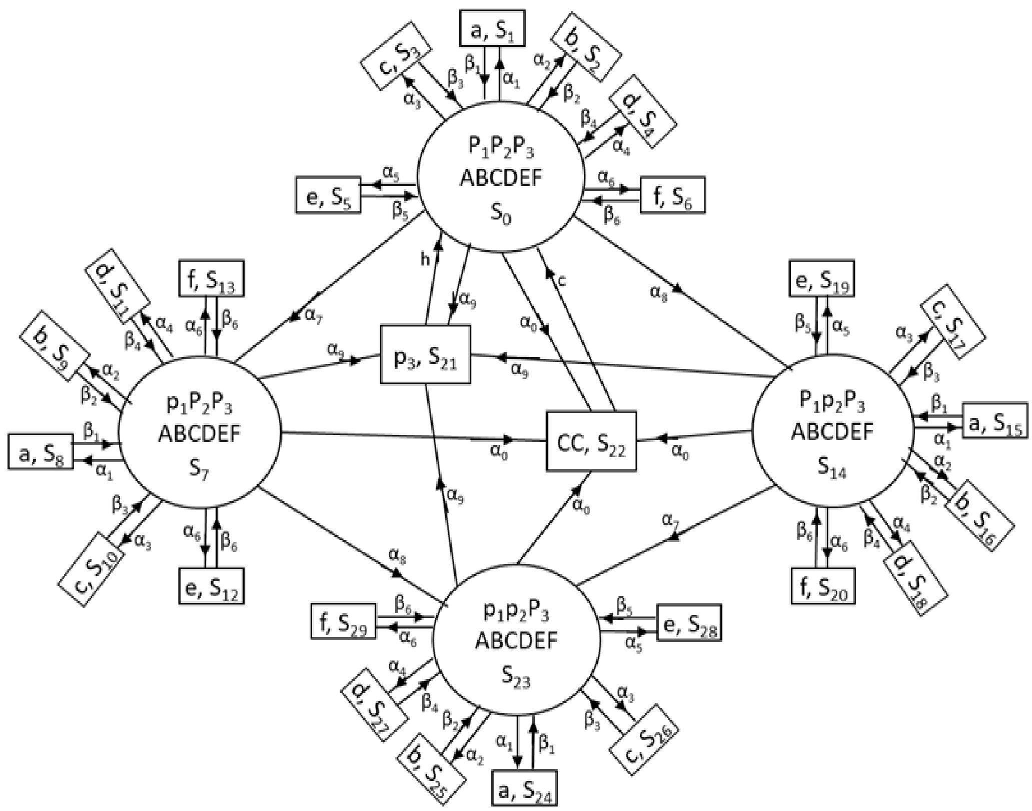


Figure 2: *Transition Diagram of System Design*

V. Transition Probabilities and Mean Sojourn Times (SMT)

Table 1 and Table 2 represents the Transition probabilities and MST for the states i, j respectively.

Table 1: *Transition Probabilities*

$q_{i,j}(t)$	$p_{ij} = q^*_{i,j}(0)$
$q_{0,i}(t) = \alpha_i e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_9+\alpha_8+\alpha_0)t}$	$p_{0,i} = \alpha_i/(\alpha_1+\alpha_5+\alpha_0+\alpha_4+\alpha_7+\alpha_8+\alpha_2+\alpha_9+\alpha_3+\alpha_6)$
$q_{0,14}(t) = \alpha_8 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_8+\alpha_9+\alpha_0)t}$	$p_{0,14} = \alpha_8/(\alpha_3+\alpha_5+\alpha_0+\alpha_4+\alpha_6+\alpha_8+\alpha_7+\alpha_9+\alpha_2+\alpha_1)$
$q_{0,21}(t) = \alpha_9 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_8+\alpha_9+\alpha_0)t}$	$p_{0,21} = \alpha_9/(\alpha_1+\alpha_5+\alpha_0+\alpha_4+\alpha_2+\alpha_8+\alpha_7+\alpha_9+\alpha_6+\alpha_3)$
$q_{0,22}(t) = \alpha_0 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_8+\alpha_9+\alpha_0)t}$	$p_{0,22} = \alpha_0/(\alpha_1+\alpha_5+\alpha_0+\alpha_4+\alpha_2+\alpha_8+\alpha_7+\alpha_9+\alpha_6+\alpha_3)$
Where i = 1 to 7	
$q_{i,0}(t) = \beta_i e^{-\beta_i t}$, $q_{7+i}(t) = \beta_{7+i} e^{-\beta_i t}$	$p_{7+i} = 1$, $p_{14+i} = 1$, $p_{21+i} = 1$
$q_{14+i}(t) = \beta_{14+i} e^{-\beta_i t}$, $q_{21+i}(t) = \beta_{21+i} e^{-\beta_i t}$	$p_{i,0} = 1$, where $1 \le i \le 6$
$q_{7,7+i}(t) = \alpha_i e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_8+\alpha_9)t}$	$p_{7,7+i} = \alpha_i/(\alpha_0+\alpha_9+\alpha_3+\alpha_8+\alpha_6+\alpha_5+\alpha_2+\alpha_4+\alpha_1)$
$q_{7,23}(t) = \alpha_8 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_8+\alpha_9)t}$	$p_{7,23} = \alpha_8/(\alpha_0+\alpha_9+\alpha_2+\alpha_4+\alpha_6+\alpha_5+\alpha_1+\alpha_8+\alpha_3)$
$q_{7,21}(t) = \alpha_9 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_8+\alpha_9)t}$	$p_{7,21} = \alpha_9/(\alpha_0+\alpha_9+\alpha_4+\alpha_8+\alpha_6+\alpha_5+\alpha_1+\alpha_3+\alpha_2)$
$q_{7,22}(t) = \alpha_0 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_8+\alpha_9)t}$	$p_{7,22} = \alpha_0/(\alpha_0+\alpha_9+\alpha_2+\alpha_1+\alpha_6+\alpha_5+\alpha_8+\alpha_4+\alpha_3)$
$q_{7+i,7}(t) = \beta_1 e^{-\beta_1 t}$	$p_{7+i,7} = 1$, where $1 \le i \le 6$
$q_{14,14+i}(t) = \alpha_i e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_9+\alpha_0)t}$	$p_{14,14+i} = \alpha_i/(\alpha_6+\alpha_2+\alpha_5+\alpha_4+\alpha_9+\alpha_1+\alpha_7+\alpha_0+\alpha_3)$
$q_{14,22}(t) = \alpha_0 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_9+\alpha_0)t}$	$p_{14,22} = \alpha_0/(\alpha_5+\alpha_2+\alpha_0+\alpha_4+\alpha_9+\alpha_3+\alpha_7+\alpha_6+\alpha_1)$
$q_{14,23}(t) = \alpha_7 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_9+\alpha_0)t}$	$p_{14,23} = \alpha_7/(\alpha_6+\alpha_2+\alpha_0+\alpha_5+\alpha_9+\alpha_1+\alpha_7+\alpha_4+\alpha_3)$
$q_{14,21}(t) = \alpha_9 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_9+\alpha_0)t}$	$p_{14,21} = \alpha_9/(\alpha_5+\alpha_2+\alpha_0+\alpha_4+\alpha_9+\alpha_1+\alpha_7+\alpha_6+\alpha_3)$
$q_{14+i,14}(t) = \beta_1 e^{-\beta_1 t}$, $q_{23+i,23}(t) = \beta_i e^{-\beta_i t}$	$p_{14+i,14} = 1$, $p_{23+i,23} = 1$, where $1 \le i \le 6$
$q_{21,0}(t) = \mathrm{h}e^{-ht}$, $q_{22,0}(t) = \mathrm{c}e^{-ct}$	$p_{21,0} = 1$, $p_{22,0} = 1$
$q_{23,22}(t) = \alpha_0 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_9)t}$	$p_{23,22} = \alpha_0/(\alpha_6+\alpha_2+\alpha_9+\alpha_4+\alpha_1+\alpha_3+\alpha_0+\alpha_6)$
$q_{23,23+i}(t) = \alpha_i e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_9)t}$	$p_{23,23+i} = \alpha_i/(\alpha_5+\alpha_2+\alpha_9+\alpha_1+\alpha_4+\alpha_6+\alpha_0+\alpha_3)$
$q_{23,21}(t) = \alpha_9 e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_9)t}$	$p_{23,21} = \alpha_9/(\alpha_6+\alpha_2+\alpha_9+\alpha_4+\alpha_1+\alpha_6+\alpha_0+\alpha_3)$

Table 2: *Mean Sojourn Times*

$R_i(t)$	$\mu_i = R_i^*(0)$
$R_0(\mathrm{t}) = e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_7+\alpha_8+\alpha_9+\alpha_0)t}$	$\mu_0 = 1/(\alpha_3+\alpha_2+\alpha_8+\alpha_1+\alpha_9+\alpha_6+\alpha_4+\alpha_7+\alpha_5+\alpha_0)$
$R_{k+i}(t) = e^{-\beta_i t}$ where $1 \le i \le 6$,	$\mu_i = 1/\beta_i$, where $1 \le i \le 6$
$R_j(\mathrm{t}) = e^{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6+\alpha_0+\alpha_8+\alpha_9)t}$	$\mu_j = 1/(\alpha_3+\alpha_2+\alpha_1+\alpha_6+\alpha_5+\alpha_4+\alpha_8+\alpha_0+\alpha_9)$
where j = 7, 14, 23	where j = 7, 14, 23
$R_{21}(t) = e^{-\mathrm{h}t}$, $R_{22}(t) = e^{-\mathrm{c}t}$	$\mu_{21} = 1/\mathrm{h}$, $\mu_{22} = 1/\mathrm{c}$

The following paragraphs outline meaning of parameters assessment, Availability of system, Expected fractional no. of inspection by repairman and busy period of server.

III. Evaluation of Path Probabilities

The change likelihood of all reachable states from base state 'ξ' = '0' are: Probabilities from state '0' to various vertices are given as

$V_{0,0} = 1,$ (1)

$V_{0,j} = (0,j) = p_{0,j}$; where $1 \le j \le 6$, (2)

$V_{0,7} = (0,7)/\{(1-L_1)(1-L_2)(1-L_3)(1-L_4)(1-L_5)(1-L_6)\}$ (3)

$V_{0,j} = (0,7,j)/\{(1-L_1)(1-L_2)(1-L_3)(1-L_4)(1-L_5)(1-L_6)(1-L_i)\}$; where $8 \le j \le 13$; $7 \le i \le 12$, (4)

$V_{0,14} = (0,14)/\{(1-L_{13})(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})\}$ (5)

$V_{0,j} = (0,14,j)/\{(1-L_{13})(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})(1-L_i)\}$; where $15 \le j \le 20$; $19 \le i \le 24$, (6)

$V_{0,21} = (0,21)+\{(0,14,21)/(1-L_{13})(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})\}+\{(0,7,21)$

$/(1-L_1)(1-L_2)(1-L_3)(1-L_4)(1-L_5)(1-L_6)\}\;)\}+\{(0,7,23,21)/(1-L_1)(1-L_2)(1-L_3)(1-L_4)$
$(1-L_5)(1-L_6)(1-L_{25})(1-L_{26})(1-L_{27})(1-L_{28})(1-L_{29})(1-L_{30})\}+\{(0,14,23,21)/(1-L_{13})$
$(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})(1-L_{25})(1-L_{26})(1-L_{27})(1-L_{28})(1-L_{29})(1-L_{30})\}$ (7)

$V_{0,22} = (0,22)+\{(0,14,22)/(1-L_{13})(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})\}+\{(0,7,22)$
$/(1-L_1)(1-L_2)(1-L_3)(1-L_4)(1-L_5)(1-L_6)\}+\{(0,7,23,22)/(1-L_1)(1-L_2)(1-L_3)(1-L_4)$
$(1-L_5)(1-L_6)(1-L_{25})(1-L_{26})(1-L_{27})(1-L_{28})(1-L_{29})(1-L_{30})\}$ (8)

$V_{0,23} = \{(0,7,23)/\,(1-L_1)(1-L_2)(1-L_3)(1-L_4)(1-L_5)(1-L_6)(1-L_{25})(1-L_{26})(1-L_{27})(1-L_{28})$
$(1-L_{29})(1-L_{30})\}+\{(0,14,23)/(1-L_{13})(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})(1-L_{25})$
$(1-L_{26})(1-L_{27})(1-L_{28})(1-L_{29})(1-L_{30})\}$ (9)

$V_{0,j} = \{(0,14,23,j)/\,(1-L_{13})(1-L_{14})(1-L_{15})(1-L_{16})(1-L_{17})(1-L_{18})(1-L_{25})(1-L_{26})$
$(1-L_{27})(1-L_{28})(1-L_{29})(1-L_{30})(1-L_i)\}+\{(0,7,23,24)//(1-L_1)(1-L_2)(1-L_3)(1-L_4)(1-L_5)(1-L_6)(1-L_{25})(1-L_{26})(1-L_{27})(1-L_{28})(1-L_{29})(1-L_{30})(1-L_i)\}$; where $24 \le j \le 29$; $31\le i \le 36$, (10)

Where L_i are cycles of level 1 and

$(1-L_j) = \{1-(7,i,7)\} = (1-p_{7,i}p_{i,7})$, where $1 \le j \le 6$; $8 \le i \le 13$, (11)
$(1-L_7) = \{1-(8,7,8)\} = (1-p_{8,7}p_{7,8})$ (12)
$(1-L_8) = \{1-(9,7,9)\} = (1-p_{9,7}p_{7,9})$ (13)
$(1-L_9) = \{1-(10,7,10)\} = (1-p_{10,7}p_{7,10})$ (14)
$(1-L_{10}) = \{1-(11,7,11)\} = (1-p_{11,7}p_{7,11})$ (15)
$(1-L_{11}) = \{1-(12,7,12)\} = (1-p_{12,7}p_{7,12})$ (16)
$(1-L_{12}) = \{1-(13,7,13)\} = (1-p_{13,7}p_{7,13})$ (17)
$(1-L_j) = \{1-(14,i,14)\} = (1-p_{14,i}p_{i,14})$; where $13 \le j \le 18$; $15 \le i \le 20$, (18)
$(1-L_{19}) = \{1-(15,14,15)\} = (1-p_{15,14}p_{14,15})$ (19)
$(1-L_{20}) = \{1-(16,14,16)\} = (1-p_{16,14}p_{14,16})$ (20)
$(1-L_{21}) = \{1-(17,14,17)\} = (1-p_{17,14}p_{14,17})$ (21)
$(1-L_{22}) = \{1-(18,14,18)\} = (1-p_{18,14}p_{14,18})$ (22)
$(1-L_{23}) = \{1-(19,14,19)\} = (1-p_{19,14}p_{14,19})$ (23)
$(1-L_{24}) = \{1-(20,14,20)\} = (1-p_{20,14}p_{14,20})$ (24)
$(1-L_j) = \{1-(23,i,23)\} = (1-p_{23,i}p_{i,23})$; where $25 \le j \le 30$; $24 \le i \le 29$, (25)
$(1-L_{31}) = \{1-(24,23,24)\} = (1-p_{24,23}p_{23,24})$ (26)
$(1-L_{32}) = \{1-(25,23,25)\} = (1-p_{25,23}p_{23,25})$ (27)
$(1-L_{33}) = \{1-(26,23,26)\} = (1-p_{26,23}p_{23,26})$ (28)
$(1-L_{34}) = \{1-(27,23,27)\} = (1-p_{27,23}p_{23,27})$ (29)
$(1-L_{35}) = \{1-(28,23,28)\} = (1-p_{28,23}p_{23,28})$ (30)
$(1-L_{36}) = \{1-(29,23,29)\} = (1-p_{29,23}p_{23,29})$ (31)

IV. Evaluation of System Parameters

The MTSF and other parameters are evaluated under steady-state conditions by using S_1 as the base state.

- Mean time to system failure (T0): Regenerative un-failed states to which the framework can travel (starting state '0'), Preceding entering any bombed state are: 'j' = 7, 0, 14, 23 taking 'ξ' = '0'.

$T_0=(V_{0,0}\mu_0+V_{0,7}\mu_7+V_{0,14}\mu_{14}+V_{0,23}\mu_{23})/(1-p_{0,7}p_{7,21}p_{21,0}-p_{0,7}p_{7,23}p_{23,21}p_{21,0}-p_{0,7}p_{7,23}p_{23,22}p_{22,0}-$
$p_{0,14}p_{14,21}p_{21,0}-p_{0,14}p_{14,22}p_{22,0}-p_{0,14}p_{14,23}p_{23,21}p_{21,0}-p_{0,14}p_{14,23}p_{23,22}p_{22,0})$ (32)

- Availability of System (A_0): The states at which the framework is accessible are 'j' = 0, 14, 7, 23 taking 'ξ' = '0' the all-out division of time for which framework is accessible is given by

$A_0 = \left[\sum_j V_{\xi,j}\,, f_j, \mu_j\right] \div \left[\sum_i V_{\xi,i}\,, f_j, \mu_i^1\right] = (V_{0,0}\mu_0+V_{0,7}\mu_7+V_{0,14}\mu_{14}+V_{0,23}\mu_{23})/D$ (33)

WhereD=$(V_{0,4}\mu_4+V_{0,2}\mu_2+V_{0,10}\mu_{10}+V_{0,8}\mu_8+V_{0,0}\mu_0+V_{0,3}\mu_3+V_{0,6}\mu_6+V_{0,9}\mu_9+V_{0,5}\mu_5+V_{0,7}\mu_7+V_{0,1}\mu_1+V_{0,13}$

$\mu_{13}+V_{0,12}\mu_{12}+V_{0,11}\mu_{11}+V_{0,14}\mu_{14}+V_{0,17}\mu_{17}+V_{0,16}\mu_{16}+V_{0,15}\mu_{15}+V_{0,18}\mu_{18}+V_{0,21}\mu_{21}+V_{0,20}\mu_{20}+V_{0,19}\mu_{19}+V_{0,22}\mu_{22}+V_{0,25}\mu_{25}+V_{0,24}\mu_{24}+V_{0,23}\mu_{23}+V_{0,26}\mu_{26}+V_{0,29}\mu_{29}+V_{0,28}\mu_{28}+V_{0,27}\mu_{27})$

- Busy Period of Server: States where server is busy are S_i, S_{7+i}, S_{14+i}, S_{23+i}, where $1 \le i \le 6$, S_{21}, S_{22} taking ξ = '0', the time server remains busy is

$$B_0=(V_{0,9}\mu_9+V_{0,4}\mu_4+V_{0,3}\mu_3+V_{0,11}\mu_{11}+V_{0,10}\mu_{10}+V_{0,1}\mu_1+V_{0,8}\mu_8+V_{0,6}\mu_6+V_{0,5}\mu_5+V_{0,13}\mu_{13}+V_{0,12}\mu_{12}+V_{0,2}\mu_2+V_{0,15}\mu_{15}+V_{0,18}\mu_{18}+V_{0,17}\mu_{17}+V_{0,16}\mu_{16}+V_{0,19}\mu_{19}+V_{0,22}\mu_{22}+V_{0,21}\mu_{21}+V_{0,20}\mu_{20}+V_{0,24}\mu_{24}+V_{0,27}\mu_{27}+V_{0,26}\mu_{26}+V_{0,25}\mu_{25}+V_{0,28}\mu_{28}+V_{0,29}\mu_{29})/D \quad (34)$$

- Expected Fractional Number of server visits by repairman: States where repairman do visit's a fresh are j =7, 14, 23 and S_i, S_{7+i}, S_{14+i}, S_{23+i}, where $1 \le i \le 6$, S_{21}, S_{23} taking 'ξ' = '0',

$$V_0 = (V_{0,7} +V_{0,14}+V_{0,21})/ (V_{0,1}\mu_2+V_{0,4}\mu_4+V_{0,3}\mu_3+V_{0,25}\mu_{25}+V_{0,10}\mu_{10}+V_{0,9}\mu_9+V_{0,8}\mu_8+V_{0,6}\mu_6+V_{0,5}\mu_5+V_{0,21}\mu_{21}+V_{0,24}\mu_{24}+V_{0,27}\mu_7+V_{0,15}\mu_{15}+V_{0,18}\mu_{18}+V_{0,17}\mu_{17}+V_{0,16}\mu_{16}+V_{0,29}\mu_{29}+V_{0,22}\mu_{22}+V_{0,12}\mu_{13}+V_{0,20}\mu_{20}+V_{0,14}\mu_{14}+V_{0,2}\mu_2+V_{0,26}\mu_{26}+V_{0,11}\mu_{11}+V_{0,28}\mu_{28}+V_{0,19}\mu_{19}) \quad (35)$$

V. Results

Particular cases of Sensitivity Analysis: Furthermore, the following paragraphs describe two Sensitivity Analysis cases and corresponding results in tabular and graphical forms.

Case 1: Sensitivity Analysis w. r. t. change in repair rates**.** Taking α_i = 0.1 ($0 \le i \le \alpha$) and varying β_1, β_2, β_3, β_4, β_5, β_6 one by one respectively at 0.75, 0.80, 0.85, 0.90, 0.95, 1.00.

Table 3: *MTSF (T_0)*

β_i	β_1	β_2	β_3	β_4	β_5	β_6	H	C
0.75	1.63964	1.63961	1.63963	1.63961	1.63960	1.63963	1.63964	1.63965
0.80	1.63965	1.63962	1.63964	1.63962	1.63961	1.63964	1.63965	1.63965
0.85	1.63966	1.63963	1.63965	1.63963	1.63962	1.63965	1.63966	1.63966
0.90	1.63967	1.63964	1.63967	1.63964	1.63963	1.63966	1.63966	1.63966
0.95	1.63968	1.63965	1.63968	1.63965	1.63964	1.63967	1.63967	1.63967
1.00	1.63969	1.63966	1.63969	1.63966	1.63965	1.63968	1.63967	1.63968

Table 4: *Availability of System (A_0)*

β_i	β_1	β_2	β_3	β_4	β_5	β_6	H	C
0.75	0.52072	0.51837	0.51631	0.51449	0.51288	0.51143	0.50065	0.50099
0.80	0.52309	0.52072	0.51865	0.51681	0.51516	0.51372	0.50284	0.50310
0.85	0.52521	0.52282	0.52072	0.51887	0.51723	0.51576	0.50479	0.50497
0.90	0.52710	0.52469	0.52258	0.52072	0.51907	0.51759	0.50654	0.50665
0.95	0.52881	0.52638	0.52426	0.52239	0.52072	0.51923	0.50811	0.50816
1	**0.53035**	0.52791	0.52578	0.52389	0.52222	0.52072	0.50953	0.50953

Table 5: *Busy Period of Server Visits (B_0)*

β_i	β_1	β_2	β_3	β_4	β_5	β_6	H	C
0.75	0.66957	0.67107	0.67237	0.67353	0.67455	**0.67547**	0.66773	0.66750
0.80	0.66807	0.66957	0.67089	0.67206	0.67310	0.67402	0.66628	0.66610
0.85	0.66673	0.66825	0.66957	0.67075	0.67179	0.67272	0.66498	0.66486
0.90	0.66553	0.66705	0.66839	0.66957	0.67063	0.67156	0.66382	0.66374
0.95	0.66444	0.66598	0.66733	0.66852	0.66957	0.67052	0.66278	0.66274
1	0.66346	0.66501	0.66637	0.66756	0.66862	0.66957	**0.66183**	**0.66183**

Table 6: *Expected Fractional Number of server visits by Repairman (V_0)*

β_i	β_1	β_2	β_3	β_4	β_5	β_6	H	C
0.75	0.49049	0.48823	0.48624	0.48449	0.48293	**0.48154**	0.48194	0.48327
0.80	0.49278	0.49049	0.48849	0.48672	0.48515	0.48375	0.48405	0.48531
0.85	0.49482	0.49252	0.49049	0.48719	0.48713	0.49572	0.48593	0.48712
0.90	0.49665	0.49432	0.49228	0.49049	0.48890	0.48747	0.48761	0.48875
0.95	0.49829	0.49595	0.49390	0.49210	0.49049	0.48906	0.48912	0.49021
1	0.49978	0.49743	0.49536	0.49355	0.49193	0.49040	0.49049	0.49153

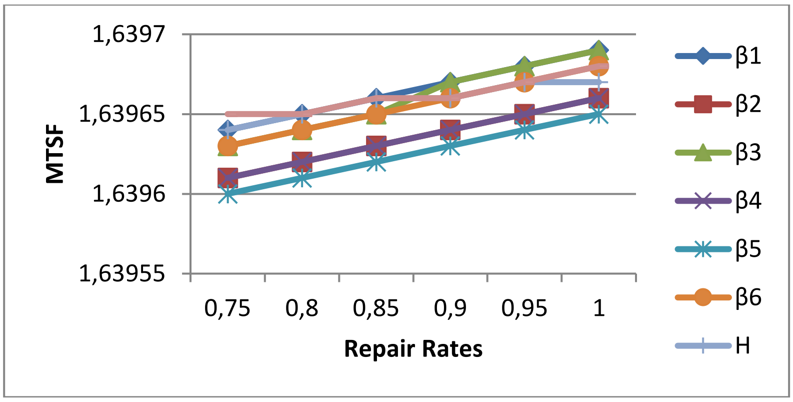


Figure 3: *Mean Time to System Failure*

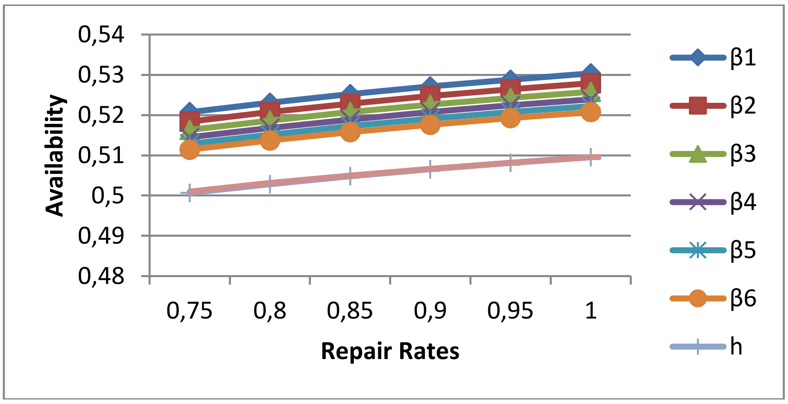


Figure 4: *Availability of System*

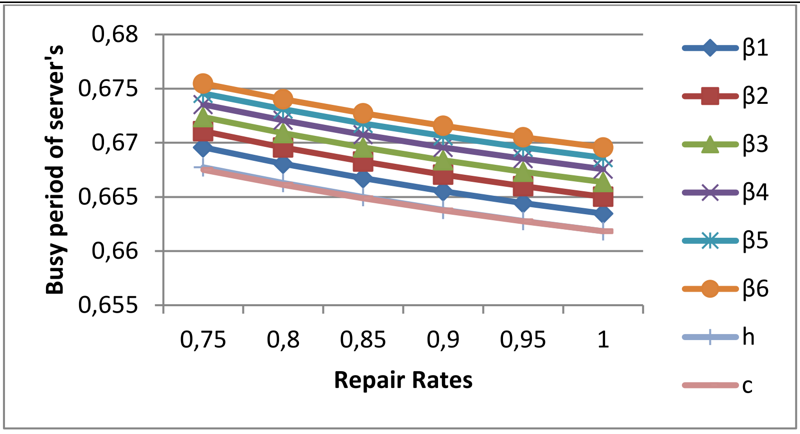


Figure 5: *Busy Period of the Server Visits*

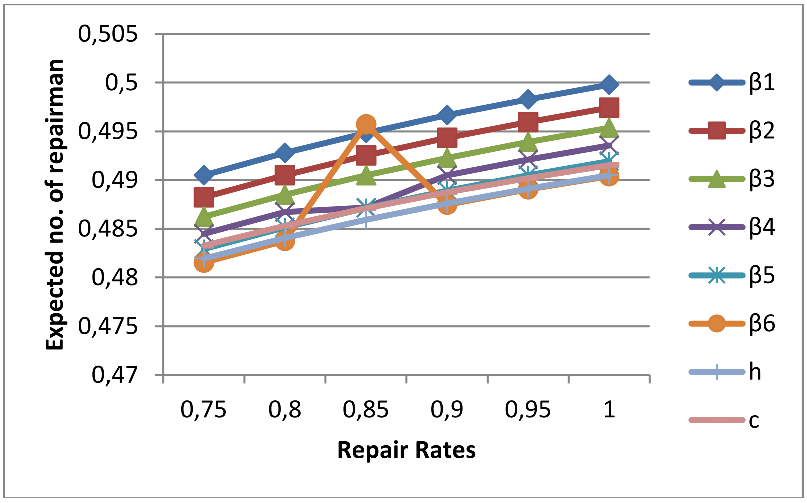


Figure 6: *Expected Fractional Number of server visits by Repairman*

Case 2: Now we consider Sensitivity Analysis case 2 with respect to change in failure rates: Fixing $\beta_i = 0.80$ $(0 \le i \le 6)$ h = 1, c = 1, $\alpha_1 = \alpha_6 = \alpha_5 = \alpha_4 = \alpha_3 = \alpha_2 = 0.01$; Taking $\alpha_i = 0.1, 0.2, 0.3, 0.4$ for i = 0, 7, 8, 9, we have

Table 7: *MTSF (T_0)*

α_i	α_0	α_6	α_7	α_8	α_9
0.1	2.88127	1.90761	3.26806	3.15184	**17.21485**
0.2	1.39829	1.35175	2.88127	3.02860	8.28176
0.3	1.32028	1.00665	2.59215	2.88127	4.64494
0.4	**0.86355**	0.74855	2.36740	2.62535	2.61925

Table 8: *Availability of System (A_0)*

α_i	α_0	α_6	α_7	α_8	α_9
0.1	0.63513	0.54051	0.77387	0.63730	**0.80021**
0.2	0.59134	0.47740	0.63513	0.63650	0.73625
0.3	0.56749	0.43190	0.62720	0.63513	0.64462
0.4	0.53866	**0.40780**	0.61278	0.63408	0.63513

Table 9: *Busy Period of Server Visits (B_0)*

α_i	α_0	α_6	α_7	α_8	α_9
0.1	0.67607	0.63074	0.57091	0.60098	**0.52002**
0.2	0.67782	0.67748	0.67607	0.65413	0.53951
0.3	0.67862	0.77832	0.71482	0.67607	0.61785
0.4	0.67980	**0.89412**	0.73151	0.70882	0.67607

Table 10: *Expected Fractional Number of Server visits by Repairman (V_0)*

α_i	α_0	α_6	α_7	α_8	α_9
0.1	0.25385	**0.13536**	0.23051	0.15806	0.20995
0.2	0.25412	0.21479	0.25385	0.17688	0.22616
0.3	0.26083	0.21648	0.25593	0.25385	0.22749
0.4	0.26222	0.21902	0.28552	0.25786	0.25385

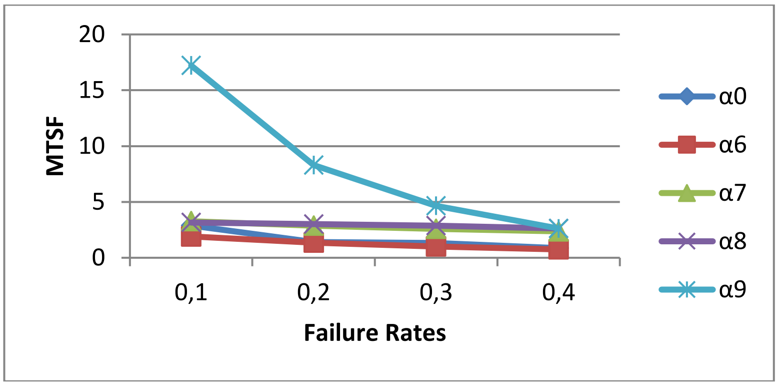


Figure 7: *MTSF*

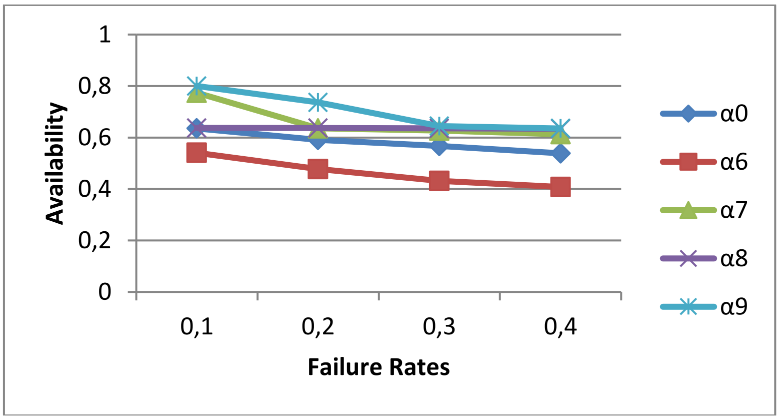


Figure 8: *Availability of System*

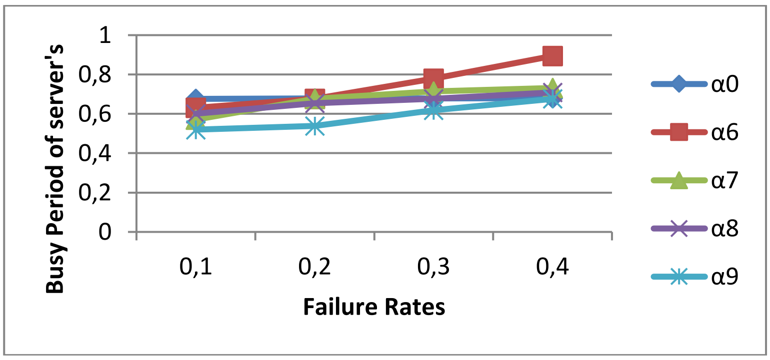


Figure 9: *Busy Period of Server Visits*

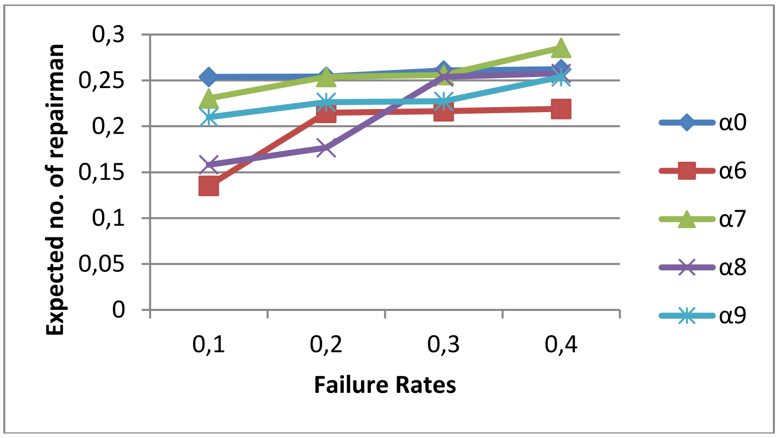


Figure 10: *Expected Fractional Number of server visits by Repairman*

VI. Discussion

Parameters related to sensitivity analysis for urea fertilizer plant are analyzed using RPGT. Effect of failure and repair rates on MTSF, availability of the system, busy period of the server, expected fractional number of server visit are discussed with the help of tables and graphs. Further from table 3 and figure 3, it observed that MTSF is independent of repair rates of various sub-units. From table 4 and figure 4, it is seen that availability increases with respect to repair rates. But there is no significance change in the value of availability of system while changing the value of repair rates. It is seen that for achieving the maximum value of A_0 repair rate of server should be maximum. For an operational system one has to minimize the busy period of the server to attain optimal level of production. It is seen from table 5 and figure 5, maximum value of repair rate of subunits leads to optimum value of the busy period. Moreover effect of repair rate of unit 'F' on the busy period of the server is more significance than other units. From table 6 and figure 6, it is seen that there is no significant change in the value of expected fraction number of server visits by repairman with the increase in repair rates of the subunits. From the table 7 and figure 7, MTSF is maximized when failure rate of higher pressure unit is minimum. MTSF is minimized when common cause failure rate is maximized. For optimum value of MTSF, failure rate of high pressure unit and common cause failure should be minimum. It is observed that availability is maximum when failure rates of high pressures unit and common cause failure rate is minimum. For an efficient system, availability should be highest, from above table 8 and Figure 8. From table 9 and figure 9, it is seen that busy period of the server increases by 62.36 % when failure rates of busy period increase from 0.1 to 0.4. From table 10 and figure 10, it is observed that the value of expected number of server's visits by repairman increased by 20.13 % when failure rates of the same and varying from 0.1 to 0.4.

VII. Conclusion

For urea fertilizer plant, In order to accomplish the ideal value of system parameters, administration may control the values of repair and failure rates of sub units. For the plant under consideration, the following conclusions are made from above research.

Case 1: Sensitivity Analysis with respect to change in repair rates (keeping failure rates constant).

- MTSF is independent of repair rates of all sub-units.
- Increase in repair rates does not have significant increase in the value of availability of system.
- In case of busy period of the server, effect of repair rate of unit 'F' is more significant as compared to other units. So repairman should be efficient in repairing the unit 'F' to minimize

the value of busy period of the server. Value of busy period is minimum when repair rate of pressure unit and common cause failure is maximum.

- No significant change in the value of expected fraction number of server visits by the repairman with change in value of repair rates of sub-units.

Case 2: Sensitivity Analysis with respect to change in failure rates (keeping repair rates constant).

- In order to have optimum value of MTSF, failure rate of high-pressure unit and common cause should be minimum.
- System availability is maximum when failure rate of high pressures unit and common cause failure rate are minimum. Availability is minimum when failure rate of sub- units are maximum.
- The optimum value of busy period is 0.52002 when the failure rate of high-pressure unit is minimum.
- The value of expected fraction number of server visits by the repairman with the increase in failure rates of the subunits.

The results obtain from above research are valuable for management to optimized the availability of plant, productions, and safety of workers. Last but not least, mathematical modeling utilizing in this paper is applicable to another manufacturing industries as well with suitable assumptions, and limitations.

References

[1] Agrawal, A., Garg, D., Kumar, A., & Kumar, R. (2021). Performance Analysis of the Water Treatment Reverse Osmosis Plant. *Reliability: Theory & Applications,* 16(3): 16-25.

[2] Dahiya, T., Garg, D., Devi, S., Kumar, R. (2021). Reliability Optimization Using Heuristic Algorithm In Pharmaceutical Plant. *Reliability: Theory & Applications*, 16(3): 195-205.

[3] Gao, Y., Zhang, F., and Li, Y. (2019). Reliability optimization design of a planar multi-body system with two clearance joints based on reliability sensitivity analysis. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 233(4):1369–1382.

[4] Garg, D., Kumar, K., and Sigh, J. (2010). Decision support system of a tab manufacturing plant. *Journal of Mechanical Engineering,* 41(1):71-79.

[5] 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.

[6] Jindal, S., Garg, R., Garg, T. K., and Garg, Y. (2019). Performance Modeling& Reliability Analysis of Demand and Supply Model. *Journal of Xi'an University of Architecture & Technology*, 11(12): 1009-1014.

[7] 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.

[8] Kumar, A., Goel, P. and Garg, D. (2018). Behaviour analysis of a bread making system. *International Journal of Statistics and Applied Mathematics,* 3(6): 56-61.

[9] Kumar, A., Goel, P., Garg, D., and Sahu A. (2017). System behavior analysis in the urea fertilizer industry. Book: Data and Analysis communications in computer and information Science, 1: 3-12.

[10] Kumar, D., & Singh, S. B. (2016). Stochastic analysis of complex repairable system with deliberate failure emphasizing reboot delay. *Communications in Statistics-Simulation and Computation,* 45(2): 583-602.

[11] Kumar, J. Kadyan, Malik, S. C. and Jindal, C. (2014). Reliability measures of a single unit system under preventive maintenance and degradation with arbitrary distributions of random variables. *Journals of reliability and statistical studies,* 7: 77 – 88.

[12] Lin., C. W. (2018). System reliability analysis of retrial machine repair systems with warm standby and a single server of working breakdown and recovery policy. *Journal of the international council on system engineering,* 15(1): 80-97.

[13] Liu, B. and Xie, L. (2020). An Improved Structural Reliability Analysis Method Based on Local Approximation and Parallelization. *Mathematics,* 8: 1-13.

[14] Mishra, S., Bhardwaj, P., and Bhadauria, N. (2016). Optimal availability analyses of break drum manufacturing system by using Markov approach. *International journal of engineering technology. Management and Applied Science,* 14(8): 32-37.

[15] Ni, P., Li, J., Hao, H., Yan, W., Du, X., and Zhou, H. (2020). Reliability analysis and design optimization of nonlinear structures. *Reliability Engineering & System safety,* 198: 1-13.

[16] Rajbala, Kumar, A. and Garg, D. (2019). Systems Modeling and Analysis: A Case Study of EAEP Manufacturing Plant. *International Journal of Advanced Science and Technology,* 28(14): 250-259.

[17] Ram, M., Manglik, M. (2016). Reliability measures analysis of an industrial system under standby modes and catastrophic failure. *International journal of operations research and information systems.*7(3): 36-37.

[18] Tahir, M. F., Haoyong, C., Mehmood, K., Ali, N. and Bhutto, J. N. (2019). Integrated Energy System Modeling of China for 2020 by Incorporating Demand Response. Heat Pump and Thermal Storage, *IEEE Access,* 7: 40095-40108.

Fractional Multi-objective Capacitated Transportation Problem with Different Membership Functions

[1]Sheema Sadia, [2]Qazi Mazhar Ali, [3]Zainab Asim, [4,*]Ahteshamul Haq

•

[1,2,4]Department of Statistics & Operations Research Aligarh Muslim University, Aligarh-202002
[3]Faculty of Commerce & Management, SGT University, Gurgaon, Haryana
[1]sadia.sheema63@gmail.com, [2]qaziali88@gmail.com, [3]asmizainab90@gmail.com,
[4,*]a.haq@myamu.ac.in *Corresponding author

Abstract

This Fractional Transportation Problem arises when an enterprise has to face the issue of maintaining a good ratio of some critical parameters. These parameters are directly concerned with product(s) transportation from sources to destination. This paper considers a multi-objective Capacitated Transportation Problem with Fractional Objectives. A fuzzy goal programming approach with different membership functions is applied to generate a different set of solutions. We also use Chebyshev's Goal Programming for obtaining the solutions. Finally, a numerical illustration is provided to validate our proposed model.

Keywords: Multi-objective programming, Quadratic membership function, Mixed constraints, Fuzzy normal membership function, Fuzzy Cauchy membership function, Fuzzy programming

1. Introduction

A transportation problem (TP) occurs when a product (or products) must be transported from multiple sources (also known as origin, supply, or capacity centres) to multiple sinks (also called destination, demand or requirement centres). The fundamental TP was devised by Alfred Hitchcock [9]. The TP with fractional objective function is known as a fractional transportation problem (FTP). Swarup [14] was the first to propose it. It is crucial in logistics, supply chain management, stock cutting problems, resource allocation problems, ship and plane routing problems, cargo loading problems, and inventory problems. The FTP arises when an enterprise faces the challenge of maintaining a good ratio between critical parameters. These parameters are directly concerned with transporting a product or products from their origin to their destination. Fractional programming can optimize actual/standard transportation costs or total return/total investment on machines delivered from factories to workshops. In linear fractional TPs with mixed constraints, Gupta *et al.* [5] presented a paradox. Gupta and Arora [6-7] and Liu [10] are two other authors written about FTPs. In general, real-world TPs are modelled with multiple, conflicting objectives.

Furthermore, combining all objective functions into a single overall utility function is difficult for the decision-maker. So it is better to formulate a multi-objective TP. The capacitated TP are the TPs with bounds on general availabilities at assets and general vacation spot requirements. It may benefit telecommunication networks, production-distribution systems, rail and concrete street systems.

The capacitated TPs have also been discussed by authors like Arora and Gupta [2] and Gupta and Bari [8]. Zadeh [15] first delivered the idea of a fuzzy set concept. Then Zimmermann [16] carried out the fuzzy set concept with a few suitable membership functions to resolve linear programming problems with numerous goal functions. Bit *et al.* [3] implemented a fuzzy programming approach with a linear membership function to resolve the multi-objective TP. El-Wahed [4] gave the idea of a fuzzy programming approach to determine the optimal compromise solution of a MOTP with a fuzzy membership function. Akkapeddi [1] discussed the quadratic membership for the multi-objective TP. Singh [13] worked on multiple objective fractional costs TP with bottleneck time and impurities. Sadia *et al.* [12] presented a fuzzy approach to obtain the solution of multi-objective capacitated FTP. Fuzzy normal and fuzzy Cauchy membership functions were used by Mon and Cheng [11]. Gupta *et al.* [17] discussed two stage transportation problem with the different types of fuzzy environments. Kamal *et al.* [18] described the parameters estimation and goodness of fit for the multi-objective transportation problem under type-2 trapezoidal fuzzy numbers. The purpose of using FTP is to make the problem more realistic. It can prove more beneficial for the decision-maker to consider the proportion of transporting cost due to the covered path and favoured path because the transportation cost may vary due to the travelled and favoured path. Likewise, the proportion of exact and standard transportation time and transporting damage cost due to covered path and favoured path are also measured.

In this paper, we have taken mixed constraints of MOCFTP with fractional type objectives. As it is a multi-objective problem and the objectives are conflicting in nature. So a compromise solution is obtained by using the fuzzy programming approach. We have tried to use three membership functions: quadratic, fuzzy normal, and fuzzy Cauchy. As far as our knowledge, these membership functions have never been used to deal with TPs. The rest of the paper is organized as follows: Assumptions, notations and formulation are discussed in Section 2. In Section 3, we have discussed the algorithm using a fuzzy optimization approach with different membership functions and Chebyshev's Goal Programming. In section 4, an example of the proposed method is illustrated. The conclusion is presented in section 5.

2. Assumptions, notations and formulation of the problem

We consider mixed constraints MOCFTP under the following notations

2.1 Notations

m	Number of origins
n	Number of destinations
a_i	Units of supply $(i = 1,2,\dots,m)$
b_j	Units of demands $j = 1,2,\dots,n$
c_{ij}	Unit transporting cost due to travelled route from the i^{th} starting point to j^{th} endpoint.
r_{ij}	Unit transporting cost due to preferred route from the i^{th} starting point to j^{th} endpoint
t_{ij}^{a}	Actual transportation time from the i^{th} starting point to j^{th} endpoint
t_{ij}^{s}	Standard transportation time from the i^{th} starting point to j^{th} endpoint
d_{ij}	Damage transporting cost due to travelled origin from the i^{th} path to j^{th} endpoint
x_{ij}	Units transported from the i^{th} starting point to j^{th} endpoint
l_{ij}	Minimum quantity transported from the i^{th} starting point to j^{th} endpoint
s_{ij}	Maximum transported quantity from the i^{th} starting point to j^{th} endpoint

2.2 Problem's statement

Consider a TP of fractional type objective function with m numbers of starting points having $a_i (i = 1,2,\dots,m)$ units of supply to be transported among n numbers of endpoints with $b_j (j = 1,2,\dots,n)$ units of demand. The problem is determining the best transportation schedule for transporting the available quantity of products to meet demand while minimizing total transportation costs, time, and damage charges.

Let x_{ij} be the number of units transported from *i^th^* starting point to the *j^th^* endpoint. The mathematical model of the MOCFTP with mixed constraints can be expressed as:

$$\text{Minimize} f_1 = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n} r_{ij}x_{ij}}$$

$$\text{Minimize} f_2 = max\left\{\frac{t_{ij}^{a} \middle| x_{ij} > 0}{t_{ij}^{s} \middle| x_{ij} > 0}\right\}$$

$$\text{Minimize} f_3 = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} d_{ij}x_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n} r_{ij}x_{ij}}$$

$$\text{subject to:} \sum_{i=1}^{m} x_{ij} \leq a_i\,; \sum_{j=1}^{n} x_{ij} \geq b_j$$

$$l_{ij} \leq x_{ij} \leq s_{ij}; x_{ij} \geq 0$$

2.3 Interpretation of objectives function

1. The proportion of unit transporting cost c_{ij} and r_{ij} due to travelled path and a preferred route respectively.

2. The proportion of the actual transportation time t_{ij}^{a} and a standard transportation time t_{ij}^{s}.

3. The proportion of unit transporting damage cost d_{ij} (loss of quantity and quality transportation) and r_{ij} due to the travelled and a preferred path, respectively.

3. Solution Approach for MOCFTP

3.1 Fuzzy Optimization Approach-Algorithm:

In order to solve the multiobjective fractional capacitated TP with mixed constraints, we use the following algorithm

Step 1:- Firstly, we will formulate the payoff matrix as:-

$$\text{Payoff Matrix} = \begin{array}{c} \\ x_{ij}^{(1)} \\ x_{ij}^{(2)} \\ \vdots \\ x_{ij}^{(k)} \end{array} \begin{array}{c} \begin{array}{ccc} f_1 & f_2 & f_3 \end{array} \\ \begin{bmatrix} f_1(x_{ij}^{(1)}) & f_1(x_{ij}^{(1)}) & f_1(x_{ij}^{(1)}) \\ f_1(x_{ij}^{(1)}) & f_1(x_{ij}^{(1)}) & f_1(x_{ij}^{(1)}) \\ \vdots & \vdots & \vdots \\ f_1(x_{ij}^{(1)}) & f_1(x_{ij}^{(1)}) & f_1(x_{ij}^{(1)}) \end{bmatrix} \end{array}$$

where, $x_{ij}^{(k)}; k = 1,2,\dots,K$ are the *k^th^* individual optimal solutions that optimize the *k^th^* objective.

Step 2:- We will apply the fuzzy approach with the following membership functions defined below:

A. ***Quadratic membership function:*** To derive the compromise solution of MOCFTP, we used fuzzy programming. The membership functions for the cost objective are: f_{kl} and f_{ku} be the achieved aspired level and the highest acceptance level of the k^{th} objective function, respectively. The membership function of the k^{th} objective function is represented as follows:

$$\mu_k(f_k) = q_{k1}f_k^2 + q_{k2}f_k + q_3$$

The membership values of the k^{th} objective function at the aspiration level and the highest acceptable level is 1 and 0, respectively. We used the equations as:

$$\mu_k(f_{kl}) = q_{k1}f_{kl}^2 + q_{k2}f_{kl} + q_3 = 1$$

$$\mu_k(f_{ku}) = q_{k1}f_{ku}^2 + q_{k2}f_{ku} + q_3 = 0$$

The above linear system of equations q_{k2} and q_{k3} are expressed in terms of q_{k1}. Thus, the quadratic membership function for the k^{th} objective function is used in the following form:

$$\mu_k^Q(F_k) = \frac{f_{ku}-f_k}{f_{ku}-f_{kl}} + q_{k1}f_k^2 - q_{k1}(f_{kl}+f_{ku})f_k + q_{k1}f_{lk}f_{lu}$$

B. ***Fuzzy Normal:*** The membership function for Fuzzy Normal will take the following form:

$$\mu_k^{FN}\{F_k\} = \begin{cases} 1i & ff_k \le f_{kl} \\ exp\left[-k\left(\frac{f_k-f_{kl}}{f_{ku}-f_{kl}}\right)^2\right] if f_{kl} < f_k < f_{ku} & \\ 0i & ff_k \ge f_{ku} and k \ge 0 \end{cases}$$

C. ***Fuzzy Cauchy:*** The membership function for Fuzzy Cauchy will take the following form:

$$\mu_k^{FC}\{F_k\} = \begin{cases} 1i & ff_k \le f_{kl} \\ \frac{1}{1+\alpha\left(\frac{f_k-f_{kl}}{f_{ku}-f_{kl}}\right)^\beta} if f_{kl} < f_k < f_{ku} & \\ 0i & ff_k \ge f_{ku} \end{cases}$$

$$\alpha \ge 0 \text{and} \beta \text{ispositiveeven}$$

Step 3: The MOCFTP with mixed constraints can now be converted into equivalent non-linear models for the above-defined membership functions as follow:

A. ***Quadratic Membership function***: The proposed model for MOCFTP with mixed constraints on applying quadratic membership function will be of the following form:

$$\begin{aligned} &\text{Minimize}\lambda \\ &\quad \text{Subjectto} \\ &\quad \frac{f_{1u}-f_1}{f_{1u}-f_{1l}} + q_{11}f_1^2 - q_{11}(f_{11}+f_{1u})f_1 + q_{11}f_{1u}f_{1l} \le \lambda \\ &\quad \frac{f_{2u}-f_2}{f_{2u}-f_{2l}} + q_{21}f_2^2 - q_{21}(f_{2l}+f_{2u})f_2 + q_{21}f_{2u}f_{2l} \le \lambda \\ &\quad \frac{f_{3u}-f_k}{f_{3u}-f_{3l}} + q_{31}f_3^2 - q_{31}(f_{3l}+f_{3u})f_3 + q_{31}f_{3u}f_{3l} \le \lambda \\ &\quad \sum_{i=1}^{m} x_{ij}\{\le/=/\ge\}a_i; \sum_{j=1}^{n} x_{ij}\{\le/=/\ge\}b_j \\ &\quad 0 \le x_{ij} \le s_{ij}; x_{ij} \ge 0; \lambda \ge 0 \end{aligned}$$

B. ***Fuzzy Normal Membership Function:*** The proposed model for MOCFTP with mixed constraints on applying fuzzy normal membership function will be of the form:

$$\text{Minimize } \lambda$$
$$\text{subject to:}$$
$$exp\left[-k\left(\frac{f_1 - f_{1l}}{f_{1u} - f_{1l}}\right)^2\right] \leq \lambda, exp\left[-k\left(\frac{f_2 - f_{2l}}{f_{2u} - f_{2l}}\right)^2\right] \leq \lambda$$
$$exp\left[-k\left(\frac{f_3 - f_{3l}}{f_{3u} - f_{3l}}\right)^2\right] \leq \lambda$$
$$\sum_{i=1}^{m} x_{ij} \leq a_i; \sum_{j=1}^{n} x_{ij} \geq b_j\,; 0 \leq x_{ij} \leq s_{ij}; x_{ij} \geq 0; \lambda \geq 0$$

We will solve it for k=1

C. ***Fuzzy Cauchy Membership Function*****:** The proposed model for MOCFTP with mixed constraints on applying fuzzy Cauchy membership function will be of the form:

$$\text{Minimize}\lambda$$
$$\text{Subjectto} \frac{1}{1 + \alpha\left(\frac{f_1 - f_{1l}}{f_{1u} - f_{1l}}\right)^{\beta}} \leq \lambda$$
$$\frac{1}{1 + \alpha\left(\frac{f_2 - f_{2l}}{f_{2u} - f_{2l}}\right)^{\beta}} \leq \lambda$$
$$\frac{1}{1 + \alpha\left(\frac{f_3 - f_{3l}}{f_{3u} - f_{3l}}\right)^{\beta}} \leq \lambda$$
$$\textstyle\sum_{i=1}^{m} x_{ij} \leq a_i; \sum_{j=1}^{n} x_{ij} \geq b_j\,; 0 \leq x_{ij} \leq s_{ij}; x_{ij} \geq 0; \lambda \geq 0$$

D. ***Chebyshev's Goal Programming:*** Chebyshev's Goal Programming is considered a particular case of the weighted Goal Programming technique. It seeks a solution that minimizes the worst deviation from each objective. The mixed constraints of the MOCFTP using Chebyshev's Goal Programming will be represented as:

$$\text{Maximize}\lambda$$
$$\text{Subjectto} f_1 + \lambda \leq f_{1u}$$
$$f_2 + \lambda \leq f_{2u}$$
$$f_3 + \lambda \leq f_{3u}$$
$$\sum_{i=1}^{m} x_{ij}\{\leq/=/\geq\}a_i; \sum_{j=1}^{n} x_{ij}\{\leq/=/\geq\}b_j$$
$$0 \leq x_{ij} \leq s_{ij}; x_{ij} \geq 0; \lambda \geq 0$$

where, the worst deviation level (λ) and aspiration levels for the upper bound is $f_{iu}(i = 1,2,3)$.

4. Numerical Illustration

A case study is discussed to demonstrate and utility of the approaches. The numerical problem of simulated data (Sadia *et al.* [12]) is presented. The discussed models are defined to solve the problem. We consider three starting points and three endpoints. The fractional transportation cost, time and damage charges (both quantity and quality damage) are represented in Table [1-3].

Table 1: Cost charges matrix

	b_1	b_2	b_3	Supply
a_1	5/3	7/4	15/13	≤ 12
a_2	8/12	17/14	12/7	= 15
a_3	19/15	10/6	13/8	≥ 20
Demand	≥ 9	= 13	≤ 21	

Table 2: Time charges matrix

	b_1	b_2	b_3	Supply
a_1	17/9	5/2	10/3	**≤ 12**
a_2	1/2	11/4	6/5	**= 15**
a_3	13/8	16/12	10/11	**≥ 20**
Demand	**≥ 9**	**= 13**	**≤ 21**	

Table 3: Damage charges matrix

	b_1	b_2	b_3	Supply
a_1	13/8	15/9	8/11	**≤ 12**
a_2	11/15	14/6	19/7	**= 15**
a_3	9/7	15/6	8/17	**≥ 20**
Demand	**≥ 9**	**= 13**	**≤ 21**	

The mixed constraints of the MOCFTP will be as follows:

$$\text{Min} f_1 = \frac{5x_{11}+7x_{12}+15x_{13}+8x_{21}+17x_{22}+12x_{23}+19x_{31}+10x_{32}+13x_{33}}{3x_{11}+4x_{12}+13x_{13}+12x_{21}+14x_{22}+7x_{23}+15x_{31}+6x_{32}+8x_{33}}$$

$$\text{Min} f_2 = \frac{13x_{11}+15x_{12}+8x_{13}+15x_{21}+14x_{22}+19x_{23}+9x_{31}+15x_{32}+8x_{33}}{8x_{11}+9x_{12}+11x_{13}+15x_{21}+6x_{22}+7x_{23}+7x_{31}+6x_{32}+17x_{33}}$$

$$\text{Min} f_3 = \frac{17x_{11}+5x_{12}+10x_{13}+x_{21}+11x_{22}+6x_{23}+13x_{31}+16x_{32}+10x_{33}}{9x_{11}+2x_{12}+3x_{13}+2x_{21}+4x_{22}+5x_{23}+8x_{31}+12x_{32}+11x_{33}}$$

$$\text{Subjectto} \sum_{j=1}^{3} x_{1j} \le 12; \sum_{j=1}^{3} x_{2j} \le 15\,; \sum_{j=1}^{3} x_{3j} \le 20$$

$$\sum_{j=1}^{3} x_{i1} \le 9; \sum_{j=1}^{3} x_{i2} \le 13\,; \sum_{j=1}^{3} x_{i3} \le 21\; 0 \le x_{11} \le 6, 0 \le x_{12} \le 7, 0 \le x_{13} \le 13, 0 \le x_{21} \le 6,$$

$$0 \le x_{22} \le 2, 0 \le x_{23} \le 13, 0 \le x_{31} \le 4, 0 \le x_{32} \le 7, 0 \le x_{33} \le 14.$$

*A. **Different membership functions for fuzzy programming approach***

The payoff matrix for $[l_{ij} = 0]$ is obtained after solving the problem as a single objective (ignoring the other objectives) using the LINGO optimization software will be as follows:

Payoff Matrix = $f_1 \quad f_2 \quad f_3$

$$\begin{array}{c} x_{ij}^{(1)} \\ x_{ij}^{(2)} \\ x_{ij}^{(3)} \end{array} \begin{bmatrix} 1.316832 & 1.16129 & 1.34472 \\ 1.37988 & 1.068410 & 1.79661 \\ 1.406433 & 1.170886 & 1.168285 \end{bmatrix}$$

$$f_{1u} = 1.406433, f_{1l} = 1.316832, f_{2u} = 1.170886,$$
$$f_{2l}^2 = 1.068410, f_{3u} = 1.79661 \text{and } f_{3l} = 1.168285$$

Individual optimum solutions are obtained by solving the above problem separately for each objective using the optimizing software LINGO in Table 4.

Table 4: Individual optimum solution

	Objectives	**Cost**	**Damage**	**Time**
	Objective Values	***1.316832***	***1.068410***	***1.168285***
Allocations	x_{11}	0	0	0
	x_{12}	4	7	6
	x_{13}	5	0	0
	x_{21}	2	6	6
	x_{22}	6	2	0
	x_{23}	7	7	9
	x_{31}	4	3	3
	x_{32}	7	4	7
	x_{33}	9	14	12

The compromise solution obtained for **Quadratic Membership Function** is as follows: $x_{11}^* = 0, x_{12}^* = 7, x_{13}^* = 2, x_{21}^* = 6, x_{22}^* = 2, x_{23}^* = 7, x_{31}^* = 4, x_{32}^* = 4, x_{33}^* = 12$

The optimal compromise solution obtained using the **Fuzzy normal Membership Function** will be as follows: $x_{11}^* = 0, x_{12}^* = 4, x_{13}^* = 1, x_{21}^* = 5, x_{22}^* = 2, x_{23}^* = 8, x_{31}^* = 4, x_{32}^* = 7, x_{33}^* = 9$

The crisp problem for fuzzy Cauchy has been obtained after setting $\alpha = 0.5$and$\beta = 2$. The compromise solution obtained for **Fuzzy Cauchy Membership Function** is as follows: $x_{11}^* = 4, x_{12}^* = 4, x_{13}^* = 4, x_{21}^* = 5, x_{22}^* = 2, x_{23}^* = 8, x_{31}^* = 4, x_{32}^* = 7, x_{33}^* = 9$

The compromise solution obtained for **Chebyshev's Goal Programming** is as follows: $x_{11}^* = 0, x_{12}^* = 7, x_{13}^* = 2, x_{21}^* = 6, x_{22}^* = 2, x_{23}^* = 7, x_{31}^* = 4, x_{32}^* = 4, x_{33}^* = 12$

5. Conclusion

This article represents the optimal compromise solution with mixed constraints for a multiobjective fractional capacitated TP. Fuzzy programming with three different membership functions viz. quadratic, fuzzy normal and fuzzy Cauchy is used to obtain a compromise solution using a fuzzy programming approach, and Chebyshev's Goal Programming is also discussed to solve the problem multiobjective fractional capacitated TP. Finally, a comparative study is done with the results obtained in the paper and the results from Sadia *et al.* [12]. The results are summarized in Table 5.

This paper proposes fuzzy programming models by applying different membership functions to solve multiobjective fractional capacitated TP. Table 5 also compares the results obtained through different procedures to obtain and compare their efficiency. The methods used in the paper can also be applied for transportation, assignment and transhipment problems.

Table 5: Compromise optimum solution

	Approach	**Membership/ Methods**	**Objective Values**		
			Cost	***Damage Charges***	***Time***
Sadia *et al.* [12]	Fuzzy Programming	Linear	1.359296	1.238494	1.389058
		Exponential	**1.349030**	1.229213	1.314935
		Hyperbolic	1.359296	1.238494	1.389058
	Goal Programming	Lexicographic, D_1distance	1.353103	1.129854	**1.237344**
Discussed method	Fuzzy Programming	Quadratic	1.349869	**1.097561**	1.257840
		Fuzzy Normal	1.371758	1.249359	1.264085
		Fuzzy Cauchy	1.359296	1.238494	1.389058
	Goal Programming	Chebyshev's	1.349869	**1.097561**	1.257840

References

[1] Akkapeddi, S. M. (2015). Fuzzy programming with quadratic membership functions for multi-objective transportation problem. *Pakistan Journal of Statistics and Operation Research*, 231-240. https://doi.org/10.18187/pjsor.v11i2.966

[2] Arora, S. R., & Gupta, K. (2012). An algorithm for solving a capacitated fixed charge bi-criterion indefinite quadratic transportation problem with restricted flow. *International Journal Of Research In IT, Management and Engineering (ISSN 2249-1619)*, *1*(5), 123-140.

[3] Bit, A. K., Biswal, M. P., & Alam, S. S. (1992). Fuzzy programming approach to multicriteria decision making transportation problem. *Fuzzy sets and systems*, *50*(2), 135-141. https://doi.org/10.1016/0165-0114(92)90212-M

[4] Abd El-Wahed, W. F. (2001). A multi-objective transportation problem under fuzziness. *Fuzzy sets and systems*, *117*(1), 27-33. https://doi.org/10.1016/S0165-0114(98)00155-9

[5] Gupta, A., Khanna, S., & Puri, M. C. (1993). A paradox in linear fractional transportation problems with mixed constraints. *Optimization*, *27*(4), 375-387. https://doi.org/10.1080/02331939308843896

[6] Gupta, K., & Arora, R. (2019). Optimum cost-time trade-off pairs in a fractional plus fractional capacitated transportation problem with restricted flow. *Investigación Operacional*, *40*(1), 46-60.

[7] Gupta, K., & Arora, S. R. (2012). Paradox in a fractional capacitated transportation problem. *International journal of research in IT, management and engineering*, *2*(3), 43-64.

[8] Gupta, N., & Bari, A. (2014). Fuzzy multi-objective capacitated transportation problem with mixed constraints. *Journal of Statistics Applications & Probability Letters*, *3*(2), 1-9. http://dx.doi.org/10.12785/jsap/JSAP021314S

[9] Hitchcock, F. L. (1941). The distribution of a product from several sources to numerous localities. *Journal of mathematics and physics*, *20*(1-4), 224-230. https://doi.org/10.1002/sapm1941201224

[10] Liu, S. T. (2016). Fractional transportation problem with fuzzy parameters. *Soft computing*, *20*(9), 3629-3636. https://doi.org/10.1007/s00500-015-1722-5

[11] Mon, D. L., & Cheng, C. H. (1994). Fuzzy system reliability analysis for components with different membership functions. *Fuzzy sets and systems*, *64*(2), 145-157.

[12] Sadia, S., Gupta, N., & Ali, Q. M. (2016). Multi-objective capacitated fractional transportation problem with mixed constraints. *Mathematical Sciences Letters*, *5*(3), 235-242.

[13] Singh, P. (2015). Multiple–objective fractional costs transportation problem with bottleneck time and impurities. *Journal of Information and Optimization Sciences*, *36*(5), 421-449.

[14] Swarup, K. (1966). Transportation technique in linear fractional functional programming. *Journal of royal naval scientific service*, *21*(5), 256-260.

[15] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.

[16] Zimmermann, H. J. (1978). Fuzzy programming and linear programming with several objective functions. *Fuzzy sets and systems*, *1*(1), 45-55.

[17] Gupta, S., Haq, A., and Ali, I.(2019) Two-stage transportation in different types of fuzzy environment: an industrial case study. *International Journal of Mathematics and computation*, 30(3), 94-118.

[18] Kamal, M., Alarjani, A., Haq, A., Yusufi, F. N. K., & Ali, I. (2021). Multi-objective transportation problem under type-2 trapezoidal fuzzy numbers with parameters estimation and goodness of fit. *Transport*, *36*(4), 317-338. https://doi.org/10.3846/transport.2021.15649

Reliability of Gas Insulated System under Electric Field Stress with Optimal Design of FGM Post Type Spacer

Akanksha Mishra [1], G. V. Nagesh Kumar[2*], D. Deepak Chowdary[3] and B. Sravana Kumar[4]

•

[1]Department of Electrical and Electronics Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, India; misakanksha@gmail.com
[2]Department of Electrical and Electronics Engineering, JNTUA CE Pulivendula, A.P, India; drgvnk14@gmail.com
[3]Department of Electrical and Electronics Engineering, Dr. L Bullayya College of Engineering, A.P, India; duvvada27@gmail.com
[4]Department of Electrical and Electronics Engineering, GITAM University, Visakhapatnam, A.P, India; sravanbali@gmail.com
*Corresponding Author

Abstract

Gas Insulated Substation (GIS) is essential for the transmission and control of power both in AC and DC electrical systems. Functionally Graded Material (FGM) technology is widely used for the design of the spacer material in the GIS to reduce the electric stress in the system. Optimal designing of the material of the spacer gradings with a particular attention to the number of gradings may prove to be very useful in reduction of the stress in the GIS at an effective cost. This paper deals with the design and development of an optimal dielectric material for the functionally graded material (FGM) spacer in a GIS. A novel optimization method has been proposed which is used for the optimization of the conductor material and the FGM epoxy spacer. The optimal value for each grading of functionally graded material spacer is determined by the proposed method. A dual-objective function is chosen for the optimization problem. The objective of the problem is to minimize the maximum field stress in addition to the standard deviation in the electric field. A post type spacer has been considered for the study. Initially, the optimization of the dielectric material is done only for 4 gradings. Gradually, the number of gradings in the FGM-spacer is increased to determine the optimal number of gradings suitable for the design.

Keywords: HVDC, Gas Insulated Substation, Functionally Graded Material, Optimization

I. Introduction

In the modern industrial world, electric power systems may easily be designated as the backbone of the world economy. A reliable power system is therefore a basic necessity in the present century. The ever-increasing load on the transmission system makes the goal rather difficult to be achieved. The privatization of power industry and rapid industrialization has increased the rapid changes in the power exchanges which may cause further instability in the transmission systems. The need for the electric power systems to become more compact and robust is at a constant rise.

The precedence of HVDC over HVAC transmission system for long distance has been established over time. Hence, the improvement of the performance and reliability of the HVDC systems is a focus of many recent researches. GIS has been associated with a considerable number of power system disturbances and failures. Hence improvement in the performance of the GIS plays an important role in increasing the reliability and stability of the power system. A lot of research is being conducted on the design aspects of HVDC power systems to improve its robustness and reliability. Wang et al [1] have proposed a method for charge dissipation from a GIS without opening the tank. Richard et al [2] have designed an efficient insulated bus pipe for shipboard applications. Sridhar et al [3] have estimated the power loss in a 420KV AC GIB by finite element method in 2-dimension. Riechert [4] has proposed the design of a compact gas Insulated System with reduced footprint on the environment. Volpov et al [5] has analyzed and formulated the SF6 spacer design for both HVAC and DC under time-varying and impulse operating conditions. Kosse et al [6] performed the design and testing for compact DC-GIS at 550 KV. The platform size of the system is expected to reduce by 10% by opting for modular designing method. Electric field stress in the power systems is the crucial factor that determines the consistency and robustness of the GIS. It has been established that the most serious distortion of electric field in a GIS occurs at the triple junction [7].

Sayed et al [8] have calculated electric field stress for different spacer types in the bus duct. An insulation compounding scheme has been proposed to reduce the flashover and stress in the AC and DC systems [9]. Zhang et al [10] have proposed the design of basin type insulators for an HVDC GIS to reduce the stress and flashover at the weak links. Researchers have studied and presented the advantage of an FGM spacer design on the electric field distribution in an GIS [11]. The effectiveness of the design has been evaluated for a multi-particle contamination [12]. Kurimoto et al [13] proposed a U-shaped distribution of ε-FGM spacer for efficient reduction in electric stress. Adari et al [14] proposed an FGM cone type spacer and studied its effectiveness under delaminated condition.

Many researchers have proposed the use of FGM post type spacer as it is effective in reduction of electric filed stress and at the same time its simple design makes it easy to manufacture and cost effective. It is very effective in reducing the field stress at the triple junction even under a protrusion and depression condition [15,16]. Ten gradings of equal dimensions have been used in the spacer and the dielectric coefficient of the material has been designated in an increasing order. Naik et al. [17] have used a di-post FGM spacer which has been found very effective in reduction of the electric field stress. The advantageous position of the spacer due to its simple design has been cited. Metwally [18] has compared the design of disc-type, conical-shaped and the post type spacer. In this work it has been cited that while post type spacer is effective in reducing the electric field stress, inclusion of metal cavities in the design limits the flexibility of the design and limits the manufacturing cost of the spacer.

In all the above methods, it has been proposed to vary the dielectric strength of the FGM-spacer in a specific pattern for various spacer designs. However, the researchers have not concentrated on the optimization aspect of the design. The proposed designs although effective may not be the optimal solution for the problem. Recently, researchers have proposed the use of optimization methods for the optimization of the ε-FGM spacer design. Qasim etal. [19] has proposed the use of PSO for the optimization of the FGM-spacer. Talaat etal. [20] have optimized the dielectric material for a cone shaped spacer using COMSOL-live link. A U-shaped permittivity distribution has been preferred in the design. Although some research has been conducted on the optimization of the FGM-spacer. In the above works, the area of research has been limited to the optimization of the FGM-material only for a fixed number of gradings. Hence, there is a scope for further research in the design of FGM spacer, to be able to qualify the GIS as properly optimized. A detailed study to determine the number of gradings and the appropriate dimensions of the same is necessary. The number and dimensions of the gradings may vary as per requirement based on the electric field strength in the zone. Also, the determination of the dielectric coefficient for each grading should be done very precisely to be able to provide an optimal solution to the problem.

In this paper, an optimization algorithm has been proposed for the GIS. Dual objectives have been chosen for the optimization, namely, reduction of the maximum electric field and the standard deviation in the field stress with equal weightage. The FGM-spacer material has been optimized for 4 gradings initially. The detailed results are observed. Progressively, the number of gradings is increased in the high stress zones and the optimization process is repeated. The process is continued until an increase in the number of gradings improves the value of the objective function. Since, the gradings have been designed as per requirement; they may not be of equal dimensions. At the same time, since the number of gradings is as per optimization requirements, so the cost of extra gradings is avoided. The results to verify the proposed methodology have been presented and analyzed in detail.

II. Mathematical Modelling of HVDC GIS

The distribution of electric field intensity, E can be determined from the Poisson's Equation [21].

$$E = -\nabla V \tag{1}$$

Where, V is the electric Potential applied in Volt (V)

The electric flux density D in the distributed area is given by

$$D = \varepsilon_0 \varepsilon_r E (C/m^2) \tag{2}$$

where, $\varepsilon_0 = 8.854 \times 10^{-12} (F/m)$ and ε_r = relative dielectric strength of the medium

We know,

$$\nabla . D = \rho_v \tag{3}$$

ρ_v = volume charge density in C/m^3

$$-\nabla . (\varepsilon_0 \varepsilon_r (\nabla V)) = \rho_v \tag{4}$$

Assuming there are no free charges, $\rho_v = 0$, we have

$$\nabla^2 V = 0 \tag{5}$$

The finite element method can be used to analyze different constructions of the GIB. In FEM, the electrostatic energy in the given space can be minimized by the following analysis.

$$W = \frac{1}{2} \int_v \varepsilon E^2 dv \tag{6}$$

In 3-D dimensions the electrostatic energy in terms of electrical potential is given by-

$$W = \frac{1}{2} \int_v \left[\varepsilon_r (\frac{\partial V}{\partial r})^2 + \varepsilon_\phi (\frac{\partial V}{\partial \phi})^2 + \varepsilon_z (\frac{\partial V}{\partial z})^2 \right] ds \tag{7}$$

Where, $\boldsymbol{\varepsilon_r}$, $\boldsymbol{\varepsilon_\emptyset}$ and $\boldsymbol{\varepsilon_z}$ are the r, ϕ, z- components of the dielectric constant.

In this case, $\left(\frac{\partial V}{\partial \emptyset}\right) = \mathbf{0}$. For a 2D simulation,

$$W = \frac{1}{2} \int_v \left[\varepsilon_r (\frac{\partial V}{\partial r})^2 + \varepsilon_z (\frac{\partial V}{\partial z})^2 \right] ds \tag{8}$$

Where, s is the bounded surface.

After further simplification,

$$\left[\frac{\partial W'}{\partial V_1}\right] ---- \left[\frac{\partial W'}{\partial V_n}\right] = 0 \tag{9}$$

$\frac{\partial w'}{\partial v}$ should be calculated at every dielectric constant domain.

III. Proposed Objective Function for the Design of the Spacer Material

When the spacer is inserted in the system, the electric field distribution becomes non-uniform. It is maximum at the junctions. A multi-objective function is chosen for optimization of spacer material.

I. Minimization of Electric Field

The electric stress distribution in the bus duct is non-uniform. It is highest at the junction of the spacer, SF6 gas and the outer coating. The objective is to reduce the maximum stress incurred by a GIS in order to reduce the occurrences of localized heating.

$$F_1(\varepsilon_i) = Max\left(E_i\left(\varepsilon_i\right)\right) \ V/m \tag{10}$$

Let,

$$Objective\ 1 = Min\ \left(F_1\left(\varepsilon_i\right)\right)$$

where, Max(E_i) is the maximum electric field for the material ε_i for each polygon of the spacer. The objective of the optimization is to choose ε of the spacer material such that the peak value electric field in the system is minimized.

Constraints:

The material of the spacer can be varied within upper and lower bounds.

$$\varepsilon_{i\ min} < \varepsilon_i < \varepsilon_{i\ max} \tag{11}$$

II. Uniform Electric Field distribution

The non-uniformity in the distribution of electric stress leads to the creation of local hot spots. To make the field distribution uniform, the standard deviation of the electric field in the system is minimized.

$$F_2(\varepsilon_i) = Min(\sigma(\varepsilon_i)) \tag{12}$$

$$\sigma(\varepsilon_i) = \sqrt{\frac{\sum_{i=1}^{n}(E_j - \mu)^2}{N}} V/m \tag{13}$$

E_j is the electric field at the nodes.
where, N = No. of samples

No. of samples is chosen such that the distance between neighboring samples is less than d_{max}

$$N:\ d_{min} < d < d_{max}$$

The multi-objective function is given by-

$$\min(F(\varepsilon_i)) = w_1 \times F_1(\varepsilon_i) + w_2 \times F_2(\varepsilon_i) \tag{14}$$

$$w_1 + w_2 = 1 \tag{15}$$

w_1 and w_2 are chosen such that $w_1 \times F_1(\varepsilon_i) \approx w_2 \times F_2(\varepsilon_i)$. Thus, equal weightage is given to both the objectives.

IV. Optimization of Dielectric Material of FGM Spacer

A reliable algorithm for optimization of the GIB is developed in this research paper. The flowchart for the proposed optimization algorithm is shown in Fig. 1. The flowchart explains the process of optimization of the dielectric material of each polygon of the post type spacer.

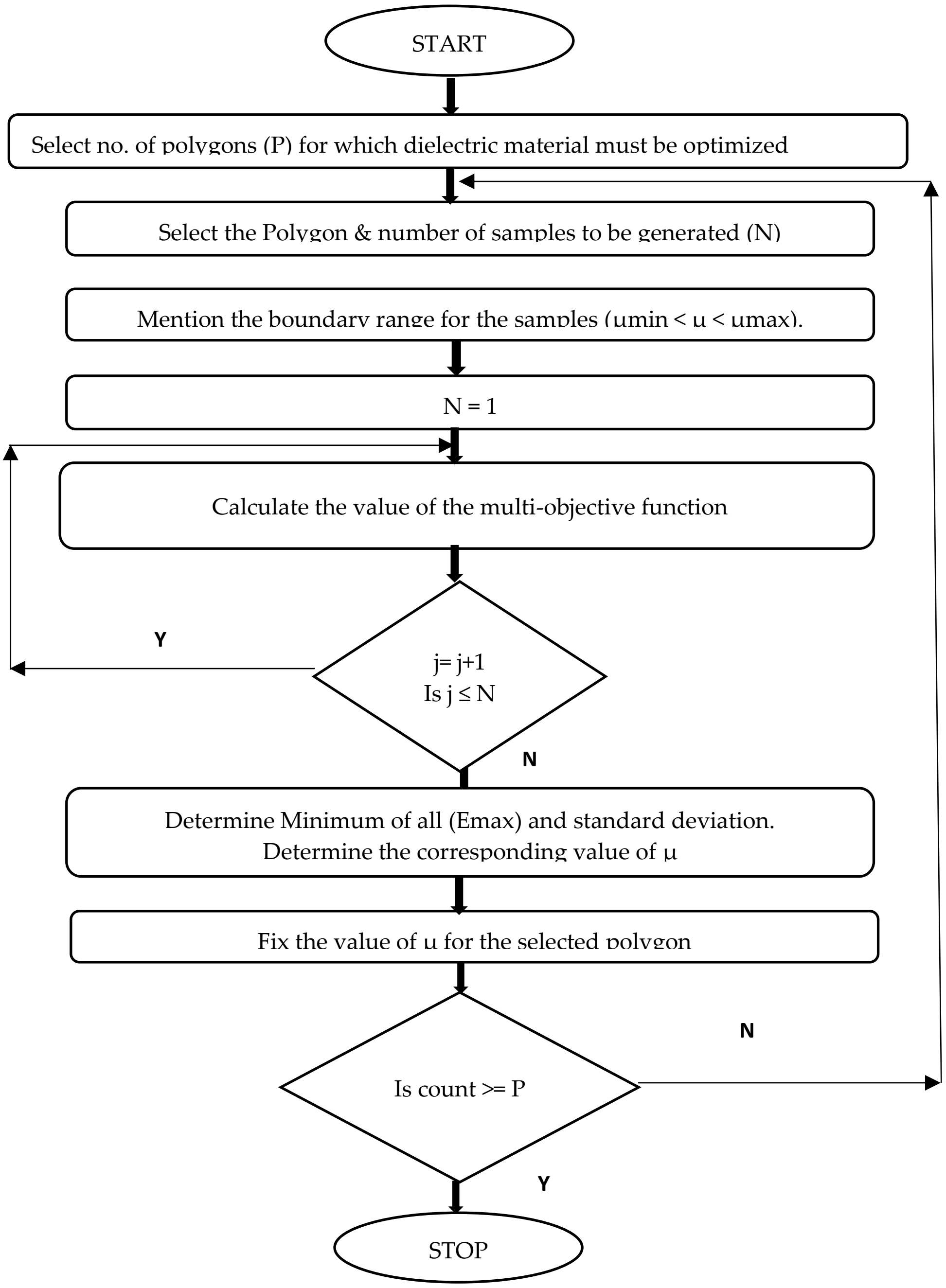


Figure 1: Flowchart depicting the optimization process for GIB post type spacer

The procedure for optimization of the spacer material is mentioned in the steps mentioned below-

Step-1 Install the post-type spacer in the GIS and optimize the dielectric strength of the material for the given multi-objective function (Fig. 1).

Step-2 Provide 4 gradings in the spacer and follow the procedure mentioned in the previous step to optimize the material of each grading.

Step-3 Increase the gradings to 6. Optimize the grading material and compare the value of the parameters and the objective function with the previous case.

Step-4 If a considerable improvement is noticed repeat the step for 8 gradings else GOTO Step-7.

Step-5 The location of the grading is chosen as per the electric field distribution in the various zones of the FGM-spacer. The zone under higher electric stress is graded into smaller zones for further optimization.

Step-6 Compare and analyze the results to obtain the final design of the FGM-Spacer.

Step-7 End Process.

III. Results

A HVDC GIS is considered for the study. The radius of the outer enclosure is 56 mm. The conductor radius is 20.4mm. The insulating medium is SF6 gas with dielectric constant of 1.06. A copper conductor is chosen for the study, with a dielectric constant of 1. The voltage applied to the conductor in this study is 1V. The values for the other voltages can be taken proportionately. The aim of the research is to develop an optimal Post-type FGM spacer for a HVDC GIS. A bi-objective function, consisting of minimization of the electric stress and the stress deviation is chosen for the optimization problem. COMSOL software has been used for the design while MATLAB software has been used for the optimization of the design.

I. Design of 4G- FGM Spacer

After the conductor radius has been optimized, the study focuses on the design of an optimal FGM post-type spacer model. Then, a post-type spacer is designed for the system. Initially, the FGM spacer has been graded into 4 equal layers, each of approximately 9mm as shown in Table 1. The material of each of the gradings of the spacer were optimized. Fig. 2 shows a sample of the optimization result for the 4^{th} grading of the spacer material. It is observed that the objective function has the minimum value at $\varepsilon = 4$. Hence, the value has been chosen for the design. Similarly, the optimization has been performed for each of the gradings of the spacer. The distribution of permittivity in the FGM spacer has been shown in Fig. 3. It is observed that the permittivity in the first grading is 4.4 while the other three gradings have a permittivity of 4 for optimal operation. The surface plots for voltage and the electric field stress distribution are shown in Fig. 4 and 5 respectively. Fig. 6 shows the line plot for the stress distribution in the FGM-spacer.

Table 1: *Permittivity distribution in 4-G FGM Spacer*

Grading No.	Distance from the Outer enclosure (mm)	Permittivity of the FGM gradings
1	0-9	4.4
2	9-18	4
3	18-27	4
4	27-36	4

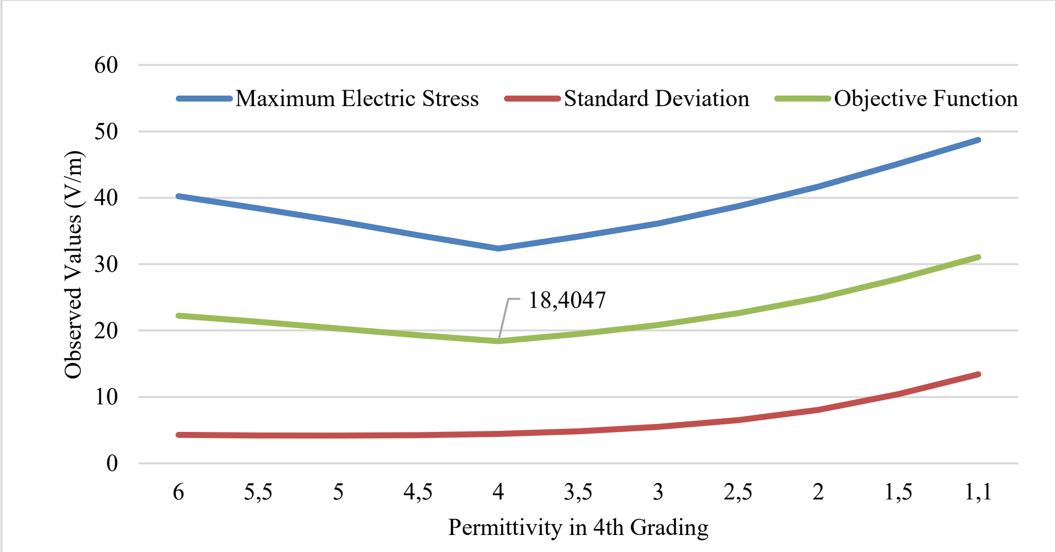


Figure 2: *Distribution of Objective function, Maximum Stress, and Standard Deviation vs permittivity for the 4th Grading of 4G-FGM Spacer*

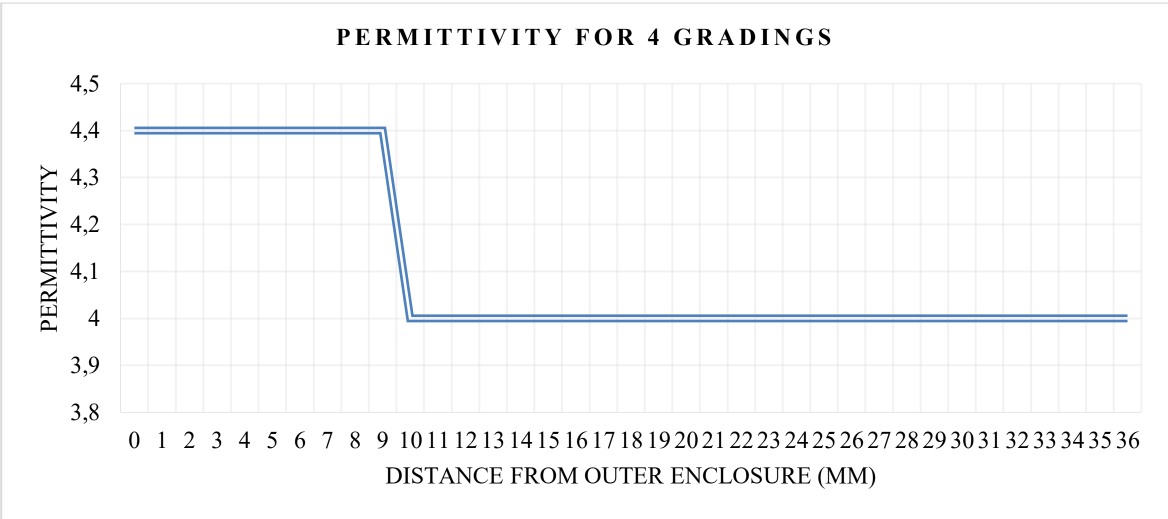


Figure 3: *Distribution of Permittivity in a 4-G FGM Spacer*

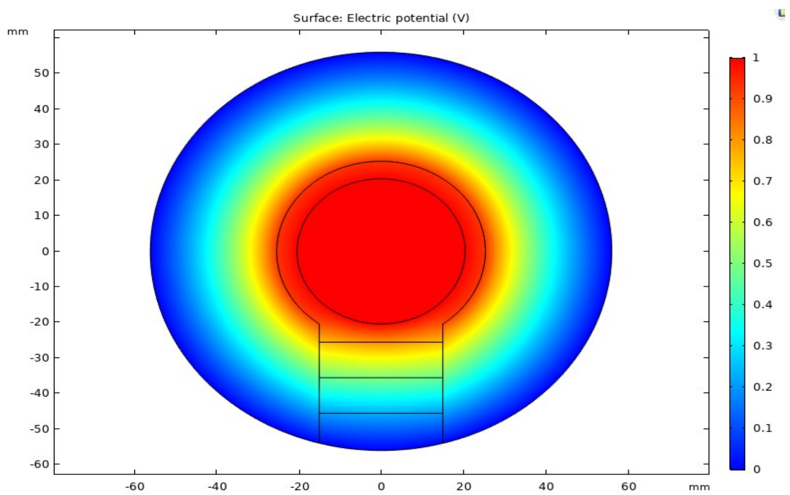


Figure 4: *Electric Potential Distribution for 4-G FGM Post Type Spacer*

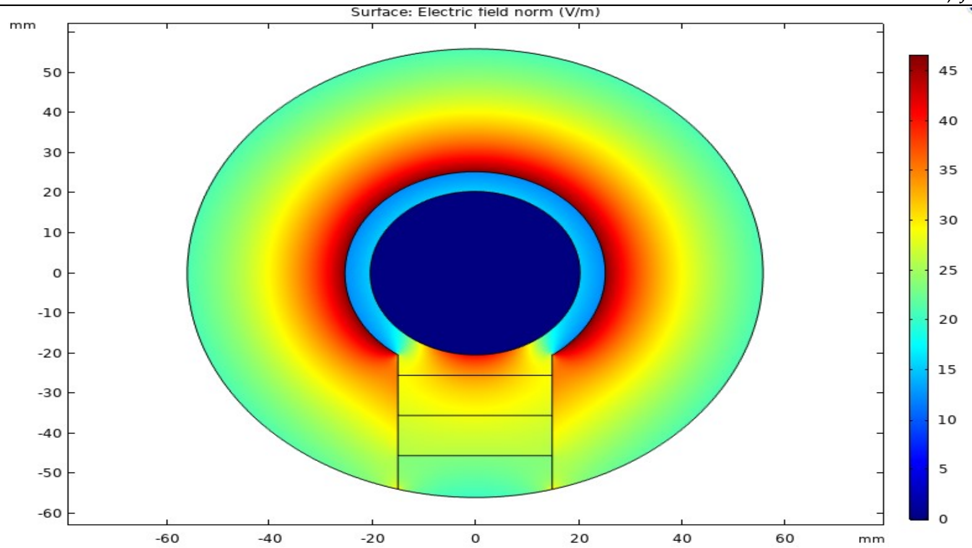


Figure 5: *Electric Field Stress Distribution Surface Plot for 4-G FGM Post Type Spacer*

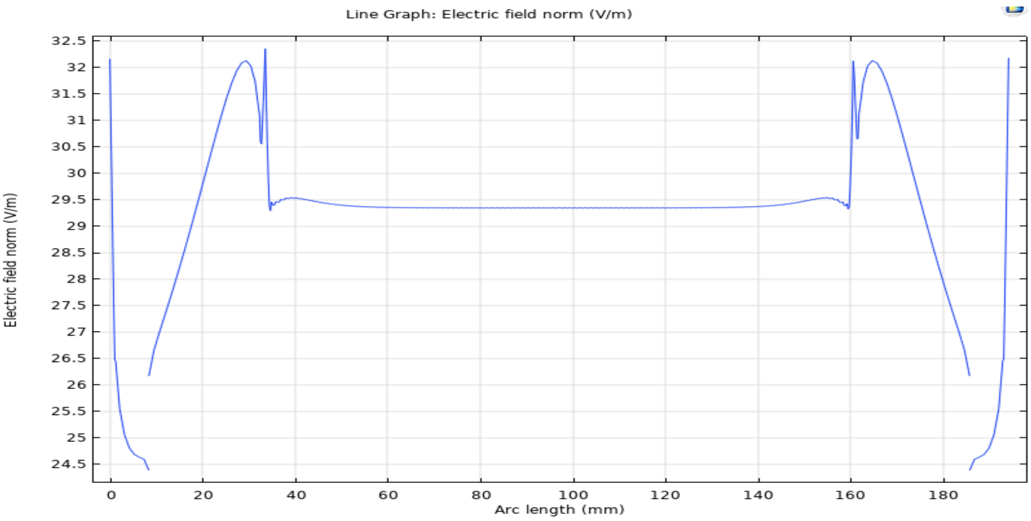


Figure 6: *Electric Field Stress Distribution Line Plot for 4-G FGM Post Type Spacer*

II. Design of 6G FGM Spacer

In a 4G FGM spacer, it is observed that the stress in grading 3 and 4 is much higher in comparison to grading 1 and 2. Hence, the layers 3 and 4 are subdivided into two more layers. The distribution of gradings with respect to distance is shown in Table 2.

Table 2: *Permittivity distribution in 6G FGM Spacer*

Grading No.	Distance from the Outer enclosure (mm)	Permittivity of the FGM gradings
1	0-9	4.8
2	9-18	3.8
3	18-22.5	3.7
4	22.5-27	4.1
5	27-31.5	4.1
6	31.5-36	3.6

The optimization of the permittivity for the 6G spacer material is performed for the chosen objective function. Fig. 7 shows the variation of objective function for various values of permittivity for the 6th grading in the 6G- FGM spacer.

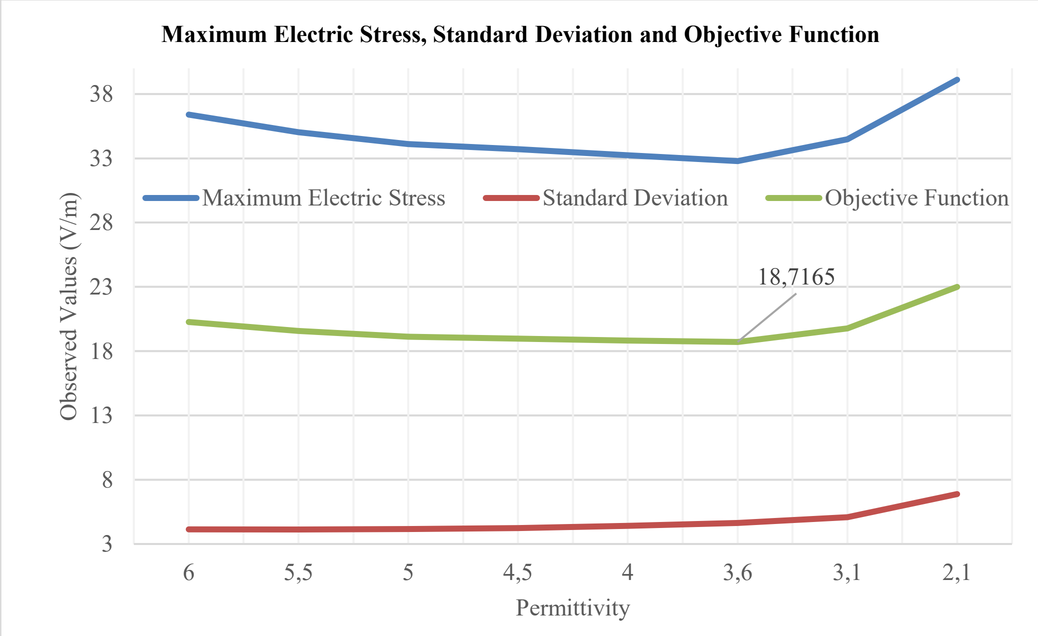


Figure 7: *Distribution of Objective function, Maximum Stress, and Standard Deviation vs permittivity for the 6th Grading of 6G-FGM Spacer*

The optimal permittivity distribution for a 6G-FGM spacer is shown in Fig. 8.

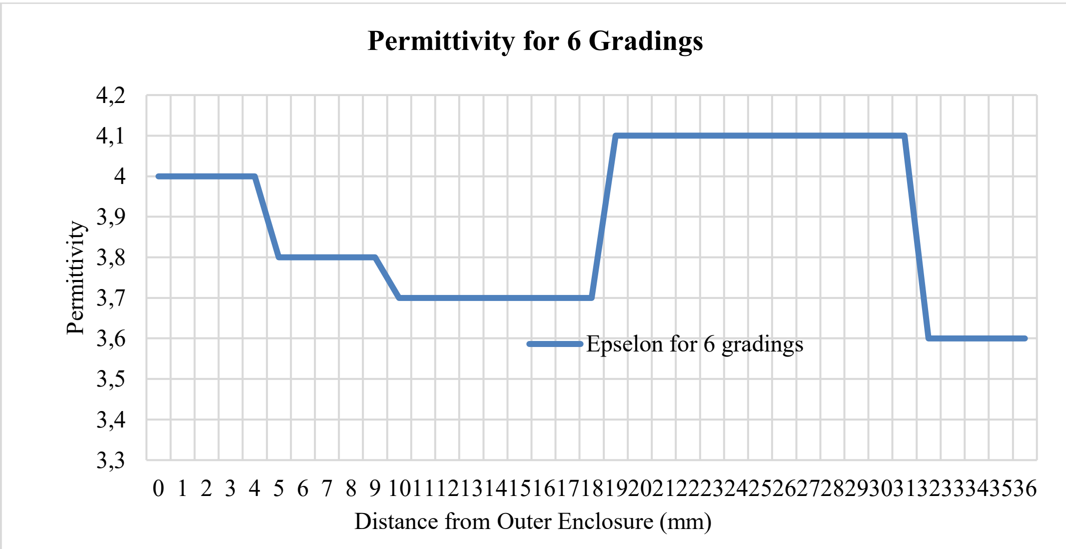


Figure 8: *Distribution of Permittivity in a 6G-FGM Spacer*

The surface plots for electric potential and electric stress for a 6G-FGM post type spacer are shown in Fig. 9 and 10 respectively. Fig. 11 shows the line plot for distribution of stress vs arc length.

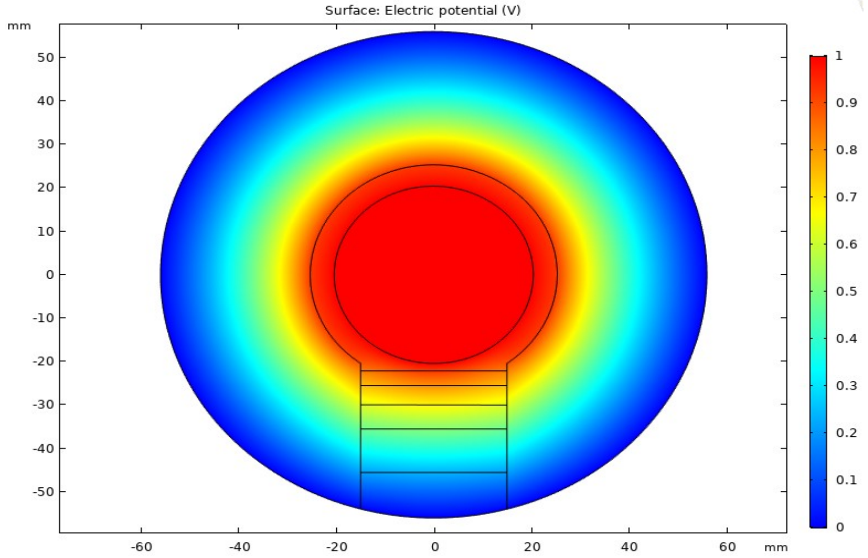


Figure 9: *Electric Potential Distribution for 6-G FGM Post Type Spacer*

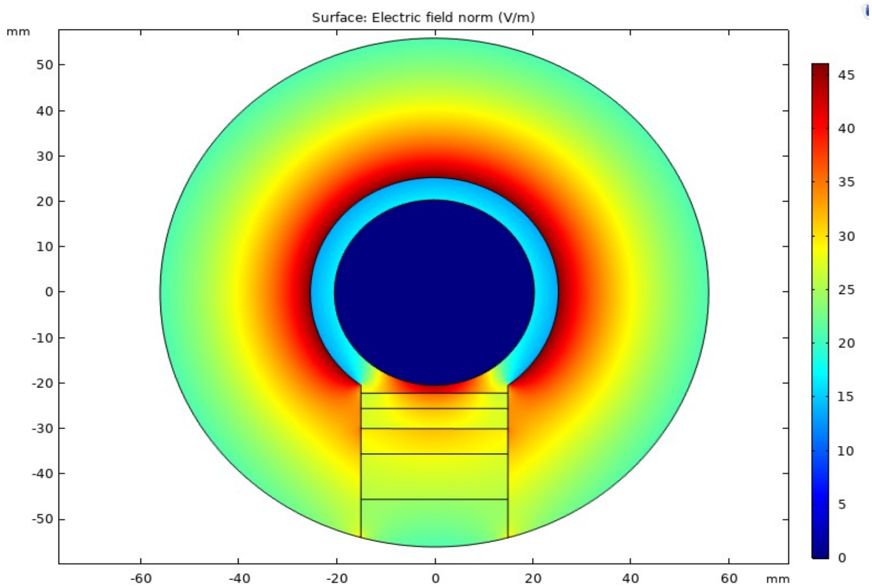


Figure 10: *Electric Field Stress Distribution Surface Plot for 6-G FGM Post Type Spacer*

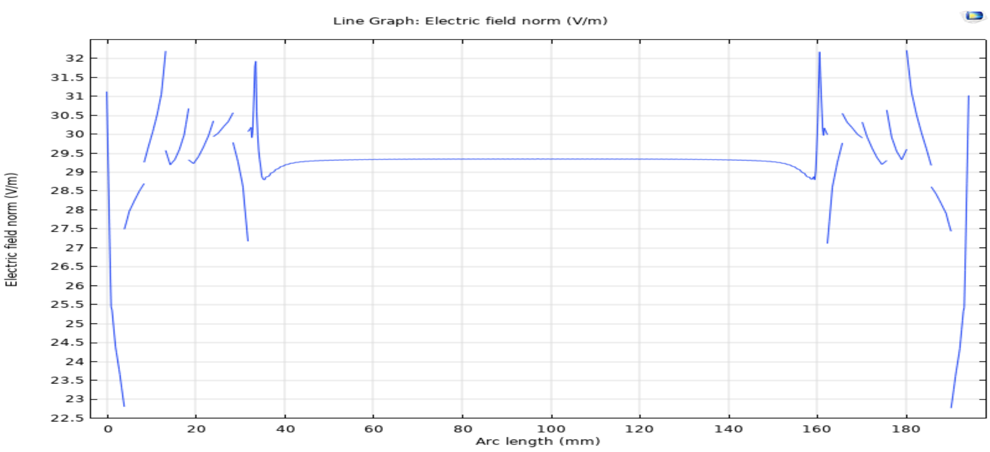


Figure 11: *Electric Field Stress Distribution Line Plot for 6G-FGM Post Type Spacer*

III. Design of 8-G FGM Spacer

The remaining two slots in the 8G-FGM spacer are subdivided to make 8 gradings in the FGM spacer as shown in Table 3. The dielectric strength of the FGM spacer is optimized for the given objective function. A sample of the optimization process is shown in Fig. 12. The optimized distribution of dielectric strength in the spacer is shown in Fig. 13.

Table 3: *Permittivity distribution in 8G-FGM Spacer*

Grading No.	Distance from the Outer enclosure (mm)	Permittivity of the FGM gradings
1	0-4.5	4.7
2	4.5-9	3.7
3	9-13.5	3.6
4	13.5-18	4.1
5	18-22.5	4.4
6	22.5-27	4.5
7	27-31.5	4.7
8	31.5-36	4

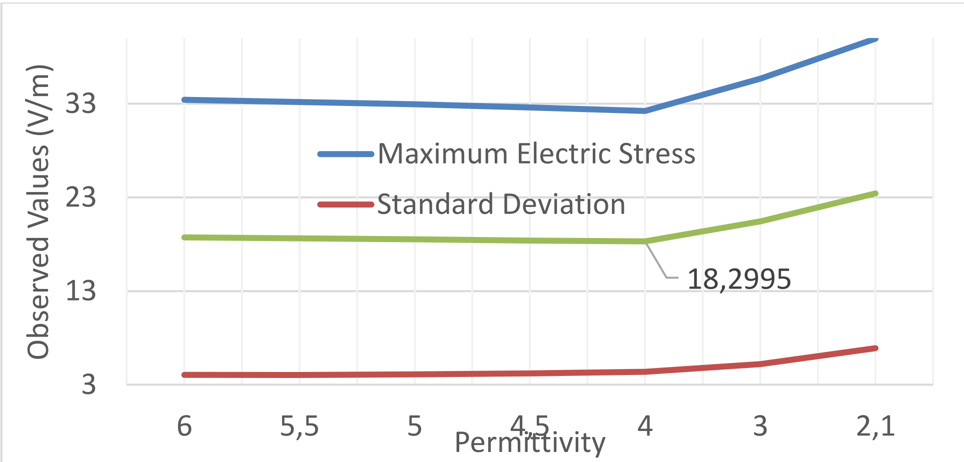


Figure 12: *Distribution of Objective function, Maximum Stress, and Standard Deviation vs permittivity for the 8th Grading of 8G-FGM Spacer*

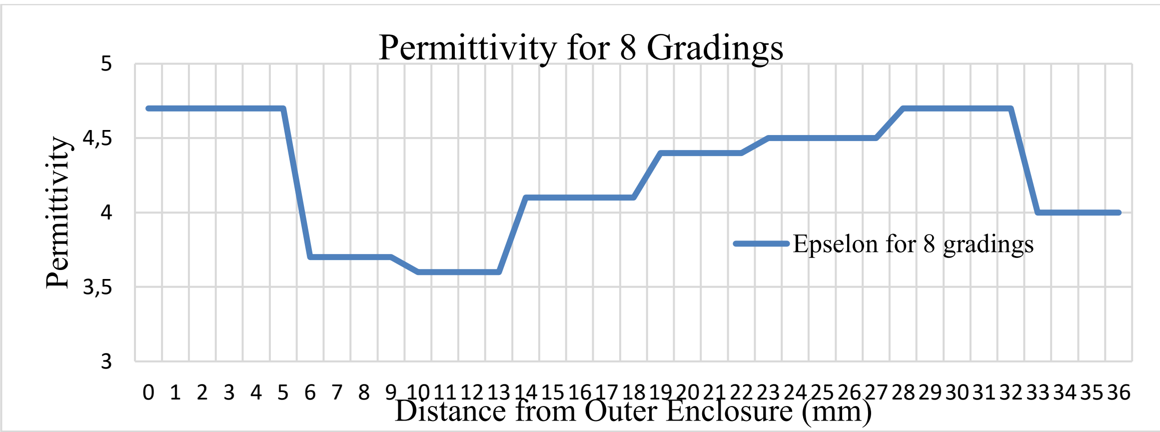


Figure 13: *Distribution of Permittivity in 8G-FGM Spacer*

The performance of the spacer is studied in terms of electric potential distribution and electric stress distribution as shown in the surface plots in Fig. 14 and Fig. 15 respectively. The line plot for electric stress vs arc length and arrow line showing the direction of the stress distribution is shown in Fig. 16 and Fig. 17 respectively. The zoomed contour plot in Fig. 18 focuses on the electric field in the stress zone. The 3D view of the electric stress distribution can be seen in Fig. 19.

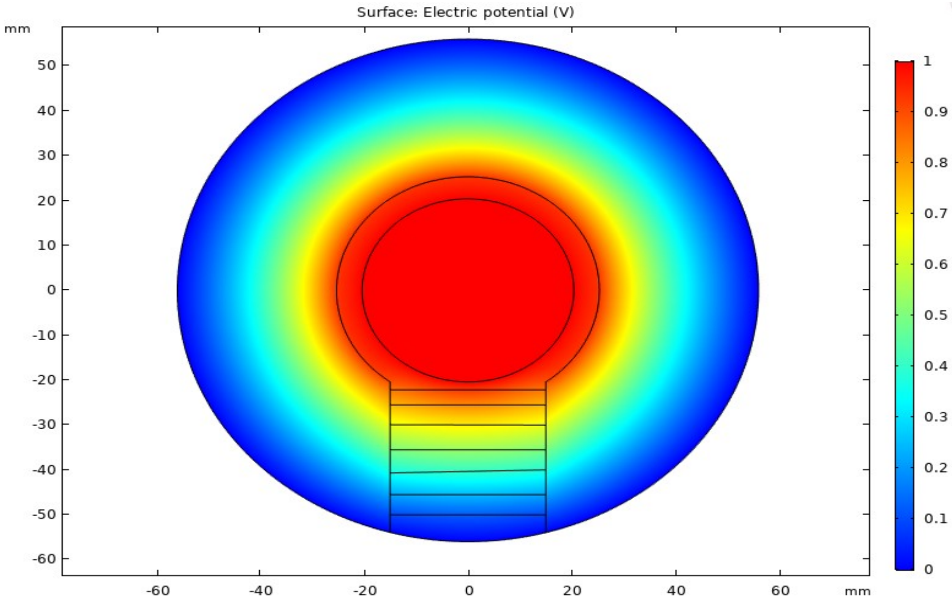


Figure 14: *Electric Field Stress Distribution Surface Plot for 8G- FGM Post Type Spacer*

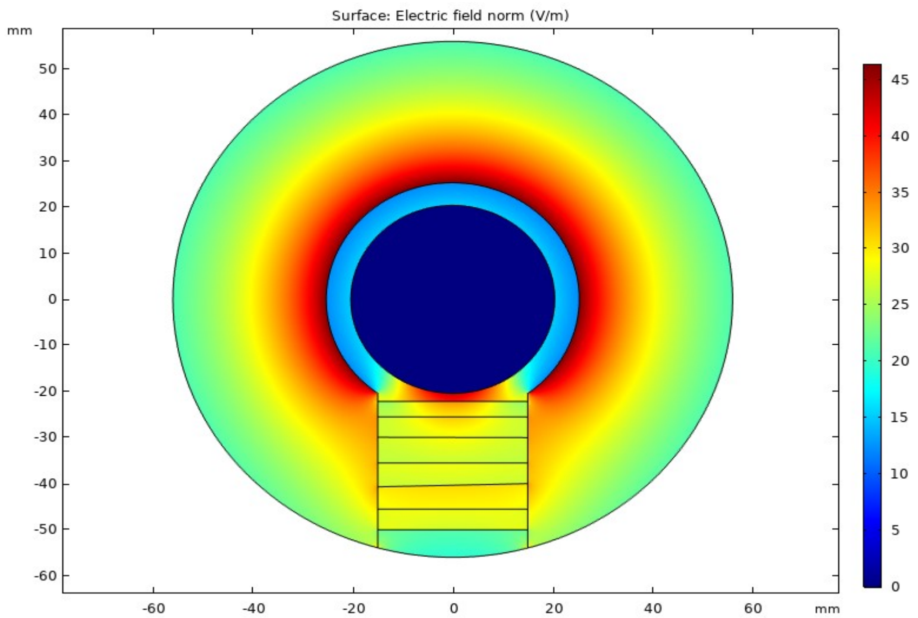


Figure 15: *Electric Potential Distribution for 6-G FGM Post Type Spacer*

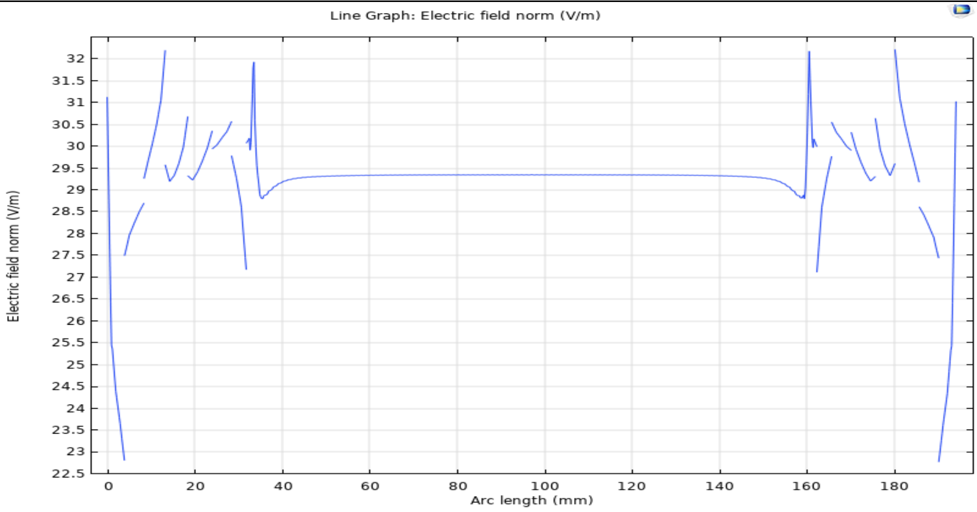


Figure 16: *Electric Field Stress Distribution Line Plot for 6-G FGM Post Type Spacer*

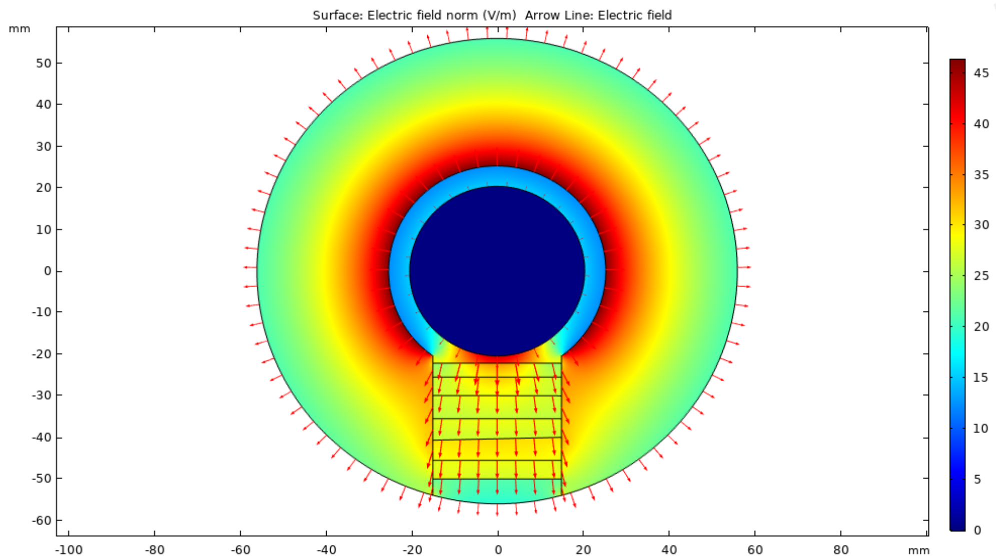


Figure 17: *Arrow line showing direction of field stress*

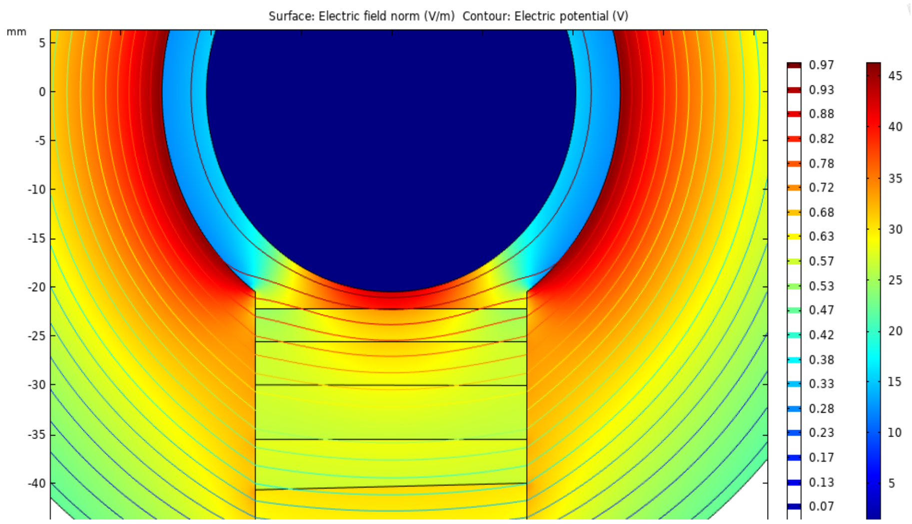
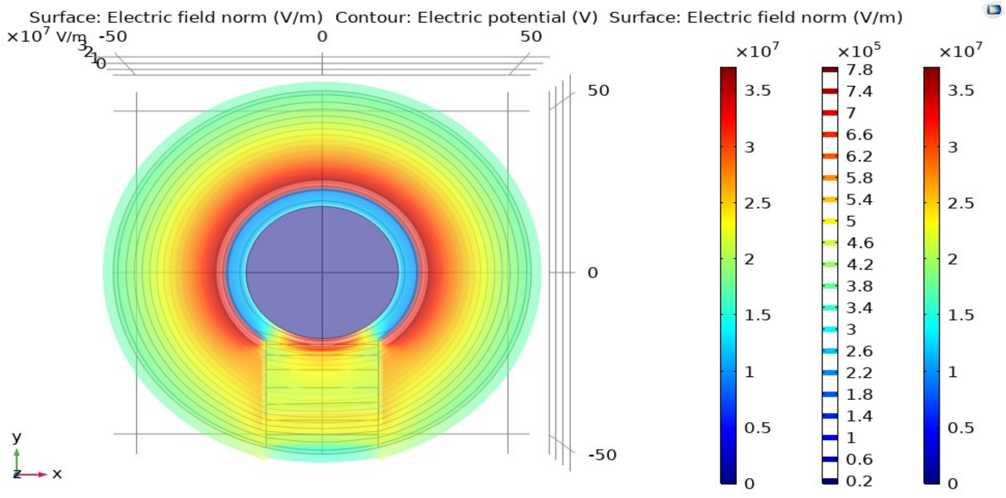
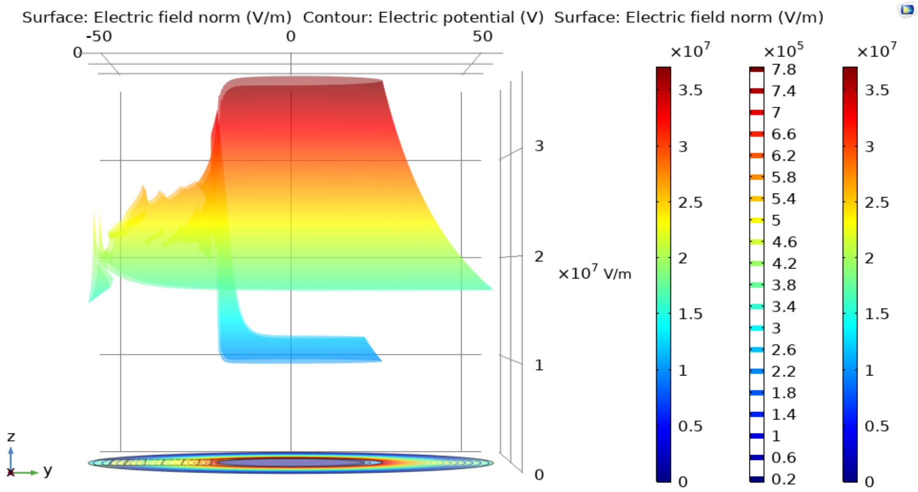


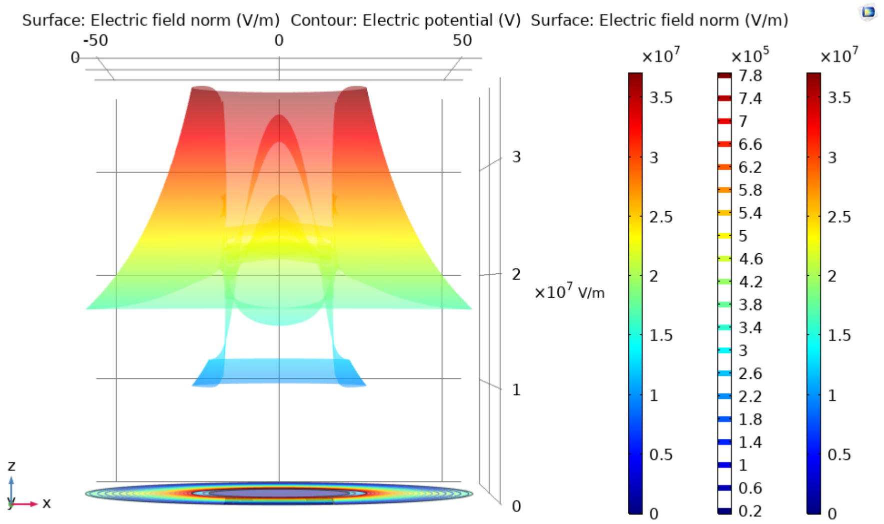
Figure 18: *Zoomed Contour Plot for Electric Stress Distribution*



(a) XY View of electric stress distribution



(b) YZ view of Electric Stress Distribution



© XZ View of Electric Stress Distribution

Figure 19: *3D View of Electric Stress Distribution*

IV. Comparative Analysis of Results

A comprehensive data of the electric field distribution can be observed in Table 4 shown below. Fig. 20 shows the variation of the dielectric strength of the material in different zones for 4G-FGM, 6G-FGM and 8G-FGM respectively.

Table 4: *Distribution of dielectric strength in FGM Spacer*

Type	0 - 4.5 mm	4.5 - 9 mm	9 - 13.5 mm	13.5 – 18 mm	18 – 22.5 mm	22.5 – 27 mm	27 – 31.5 mm	31.5 – 35.5 mm
8 - Grad	4.7	3.7	3.6	4.1	4.4	4.5	4.7	4
6 - Grad	4.8		3.8		3.7	4.1	4.1	3.6
4 - Grad	4.4		4		4		4	
0 - Grad	4.1							

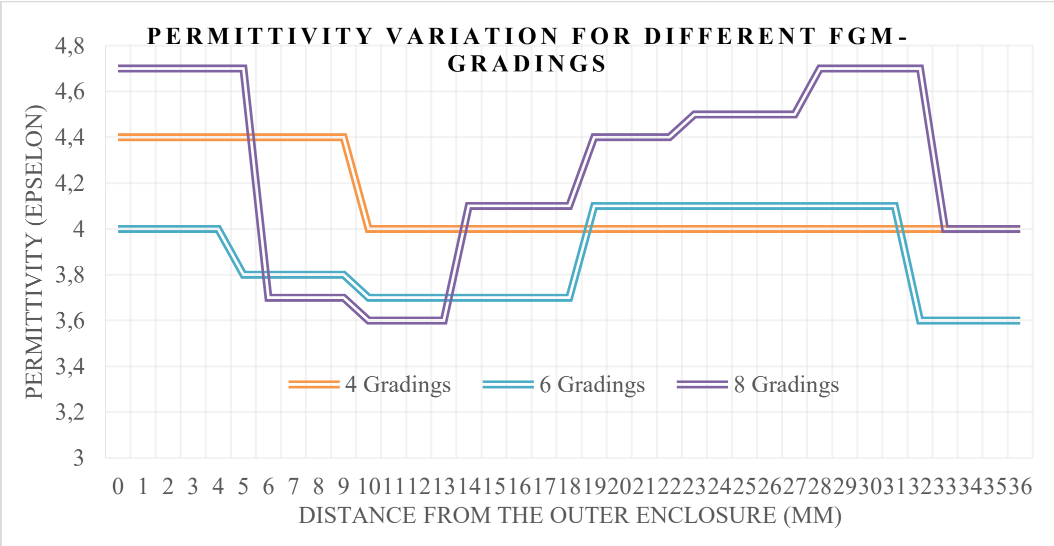


Figure 20: *Distribution of dielectric strength in the FGM-Spacer*

The comparison of the study for different gradings is presented in Table 4. Since, the average of the dielectric strength of the material in each case of FGM spacer is nearly equal to 4.1. The results obtained for different grading types are represented in Table 5. The minimum value for objective function is 18.29 which is obtained in case of 8G FGM spacer.

Table 5: *Comparison of Gradings*

S. No.	Grading Case	Maximum Electric Stress	Standard Deviation	Objective Function
1	0G; Ɛ = 4	34.4184	4.2164	19.3174
2	4G	32.3525	4.4569	18.4047
3	6G	32.7852	4.6477	18.7165
4	8G	32.2214	4.3777	18.2995

Grading high (GH), grading low (GL) and grading U (GU) are the standard methods of distribution of dielectric strength in the spacer material. The performance of the proposed design has been compared with each of the existing methods in Table 6. The distribution of the dielectric material in each of the above methods is shown in Fig. 21.

The average value of the dielectric strength in each of the cases has been maintained nearly equal to maintain uniformity.

Table 6: *Comparison with Existing Methods*

S. No.	Parameter	Proposed 8G FGM	GH-8G FGM	GL-8G FGM	GU-8G FGM
1	Maximum Electric Stress	32.2214	38.6171	35.8664	33.417
2	Standard Deviation	4.3777	4.2556	5.28	4.1066
3	Objective Function	18.2995	21.4363	20.5737	18.76

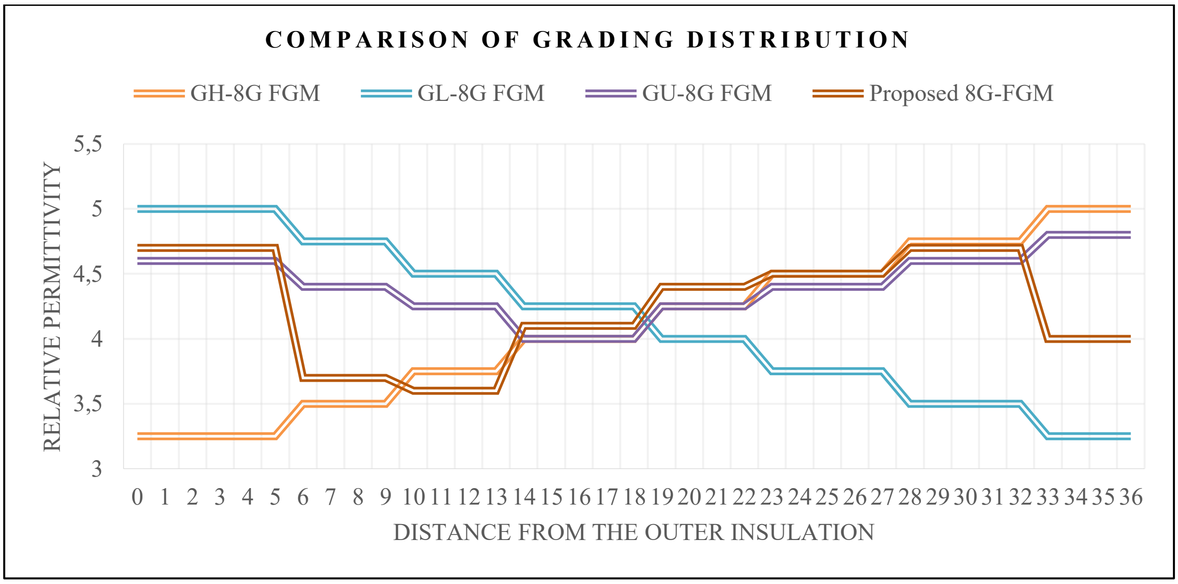


Figure 21: *Distribution of gradings in the FGM spacer for different configurations.*

V. Conclusion

An optimum design of GIS spacer can increase the reliability of the electric transmission system. In this paper, the FGM GIS spacer is designed and the results have been studied. It is observed that

- The FGM method reduces the electric stress in the spacer. There is a marked reduction in the values of the objective in FGM cases.
- The distribution of stress in the post-type spacer has been studied for each of category of FGM grading.
- All the FGM gradings have shown a satisfactory performance in reduction of the maximum electric stress and objective function. However, the 8G- FGM has been most effective in the reduction of the electric stress and the objective function.
- The comparison of the results with existing FGM methods shows that the proposed method helps to achieve a lower objective function value. Thus, the performance of the proposed FGM is better in comparison to the state of art methods.

References

[1] Feng W., Fangwei L., Lipeng Z., She C., Chuanyang L., and Yi X. (2021). Short-time X-ray Irradiation as a Non-contact Charge Dissipation Solution for Insulators in HVDC GIS/GIL, *IEEE Transactions on Dielectrics and Electrical Insulation*, 28(2): 704-709.

[2] Richard W., Moni I., Ruth H. P., Charles S. (2007). Insulated Bus Pipe (IBP) for Shipboard Applications, *IEEE Electric Ship Technologies Symposium*.8:122-129.

[3] Sridhar C., Venkatesh S. (2013). Power Loss Estimation in a 420kV Gas Insulated Busbar (GIB) by Theoretical and Simulation Techniques – A Comparison, *2013 IEEE 1st International Conference on Condition Assessment Techniques in Electrical Systems (CATCON)*, India,213-217.

[4] Uwe R. (2020). Compact High Voltage Direct Current Gas-insulated Systems, *IEEE PES Transmission and Distribution Conference and Exposition*, Chicago,1-4.

[5] Evgeni V. (2004). Dielectric Strength Coordination and Generalized Spacer Design Rules for HVAC & DC SF Gas Insulated Systems, *IEEE Transactions on Dielectrics and Electrical Insulation*, 11(6):949-963.

[6] Maria K., Karsten J., Mark K., Dejun L, (2019). Overview of development, design, testing and application of compact gas-insulated DC systems up to ±550 kV, *Global Energy Interconnection*, 2(6): 567-577.

[7] Boxue D., Jin L. and Hucheng L.(2019). Electrical Field Distribution along HVDC GIL Spacer in SF6/N2 Gaseous Mixture, *Electric Power Conversion*, 1-19.

[8] Sayed A. W., Turky G.M., Shabayek D. M. (2014). Electric Field Stress Calculation for different spacer types in Bus Duct, *International Journal of Scientific Research Engineering & Technology (IJSRET)*, 3(6): 958-963.

[9] C. Li, B. Liu, J. Wang, R. Gong etal(2019). Novel HVDC Spacers in GIS/GIL by Adaptively Controlling Surface Charges – Insulation Compounding Scheme, *2nd International Conference on High Voltage Engineering and Power Systems (ICHVEPS)*, Indonesia,38-44.

[10] Zhang B., Zhong J., Han G. (2020). Modification of HVDC GIS/GIL Basin Insulators Based on Electrical and Mechanical Collaborative Design, *2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, China, 1-5.

[11] Syarif H., Fransiskus D., Umar K. (2016). Electric Field Optimization on 150 kV GIS Spacer using Functionally Gradient Material, *2nd International Conference of Industrial, Mechanical, Electrical, Chemical Engineering (ICIMECE)*,Indonesia, 254-259.

[12] Abd Allah M. A., Sayed A. W., Youssef A. A. (2013) Effect of Functionally Graded Material of Disc Spacer with Presence of Multi-Contaminating Particles on Electric Field inside Gas Insulated Bus Duct, *International Journal of Electrical and Computer Engineering (IJECE)*, 3(6): 831-848.

[13] Muneaki K., Masahiro H., Hitoshi O. (2010). Application of Functionally Graded Material for Reducing Electric Field on Electrode and Spacer Interface, *IEEE Transactions on Dielectrics and Electrical Insulation*, 17(1):256-263.

[14] Jagadeesh A., Venkata Siva Krishna R. G., Venkata Nagesh Kumar G. (2018). Mitigation of field stress with metal inserts for cone type spacer in a gas insulated busduct under delamination, *Engineering Science and Technology*, 21(5): 850–861.

[15] Janaki P., Karthick N. & Nagesh Kumar G. V., (2020). Study of electrostatic field effect in a three-phase gas-insulated busduct with FGM spacer under the effect of protrusion, *Journal of Electromagnetic Waves and Applications*, 34(16): 2107–2129.

[16] Janaki P., Nagesh Kumar G. V., Deepak Chowdary D., Sravana Kumar B.(2020). Study of electric field stress on the surface contour and at the triple junction in three phase GIS with FGM spacer under the depression defect, *International Journal of Emerging Electric Power Systems*, 21(5):20200080.

[17] Gopichand Naik M., Amarnath J., and Kamakshiah S.,(2012) Computation of Electric Field for FGM Spacer Using Di-Post Insulator in GIS, *International Journal of Electrical Engineering*, 5(4): 465-474.

[18] Metwally I. A., (2012). Reduction of electric-field intensification inside GIS by controlling spacer material and design, *Journal of Electrostatics*, 70(2): 217-224.

[19] Syed Abdullah Q. and Nandini G. (2015). Use of Particle Swarm Optimization in the Computation of an optimal permittivity distribution in Functionally Graded Material insulators, *International Conference on Condition Assessment Techniques in Electrical Systems (CATCON)*, India, 184-188.

[20] Talaata M., El-Zein A., Amin M. (2017). Developed optimization technique used for the distribution of U-shaped permittivity for cone type spacer in GIS, *Electric Power Systems Research,* 163:754-766.

[21] William H. Hayt Jr., Engineering Electromagnetics, fifth ed., McGraw-Hill Book Company, 1989.

A GENERALIZED APPROACH IN MULTIPROCESSOR ENVIRONMENT USING REGRESSION TYPE ESTIMATOR AND COST ANALYSIS

Sarla More, Diwakar Shukla

•

Department of Computer Science and Applications
Dr. Harisingh Gour University Sagar (MP) India
sarlamore@gmail.com, diwakarshukla@rediffmail.com

Abstract

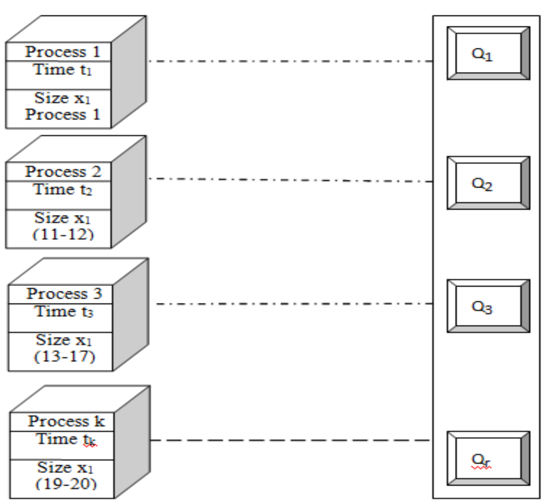
Consider a multi-processors computer system consisting of a ready queue of different jobs to be executed/processed. Lottery scheduling is fair enough to schedule the resources for each and every job. The research idea assumes condition where one can observe some processes to be fully executed; some partially executed few blocked/suspended/ terminated, after sudden system breakdown. An estimation strategy has been designed for the estimation of the total time required to process all these types of processes (processed, partially processed and blocked processes). How much time is required to process the remaining in any hazardous situation? A regression type estimator of sampling theory is used to perform this task. This remaining time estimation technique deals with the backup cost and recovery management as well. Sampling techniques are used in proposed approach for the testing purpose and a simulation has been performed. Another tool adopted is the confidence intervals which are calculated and gives proper précised values in comparison to the true mean for the total remaining time. The linear, square root and square cost function model are adopted for the calculation of backup cost and recovery management. In addition some auxiliary information is also incorporated in the form of size measure of the processes which is an effective approach to calculate the complete remaining time of the processes in multiprocessor environment. The purpose of the proposed research has been served effectively as one can observe the results of disaster and recovery management of the computer system.

Keywords: Ready Queue, Lottery scheduling, Multiprocessors, Simulation, Random Sampling, Estimation, Confidence Interval, Jobs(Processes), Size measure, Estimator

I. Introduction

In the scenario of cloud computing, ready queue is a setup among many servers and processors. For optimal resource allocation there exists several priority scheduling methodologies in the literature of scheduling schemes. In same way lottery scheduling scheme works on randomness of selection of process and distribution of resources providing fair chances. A random number is generated by processors in multiprocessor environment and some token numbers are assigned to each of the process. The execution of process depends upon the condition when the token number of a process is matches with the token number of the processor. The process which has the highest number of tokens has the chance to be allocated the resource for execution of the task. The jobs waiting in the queue always have the chance to be allocated the resource. lottery scheduling maintains the fairness between processes and gives equal chance to each and every process to be allocated the resource. Due to this reason Lottery scheduling is also known as starvation free scheme. In multiprocessor cloud based environment working of Lottery scheduling scheme is similar to draw a random sample through the sampling technique. The remaining time parameter estimation of the ready queue can be executed using the sampling techniques. A job in the ready queue has its process ID, the CPU time(in terms of bytes) as well as the process size (in terms of bytes). With the use of information of process size, it is expected to estimate better the unknown parameter. This paper exploit the approach of use of size measure information for efficient prediction.

Let (t_1, x_1), (t_2, x_2), (t_3, x_3)………..(t_i, x_i)........(t_k, x_k) be the time consumed by i^{th} process in the waiting queue having size measure x_i. Further let Q_1, Q_2, Q_3, ...Q_r be the r processors (r < k) in a computer system who generate random numbers to select processes for resource allocation. Figure 1 describes the general setup of multiprocessors and ready queue. The Figure 2 and 3 are showing the same but in the classified and categorized manner.



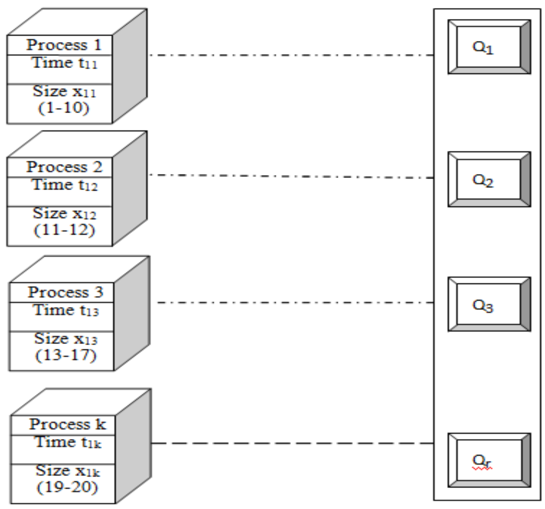


Figure 1: *Ready queue with waiting Processes and Multiprocessor*, **Figure 2:** *Small size processes and Multiprocessors*

This paper takes into account the approach of [4] but adds additional feature of partially processed, blocked processes and size measure of processes for time estimation. All these features are under assumption that the multiprocessor computer system fails at an instant due to unavoidable reasons and backup/recovery management is required. How much the backup cost is needed while sudden breakdown is a question of interest and can be predicted by using the suggested methodology of this paper.

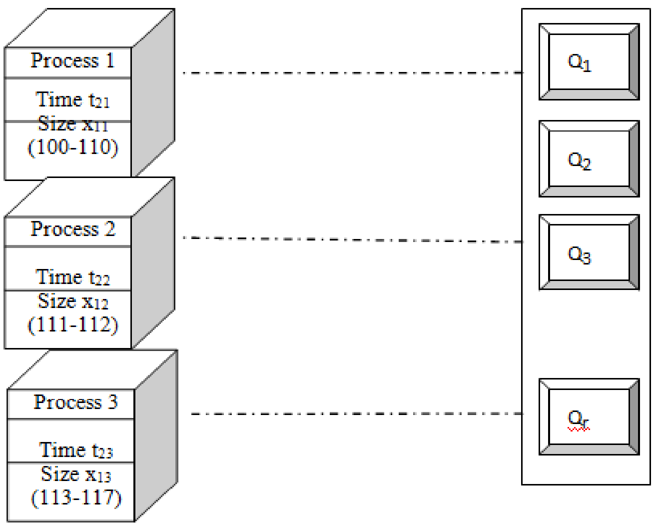


Figure 3: *Big size processes and Multiprocessors*

II. A Review

The priority scheduling is used when any of the jobs is to prefer over others in the waiting queue. Lottery scheduling is one such similar [8] where the job having highest number of tickets has the high chance of being allocated the desired resource. In Linux kernel setup, the lottery scheduling is useful [18] and it could be utilized as a framework [5, 7] for applying the sampling techniques. The similar job group formation scheme for mean time estimation of a ready queue [6] came into picture using lottery scheduling. A review on ready queue mean estimation [3] has opened up avenues for developing new methods in this area. The lottery scheduling types and model based utilization [16, 17] exists in literature as hybrid multilevel structure using Markov chain model along with analysis and chance based prediction. A sample can be used as a suitable input source for mean value prediction [9, 11, 16]. Many various sampling methodologies exist [10, 13, 14] who are comparatively better over another. The best method of selection among them [15] is always possible for precise prediction of unknown parameter. For missing data, the imputation techniques are popular who to replace the non-responding units [19, 20, 21] by known values. Some of most popular imputation methods are mean imputation, deductive imputation, mean imputation within classes, deductive imputation within classes, hot deck imputation, cold deck imputation etc. ([22, 23, 24, 25]). The content of this paper follows idea of [5] and [4] and uses them as input sources in order to resolve the issue of remaining time estimation in presence of sudden breakdown of the system. The contribution in [26] has opened up avenues to think for the use of size measure of processes.

I. Remaining Time Estimation Problem

Let there are finite number of N processes in a ready queue and n ($n < N$) have been processed completely before the system breakdown, obviously (N-n) are still in waiting to get signal for resource allocation. One can assume that n processes are just like a random sample selected from ready queue of size N using lottery scheduling. If θ is mean time obtained through sample then remaining total time estimate is $\Delta = [(N-n)\ \theta]$ which is an unknown quantity. For numbers 'c' and 'd', if Δ is predicted as $\Delta \in (c, d)$ who is an interval containing Δ with very high probability, then $\Delta_1 = [(N-n)\ c]$ is lowest, $\Delta_2 = [(N-n)\ d]$ is upper expected remaining time. If highest expected time is precisely estimated then it could be used for backup management during system failure. The efficient estimation of this expected range is a problem which is chosen in this paper for strategy formation in the multiprocessor setup with the consideration of multiple real life possibilities.

II. Confidence Interval (CI)

A confidence interval is a kind of predictive range for catching of unknown parameter. The feature of a confidence interval is that it contains the true value with 95% precision. Let P[A] denotes the probability of happening of an event A. In statistical theory, contains for any two real numbers a', b', the 95% confidence interval is defined as P [a' < true unknown value < b'] = 0.95. It could be interpreted as chance of being true value within a', b' is 95 percent. The length of confidence interval is a tool for measure of betterment. It is a difference of lower limit and upper limit. Let there are m different confidence intervals of length ($l_1, l_2, l_3, l_4 \dots l_m$) who all catch the true value than an efficiency measure is: Best Confidence Interval = min [$l_1, l_2, l_3, l_4 \dots l_m$]

III. Simulated Cost Aspect

Let C_0 be the fixed cost and C_1 be the cost per unit predicted time. If δ_1 is the minimum and δ_2 is the maximum remaining time after the occurrence of breakdown than

(a) Linear cost function is total cost $(T_c)_{1A} = C_0 + C_1 * \delta_1$ and $(T_c)_{2A} = C_0 + C_1 * \delta_2$
(b) Square root cost function $(T_c)_{1B} = C_0 + C_1 \sqrt{\delta_1}$ and $(T_c)_{2B} = C_0 + C_1 \sqrt{\delta_2}$

(c) Squared cost function is $(T_c)_{1C} = C_0 + C_1 * \delta_1^2$ and $(T_c)_{2C} = C_0 + C_1 * \delta_2^2$

Overall average cost = [Linear cost + Square root cost + Squared cost] / 3

The average cost is likely to incur in the recovery management of resources after the system breakdown. Averaging over linear, squared function and square-root function is taken to control the sampling fluctuations due to lottery scheduling sample.

IV. Sample based Estimation Method

Let $(Y_1, X_1), (Y_2, X_2), (Y_3, X_3)$......... (Y_N, X_N) be the data of totality of size N where Y is variable of main interest and X is the support correlated information. For example, the Y may be expenditure of army officers in a country while x is income data which is known from the salary register of organization/head quarter. The mean of population is $\overline{Y} = (1/N) \sum Y_i$ and $\overline{X} = (1/N) \sum X_i$

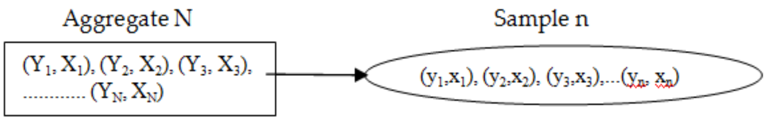


Figure 4: *Sample selection from Aggregate (n<N)*

A sample of size n (n<N) is drawn randomly from N by simple random sampling without replacement method. Sample values are $(y_1, x_1), (y_2, x_2), (y_3, x_3) \ldots (y_n, x_n)$.

Sample mean are $\overline{y} = (1/n) \sum y_i$ and $\overline{x} = (1/n) \sum x_i$

The objective is to estimate unknown parameter $\overline{Y}$ using known $\overline{X}$ along with sample means $\overline{y}$ and $\overline{x}$. Some well known estimators are:

- Sample mean estimator: $\overline{y}$
- Ratio-estimator: $\overline{y_r} = \overline{y}\,(\overline{X}/\overline{x})$
- Difference estimator: $\overline{y_d} = \overline{y} + d\,(\overline{X} - \overline{x})$

III. Motivation

Earlier contributions (specially [4], [5]) were under assumption that processes who exist in a multiprocessors system are completed before sudden failure. But this is not a practical reality. Since some jobs may complete, some may partially processed and some may blocked by the processors [see figure 4]. The processed and unprocessed case was considered in [4] [see figure (6)]. This paper extends the approach of [4] and [26] by applying the tools of random imputation method against the blocked processes.

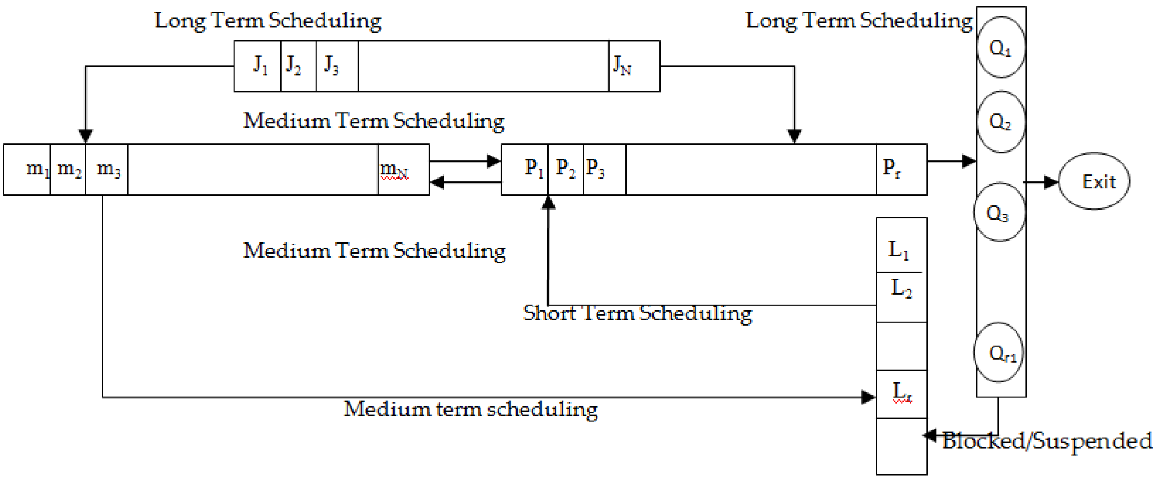


Figure 5: *Ready Queue Processing under Lottery Scheduling (due to [6])*

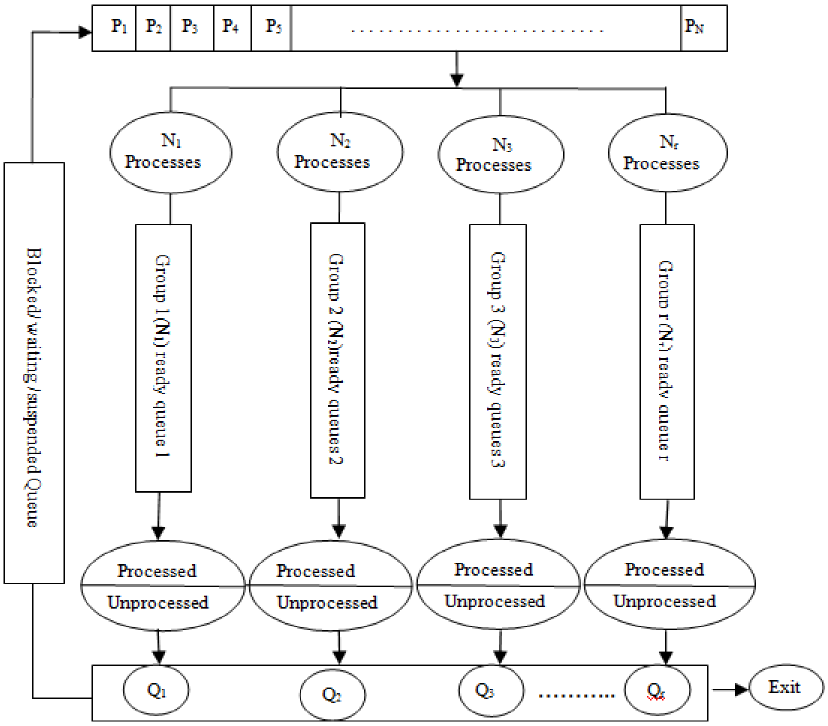


Figure 6: *Setup of ready queue and multiprocessor environment (due to[23])*

IV. Proposed Generalized Computational Setup

Assume the existence a virtual sampled ready queue in a computer system having multiprocessors environment. Some jobs are randomly selected using lottery scheduling from the ready queue and placed in the sampled ready queue from top to bottom in the sequential manner of their selection. Processors are assigned processes in the ordered manner from top to bottom of the virtual sampled ready queue. Figure 5 shows basic setup of this approach but without the size measure while figure 5 shows the earlier approaches [4], [5], [6], [7]. Moreover, figure 6 reveals the special case when all sample units processed before the occurrence of breakdown.

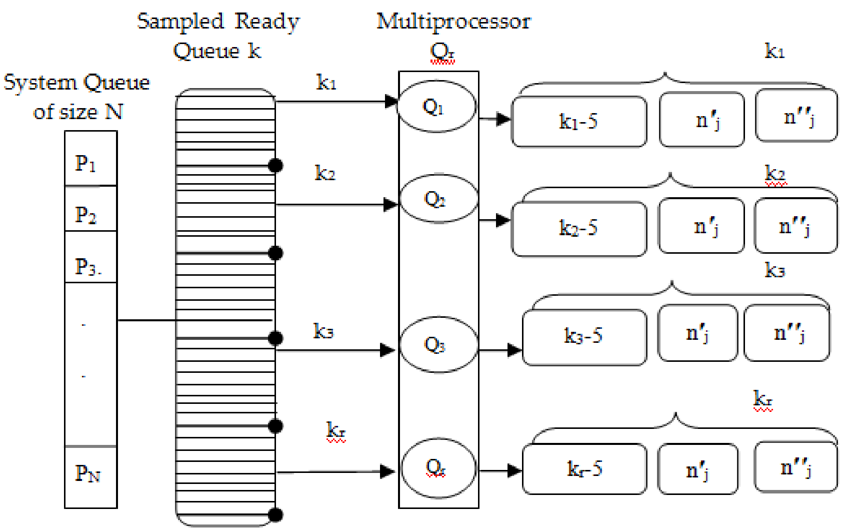


Figure 7: *Sampled Ready Queue Processing Time Estimation setup without size measure*

V. Generalized Assumption and Model

As per figure 7, let the selection of processes is according to lottery scheduling. The process who selects first is placed at the top of the virtual queue who is segment or group of processes likely to allocate to the multi-processors.

1. Assume r processors and a ready queue of N processes in a system like denoted as [P_1, P_2, P_3.........P_N] waiting for allocation of resources.
2. The selection of process for resource allocation is on priority basis using lottery scheduling.
3. If all N are processed completely, time consumed are [t_1, t_2, t_3 ….t_N] who has known size measure [x_1, x_2, x_3 ….x_N].
4. Overall ready queue mean time $\bar{t}=\frac{1}{N}\sum_{i=1}^{N}t_i$, mean size measure $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i$ mean squares $S_t^2=\frac{1}{N-1}\sum_{i=1}^{N}(t_i-\bar{t})^2$, $S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2$.
5. The P_i of known size X_i consumes time t_i(i = 1,2,3,……N) when all assumed processed.
6. Consider r multiprocessors Q_1, Q_2, Q_3…..Q_r, (r < N) and time consumed by the i[th] process in the j[th] processor is t_{ij} with corresponding size measures x_{ij} (j = 1,2,3,……r)
7. The unknown total completion time of ready queue is $N\bar{t}$, which is an unknown quantity. This paper is focused to estimate such using sampling methodology. Lottery scheduling is a tool for such estimation where process P_i has a bunch of token numbers and Q_j generates a random number. A process who receives the random number gets the desired resource from Q_j. This scheduling produces a random sample.
8. A virtual ready queue of size k (k < N, k>3r) exists to store sequentially the records of randomly selected k processes from N. The j[th] segment of virtual sampled queue is k_j($k=\sum_{j=1}^{r}k_j$), who is allocated to the j[th] processor Q_j in sequential manner.
9. In sample, let sx_{jl} denotes the file size measure and st_{jl} denotes time consumed by i[th] process in Q_j (l = 1,2,3,...k_j) when all processed completely who are included in the sample of size k.
   - Sample mean of time $\bar{st}=\frac{1}{k}\sum_{j=1}^{r}\sum_{l=1}^{k_j}st_{jl}$
   - Sample mean square of time, $(es)_t^2=\frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(st_{jl}-\bar{st})^2$
   - Sample mean of size, ($\bar{sx}$)= $\frac{1}{K-1}\sum_{j=1}^{k_j}\sum_{l=1}^{k_j}(sx_{jl})$
   - Sample mean square of size, $(es)_x^2=\frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(sx_{jl}-\bar{sx})^2$

i. The term $\bar{st}$, $\bar{sx}$, $(es)_t^2$, $(es)_x^2$ hold when system runs without failure.

10. Assume system breakdown occurs at the time instant T and there are (k_j – n'_j – n''_j) processes completed in Q_j, but n'_j remain partially processed and n''_j remain unprocessed (blocked). This is an assumed generalized model shown in figure 7. Define $g=\sum_{j=1}^{r}n'_j$ and $u=\sum_{j=1}^{r}n''_j$
11. Let $(st')_{jl}$ is time consumed by the l[th] process in the processor Q_j [l =1, 2, 3... (k_j – n'_j – n''_j)],who is among those processed completely before the occurrence of T.
12. Some sample mean related measures are:
    - Sample mean of (k_j – n'_j – n''_j) process, $(\bar{st}')_j=\frac{1}{(k_j-n'_j-n''_j)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(st'_{jl})$
    - Sample mean square, $(es')_j^2=\frac{1}{(k_j-n'_j-n''_j-1)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(st'_{jl}-(\bar{st}')_j)^2$
    - Similar is for size measure also as (sx'_{jl}) represents size of l[th] process who is in Q_j before T.
    - Sample mean,$(\bar{sx}')_j=\frac{1}{(k_j-n'_j-n''_j)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(sx'_{jl})$
    - $(\bar{sx})_j=\frac{1}{(k_j-n'_j-n''_j)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(sx'_{jl})$is sample mean of all k_j known values related to x in j[th] segment of ready queue.
    - Sample mean square, $(ex')_j^2=\frac{1}{(k_j-n'_j-n''_j-1)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(sx'_{jl}-(\bar{sx}')_j)^2$
    - Sample Covariance, $(es'x')_j=\frac{1}{(k_j-n'_j-n''_j-1)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(st'_{jl}-(\bar{st}')_j)(sx'_{jl}-(\bar{sx}')_j)$

13. Assume t_m^* is partially processed time of a process in Q_j (j = m =1,2,3....r) whose sample mean under T is
14. $(\bar{t}^*/T) = \frac{1}{r}\sum_{m=1}^{r} t_m^*$, Variance $(\bar{t}^*/T) = V(\bar{t}^*/T) = \left(\frac{1}{g} - \frac{1}{N-k+g}\right) S_T^2$, where S_T^2 is the conditional ready queue mean square of the remaining un-sampled part [N- k + g] expressed as: $S_T^2 = \frac{1}{(N-k+g-1)}\sum_{i=i}^{N-k+g}(t_i - \bar{t}_T)^2$ where $\bar{t}_T = \frac{1}{N-k+g}\sum_{i=1}^{N-k+g}(t_i)$ where $g = \sum_{j=1}^{r} n'_j$
15. Herein to mention that S_T^2 and $\bar{t}_T$ contain time t_i only from non-sampled processes (N-k) of the main ready queue with the addition of those g who partially processed. For such, the size converts from N into (N - k + g) and only those processes are the part of $\bar{t}_T$ and S_T^2 who are in (N – k + g).
16. The u blocked processes are imputed by Random Imputation Method using random selection of a process among $(k_j - n'_j - n''_j)$ relating to Q_j. Let from Q_j all random imputed time are denoted as t_m^{**}.
    - Sample mean of all random imputed time, $\bar{t}^{**} = \frac{1}{u}\sum_{m=1}^{u} t_m^{**}$
    - Variance of imputation under T,$V(\bar{t}^{**}/T) = \left(\frac{1}{u} - \frac{1}{k}\right)(es)^2$, u < k.
17. Sample based estimate of $(es)^2$ can be obtained by using all k values of time consumption in sample including the partially processed time t_m^* and imputed time value t_m^{**}. It is denoted as $(es^*)^2$ and mathematically expressed as $(es^*)^2 = \frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(st^*_{jl} - \overline{st}^*)^2$ where (st^*_{jl}) and $\overline{st}^*$ include completely processed time st^*_{ij}, partially processed t_m^* and imputed t_m^{**}.
18. The sample estimate of S_T^2 is $(es')^2 = \frac{1}{g-1}\left[\sum_{m=1}^{g}(t_m^* - \bar{t}^*)^2\right]$
19. Bias of estimation strategy is assumed negligible wherever appears and applicable in mathematical expressions

I. Computational Set-up

Aim is to compute the remaining ready queue processing time after occurrence of sudden failure of system at time instant T. This is subject to condition that r processes are partially processed, r is unprocessed (blocked) and remaining fully completed. Blocked and partially processed are n_j' and n_j'' from every Q_j and known size measures are the part of computation. Some frequently used symbols for process time t and process size measure X are as under:

$$\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i = \frac{1}{N}\sum\sum t_{ij} \quad (1)$$
$$\bar{t}^* = \frac{1}{g}\sum_{m=1}^{g} t_m^* \quad (2)$$
$$\bar{t}^{**} = \frac{1}{u}\sum_{m=1}^{u} t_m^{**} \quad (3)$$
$$(\overline{st}')_j = \frac{1}{(k_j - n'_j - n''_j - 1)}\sum_{j=1}^{(k_j - n'_j - n''_j)}(st'_{jl}) \quad (4)$$
$$(\overline{sx}')_j = \frac{1}{(k_j - n'_j - n''_j - 1)}\sum_{j=1}^{(k_j - n'_j - n''_j)}(sx'_{jl}) \quad (5)$$
$$(\overline{sx})_j = \frac{1}{(k_j)}\sum_{l=1}^{k_j}(sx'_{jl}) \quad (6)$$
$$(es')_j^2 = 1/(k_j - n'_j - n''_j - 1)\sum_{l=1}^{(k_j - n'_j - n''_j)}(st'_{jl} - (\overline{st}')_j)^2 \quad (7)$$
$$(ex')_j^2 = 1/(k_j - n'_j - n''_j - 1)\sum_{l=1}^{(k_j - n'_j - n''_j)}(sx'_{jl} - (\overline{sx}')_j)^2 \quad (8)$$
$$(es'x')_j = \frac{1}{(k_j - n'_j - n''_j - 1)}\sum_{l=1}^{(k_j - n'_j - n''_j)}\left(st'_{jl} - (\overline{st}')_j\right)\left(sx'_{jl} - (\overline{sx}')\right)_j \quad (9)$$
$$(es^*)^2 = \frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(st^*_{jl} - \overline{st}^*)^2 \quad (10)$$
$$\bar{t}_{rj} = [(\overline{st}')_j + d_j\{(\overline{sx})_j - (\overline{sx}')_j\}],\ d_j \text{ being constant, } (0 < d_j < \infty) \quad (11)$$

Note: The $\bar{t}_{rj}$ is a Difference type estimator as stated in subsection IV of section II.

II. Estimation Strategy

The sample based proposed estimation strategy for mean time is:

$$(t_{mean}/T) = €_1 \left[\sum_{j=1}^{r} w_j \; (\bar{t}_{rj}/T) \right] + €_2 \; (\bar{t}^{*}/T) + (1 - €_1 - €_2)(\bar{t}^{**}/T) \tag{12}$$

with condition that $\sum_{p=1}^{3} €_p = 1$ and $€_p$ denotes constants to be determine suitability and $w_j = (k_j/k)$ is known weight ($\sum w_j = 1$). With the help of Cochran [16; see page 166, page 27, 29] for t_{mean}, the expected value E[.] is expressed as:

$$E[t_{mean}/T] = E\left[€_1 \left[\sum_{j=1}^{r} w_j \left(\bar{t}_{rj}/T\right)\right] + €_2 \; (\bar{t}^{*}/T) + (1 - €_1 - €_2)(\bar{t}^{**}/T)\right]$$

$$= €_1 \left[\sum_{j=1}^{r} W_j E\;(\bar{t}_{rj}/T)\right] + €_2 E\,(\bar{t}^{*}/T) + (1 - €_1 - €_2) E\,(\bar{t}^{**}/T)] \tag{13}$$

$\neq \bar{t}$ which shows estimator (t_{mean}/T) is biased.

III. Mean Squared Error

Let MSE (.), V (.) and B (.) denote mean squared error, variance and bias respectively. One can express

MSE (t_{mean}/T) = Variance (t_{mean}/T) + [Bias $(t_{mean}/T)]^2$ which holds in general. Assume the bias is small, therefore negligible (as in assumption no. 16)

$$MSE\,(t_{mean}/T) = \text{Variance}\,(t_{mean}/T) = €_1^2 \left[\sum_{j=1}^{r} w_j^2 \, V(\bar{t}_{rj}/T)\right] + €_2^2 V\,(\bar{t}^{*}/T) + (1 - €_1 - €_2) V\,(\bar{t}^{**}/T)]$$

$$= €_1^2 \left[\sum_{j=1}^{r} \left(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\right) w_j^2 \{(es')_j^2 + d_j^2 (ex')_j^2 - 2d_j (es'x')_j\}\right] + €_2^2 \left[\left(\frac{1}{g} - \frac{1}{N-k+g}\right) s_T^2\right] + (1 - €_1 - €_2)^2 \sum_{j=1}^{r} \left(1 - \frac{1}{(k_j - n'_j - n''_j)}\right) w_j (es')_j^2]$$

(as per Cochran[12] page 24, page 29 and page 164) (14)

The expressions P, Q, R are in the sample based estimate form of population parameters

Let $P = \sum_{j=1}^{r} \left(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\right) w_j^2 \{(es')_j^2 + d_j^2 (ex')_j^2 - 2d_j (es'x')_j\}$,

$Q = \left(\frac{1}{g} - \frac{1}{N-k+g}\right) s_T^2$

$R = \sum_{j=1}^{r} \left(1 - \frac{1}{(k_j - n'_j - n''_j)}\right) w_j^2 (es')_j^2$

The above expression is re-written as:

$V[t_{mean/T}] = [€_1^2 P + €_2^2 Q + (1 - €_1 - €_2)^2 R]$ ignoring the covariance terms due to independency. For optimum variance, differentiate $V[t_{mean}/T]$ with respect to $€_1$ and $€_2$ and equate to zero, one gets

$$(€_1)_{opt} = (QR) / [PQ+PR+QR] = QM \tag{15}$$

$$(€_2)_{opt} = PQ / [PQ+PR+QR] = PM \text{ where } M = R / [PQ+PR+QR] \tag{16}$$

One can differentiate the variance expression by d_j also to get optimum value which is $(d_j)_{opt} = [(es'x')_j / (ex')_j^2]$ Substituting optimum choices in expression, the optimum variance is:

$$V[t_{mean}/T]_{opt} = (€_1)^2_{opt} P + (€_2)^2_{opt} Q + (1 - (€_1)_{opt} - (€_2)_{opt})^2 R] \text{ with } (d_j)_{opt} \tag{17}$$

VI. Numerical Illustration

Consider the 150 processes with processed CPU time whose details are in table 1 with assumption that all 150 processes have been completed.

Table 1: *System Ready Queue Processes with time (N = 150)*

Process	J_{1}	J_{2}	J_{3}	J_{4}	J_{5}	J_{6}	J_{7}	J_{8}	J_{9}	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}
CPU Time	30	20	42	45	59	35	25	48	50	60	32	55	62	47	69
Process Size	41	71	103	142	316	82	199	163	220	127	76	192	251	52	133
Process	J_{16}	J_{17}	J_{18}	J_{19}	J_{20}	J_{21}	J_{22}	J_{23}	J_{24}	J_{25}	J_{26}	J_{27}	J_{28}	J_{29}	J_{30}
CPU Time	34	24	44	70	57	65	38	84	101	66	80	90	92	111	85
Process Size	318	202	106	181	242	148	46	252	136	222	261	97	109	271	116
Process	J_{31}	J_{32}	J_{33}	J_{34}	J_{35}	J_{36}	J_{37}	J_{38}	J_{39}	J_{40}	J_{41}	J_{42}	J_{43}	J_{44}	J_{45}
CPU Time	61	52	72	75	89	67	51	78	80	91	63	86	93	77	99
Process Size	172	243	253	262	83	203	183	166	219	193	223	272	281	301	289
Process	J_{46}	J_{47}	J_{48}	J_{49}	J_{50}	J_{51}	J_{52}	J_{53}	J_{54}	J_{55}	J_{56}	J_{57}	J_{58}	J_{59}	J_{60}
CPU Time	64	54	74	100	87	95	68	114	131	96	110	123	122	141	49
Process Size	205	244	223	254	146	263	53	218	273	139	282	302	173	309	290
Process	J_{61}	J_{62}	J_{63}	J_{64}	J_{65}	J_{66}	J_{67}	J_{68}	J_{69}	J_{70}	J_{71}	J_{72}	J_{73}	J_{74}	J_{75}
CPU Time	118	81	102	105	119	97	88	108	110	121	240	113	122	107	129
Process Size	313	194	153	255	225	169	206	264	58	274	283	303	184	291	216
Process	J_{76}	J_{77}	J_{78}	J_{79}	J_{80}	J_{81}	J_{82}	J_{83}	J_{84}	J_{85}	J_{86}	J_{87}	J_{88}	J_{89}	J_{90}
CPU Time	94	73	104	130	117	234	98	237	161	126	143	236	152	171	233
Process Size	207	246	228	360	256	275	217	265	226	195	284	292	304	300	280
Process	J_{91}	J_{92}	J_{93}	J_{94}	J_{95}	J_{96}	J_{97}	J_{98}	J_{99}	J_{100}	J_{101}	J_{102}	J_{103}	J_{104}	J_{105}
CPU Time	120	112	132	135	149	125	115	138	140	150	122	232	152	137	159
Process Size	247	79	208	276	285	257	56	293	266	187	305	178	310	299	215
Process	J_{106}	J_{107}	J_{108}	J_{109}	J_{110}	J_{111}	J_{112}	J_{113}	J_{114}	J_{115}	J_{116}	J_{117}	J_{118}	J_{119}	J_{120}
CPU Time	124	114	134	160	147	155	128	174	191	156	170	180	182	201	175
Process Size	277	286	211	248	227	294	157	258	229	267	196	298	188	306	270
Process	J_{121}	J_{122}	J_{123}	J_{124}	J_{125}	J_{126}	J_{127}	J_{128}	J_{129}	J_{130}	J_{131}	J_{132}	J_{133}	J_{134}	J_{135}
CPU Time	235	142	162	165	179	151	145	168	171	238	152	175	189	167	241
Process Size	287	278	295	197	249	307	268	311	213	350	112	314	259	297	230
Process	J_{136}	J_{137}	J_{138}	J_{139}	J_{140}	J_{141}	J_{142}	J_{143}	J_{144}	J_{145}	J_{146}	J_{147}	J_{148}	J_{149}	J_{150}
CPU Time	154	144	164	190	177	185	158	204	221	186	200	210	212	231	209
Process Size	214	250	260	279	288	296	308	269	312	245	317	198	319	315	239

Table 2: *Descriptive Statistics of Table 1*

S. No.	Parameters Name	Calculated value
1	Number of Processes N	150
2	Mean time ($\bar{t}$)	122.56

I. Case-I: where each sample size k=40, and $d_j = 0$ ($d_1 = 0$, $d_2 = 0$, $d_3 = 0$)

Table 3: *Calculation for Sample No. 1*

k_1:16	k_2:13	k_3:11
{(J_{01}),(30),(41)},{(J_{31}),(61),(172)}, {(J_{61}),(118),(313)},{(J_{91}),(120),(247)}, {(J_{121},(235),287)},{(J_{63}),(102),(153)}, {(J_{32}),(52),(243)},{(J_{62}),(81),(194)}, {(J_{92}),(112),(79)},{(J_{122}),(142),(278)}, {(J_3),(42),103)},{(J_{33}),(72),(253)}, {(J_{141}),(185),(296)},{(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)}	{(J_{49}),(100),(254)},{(J_{34}),(75),(262)}, {(J_{64}),(105),(255)},{(J_{94}),(135),(276)}, {(J_{124}),(165),(197)},{(J_{135}),(241),(230)} {(J_{35}),(89),(83)},{(J_{65}),(119),(225)}, {(J_{95}),(149),(285)},{(J_{150}),(209),(239)}, {(J_{99}),(140),(266)},{(J_{143}),(204),(269)}, {(J_{116}),(170),(196)}	{(J_{29}),(111),(271)},{(J_{59}),(141),(309)} {(J_{28}),(92),(109)},{(J_{96}),(125),(257)} {(J_{119})(201)(306)},{(J_{149})(231)(315)}, {(J_{142}),(158),(308)},{(J_{97}),(115),(56)}, {(J_{108}),(134),(211)},{(J_{112})(128)(157)}, {(J_{120}), (175), (270)}
$n_i' = 2$, $n_i'' = 3$	$n_i' = 2$, $n_i'' = 2$	$n_i' = 2$, $n_i'' = 3$
Partial Processed ={(J_{33}),(72)(253)}{(J_{141}),(185),(296)} (Processed=50 unprocessed=22) (Processed=90 unprocessed=95) Blocked = {(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)} Blocked replaced α_1' = {(J_{91}),(120),(247)} α_2' ={(J_{32}),(52),(243)} α_3' = {(J_{01}),(30),(41)} [$\bar{st}_1'$ = 99.54, from eq.(4.4), $(es')_1^2$ = 3330.87, from eq.(4.7)], [$\bar{sx}_1$=3583/16 = 223.94, from eq.(4.5) $\bar{sx}_1'$=2110/11=191.81 from eq. (4.6)], [$(ex')_1^2$ =8210.36, from eq.(4.8)] [$(es'x')_1$ = 3230.60, from eq.(4.9)]	Partial Processed={(J_{150})(209)(239)} {(J_{99}),(140),(266)} (Processed=120, unprocessed=89) (Processed=90, unprocessed=50), Blocked={(J_{143}),(204),(269)}, {(J_{116}),(170),(196)} Blocked replaced β_1' ={(J_{64}),(105),(255)} β_2' ={(J_{135}),(241),(230)} [$\bar{st}_2'$ = 130.88, from eq. (4.4), $(es')_2^2$ = 2534.61 from eq.(4.7)] [$\bar{sx}_2$ =3149/13 =242.23, from eq.(4.5), $\bar{sx}_2' = \frac{2067}{9} =$ 229.66, from eq. (4.6), $(ex')_2^2$ =3761,from eq.(4.8)] [$(es'x')_2$ = 387.45, from eq.(4.9)]	Partial Processed = {(J_{142})(158)(308)} {(J_{97})(115)(56)} (Processed=110unprocessed=48), (Processed=65 unprocessed=55), Blocked={(J_{108}),(134),(211)}, {(J_{112})(128)(157)},{(J_{120})(175)(270)} Blocked replaced γ_1' = {(J_{119})(201)(306)} γ_2' = {(J_{59}),(141),(309)} γ_3' ={(J_{29}),(111),(271)} [$\bar{st}_3'$ = 150.16, from eq.(4.4), $(es')_3^2$ = 2950.56 from eq.(4.7)] [$\bar{sx}_3$ =2641/11=240.09,from eq.(4.5) $\bar{sx}_3' = \frac{1567}{6} = 261.16$, from eq. (4.6) [$(ex')_3^2$ = 6092.96, from eq.(4.8)] [$(es'x')_3$ = 2952.56, from eq.(4.9)]

$\bar{t}^*$ = (50+90+120+90+110+65)/6 = 87.5

$\bar{t}^{**}$ = $(\alpha' + \beta' + \gamma')/8$ = (120+52+30+105+241+201+141+111) / 8 = 125.13

Estimated [s_T^2 = 2,204.16] (using point 15) S_T^2 is $(es')^2 = \frac{1}{g-1}[\sum_{m=1}^{g}(t_m^* - \bar{t}^*)^2]$

[(50-**87.5**)^2 +(90-**87.5**)^2 +(190-**87.5**)^2 +(110-**87.5**)^2 +(140-**87.5**)^2+(95-**87.5**)^2]/5

=[4,333.58+1,167.58+4,117.78+250.58+200.78+950.48]= 2,204.16

Let $P = \sum_{j=1}^{r}\left(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\right) w_j^2 \{(es')_j^2 + d_j^2 (ex')_j^2 - 2d_j (es'x')_j\}$, $Q = \left(\frac{1}{g} - \frac{1}{N-k+g}\right) s_T^2$

$R = \sum_{j=1}^{r}\left(1 - \frac{1}{(k_j - n'_j - n''_j)}\right) w_j^2 (es')_j^2$

$P = \left(\frac{1}{16-2-3} - \frac{1}{40}\right)(0.4)^2 * \{3330.87\} + \left(\frac{1}{13-2-2} - \frac{1}{40}\right)(0.33)^2 \{2534.61\} + \left(\frac{1}{11-2-3} - \frac{1}{40}\right)(0.28)^2 \{2950.56\}$

= 0.0659 *0.16*3330.87+ 0.0861*0.1089*2534.61+0.1416*0.0784*2950.56= 91.64

$Q = \left(\frac{1}{3} - \frac{1}{150-40+3}\right)$ 2,204.16 = 0.3245 *2,204.16 = 715.25

$R = \left(1 - \frac{1}{16-2-3}\right)$ (0.4)²*3330.87+ $\left(1-\frac{1}{13-2-2}\right)$ (0.33)² *2534.61 +$\left(1-\frac{1}{11-2-3}\right)$ (0.28)² *2950.56

= 0.9091 *0.16*3330.87+ 0.8889*0.1089*2534.61 +0.8334*0.0784*2950.56 = 922.63

Calculation of mean and Variance V[t_{mean}/T] at d_j = 0 (for all j = 1,2,3)

$(€_1)_{opt}$= (QR) / [PQ+PR+QR] = QM = 715.25*922.63/[91.64*715.25+91.64*922.63+715.25*922.63]

= 659911.1075/810006.4307 = 0.8147

$(€_2)_{opt}$= PQ/ [PQ+PR+QR] = PM = 91.64*715.25/[91.64*715.25+91.64*870.50+715.25*870.50]

= 65545.51/ 810006.4307= 0.0809

$(t_{mean}/T) = (€_1)_{opt}[\sum_{j=1}^{r} w_j\ \bar{t}_{rj}] + (€_2)_{opt}\ (\bar{t}^{*}) + (1- (€_1)_{opt} - (€_2)_{opt})\ (\bar{t}^{**})$

$\bar{t}_{rj} = [(\overline{st}')_j + d_j\{(\overline{sx})_j - (\overline{sx}')_j\}]$,

$\bar{t}_{rj}$= [0.4*99.54+0*(223.94-191.81)]+[0.33*130.88+0*(242.45-229.66)] +[0.28*150.16+0*(240.09-261.16)]

= 39.82+43.19+42.04 = 125.05

(t_{mean}/T) = 0.8147*125.05+0.0809***87.5**+0.1044***125.13 = 122.02**

$V[t_{mean}/T] = (€_1)^2_{opt} P + (€_2)^2_{opt} Q + (1- (€_1)_{opt} - (€_2)_{opt})^2 R$

$V[t_{mean}/T]$ = [(0.8147)² *91.64+ (0.0809)²*715.25 + 0.0108*922.63] = 60.82+4.68+9.96 = 75.46

The 95% confidence intervals for $\bar{t}$, $P\,[(t_{mean}/T) \pm 1.96\sqrt{[V(t_{mean}/T)]}] = 0.95$

= 122.02± 1.96$\sqrt{75.46}$ = 122.02 ± 17.02 = (104.99, 139.04)

Table 4: *Estimated Sample Mean, Variance and Confidence Interval (CI) of Ten Random Samples*

Case-I: At $(€_1)_{opt}$, $(€_2)_{opt}$, $d_j = 0$ ($d_1 = 0$, $d_2 = 0$, $d_3 = 0$) where True mean = 122.51				
S.No.	Estimated Sample Mean	V[t_{mean}/T]	95% Confidence Interval (CI)	CI Length
1	122.02	75.46	(104.99, 139.04)	34.05
2	134.58	64.83	(118.79, 150.36)	31.57
3	117.56	74.36	(100.66, 134.46)	33.80
4	113.89	48.45	(100.25, 127.53)	27.28
5	127.00	85.37	(108.89, 145.11)	36.22
6	119.27	46.42	(105.92, 132.62)	26.70
7	123.39	45.41	(110.18, 136.60)	26.42
8	113.12	97.36	(93.78, 132.46)	38.68
9	115.01	53.05	(100.73, 129.28)	28.55
10	120.21	60.91	(104.91, 135.51)	30.60
			Average Length (3138/10) = 31.38	

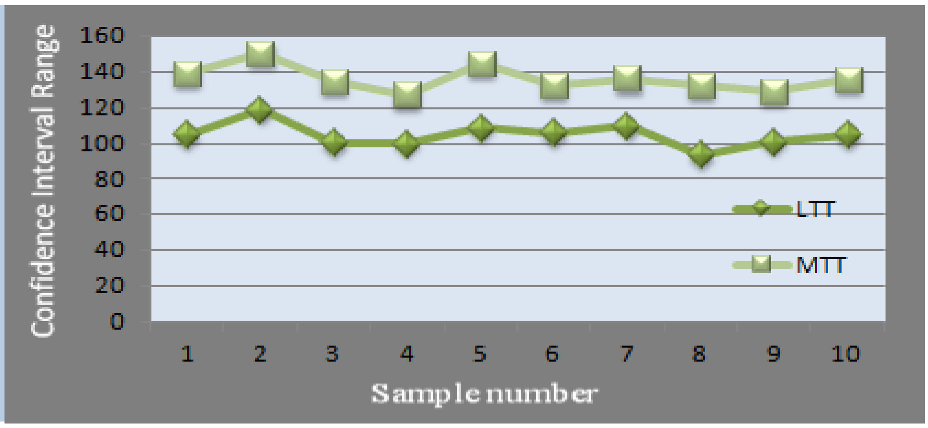


Figure 8: *Graphical representation of Confidence Interval range of Ten Random Samples for Case-I of Table 4 (X-axis has sample number as shown in table 4)*

II. Case-II: where each sample size k=40, and $(d_{opt})_j = (es'x')_j \,/\, (ex')_j^2$

Table 5: *Calculation for Sample No. 1*

k_1:16	**k_2:13**	**k_3:11**
{(J_{01}),(30),(41)},{(J_{31}),(61),(172)}, {(J_{61}),(118),(313)},{(J_{91}),(120),(247)}, {(J_{121},(235),287)},{(J_{63}),(102),(153)}, {(J_{32}),(52),(243)},{(J_{62}),(81),(194)}, {(J_{92}),(112),(79)},{(J_{122}),(142),(278)}, {(J_3),(42),103)},{(J_{33}),(72),(253)}, {(J_{141}),(185),(296)},{(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)}	{(J_{49}),(100),(254)},{(J_{34}),(75),(262)}, {(J_{64}),(105),(255)},{(J_{94}),(135),(276)}, {(J_{124}),(165),(197)},{(J_{135}),(241),(230)} {(J_{35}),(89),(83)},{(J_{65}),(119),(225)}, {(J_{95}),(149),(285)},{(J_{150}),(209),(239)}, {(J_{99}),(140),(266)},{(J_{143}),(204),(269)}, {(J_{116}),(170),(196)}	{(J_{29}),(111),(271)},{(J_{59}),(141),(309)} {(J_{28}),(92),(109)},{(J_{96}),(125),(257)} {(J_{119})(201)(306)},{(J_{149})(231)(315)}, {(J_{142}),(158),(308)},{(J_{97}),(115),(56)}, {(J_{108}),(134),(211)},{(J_{112})(128)(157)}, {(J_{120}), (175), (270)}
$n_i' = 2$, $n_i'' = 3$	**$n_i' = 2$, $n_i'' = 2$**	**$n_i' = 2$, $n_i'' = 3$**
Partial Processed = {(J_{33}),(72),(253)}, {(J_{141}),(185),(296)} (Processed=50, unprocessed=22), (Processed=90, unprocessed=95),	Partial Processed={(J_{150}),(209),(239)}, {(J_{99}),(140),(266)} (Processed=120, unprocessed=89) (Processed=90, unprocessed=50),	Partial Processed = {(J_{142}),(158),(308)},{(J_{97})(115)(56)} (Processed=110, unprocessed=48), (Processed=65, unprocessed=55),
Blocked = {(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)} Blocked replaced α_1'={(J_{91}),(120),(247)},α_2'={(J_{32}),(52),(243)} α_3' = {(J_{01}),(30),(41)}	Blocked={(J_{143}),(204),(269)}, {(J_{116}),(170),(196)} Blocked replaced β_1' ={(J_{64}),(105),(255)} β_2' ={(J_{135}),(241),(230)}	Blocked={(J_{108}),(134),(211)}, {(J_{112})(128)(157)},{(J_{120})(175)(270)} Blocked replaced γ_1'={(J_{119})(201)(306)},γ_2'={(J_{59}),(141),(309)} γ_3' ={(J_{29}),(111),(271)}
[$\bar{st}_1'$ = 99.54, from eq.(4.4), $(es')_1^2$ = 3330.87, from eq.(4.7)],[$\bar{sx}_1$=3583/16 = 223.94, from eq.(4.5), $\bar{sx}_1'$=2110/11=191.81from eq. (4.6)], [$(ex')_1^2$ =8210.36, from eq.(4.8)] ,[$(es'x')_1$ = 3230.60, from eq.(4.9)],$(d_{opt})_1 = (es'x')_1 \,/\, (ex')_1^2$ = 3230.60/8210.36 =0.3935	[$\bar{st}_2'$ = 130.88, from eq. (4.4), $(es')_2^2$ = 2534.61 from eq.(4.7)],[$\bar{sx}_2$ =3149/13 =242.23, from eq.(4.5),$\bar{sx}_2'$= $\frac{2067}{9}$ = 229.66, from eq. (4.6), $(ex')_2^2$ =3761,from eq.(4.8)],[$(es'x')_2$ = 387.45, from eq.(4.9)],$(d_{opt})_2$= $(es'x')_2 \,/\, (ex')_2^2$ = 387.45/3761 = 0.1030	[$\bar{st}_3'$ = 150.16, from eq.(4.4),$(es')_3^2$ = 2950.56 from eq.(4.7)] ,[$\bar{sx}_3$ =2641/11=240.09,from eq.(4.5),$\bar{sx}_3'$= $\frac{1567}{6}$ = 261.16, from eq. (4.6),[$(ex')_3^2$ = 6092.96, from eq.(4.8)],[$(es'x')_3$ = 2952.56, from eq.(4.9)],$(d_{opt})_3$= $(es'x')_3 \,/\, (ex')_3^2$ =2952.56/6092.96 = 0.48

$\bar{\boldsymbol{t}}^*$ = (50+90+120+90+110+65)/6 = 87.5

$\bar{\boldsymbol{t}}^{**}$ = (α' + β' + γ')/8 = (120+52+30+105+241+201+141+111) / 8 = 125.13

Estimated [s_T^2 = 2,204.16] (using point 15) S_T^2 is $(es')^2 = \frac{1}{g-1}\left[\sum_{m=1}^{g}(t_m^* - \bar{t}^*)^2\right]$

[(50-**87.5**)^2 +(90-**87.5**)^2 +(190-**87.5**)^2 +(110-**87.5**)^2 +(140-**87.5**)^2+(95-**87.5**)^2]/5

=[4,333.58+1,167.58+4,117.78+250.58+200.78+950.48]= 2,204.16

Let $P = \sum_{j=1}^{r}\left(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\right) w_j^2 \{(es')_j^2 + d_j^2\,(ex')_j^2 - 2d_j\,(es'x')_j\}$, $Q = \left(\frac{1}{g} - \frac{1}{N-k+g}\right) s_T^2$

$R = \sum_{j=1}^{r}\left(1 - \frac{1}{(k_j - n'_j - n''_j)}\right) w_j^2 (es')_j^2$

$P = \left(\frac{1}{16-2-3} - \frac{1}{40}\right)$ (0.4)^2 *{3330.87+0.39*0.39*8210.36 - 2*0.39*3230.60}+ $\left(\frac{1}{13-2-2} - \frac{1}{40}\right)$ (0.33)^2 {2534.61+0.10*0.10*3761- 2*0.10*387.45}+ $\left(\frac{1}{11-2-3} - \frac{1}{40}\right)$ (0.28)^2 {2950.56 +0.48*0.48*6092.96- 2*0.48*2952.56}

= 0.0659 *0.16*2059.79+ 0.0861*0.1089*2494.73+0.1416*0.0784*1519.92 = 61.98

$Q = \left(\frac{1}{3} - \frac{1}{150-40+3}\right)$ 2,204.16 = 0.3245 *2,204.16 = 715.25

$R = \left(1-\frac{1}{16-2-3}\right)(0.4)^2*3330.87+\left(1-\frac{1}{13-2-2}\right)(0.33)^2*2534.61+\left(1-\frac{1}{11-2-3}\right)(0.28)^2*2950.56$

$= 0.9091*0.16*3330.87+0.8889*0.1089*2534.61+0.8334*0.0784*2950.56 = 922.63$

Calculation of mean and Variance V[t_{mean}/T] at $d_j = (d_{opt})_j$

$(€_1)_{opt}$= (QR) / [PQ+PR+QR] = QM = 715.25*922.63/[61.98*715.25+61.98*922.63+715.25*922.63]

= 659911.1075/761426.9099= 0.8666

$(€_2)_{opt}$= PQ/ [PQ+PR+QR] = PM = 61.98*715.25/[61.98*715.25+61.98*870.50+715.25*870.50]

= 44331.195/ 761426.9099 = 0.0582

$(t_{mean}/T) = (€_1)_{opt}\left[\sum_{j=1}^{r} w_j\, \bar{t}_{rj}\right] + (€_2)_{opt}\,(\bar{t}^{*}) + \left(1-(€_1)_{opt}-(€_2)_{opt}\right)(\bar{t}^{**})$

$\bar{t}_{rj} = [(\overline{st}')_j + d_j\{(\overline{sx})_j - (\overline{sx}')_j\}]$,

$\bar{t}_{rj}$= [0.4*99.54+0.39*(223.94-191.81)]+[0.33*130.88+0.10*(242.45-229.66)] +[0.28*150.16+0.48*(240.09-261.16)] = [0.4*99.54+12.64]+[0.33*130.88+2.11] +[0.28*163.33-45.50] = 52.45+44.47+31.93 = 128.85

(t_{mean}/T) = 0.8666*128.85+0.0582***87.5**+0.0752***125.13 = 126.16**

$V[t_{mean}/T] = (€_1)^2_{opt}\,P + (€_2)^2_{opt}\,Q + (1-(€_1)_{opt}-(€_2)_{opt})^2 R$

V[t_{mean}/T] = [(0.8666)^2 *61.98+ (0.0582)^2*715.25 + 0.0056*922.63] = 46.54+2.42+5.17 = 54.13

The 95% confidence intervals for $\bar{t}$, $P\left[(t_{mean}/T) \pm 1.96\sqrt{[V(t_{mean}/T)]}\right] = 0.95$

= **126.16** ± $1.96\sqrt{54.13}$ = **126.16** ± 14.42 = (111.74, 140.58)

Table 6: *Estimated Sample Mean, Variance and Confidence Interval (CI) of Ten Random Samples*

Case-II: At $(€_1)_{opt}$, $(€_2)_{opt}$, $(d_{opt})_j = (es'x')_j / (ex')_j^2$ where True mean = 122.51				
S.No.	Estimated Sample Mean	V[t_{mean}/T]	95% Confidence Interval (CI)	CI Length
1	**126.16**	54.13	(111.74, 140.58)	28.84
2	130.78	39.24	(118.50, 143.06)	24.56
3	125.24	48.98	(111.52, 138.96)	27.44
4	124.84	45	(111.70, 137.99)	26.29
5	128.89	53.58	(114.54, 143.24)	28.7
6	140.30	100.86	(120.62, 159.98)	39.36
7	125.99	29.81	(115.29, 136.69)	21.4
8	110.79	77.25	(93.56, 128.02)	34.46
9	128.36	50.01	(114.50, 142.22)	27.72
10	128.07	38.42	(115.92, 140.22)	24.3
Average Length (28307/10) = 28.30				

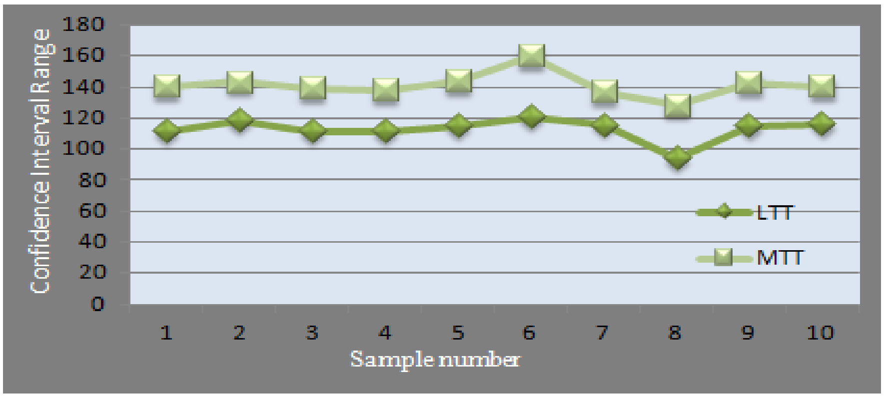


Figure 9: *Graphical representation of Confidence Interval range of Ten Random Samples for Case-II of Table 6 (X-axis has sample number as shown in table 6)*

Table 7: *Comparison between Case-I and Case-II*

S. NO	CASE-I $d_j = 0$ ($d_1=0$, $d_2=0$, $d_3=0$) 95% Confidence Interval	Length	CASE-II $(d)_j = (d_{opt})_j$ 95% Confidence Interval	Length
1.	(104.99, 139.04)	34.05	(111.74, 140.58)	28.84
2.	(118.79, 150.36)	31.57	(118.50, 143.06)	24.56
3.	(100.66, 134.46)	33.8	(111.52, 138.96)	27.44
4.	(100.25, 127.53)	27.28	(111.70, 137.99)	26.29
5.	(108.89, 145.11)	36.22	(114.54, 143.24)	28.7
6.	(105.92, 132.62)	26.7	(120.62, 159.98)	39.36
7.	(110.18, 136.60)	26.42	(115.29, 136.69)	21.4
8	(93.78, 132.46)	38.68	(93.56, 128.02)	34.46
9.	(100.73, 129.28)	28.55	(114.50, 142.22)	27.72
10.	(104.91, 135.51)	30.6	(115.92, 140.22)	24.3
	Average Length (3138/10)	**31.38**	**Average Length (2830/10)**	**28.30**

Table 8: *Case-I: Cost aspect when C_0 = 100 units, C_1 = 10 units*

S. NO	C.I Lower Limit	C I Upper Limit	δ_1	δ_2	Total cost $(T_c)_{1A}$	$(T_c)_{1B}$	$(T_c)_{1C}$	Total cost $(T_c)_{2A}$	$(T_c)_{2B}$	$(T_c)_{2C}$
1	104.99	139.04	11,548.90	15,294.40	115589	1174.65	1333771012	153044	1336.7053	2339186814
2	118.79	150.36	13,066.90	16,539.60	130769	1243.10	1707438856	165496	1386.063762	2735583782
3	100.66	134.46	11,072.60	14,790.60	110826	1152.26	1226024808	148006	1316.166107	2187618584
4	100.25	127.53	11,027.50	14,028.30	110375	1150.11	1216057663	140383	1284.411246	1967932109
5	108.89	145.11	11,977.90	15,962.10	119879	1194.43	1434700984	159721	1363.412047	2547886464
6	105.92	132.62	11,651.20	14,588.20	116612	1179.40	1357504714	145982	1307.816211	2128155892
7	110.18	136.6	12,119.80	15,026.00	121298	1200.89	1468895620	150360	1325.805857	2257806860
8	93.78	132.46	10,315.80	14,570.60	103258	1115.66	1064157396	145806	1307.087404	2123023944
9	100.73	129.28	11,080.30	14,220.80	110903	1152.63	1227730581	142308	1292.509958	2022311626
10	104.91	135.51	11,540.10	14,906.10	115501	1174.24	1331739180	149161	1320.905402	2221918272
		Average			115501	1173.74	1336802081	150026.7	1324.088329	2253142435

NOTE 8.1: Overall average cost by lower limit = (115501 +1173.743546+ 1336802081)/3
= 445639585.25 units

NOTE 8.2: Overall average cost by upper limit = (150026.7 + 1324.088329+ 2253142435)/3
= 751097928.59 units

Table 9: *Case- II: Cost aspect when C_0 = 100 units, C_1 = 10 units*

S. NO	C.I Lower Limit	C I Upper Limit	δ_1	δ_2	Total cost			Total cost		
			$\delta 1$	$\delta 2$	$(T_c)_{1A}$	$(T_c)_{1B}$	$(T_c)_{1C}$	$(T_c)_{2A}$	$(T_c)_{2B}$	$(T_c)_{2C}$
1	111.74	140.58	12,291.40	15,463.80	123014	1208.665865	1510785240	154738	1343.535283	2391291204
2	118.5	143.06	13,035.00	15,736.60	130450	1241.709245	1699112350	157466	1354.456057	2476405896
3	111.52	138.96	12,267.20	15,285.60	122772	1207.573925	1504842058	152956	1336.349465	2336495774
4	111.7	137.99	12,287.00	15,178.90	122970	1208.46741	1509703790	151889	1332.026785	2303990152
5	114.54	143.24	12,599.40	15,756.40	126094	1222.47049	1587448904	157664	1355.244996	2482641510
6	120.62	159.98	13,268.20	17,597.80	132782	1251.87673	1760451412	176078	1426.566998	3096825748
7	115.29	136.69	12,681.90	15,035.90	126919	1226.139423	1608305976	150459	1326.209607	2260782988
8	93.56	128.02	10,291.60	14,082.20	103016	1114.475234	1059170406	140922	1286.684457	1983083668
9	114.5	142.22	12,595.00	15,644.20	126050	1222.274476	1586340350	156542	1350.767764	2447410036
10	115.92	140.22	12,751.20	15,424.20	127612	1229.212115	1625931114	154342	1341.942028	2379059556
			Average value		124167.9	1213.286491	1545209160	155305.6	1345.378344	2415798653

NOTE 9.1: Overall average cost by lower limit = (124167.9+1213.286491+ 1545209160)/3
= 515111513.72 units

NOTE 9.2: Overall average cost by upper limit = (155305.6+ 1345.378344+ 2415798653)/3
= 805318434.65 units

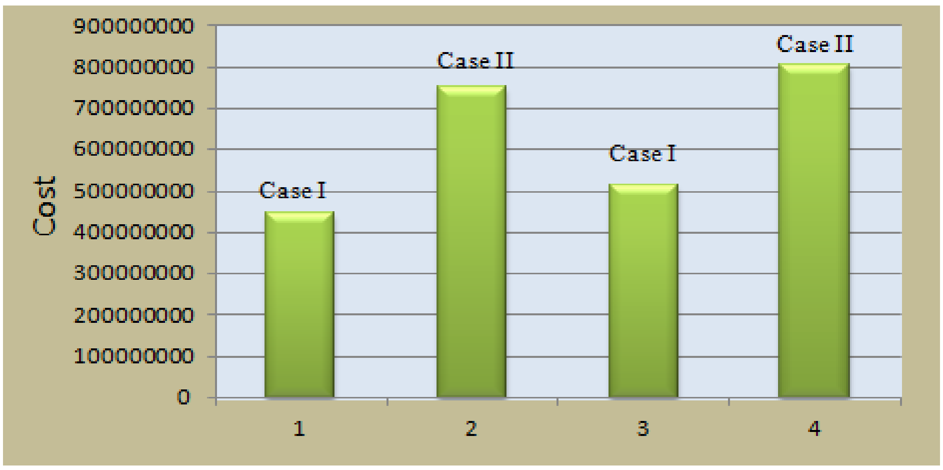


Figure 10: *Pair of graph lines for Case-I and Case-II*

VII. Discussion

In Section VI the data description is in Table 1 and 2 where 150 processes are presented assuming all finished before T. Their total processing time and size process measures are noted. The proposed estimate t_{mean} has unknown constants €$_1$, €$_2$ and d whose suitable values need to be obtained for obtaining a best estimate. Two cases are considered herein as

Case I: €$_1$ = (€$_1$)$_{opt}$, €$_2$ = (€$_2$)$_{opt}$, and $d_1 = 0$, $d_2 = 0$, $d_3 = 0$.

This case indicates for no use of size measure in the estimation strategy at the optimum choice of €$_1$ and €$_2$. The average confidence interval length, under Case-I is 31.38 as evident form table 3. The lowest predicted total remaining time is 11540.1 units while highest is 14991.9 units (table 10). Average cost consumption for lowest estimated time is 445639585.25 units and at highest time level it is 751097928.59 units (table 8).

Case II: €$_1$ = (€$_1$)$_{opt}$, €$_2$ = (€$_2$)$_{opt}$, and $d_1 = (d_{opt})_1$, $d_2 = (d_{opt})_2$, $d_3 = (d_{opt})_3$

This case contains choice of all constants at the optimum level and size measure information x has also been used. The impact of using the support information seems positive since the average length reduced to 30.53 in this case with respect to Case-I while simulated over 10 samples. Figure 9 also reveals for more condensed pair of graph lines for Case-II. Lowest predicted remaining time is 12406.9 units and highest is 15521 units (Table 10). Average cost likely to consume is 515111513.72 units as minimum whereas 805318434.65 units as highest (Table 9).

The percentage relative efficiency of Case-II with respect to Case-I is 9.82 % which supports the use of size measure in estimation (Table 2). The highest cost by Case-I and lowest by Case-II are the recommended cost required for infrastructure creation for backup management (Figure 10).

Table 10: *Ten Sample average Confidence Interval and estimated total Remaining time of processing for Recovery Management*

	Case-I (Without size measure)	**Case-II (With size measure)**	**True Value**
Average Interval (Over 10 samples)	(104.91 - 136.29)	(112.79 - 141.10)	122.51
CI Length	31.38	28.30	
Lowest Predicted Remaining time	(N-k)* 104.91 = 11540.1 units	(N-k)* 112.79 = 12406.9 units	------
Highest Predicted Remaining time	(N-k)* 136.29= 14991.9 units	(N-k)* 141.10 = 15521 units	

$$\textbf{Percentage Relative Efficiency (PRE)} = \left[\frac{[\text{Length of CI of case}-\text{I}]-[\text{Length of CI of other cases}]}{\text{Length of CI of case}-\text{I}} \right] \text{X } 100$$

Table 11: *Percentage Relative Efficiency (PRE)*

Case-II with respect to Case-I
PRE = 9.82 %

VIII. Conclusion

In case when the sudden breakdown occurs in a multiprocessor computer system this paper represents an idea of calculating the ready queue remaining processing time. The paper assumes that $(k_j - n_j' - n_j'')$ processes are completely finished before breakdown, n_j' are partially processed and n_j'' are blocked by j^{th} processor. Under this an estimation strategy is proposed for estimating the total remaining time of jobs to be processed in waiting ready queue. The proposed generalized strategy contains constants whose optimum values are derived and used. Two cases are compared where the first case is having no consideration of size measure of jobs in waiting queue whereas the second case considers the additional features of size measure of processes. The confidence interval is used as a tool for predicting about the unknown with 95% accuracy. Three cost functions are suggested for predicting about the backup infrastructure cost needed for recovery management after system breakdown. The proposed methodology under Case-II performs better than Case-I by comparing the length of confidence intervals. The highest predicted remaining time under ten considered samples is 15521 units, under Case-II. Moreover, the Case-II is 9.82 % more efficient than Case-I. The average cost required for recovery after occurrence of failure is also lower in Case-II. Overall it is found that the suggested estimation strategy is effective for predicting the remaining total time with high efficiency. The suggested is a new methodological approach for predicting the unknown using sampling methodology in the multiprocessor environment. Proposed advocates for the use of size measure of processes, if available for predicting unknown parameters like remaining time of a ready queue.

References

[1] More Sarla, and Shukla Diwakar, (2020). Some new methods for ready queue processing time estimation problem in a multiprocessor environment, *Social Networking and Computational Intelligence, Lecture notes in Networks and Systems, Springer, Singapore*, available at doi.org/10.1007/978-981-15-2071-6_54, Vol. 100, pp 661-670.

[2] More, Sarla and Shukla Diwakar, (2019). Analysis, and extension of methods in ready queue processing time Estimation in Multiprocessor Environment, *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity university Rajasthan, Jaipur-India*, available at SSRN: https://ssrn.com/ abstract = 3356312 or https:// dx.doi.org/ 10.2139/ SSRN 3356312, pp 1558-1563.

[3] More, Sarla and Shukla Diwakar, (2018). A review on ready queue processing time estimation problem and methodologies used in multiprocessor environment, *International Journal of Commuter Science and Engineering,* available at https://doi.org/10.26438/ijcse/v6i5.11511155, Vol.6, Issue 5, pp 1186-1191.

[4] Diwakar Shukla and Sarla More, (2020). Modified group lottery scheduling algorithm for ready queue mean time estimation in multiprocessor environment, *Reliability: Theory & Applications (RT&A),* Vol. 15,No 4(59), pp 69-85.

[5] Shukla Diwakar, Jain Anjali and Choudhary Amita, (2010). Prediction of ready queue processing time in multiprocessor environment using lottery scheduling (ULS), *International Journal of Commuter Internet and Management*, Vol.18, No.3, pp 58-65.

[6] Shukla Diwakar, Jain Anjali and Choudhary Amita, (2010). Estimation of ready queue Processing time under usual group lottery scheduling (GLS) in multiprocessor environment, *International Journal of Commuter Applications*, Vol.8, No.14, pp 39-45.

[7] Shukla Diwakar, Jain Anjali, and Choudhary Amita, (2010). Estimation of ready queue processing time under SL scheduling scheme in multiprocessors environment, *International Journal of Computer Science and Security*, Vol. 4, Issue 1, pp 74-81.

[8] Carl A. Waldspurger and E William Weihl, (1994). Lottery Scheduling: Flexible proportional share resource management, *The 1994 Operating systems design and implementation conference (OSDI '94), Monterey, California*.

[9] Johnnie Daniel, (2011). Sampling Essentials: Practical Guidelines for Making Sampling Choices, *Sage Publication*.

[10] Paul S. Levy and Stanley Lemeshow, (2008). Sampling of Populations: Methods and Applications, *Wiley Series in Survey Methodology,* Volume 543.

[11] Sampath, S., (2005). Sampling Theory and Methods, *Alpha Science International Publication.*

[12] Cochran, W.G, (2005). Sampling Technique, *Wiley Eastern publication, New Delhi*.

[13] Poduri S. R. S. Rao, (2000). Sampling Methodologies with Applications, *Texts in Statistical Science, Chapman and Hall/CRC Press.*

[14] Ranjan K. Som, (1995). Practical Sampling Techniques, *Second Edition Statistics: A Series of Textbooks and Monographs, CRC Press.*

[15] Steven K. Thompson, (1992). Sampling, *Wiley Series in Probability and Statistics*, Volume 272.

[16] Nurbek Saparkhojayev, Yermek Nugmanov, Amy Apon, Mamyrbek Beysenbi, (2013). Dynamic Lottery Scheduling, *AWERProcedia Information Technology & Computer Science,* Vol 03 3rd World Conference on Information Technology (WCIT-2012), pp1310-1318.

[17] J. Prassanna and Neelanarayanan Venkataraman (2019). Adaptive regressive holt–winters workload prediction and firefly optimized lottery scheduling for load balancing in cloud, *Springer Science+Business Media, LLC, part of Springer Nature, Wireless Networks* https://doi.org/10.1007/s11276-019-02090-8

[18] Hala ElAarag, David Bauschlicher, and Steven Bauschlicher, (2011). Simulation-Based Comparison Of Scheduling Techniques In Multiprogramming Operating Systems on Single and Multi-Core Processors, *The Journal of Computing Sciences in Colleges*, Volume 27, Number 2 Papers of the Twentieth Annual CCSC Rocky Mountain Conference October 14-15, Utah Valley University Orem, Utah

[19] Emily Berg, Jae-Kwang Kim, Chris Skinner, (2016). Imputation Under Informative Sampling, *Journal of Survey Statistics and Methodology*, Volume 4, Issue 4, https://doi.org/10.1093/jssam/smw032, pp 436–462.

[20] Graham Kalton & Leslie Kish, (2007). Some efficient random imputation methods, *Communications in Statistics - Theory and Methods* , Volume 13, - Issue 16, pp 1919-1939.

[21] David A. Binder and Weimin Sun, Frequency Valid Multiple Imputation for Surveys with a Complex Design Statistics Canada Business Survey Methods Division, Statistics Canada, Ottawa, ON, Canada K IA 0T6

[22] J. K. Kim, S. Yang (2017). A note on multiple imputation under complex sampling, *Biometrika*, Volume 104, Issue 1, pp 221–228,

[23] Michael R. Elliott (2021). Weighted Dirichlet Process Mixture Models to Accommodate Complex Sample Designs for Linear and Quantile Regression, *Journal of official Statistics*, http://dx.doi.org/10.2478/JOS-2021-0004, Vol. 37, No. 1, pp. 71–95

[24] Dimitris Bertsimas, Colin Pawlowski, and Ying Daisy Zhuo, (2018). From Predictive Methods to Missing Data Imputation: An Optimization Approach, *Journal of Machine Learning Research* 18, 1-39.

[25] B. K. Singh and Upasana Gogoi, (2017). Estimation of Population mean using Exponential Dual to Ratio Type Compromised Imputation for Missing data in Survey Sampling, *Journal of Statistics Applications & Probability An International Journal*, Vol. 6, No. 3, 515-522

[26] Sarla More and Diwakar Shukla (2021). Sampled Ready queue processing time estimation using size measure information in multiprocessor environment, *Reliability: Theory and application (RT&A)*, No 3(63), vol.16, pp 63-80.

IMPROVEMENT OF MANAGEMENT METHODS FOR THE OPERATIONAL RELIABILITY OF DISTRIBUTED ENERGY FACILITIES

Farhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A.
•
Azerbaijan Scientific-Research and Design-Prospecting Institute of Energetic
e-mail: elmeht@rambler.ru

Abstract

Improving the management of the technical condition of equipment, devices and installations, the service life of which exceeds the standard value, is one of the most important problems of state security. Today, the relative number of such equipment already exceeds 60%. The results of the analysis of literature data on this problem presented, which confirm its relevance and significance. It is important to note that these findings apply to not only electrical power systems, but many other production systems as well. The main difficulties in solving the analyzed problem, first of all, the paucity of statistical data characterizing the reliability of work, their multidimensional and random nature. The authors propose to solve this problem by moving from average annual reliability indicators to average monthly indicators of operational reliability. A brief description of the solution of individual tasks of this problem for overhead power lines is given, which together represent a new methodology for managing the technical condition of distributed type objects. Science-intensive, cumbersome and labor-intensive calculation algorithms determine the expediency of the transition to intelligent systems. At the same time, the management of the electric power system and its individual production enterprises will monthly receive specialized forms indicating recommendations that optimize the increase in the reliability of overhead power transmission lines by restoring wear and tear.

Keywords. Operational reliability, technical condition, automated control, risk-based approach, work efficiency, overhead power transmission lines.

I. Introduction

One of the main problems of electric power systems (hereinafter - EPS) is the improvement of the management of the technical condition of equipment, devices and installations (hereinafter - objects). The system of preventive maintenance (hereinafter - PM) traditionally used for management no longer meets the requirements, or rather, it becomes insufficient. This discrepancy arose, first of all, because the relative number of objects of the same type, the service life of which exceeds the standard value, is systematically increasing in the EPS. So back in 2012, in [1] noted that the EPS of Russia need a serious modernization of fixed assets and the replacement of almost 50% of physically and morally obsolete equipment

In [2], noted that the service life of about two-thirds of nuclear reactors exceeds the standard value, and aging management programs be introduced. The aging of objects is accompanied by an increase in the number and severity of accidents, the consequences of which not only lead to large material costs, but also to environmental violations, injury and death of personnel, and violations of state security.

At the same time, the mandatory replacement of the main facilities, the service life of which exceeds the standard value with a new one, not only creates great economic difficulties, but also inexpedient. And, first of all, because the technical condition (hereinafter - TC) of the main objects in most cases depends not only on the calendar service life, but also on the operating conditions and, above all, on the load.

II. Features of Control Methods for TS of Hazardous Production Facilities

Along with an increase in the relative number of facilities whose service life exceeds the standard value, an increase in the number of unacceptable accidents, the number of publications that form the main comments on the methods of managing the TC of production facilities based on a risk-based approach (hereinafter - RBA). Below are a number of opinions and recommendations on the results of applying these methods for the period from 2011 to 2020, which cannot be disagree. But, first of all, we will agree that by "method" we will understand a "tool" for solving the task, by "approach" - the choice of a certain method, and by methodology - an objective sequence of methods for solving the problem under study. This clarification of terms is formulated on the basis of familiarization with a number of scientific studies on their difference. The most important recommendations, in our opinion, include the following.

In [3], the scientific foundations for ensuring the safety of production facilities considered. It is noted that:

❖ absolute safety of hazardous production facilities (hereinafter - HPF) cannot be ensured in principle;

❖ the risk of HPF TC assessment is a quantitative measure of the risk of unacceptable events occurring in case of HPF failures;

❖ the main task of HPF managers is to maintain the normatively established permissible level of danger (risk);

❖ despite the variety of recommendations on the topic "safety of HPFs", they are of a declarative nature;

❖ the main directions for improving the HPF safety system are:

– creation of an approach to timely (operational) accounting of the impact of HPF TC on safety;

– taking into account the multifactorial impact on the safety of HPFs.

In [4], it is noted that in order to ensure the safety of HIFs of the oil and gas complex, the solution of problems related to the prevention of possible emergencies and the minimization of technological and environmental risks is becoming increasingly important. Currently existing methods for assessing industrial safety risks do not take into account the constantly changing in time non-stationary random nature of risks

In [5], noted that risk analysis is a rapidly developing interdisciplinary scientific direction in which:

❖ the conceptual apparatus of risk analysis has not yet been formulated and differs significantly by industry;

❖ a very large proportion of the quantitative reliability indicators used and the imperfection of the methodology for their use (formation of integral reliability indicators);

❖ imperfection of the existing methodological base. All calculations recommended to be carried out in point setting.

In [6], noted that the disadvantages of existing methods for quantitative risk assessment are:

❖ assessment of the probability of occurrence of an event is carried out on a limited amount of initial statistical data;

❖ despite the cumbersome and science-intensive methods, the calculations are performed manually;

❖ does not take into account the multidimensional nature of the risk.

In [7], noted that quantitative estimates of the frequency of occurrence of failures of the same type of objects differ from each other by two to four orders of magnitude. An analysis of the reasons for this discrepancy shows:

- the volume of initial data is insufficient;
- when performing calculations, a number of unacceptable assumptions are made;
- the presence of significant errors in the calculations of reliability indicators due to insufficient qualifications of performers;
- the more indicators that characterize the reliability of HPF, the greater the spread of calculation results.

III. Peculiarities of Normative Methods of Control of TC HPF EPS

The normative methods of managing the TC HPF EPS include, first of all, [8,9]. Unfortunately, the shortcomings of the HPF security methods noted above in [8,9] have not been eliminated. Yes, they not taken into account, despite the fact that the experience of the practical use of these methods discussed at scientific and practical conferences on the topic “Control of the technical condition of the equipment of electric power facilities”, held in 2018 and 2019.

At the same time, noted in [12] that:

- the probability of an event occurring in the future is determined based on the frequency of occurrence of this event in the past. This does not take into account the possibility of adjusting actions that transform the flow of events;
- there is an insufficient volume and heterogeneity of the sample of statistical data on failures;
- indicators such as the number, frequency and average severity of accidents are completely unfounded;
- statistics of past years (for 3-5 years) in full accordance with the rules of probability theory and mathematical statistics is unsuitable, as it is non-random and heterogeneous.

In [13] it is noted:

- in the foreseeable future, energy companies will have to solve the problem of ensuring reliability in the face of high equipment wear and lack of resources;
- along with the systems of preventive maintenance and repair according to the TC, it is advisable to introduce a repair system based on the RBA.
- Existing methods for predicting TC HPF based on RBA do not cover power lines and devices with a voltage of 0.4-10 kV. But these transmission lines make up about 90% of the length of networks, belong to the distribution electrical network (hereinafter - DEN) and are sources of 80% of accidents;
- methods [8] and [9] are actively used in the EPS of the Russian Federation;
- discrete index of technical condition (hereinafter - ITC) serves as the basis for ranking EPS objects (it does not take into account the random nature of the initial data);
- for the management of operational repair activities, the calculated ITC is unacceptable.
- there are significant gaps in the regulation of RBA

IV. Recommended Approaches to Managing the Technical Condition of Overhead Power Transmission Lines

EPS objects classified into objects of continuous (transformers) and discrete (switches) action, concentrated (power units) and distributed power transmission lines. This difference causes the difference in the number of indicators of their reliability and methods of evaluation. Considering that more than 80% of accidents in EPS are associated with DEN [13], we will consider the solution to this problem for overhead power transmission lines (hereinafter - OPTL).

The foregoing allows us to conclude that the improvement of the management of the TC OPL EPS provides for the possibility of objectively solving a number of tasks, which, first of all, include:
1. Ranking of distribution network enterprises (hereinafter - DNE) according to the degree of aging (hereinafter - DA) of the OPTL and recognition of DNE, the DA of the OPTL of which is the largest;
2. Evaluation of objective indicators of reliability of high OPTL;
2.1. Overcoming the subjectivity of selective survey;
3. Accounting for the random nature of estimates of operational reliability indicators and differences in the degree of technical use of high-voltage transmission lines when:
❖ comparison of estimates in the settlement and previous months;
❖ ranking of the DNE;
❖ recognition of "weak links"
4. Accounting for the random and multidimensional nature of the operational reliability of the OPTL EPS when:
❖ ranking of the DNE;
❖ recognition of "weak links"
5. Assessment of the objectivity of recommended methods and algorithms
6. Minimization of the risk of an erroneous decision with the methodological support of the management of the EPS and DNE.

First of all, let's clarify our attitude to the adopted system of designing power supply systems. And, in particular, to ensure the reliability of power supply, taking into account the category of consumers and the requirements of the Electrical Installation Rules.

Many years of experience in the design and operation of power, supply system’s indicates that in the vast majority of cases, the commissioned power supply system’s meet the requirements. Therefore, there is no reason to doubt the infallibility of design methods. At the same time, the risk of occurrence of unacceptable events exists (recall that absolute safety is excluded [3], although it is insignificant.

Consequently, the meaning of the control of the TC of the power supply system is not at all in assessing the changes in the consequences of a power supply failure. Possible consequences are known from the project documentation. *Not a change in the size of the consequences causes a change in the reliability of objects of power supply systems, but vice versa. Changing the reliability indicators changes the size of the possible consequences.* Thus, when analyzing the TC of power supply facilities, it is sufficient to assess the significance of changes in estimates of the reliability indicators of objects. But the significance of changes in the reliability indicators of specific objects, as a rule, cannot be determined, since there could be no failures, or there are so few of them that quantitative estimates are unreliable. If we also take into account that there are thousands and even tens of thousands of such objects in the EPS, then the cumbersome and laborious calculations also cause a high risk of an erroneous decision and the undoubted advantages of the intuitive approach of specialists.

But you can approach the problem in a slightly different way. Let us summarize some of the results obtained by us and published in 2021, results of solving the problem for OPTL.

In [14], quantitative estimates of DA OPTL are analyzed (see paragraph 1). Shown, that the estimate of the relative number of OPTL, the service life of which exceeds the standard value, used in practice, is unacceptable for comparing and ranking DA OPTL DNE EPS. A method and algorithm for solving this problem developed.

In [15] (see paragraphs 2.1 and 2.2) noted, that it is advisable to evaluate the TC OPTL by operational indicators of the reliability of work. The average monthly values of the probability of automatic emergency shutdowns of the OPTL calculated from the data of operational logs. The results of an objective classification make it possible to identify "weak links" and thereby recognize the OPTL that reduce the efficiency of the EPS to the greatest extent. Reducing the risk of an erroneous decision during the survey of the TC OPTL is achieved by the recommended method of forming a representative sample of the elements of these OPTL to be tested.

In [16] (see paragraph 3), a new method and algorithm for the operational assessment and comparison of the reliability indicators of OPTL is recommended. A distinctive feature of the calculation method is taking into account the degree of technical use of OPTL. The algorithm for estimating operational reliability indicators developed for the specific number of emergency automatic shutdowns of OPTL, for stable failures, the probability of stable failure and the technical utilization factor. Comparison of estimates for two adjacent operational intervals is supposed to be carried out on the basis of the boundary values of fiducial intervals

In [17] (see paragraph 4), a new algorithm for assessing the feasibility of classifying multidimensional data on failures and downtime of OPTL was developed. A method and algorithm for estimating the operational integral indicator of reliability, the physical essence of which is adequate to the wear of the TC, developed. A method and algorithm for comparing integral indicators of operational reliability developed, taking into account errors of the first and second kind. A method and algorithm for ranking indicators of operational reliability has been developed, which makes it possible to identify OPTL that require prompt survey.

In [18] (see paragraph 5), on the example of statistical data on failures of the main 110 and 220 kV OPTL, the results of assessing and comparing operational reliability of operation are given. The purpose of manual analysis is to recognize the varieties of signs for which the specific number of stable failures will exceed this indicator to the greatest extent for the set of OPTL. The identified significant varieties of signs completely coincided with the results of choosing the least reliable OPTL based on an intuitive approach, which confirms the correctness of the recommended method of quantitative analysis.

In [19], a method and algorithm for automated comparative analysis (benchmarking) of the operational reliability of DEN objects is proposed. The transition from an intuitive approach to a quantitative approach to the operational recognition of “weak links” among thousands of objects of the same type made it possible to reduce the risk of an erroneous decision when organizing their maintenance and repair.

For illustrative purposes, Fig. 1 shows an enlarged block diagram of the algorithm, which, based on the methods described in [14-19], allows you to automate the process of generating recommendations to improve the operational reliability of the OPTL EPS.

The results of the monthly automated monitoring of the operational reliability of the operation of the OPTL and recommendations for improving the efficiency of work formalized in the form of special forms and provide information and methodological support for the technical management of the EPS and each ENE. For illustrative purposes, Fig. 2 shows one of the variants of these forms.

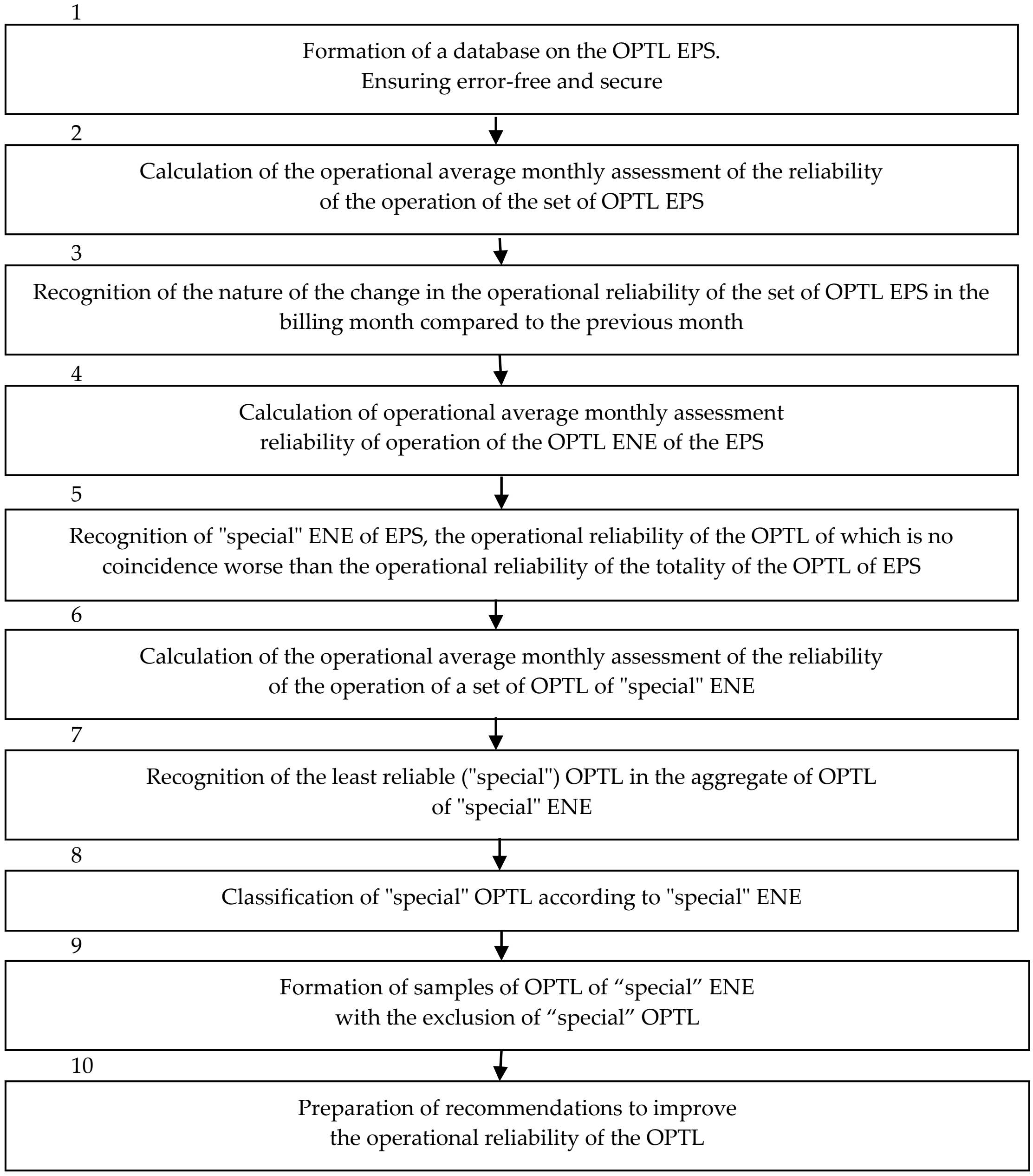


Fig. 1. An enlarged block diagram of the algorithm for improving the operational reliability of the OPTL EPS

Date of analysis
05 may 2021 year

EPS Chief Engineer
____

Information
on the operational reliability of OPTL
in the month of April 2021

1. Compared to the previous month, the operational reliability of the OPTL EPS has ***decreased***;
2. Table A shows the average monthly estimates of the operational reliability of the operation of the OPTL ENE EPS, arranged in descending order.

Table A. Information about the operational reliability of the OPTL ENE EPS

№	Name ENE	Rating ENE	№	Name ENE	Rating ENE
1	***ENE 5***	***Good***	6	***ENE 9***	***Satisfactory***
2	***ENE 3***	***Good***	7	***ENE 6***	***unsatisfactory***
3	***ENE 7***	***Good***	8	***ENE 2***	***unsatisfactory***
4	***ENE 8***	***Good***	9	***ENE 4***	***unsatisfactory***
5	***ENE 1***	***Satisfactory***	10	***ENE***	***unsatisfactory***

3. Table B1 shows the passport data of the OPTL EPS, the operational reliability of which, as well as the rating of their ENE, is ***unsatisfactory***.

Table B1.

№	Name		Voltage class, kV.	Service life, year	Length, км	Execution
	ENE	OPTL				

4. Table B2 shows the passport data of the OPTL EPS, the operational reliability rating of which is unsatisfactory, and the service life does ***not exceed*** the standard value.

Table B2.

№	Name		Voltage class, kV.	Service life, year	Length, км	Execution
	ENE	OPTL				

5. Recommendations for improving operational reliability:
5.1. Carry out a selective survey of the OPTL indicated in Table. B1
5.2. Establish and eliminate the causes of automatic shutdowns of OPTL, given in Table. B2.
5.3. Arrange for the personnel of the ENE indicated in Table. B1, advanced training courses on the problem of maintenance of aging OPTL.

Fiq. 2. The results of the analysis of the average monthly operational reliability of the OPTL EPS

RT&A, No 2 (68)
Volume 17, June 2022

Conclusion

1. The relevance of improving the methods of operating reliability management confirmed not only for EPS, but also for many other production systems.
2. The solution of this problem proposed to carry out by the transition to the analysis of integral indicators of operational reliability.
3. Methods for calculating the integral indicators of operational reliability, taking into account the random nature of quantitative estimates, assessing the feasibility of classifying data according to varieties of comparison features and ranking the estimates form the methodology of the operational reliability management system in EPS.
4. These methods based on a fiducial approach, simulation modeling of representative samples, taking into account errors of the first and second kind
5. The science-intensive, cumbersome and labor-intensive manual calculation, the high risk of subjective errors and prerequisites necessitate the transition to intelligent control systems for objects whose service life exceeds the standard value.
6. Automated management of operational reliability increases the importance of an intuitive approach in the implementation of methodological recommendations, which limits the risk of an erroneous decision.

References

[1] Basic provisions of the concept of an intelligent power system with an active-adaptive grid. M, FSKES, 2012
[2] Safe lifetime management of nuclear power plants INSAG-14 (Report of the International Nuclear Safety Group)/IAEA, Vienna, 2014, 47 p.
[3] Kuzmin A.A. Organization of ensuring the safety of a hazardous production facility based on a quantitative assessment of the state of its safety system. "Technology of civil security", 2011, volume 8 No. 4 (30), pp. 72-79
[4] Aburakhmanov N.Kh., Shaibakov R.A., Markov A.B. Information models for managing the risk minimization of potentially hazardous production facilities//GIAB 2015, No.57, URL, pp. 729-736.
[5] Kolesnikov F.Yu. Main problems of accident risk analysis methodology. Fire safety, 2016, vol.25 No 2, pp. 4-8.
[6] Lisanov M.V., Pecherkin A.S., Shvyryayev A.A. Methodological support and problems of accident risk analysis at hazardous production facilities of the oil and gas complex - M. "Improving the reliability and safety of gas industry facilities" No.1 (29), 2017, pp. 179-184.
[7] Lisanov M.V., Sumskoy S.I., Shvyryayev A.A. Uncertainty in the quantitative assessment of the risk of accidents at oil and gas facilities. M: "Improving the reliability and safety of gas industry facilities", No2 (34), 2018, pp. 125-135.
[8] Methodology for assessing the technical condition of the main technological equipment and transmission lines of power stations and electrical networks. https//meganorm.ru>Index2. Approved by the Ministry of Energy of the Russian Federation in 2017. Appendix to Order No. 676. Year of update 2021 30 s.
[9] Guidelines for calculating the probability of failure of a functional unit and a unit of main technological equipment and assessing the consequences of such a failure https//meganorm.ru.>Index2. Ministry of Energy of the Russian Federation, 2019 Appendix to Order No. 123, 35 p.
[10] V Scientific-practical conference "Control of the technical condition of electric power facilities"./Electricity. Transfer and distribution. – 2018
[11] VI Scientific-practical conference "Control of the technical condition of electric power facilities"./Electricity. Transfer and distribution. – 2019
[12] Fedorets A.G. Application of modern risk management methodology in safety and labor protection management systems. / Safety and labor protection. - 2018, No.1 p.1-18.
[13] Antonenko I.N. Analysis of the risk-oriented approach to the management of physical assets. M. "Industrial Energy" No.10, 2020, pp. 57-60.
[14] Farkhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A., Nazarov A.A. Automated analysis of the service life of overhead transmission lines of electric power systems. Minsk, Energy. Izv. higher textbook institutions and energy. associations of the CIS. Energy Problems Vol. 64, No. 5, 2021, pp. 435-445. DOI: 10.21122/1029-7448-2021-64-5-435-445
[15] Farhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A., Nazarov A.A. Assessment of the technical condition of overhead high-voltage power lines at the stage of their aging. M:, Reliability and safety of energy. Volume 14, No 2, 2021, pp. 100-107. DOI: 10.24223/1999-5555-2021-14-2-100-107
[16] Farkhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A., Nazarov A.A. Quantitative assessment of the operational reliability of overhead power lines "M: Electric stations, No.2, 2021, pp. 28-34. DOI: 10.34831/EP.2021.1081.8.004
[17] Farhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A. Estimation, comparison and ranking of operational reliability indicators of overhead transmission lines of electric power systems. Journal: Reliability: Theory&applications. RT&A (Vol.4 No (65)) 2021, Desember, USA, p.186-197. DOI: 10.24412/1932-2321-2021-465-186-196

[18] Farhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A. Quantitative analysis of operational reliability of overhead transmission lines. M:. Operational management in the electric power industry, No.5, 2021, p.16-24

[19] Farhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A. Improving the operational management of the efficiency of distribution networks. M:, Industrial Energy No.12, 2021, pp. 16-23. DOI: 10.34831/EP.2021.77.85.002

RELIABILITY CRITERIA FOR DESIGNING LIFE TEST SAMPLING INSPECTION PLANS BASED ON LOMAX DISTRIBUTION

R. Vijayaraghavan[1] and A. Pavithra[2]

•

Department of Statistics, Bharathiar University
Coimbatore 641 046, Tamil Nadu, INDIA

[1]vijaystatbu@gmail.com, [2]pavistat95@gmail.com

Abstract

Acceptance sampling plays an important role in ensuring the quality of the products manufactured by the industrial production processes. Sampling inspection plans by attributes are adopted for taking decisions about the lots submitted for inspection. Such procedures are employed for sentencing individual lots or batches or lots in continuous stream. Reliability sampling is s specific inspection procedure which is used to decide whether the submitted lot or batch is acceptable or non-acceptable based on life tests. In reliability sampling, the lifetime of the items randomly drawn from the lot is considered as a random variable which follows a continuous probability distribution. In this paper, designing of single sampling plans for life tests is considered under the assumption that the lifetime random variable follows a Lomax distribution. Reliability criteria for designing life test plans when lot quality is evaluated in terms of mean life, median life, hazard rate and reliability life are proposed. Conversion factors for adapting acceptable quality levels to life and reliability testing under the assumption of Lomax distribution are determined and suitable illustrations are provided.

Keywords: Acceptable quality level, Consumer's risk, Lomax distribution, Operating characteristic function, Producer's risk, Reliability sampling, Single sampling plan.

1. Introduction

Sampling inspection is a product control strategy that decides whether a lot should be accepted or rejected based on the information obtained by the inspection of random sample(s) drawn from the submitted lot(s). Sampling inspection procedures are generally classified according to the nature of the quality characteristics, *viz*., measurable and non-measurable characteristics. When the quality characteristics are non-measurable, but are classified into go or no-go basis, such as good or bad, non-conforming or conforming, etc., the sampling inspection procedures are termed as attribute sampling. When the quality characteristics are measurable on a continuous scale, the corresponding sampling inspection procedures are called variables sampling, which are devised under the implicit assumption that the quality characteristic is a continuous random variable following a specific probability distribution. Reliability sampling plans, also termed as life test sampling plans, are operationally attributes sampling procedures, but involve lifetime of the components or items as a random variable which is distributed according to a specific continuous type probability distribution, such as the exponential, Weibull, lognormal, gamma distributions, etc. The lifetime of the components or items is observed by placing the sampled items under the test, called life test, which is defined as the process of evaluating the lifetime of the items through experiments. The literature in product control provides the importance of various continuous probability distributions like exponential, Weibull, lognormal and gamma distributions as well as several compound distributions for modeling lifetime data in the studies relating to the design and evaluation of reliability sampling plans.

The earlier works, which laid the foundation for the expansion of several types of sampling plans, would include the theory of reliability sampling proposed and developed by various authors. One may refer [1] – [8] for the basic notions and terminologies of sampling inspection for life tests. The literature in statistical product control provides significant studies relating to the construction of life test sampling plans employing exponential, Weibull, lognormal and gamma distributions as well as several compound distributions for modeling lifetime data. A detailed account of the properties, methods of construction and performance of such plans is provided in [9] – [25], and the recent advances in the theory of life test sampling plans are discussed in [26] – [30].

Lomax distribution, introduced in [31], is a heavy-tailed probability distribution and is considered as Pareto Type II distribution. It has a wide range of applications in many fields which include business, economics, actuarial, medical and biological sciences. It has been proved to be much useful in reliability and life testing studies and in survival analysis. Properties of Lomax distribution and its extended form can be seen in [32] – [35]. In this paper, a specific life-test sampling plan is devised with reference to the life-time quality characteristic, which is modeled by Lomax distribution. A procedure for the selection of such plans indexed by acceptable and unacceptable mean life is evolved. Three different criteria for designing life-test plans when lot quality is evaluated in terms of mean life, hazard rate and reliability life are proposed. Factors for adapting acceptable quality level to life and reliability testing under the assumption of Lomax distribution are also illustrated.

2. Life Test Sampling Inspection Plans Based on Lomax Distribution

Sampling plans for life tests include a set of sampling procedures and rules for deciding whether to accept or reject a large number of items based on the sampled lifetime information about the items. Sampled items are tested for a set period of time under such plans. When all units are tested to failure, the standard plans can be used to compare the performance to the specified requirements, and the results can be used in an attribute sampling plan when the lifetimes are tested and the distributional assumption of the quality characteristics is fulfilled. Further, the number of failures which occur before a required time can be used with standard attributes plans in determining the disposition of the material. (See, [9]).

A typical life test sampling plan can be formulated in the following manner: Suppose, n items are placed for a life test and the experiment is stopped at a predetermined time, T. The number of failures occurred until the time point T is observed, and let it be d. The lot is accepted if d is less than or equal to the acceptance number, say, c; otherwise, it is rejected. Thus, the life test sampling plan is represented by n, the number of units on test, and the maximum allowable number of failures, c, called the acceptance number. Life tests, terminated before all units have failed, may be classified into two types, namely, failure terminated and time terminated. In a failure terminated life test, a given sample of n items is tested until the specified number of failures occurs and then the test is terminated. In time terminated life test, a given sample of n items is tested until a pre-assigned termination time, t, is reached and then the test is terminated.

Generally, these tests may be defined with reference to the specifications given in terms of one of the characteristics such as (i) the mean life, that is, the expected life of the product, (ii) the median life. (iii) the hazard rate, that is, the instantaneous failure rate at some specified time, t, and (iv) the reliable life, that is, the life beyond which some specified proportion of items in the lot will survive. One of the significant features of a life test plan is that it involves a random characteristic, called lifetime or time to failure, which can be more adequately described most often by skewed distributions. Application of continuous-type of distributions such as normal, exponential, Weibull, gamma and lognormal for lifetime variables in the studies concerned with the design and evaluation of life test sampling plans has been provided in the literature of sampling inspection, demonstrating the significant contributions of those distributions in the development of life test sampling plans. The Lomax distribution, one of the lifetime distributions, is now considered as the lifetime distribution for the design and evaluation of life-test sampling plan.

3. Lomax Distribution

Let T be a random variable representing the lifetime of the components. Assume that T follows Lomax distribution. The probability density function and the cumulative distribution function of T are, respectively, defined by

$$f(t;\theta,\lambda)=\frac{\lambda}{\theta}\left(1+\frac{t}{\theta}\right)^{-(\lambda+1)}, t>0,\theta>0,\lambda>0 \tag{1}$$

and $$F(t;\theta,\lambda)=1-\left(1+\frac{t}{\theta}\right)^{-\lambda}, t>0,\theta>0,\lambda>0, \tag{2}$$

where λ and θ are the shape and scale parameters, respectively.

The mean life, the median life, the reliability function and hazard function for specified time t under Lomax distribution are. Respectively, given by

$$\mu=\frac{\theta}{\lambda-1},\ for\ \lambda>1, \tag{3}$$

$$\mu_d=\theta\left(\sqrt[\lambda]{2}-1\right), \tag{4}$$

$$R(t;\theta,\lambda)=\left(1+\frac{t}{\theta}\right)^{-\lambda}, t>0,\theta>0,\lambda>0 \tag{5}$$

and $$Z(t;\theta,\lambda)=\frac{\lambda}{\theta}\left(1+\frac{t}{\theta}\right)^{-1}, t>0,\theta>0,\lambda>0. \tag{6}$$

The reliability life is the life beyond which some specified proportion of items in the lot will survive. The reliability life associated with Lomax Distribution is defined and denoted by

$$\rho(t;\theta,\lambda)=\theta\left(R^{-1/\lambda}-1\right), \tag{7}$$

where $\underline{R}$ is the proportion of items surviving beyond life ρ. The proportion, p, of product failing before time t, is defined by the cumulative probability distribution of T and is expressed by

$$p=P(T\leq t)=F(t;\theta,\lambda). \tag{8}$$

4. Application of Lomax Distribution in Reliability Sampling

The techniques for determining life test sampling plans based on Weibull distribution with mean life, hazard rate, and reliability life serving as reliability criteria for the submitted lots are discussed in [36] – [38]. The dimensionless quantities, *viz.*, $t/\mu\times100$, $tZ(t)\times100$ and $t/\rho\times100$, referred to as conversion factors, for determining the life test sampling plans under the reliability criteria are introduced in [39]. Analogous approaches are discussed, here, to construct the life test sampling plans using Lomax Distribution as the lifetime distribution for the lifetime quality characteristic.

The mean life criterion is determined by calculating the ratio $t/E(t)=t/\mu$, which is associated with the proportion, p, of products that fail before reaching the termination time t. Acceptable mean life and unacceptable mean life, which are associated with the producer's risk and the consumer's risk, are the two typically stipulated requirements when dealing with a life test sampling plan in practice for providing protection to the producer and the consumer, respectively. A desired sampling plan can be determined with the specification of these indices. The quality levels, corresponding to acceptable and unacceptable mean life, are defined by p_1 and p_2 with associated risks, where p_1 is the acceptable proportion of the lot failing before the specified time, t, and p_2 is the unacceptable proportion of the lot failing before the specified time, t. Based on mean life, median life, hazard rate, and reliability life as the criteria for life test plans under the Lomax Distribution, conversion factors are obtained. These conversion factors are used for deriving the plans satisfying the requirements.

When the test termination time is defined, the conversion factors can also be used to calculate the mean life, median life, hazard rate, and reliability life, and vice versa. The appropriate conversion factors in terms of percentages are computed and are provided in Table 1 through to 6. Numerical illustrations for demonstrating the use of tables for determining the operating characteristics of a given plan and finding the parameters of the single sampling plan satisfying the requirements in terms of acceptable quality level (acceptable mean life) and limiting quality level (unacceptable mean life) are provided in the following subsections.

4.1 Numerical Illustration

Under life testing experiments for ascertaining the reliability of components, an industrial practitioner desires to use a single sampling plan by attributes satisfying the requirements $(p_1 = 0.007, \alpha = 0.05)$ and $(p_2 = 0.05, \beta = 0.10)$. The past experimental results on the components produced by the industry have shown that the life time of the components follows Lomax distribution specified by the shape parameter $\lambda = 1.5$. For the specified requirements, the parameters of an optimum sampling plan is determined as n = 105 and c = 2 using the searching algorithm given in [40]. It is assumed to employ a test termination time of 250 hours and to count the number of failures over the span of 250 hours. Under the given conditions, the operating characteristics in terms of mean life are obtained using the operating characteristic function of the single sampling plan by attributes and provided in Table 7 along with the values of $k = t / \mu \times 100$ and $\mu = t / k \times 100$, where t = 250 hours.

4.2 Numerical Illustration

Suppose that a single sampling plan by attributes with parameters n and c is to be defined when the requirements are specified in terms of acceptable mean life of 200 hours and unacceptable mean life of 70 hours with the associated producer's and consumer's risks of 5% and 10%, respectively. Assume that the individual items are to be tested for 3 hours and that the lifetime of the items is distributed as Lomax Distribution with the shape parameter fixed as $\lambda = 2$. Then, at the specified levels, the values of k are determined as follows:

$$k_1 = t/\mu \times 100 = (3/200) \times 100 = 1.5$$

$$k_2 = t/\mu \times 100 = (3/70) \times 100 = 4.286$$

Entering Table 1 with these values, one obtains the proportions, $p_1 = 0.03$ and $p_2 = 0.08$, of product failing before the specified time t corresponding to the acceptable mean life and unacceptable mean life, respectively. The operating ratio, which is the measure of discriminating good and bad lots of items, is defined by OR = 0.08/ 0.03 = 2.67, corresponding to which a single sampling plan can be chosen from [9] as (n = 159, c = 8) or from [41] as (n =157, c = 8).

In a similar manner, while Tables 2 and 3 can be used to determine conversion factors so as to obtain the life test plans and the corresponding median life and hazard rate, Tables 4 through to 6 can be utilized for obtaining reliability life for the specified values of R, viz., 0.90, 0.95 and 0.99, respectively.

4.3 Numerical Illustration

The acceptable mean life under the life test sampling plans based on Lomax distribution can be determined using the ratio $k = t / \mu \times 100$ for any specified value of acceptable quality level, AQL, shown in [42]. When AQL is specified as 3 percent with 95 percent acceptance probability, the test termination time is given as t = 25 hours and the shape parameter is fixed as $\lambda = 1.5$, the average or expected mean life, μ, is determined as $\mu = t/k \times 100 = (25/1.026) \times 100 = 2436.6$ hours, which can be considered as an acceptable mean life. Accordingly, if a lot consisting of items which have the acceptable mean life specified at 2436.6 hours, the probability of acceptance of the lot would be 95%. Corresponding to the fixed value of AQL = 3 percent with 95 percent acceptance probability, the conversion factors for median life, hazard rate and reliability life criteria for the case in which the shape parameter is specified as $\lambda = 1.5$ under Lomax Distribution are given in the following table:

Criterion	Conversion Factor	Value of the Factor	AQL
Percent Nonconforming as per MIL - STD -105E	$p \times 100$	3	0.03
Mean Life	$k = t/\mu \times 100$	1.026	2436.6
Median Life	$k = t/\mu_d \times 100$	3.492284	715.9
Hazard Rate at 25 hours	$tz(t) \times 100$	3.015203	0.001206
Reliable Life (R = 0.90)	$t/\rho \times 100$	28.19134	88.7
Reliable Life (R = 0.95)	$t/\rho \times 100$	58.96959	42.4
Reliable Life (R = 0.99)	$t/\rho \times 100$	305.1403	8.2

It can be noted from the above table that when the proportion, p, of products that fail before reaching the termination time, i.e., t = 25 hours is specified as the acceptable level of 3 percent, the median life of the components is 715.9 hours, 90 percent of the components will survive beyond 88.7 hours, 95 percent of the components will survive beyond 42.4 hours and 99 percent of the components will survive beyond 8.2 hours.

Table 1. Values of $t/\mu \times 100$ Based on Lomax Distribution for Specified Values of λ

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
1	0.201817	0.336136	0.431968	0.503781	0.604234	0.671146
2	0.407337	0.677979	0.870847	1.015254	1.217073	1.351392
3	0.616667	1.025686	1.316821	1.534616	1.838731	2.040957
4	0.829918	1.379418	1.770079	2.062073	2.469426	2.740067
5	1.047205	1.739346	2.230818	2.597835	3.109387	3.448954
6	1.268648	2.105644	2.699241	3.142125	3.758848	4.167861
7	1.494373	2.478495	3.175561	3.695169	4.418056	4.897038
8	1.72451	2.858088	3.659998	4.257207	5.087262	5.636744
9	1.959193	3.244622	4.152782	4.828484	5.76673	6.38725
10	2.198566	3.638299	4.654149	5.409256	6.456732	7.148834
11	2.442774	4.039335	5.164347	5.999788	7.157552	7.921784
12	2.691971	4.447953	5.683634	6.600358	7.869484	8.706402
13	2.946319	4.864383	6.212277	7.211253	8.592833	9.502999
14	3.205984	5.288869	6.750557	7.832773	9.327917	10.3119
15	3.471141	5.721661	7.298763	8.465229	10.07507	11.13344
16	3.741973	6.163023	7.857198	9.108945	10.83462	11.96797
17	4.018672	6.613231	8.426179	9.76426	11.60695	12.81585
18	4.301438	7.072571	9.006036	10.43153	12.39241	13.67746
19	4.590479	7.541343	9.597112	11.11111	13.19139	14.5532

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
20	4.886016	8.01986	10.19977	11.8034	14.00431	15.44347
21	5.188277	8.508452	10.81438	12.50879	14.83158	16.3487
22	5.497506	9.007462	11.44134	13.2277	15.67364	17.26935
23	5.813954	9.51725	12.08106	13.96058	16.53095	18.20587
24	6.137887	10.03819	12.73397	14.70787	17.404	19.15874
25	6.469584	10.57069	13.40052	15.47005	18.29327	20.12848
50	18.52753	29.37005	36.44957	41.42136	47.92619	51.98421
60	27.03458	42.10079	51.60637	58.11389	66.40499	71.44177
70	40.50026	61.57216	74.22759	82.57418	92.79668	98.76032
80	65.59747	96.20089	113.1363	123.6068	135.5481	141.9952
90	132.7393	182.0794	204.5695	216.2277	226.7829	230.8869

Table 2. Values of $t/\mu_d \times 100$ Based on Lomax Distribution for Specified Values of λ

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
1	1.089282	1.144486	1.185111	1.216236	1.260759	1.291057
2	2.19855	2.308402	2.389184	2.451041	2.539475	2.599621
3	3.328382	3.492284	3.61272	3.704892	3.836589	3.92611
4	4.479376	4.696683	4.856241	4.978283	5.152561	5.270959
5	5.652156	5.922175	6.120285	6.271729	6.487866	6.634618
6	6.847369	7.169357	7.405412	7.58576	7.842995	8.01755
7	8.065691	8.438851	8.712204	8.920928	9.218458	9.42024
8	9.307824	9.731301	10.04127	10.27781	10.61479	10.84319
9	10.5745	11.04738	11.39323	11.65699	12.03252	12.28691
10	11.86648	12.38779	12.76873	13.0591	13.47224	13.75193
11	13.18456	13.75325	14.16847	14.48477	14.93453	15.23883
12	14.52958	15.14452	15.59314	15.93467	16.42001	16.74817
13	15.90238	16.56239	17.04349	17.40951	17.92931	18.28055
14	17.30389	18.00769	18.52026	18.90999	19.46309	19.8366
15	18.73505	19.48128	20.02428	20.43687	21.02205	21.41696
16	20.19683	20.98404	21.55635	21.99094	22.60689	23.02231
17	21.69028	22.51692	23.11736	23.57301	24.21838	24.65335
18	23.21647	24.08089	24.70821	25.18393	25.85728	26.3108
19	24.77653	25.67698	26.32984	26.82459	27.5244	27.99542
20	26.37166	27.30625	27.98323	28.49593	29.22058	29.708
21	28.00308	28.96982	29.66943	30.19889	30.94671	31.44936
22	29.6721	30.66887	31.3895	31.9345	32.70371	33.22037
23	31.38009	32.40461	33.14459	33.70382	34.49253	35.02192
24	33.12847	34.17833	34.93586	35.50793	36.31417	36.85492
25	34.91876	35.99138	36.76455	37.34802	38.16968	38.72038
50	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000
60	145.9157	143.346	141.5829	140.2993	138.5568	137.4297
70	218.595	209.6426	203.6446	199.3517	193.6242	189.9814
80	354.054	327.5476	310.3914	298.4132	282.8268	273.1506
90	716.4437	619.9492	561.2397	522.0199	473.1921	444.1481

Table 3. Values of $tz(t)\times 100$ Based on Lomax Distribution for Specified Values of λ

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
1	1.001004	1.001674	1.002153	1.002513	1.003016	1.003352
2	2.004032	2.006727	2.008654	2.010101	2.01213	2.013483
3	3.00911	3.015203	3.019566	3.022844	3.027441	3.03051
4	4.016262	4.027152	4.034955	4.040821	4.049051	4.054551
5	5.025514	5.04262	5.054887	5.064113	5.077068	5.085729
6	6.036893	6.061658	6.079431	6.092805	6.111597	6.124167
7	7.050427	7.084316	7.108656	7.126985	7.152751	7.169998
8	8.066143	8.110646	8.142635	8.166739	8.200644	8.223352
9	9.084069	9.140701	9.181444	9.21216	9.255394	9.284369
10	10.10424	10.17454	10.22516	10.26334	10.31712	10.35319
11	11.12667	11.21221	11.27385	11.32038	11.38595	11.42995
12	12.15141	12.25377	12.3276	12.38337	12.46201	12.51481
13	13.17847	13.29929	13.3865	13.45242	13.54543	13.60792
14	14.20791	14.34883	14.45064	14.52763	14.63635	14.70944
15	15.23973	15.40243	15.52008	15.60911	15.73491	15.81953
16	16.27399	16.46018	16.59493	16.69697	16.84125	16.93836
17	17.31072	17.52214	17.67528	17.79133	17.95552	18.06611
18	18.34994	18.58836	18.76123	18.89230	19.07787	19.20295
19	19.39171	19.65893	19.85286	20.00000	20.20847	20.34908
20	20.43605	20.73392	20.95029	21.11456	21.34748	21.50467
21	21.483	21.81339	22.0536	22.23611	22.49505	22.66994
22	22.5326	22.89743	23.16293	23.36478	23.65138	23.84508
23	23.58491	23.98612	24.27836	24.50071	24.81664	25.03031
24	24.63995	25.07952	25.40002	25.64404	25.99101	26.22584
25	25.69776	26.17773	26.52803	26.79492	27.17469	27.43191
50	53.20635	55.50592	57.23373	58.57864	60.53543	61.88984
60	64.94378	68.56747	71.33223	73.5089	76.71379	78.95812
70	77.29026	82.77893	87.04709	90.45549	95.54978	99.17011
80	90.50676	98.70072	105.2368	110.5573	118.6736	124.5589
90	105.1888	117.6835	128.0528	136.7544	150.4732	160.7523

Table 4. Values of $t/\rho \times 100$ when R = 0.90 Based on Lomax Distribution for Specified Values of λ

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
1	9.179481	9.238821	9.281348	9.313323	9.358197	9.388186
2	18.52739	18.6345	18.7112	18.76883	18.84968	18.90367
3	28.04859	28.19134	28.29348	28.37019	28.47773	28.54951
4	37.74813	37.91381	38.03227	38.12118	38.24575	38.32885
5	47.63126	47.80655	47.93179	48.02573	48.15727	48.24497
6	57.70344	57.87439	57.99643	58.08791	58.21594	58.30125
7	67.97035	68.12234	68.23074	68.31197	68.42555	68.50119
8	78.43793	78.5556	78.63946	78.70226	78.79002	78.84843
9	89.11233	89.1796	89.2275	89.26335	89.31342	89.34673
10	99.99997	99.99998	99.99998	99.99997	99.99998	99.99997
11	111.1076	111.0226	110.9622	110.917	110.8541	110.8122
12	122.4421	122.2536	122.1197	122.0197	121.8803	121.7877
13	134.0109	133.6993	133.4782	133.3132	133.0833	132.9307
14	145.8215	145.3665	145.0438	144.8031	144.468	144.2459
15	157.882	157.2619	156.8227	156.4952	156.0397	155.7378
16	170.2006	169.3929	168.8213	168.3955	167.8035	167.4114
17	182.786	181.7671	181.0466	180.5102	179.765	179.2718
18	195.6474	194.3922	193.5055	192.8458	191.93	191.3243
19	208.7942	207.2766	206.2055	205.4092	204.3044	203.5744
20	222.2365	220.4288	219.1543	218.2074	216.8947	216.0278
21	235.9846	233.8579	232.3599	231.2478	229.7072	228.6904
22	250.0496	247.5734	245.8309	244.5383	242.7488	241.5687
23	264.443	261.5851	259.5761	258.0868	256.0266	254.669
24	279.1768	275.9034	273.6046	271.9018	269.548	267.9981
25	294.2638	290.5392	287.9263	285.9922	283.3208	281.5631
50	842.7096	807.2465	783.1628	765.7495	742.2667	727.1703
60	1229.646	1157.156	1108.825	1074.341	1028.461	999.3483
70	1842.121	1692.333	1594.869	1526.535	1437.208	1381.488
80	2983.647	2644.116	2430.87	2285.098	2099.329	1986.27
90	6037.54	5004.519	4395.421	3997.365	3512.348	3229.714

Table 5. Values of $t/\rho \times 100$ when R = 0.95 Based on Lomax Distribution for Specified Values of λ

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
1	19.27196	19.32542	19.36365	19.39236	19.43257	19.4594
2	38.89754	38.97895	39.03713	39.08078	39.1419	39.18266
3	58.88693	58.96959	59.02862	59.07288	59.13483	59.17612
4	79.25073	79.3067	79.34663	79.37656	79.4184	79.4463
5	99.99998	99.99997	99.99998	99.99998	99.99998	99.99999
6	121.1461	121.0595	120.9978	120.9516	120.8871	120.8442
7	142.7011	142.4958	142.3496	142.2403	142.0876	141.9861
8	164.6774	164.3197	164.0653	163.8751	163.6098	163.4334
9	187.0879	186.5426	186.1551	185.8656	185.4619	185.1938
10	209.946	209.1762	208.6297	208.2216	207.6529	207.2754
11	233.266	232.233	231.5002	230.9533	230.1917	229.6865
12	257.0625	255.7255	254.778	254.0714	253.0879	252.436
13	281.3507	279.6673	278.4753	277.5869	276.3513	275.5327
14	306.1467	304.0722	302.6046	301.5115	299.9921	298.9862
15	331.4672	328.9547	327.1788	325.857	324.0209	322.8062
16	357.3296	354.3299	352.2115	350.6359	348.4488	347.0027
17	383.7522	380.2136	377.717	375.8613	373.2872	371.5865
18	410.7541	406.6223	403.71	401.5468	398.5482	396.5683
19	438.3553	433.5734	430.206	427.7065	424.2441	421.9597
20	466.5768	461.0847	457.221	454.3551	450.3881	447.7725
21	495.4405	489.1753	484.7719	481.5081	476.9936	474.0191
22	524.9694	517.8648	512.8765	509.1816	504.0749	500.7125
23	555.1877	547.174	541.5529	537.3925	531.6466	527.8662
24	586.1208	577.1245	570.8206	566.1584	559.7242	555.4942
25	617.7953	607.739	600.6999	595.4977	588.324	583.6111
50	1769.236	1688.568	1633.911	1594.456	1541.338	1507.245
60	2581.594	2420.495	2313.338	2237.011	2135.629	2071.404
70	3867.462	3539.96	3327.371	3178.576	2984.404	2863.486
80	6264.051	5530.865	5071.519	4758.068	4359.318	4117.05
90	12675.58	10468.27	9170.156	8323.379	7293.492	6694.403

Table 6. Values of $t/\rho \times 100$ when R = 0.99 Based on Lomax Distribution for Specified Values of λ

p%	Shape Parameter, λ					
	1.25	1.50	1.75	2.00	2.50	3.00
1	100.0001	100.0001	100.0001	100.0001	100.0001	100.0001
2	201.8351	201.698	201.6002	201.5269	201.4244	201.3561
3	305.5578	305.1403	304.8427	304.6198	304.3081	304.1006
4	411.2234	410.3754	409.7714	409.3192	408.6875	408.2672
5	518.8889	517.4535	516.4318	515.6675	514.6003	513.8907
6	628.6139	626.4268	624.8713	623.7084	622.0856	621.0071
7	740.4604	737.3495	735.1389	733.4872	731.1837	729.6538
8	854.4928	850.2781	847.2856	845.051	841.9366	839.8694
9	970.7784	965.2712	961.3644	958.4489	954.3879	951.6942
10	1089.387	1082.390	1077.43	1073.731	1068.583	1065.169
11	1210.392	1201.698	1195.541	1190.952	1184.568	1180.338
12	1333.869	1323.261	1315.755	1310.164	1302.392	1297.245
13	1459.898	1447.149	1438.135	1431.426	1422.105	1415.938
14	1588.561	1573.433	1562.747	1554.797	1543.761	1536.463
15	1719.947	1702.188	1689.656	1680.339	1667.413	1658.872
16	1854.144	1833.493	1818.933	1808.116	1793.119	1783.215
17	1991.248	1967.429	1950.651	1938.195	1920.938	1909.549
18	2131.358	2104.082	2084.887	2070.647	2050.931	2037.928
19	2274.578	2243.541	2221.721	2205.544	2183.162	2168.412
20	2421.016	2385.900	2361.235	2342.962	2317.699	2301.062
21	2570.786	2531.255	2503.517	2482.981	2454.611	2435.941
22	2724.009	2679.710	2648.658	2625.685	2593.972	2573.116
23	2880.808	2831.371	2796.752	2771.159	2735.856	2712.656
24	3041.316	2986.351	2947.900	2919.496	2880.344	2854.634
25	3205.672	3144.767	3102.206	3070.789	3027.518	2999.124
50	9180.371	8737.558	8438.036	8222.095	7931.735	7745.597
60	13395.61	12524.94	11946.82	11535.54	10989.96	10644.75
70	20067.84	18317.65	17183.60	16390.89	15357.76	14715.19
80	32503.47	28619.66	26190.94	24535.82	22433.07	21157.14
90	65772.19	54168.42	47357.61	42920.97	37532.35	34401.93

Table 7. Operating Characteristics of the Life Test Sampling Plan

Based on Mean Life Criterion
(n =105, c = 2 and λ = 1.5)

p	$P_a(p)$	$\lambda = 1.5$	
		$k = t/\mu \times 100$	$\mu = 250/k \times 100$
0.010	0.911201	0.336136	74374.66
0.015	0.790632	0.506334	49374.48
0.020	0.649366	0.677979	36874.30
0.025	0.510198	0.851089	29374.12
0.030	0.386710	1.025686	24373.94
0.035	0.284587	1.201788	20802.33
0.040	0.204328	1.379418	18123.58
0.045	0.143657	1.558597	16040.07
0.050	0.099187	1.739346	14373.22
0.055	0.067404	1.921687	13009.40
0.060	0.045163	2.105644	11872.85
0.070	0.019543	2.478495	10086.77
0.080	0.008106	2.858088	8747.105
0.090	0.003242	3.244622	7705.059
0.100	0.001256	3.638299	6871.342

5. Conclusion

A procedure for deriving single sampling plans for life tests is described under the condition that the lifetime quality characteristic is modeled by a Lomax distribution. The tables for the determining the sampling plans when lot quality is evaluated using four criteria, namely, mean life, median life, hazard rate and reliability life are constructed fixing a set of values of the shape parameter. Practitioners can generate the necessary sampling plans for other values of the shape parameter as per their requirements using the procedure described in this paper. Conversion factors are also included to assist in the calculation of the mean life, median life, and hazard rate at a certain test termination time and vice versa. The factors for adapting acceptable quality level as the index to mean life, median life, hazard rate and reliability life are also provided.

6. Acknowledgment

The authors are grateful to Bharathiar University, Coimbatore for providing necessary facilities to carry out this research work. The second author is indebted to the Department of Science and Technology, India for awarding the DST-INSPIRE Fellowship under which the present research has been carried out.

References

[1] Epstein, B. (1960a). Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential, Part I, *Technometrics*, **2**, pp. 83 - 101.

[2] Epstein, B. (1960b). Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential, Part II, *Technometrics*, **2**, pp. 167 - 183.

[3] Handbook H-108. (1960). Sampling Procedures and Tables for Life and Reliability Testing, Quality Control and Reliability, Office of the Assistant Secretary of Defense, US Department of Defense, Washington, D.C.

[4] Goode, H. P., and Kao, J. H. K. (1961). Sampling Plans Based on the Weibull Distribution, *Proceedings of the Seventh National Symposium on Reliability and Quality Control,* Philadelphia, PA, pp. 24 - 40.

[5] Goode, H. P., and Kao, J. H. K. (1962). Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Hazard Rate Criterion), *Proceedings of the Eight National Symposium on Reliability and Quality Control*, Washington, DC, pp. 37 - 58.

[6] Goode, H. P., and Kao, J. H. K. (1964). Hazard Rate Sampling Plans for the Weibull Distribution, *Industrial Quality Control*, **20**, pp. 30 - 39.

[7] Gupta, S.S. and Groll, P. A. (1961). Gamma Distribution in Acceptance Sampling Based on Life Tests, *Journal of the American Statistical Association*, **56**, pp. 942 - 970.

[8] Gupta, S. S. (1962). Life Test Sampling Plans for Normal and Lognormal Distributions, *Technometrics*, **4**, pp. 151 - 175.

[9] Schilling, E. G., and Neubauer, D. V. (2009). *Acceptance Sampling in Quality Control*, Chapman and Hall, New York, NY.

[10] Wu, J. W., and Tsai, W. L. (2000). Failure Censored Sampling Plan for the Weibull Distribution, *Information and Management Sciences*. **11**, pp. 13 - 25.

[11] Wu, J. W., Tsai, T. R., and Ouyang, L. Y. (2001). Limited Failure-Censored Life Test for the Weibull Distribution, *IEEE Transactions on Reliability*, **50**, pp. 197 - 111.

[12] Kantam, R.R.L., Rosaiah, K., and Rao, G.S. (2001). Acceptance Sampling Based on Life Tests: Log-Logistic Models, *Journal of Applied Statistics,* **28**, pp. 121 - 128.

[13] Jun, C.-H., Balamurali, S., and Lee, S.-H. (2006). Variables Sampling Plans for Weibull Distributed Lifetimes under Sudden Death Testing, *IEEE Transactions on Reliability*, **55**, pp. 53 - 58.

[14] Tsai, T.-R., and Wu, S.-J. (2006). Acceptance Sampling Based on Truncated Life-tests for Generalized Rayleigh Distribution, *Journal of Applied Statistics*, **33**, pp. 595 - 600.

[15] Balakrishnan, N., Leiva, V., and Lopez, J. (2007). Acceptance Sampling Plans from Truncated Life-test Based on the Generalized Birnbaum - Saunders Distribution, *Communications in Statistics - Simulation and Computation*, **36**, pp. 643 - 656.

[16] Aslam, M., and Jun, C.-H. (2009a). A Group Acceptance Sampling Plan for Truncated Life Test having Weibull Distribution, *Journal of Applied Statistics*, **39**, pp. 1021 - 1027.

[17] Aslam, M., and Jun, C.-H. (2009b). Group Acceptance Sampling Plans for Truncated Life Tests Based on the Inverse Rayleigh Distribution and Log-logistic Distribution, *Pakistan Journal of Statistics,* **25**, pp. 107 - 119.

[18] Aslam, M., Kundu, D., Jun, C.-H., and Ahmad, M. (2011). Time Truncated Group Acceptance Sampling Plans for Generalized Exponential Distribution, *Journal of Testing and Evaluation,* **39**, pp. 968 - 976.

[19] Kalaiselvi, S., and Vijayaraghavan, R. (2010). Designing of Bayesian Single Sampling Plans for Weibull-Inverted Gamma Distribution, *Recent Trends in Statistical Research*, Publication Division, M. S. University, Tirunelveli, pp. 123 - 132.

[20] Kalaiselvi, S., Loganathan, A., and Vijayaraghavan, R. (2011). Reliability Sampling Plans under the Conditions of Rayleigh – Maxwell Distribution – A Bayesian Approach, *Recent Advances in Statistics and Computer Applications*, Bharathiar University, Coimbatore, pp. 280 - 283.

[21] Loganathan, A., Vijayaraghavan, R., and Kalaiselvi, S. (2012). Recent Developments in Designing Bayesian Reliability Sampling Plans – An Overview, *New Methodologies in Statistical Research*, Publication Division, M. S. University, Tirunelveli, pp. 61 - 68.

[22] Hong, C. W., Lee, W. C., and Wu, J. W. (2013). Computational Procedure of Performance Assessment of Life-time Index of Products for the Weibull Distribution with the Progressive First-failure Censored Sampling Plan, *Journal of Applied Mathematics,* Article ID 717184, **2012**, pp. 1 - 13.

[23] Vijayaraghavan, R., Chandrasekar, K., and Uma, S. (2012). Selection of sampling inspection plans for life test based on Weibull-Poisson Mixed Distribution, *Proceedings of the International Conference on Frontiers of Statistics and its Applications,* Coimbatore, pp. 225 - 232.

[24] Vijayaraghavan, R., and Uma, S. (2012). Evaluation of Sampling Inspection Plans for Life Test Based on Exponential-Poisson Mixed Distribution, *Proceedings of the International Conference on Frontiers of Statistics and its Applications*, Coimbatore, pp. 233 - 240.

[25] Vijayaraghavan, R., and Uma, S. (2016). Selection of Sampling Inspection Plans for Life Tests Based on Lognormal Distribution, *Journal of Testing and Evaluation*, **44**, pp. 1960 - 1969.

[26] Vijayaraghavan, R., Sathya Narayana Sharma, K., and Saranya, C. R. (2019a). Evaluation of Sampling Inspection Plans for Life-tests Based on Generalized Gamma Distribution*, International Journal of Scientific Research in Mathematical and Statistical Sciences*, **6**, pp. 138 - 146.

[27] Vijayaraghavan, R., Saranya., C. R., Sathya Narayana Sharma, K. (2019b). Life Test Sampling Plans Based on Marshall – Olkin Extended Exponential Distribution, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, **6**, pp.131-139.

[28] Vijayaraghavan, R., Saranya., C. R., and Sathya Narayana Sharma, K. (2021). Reliability Single Sampling Plans under the Assumption of Burr Type XII Distribution, *Reliability: Theory & Applications*, 16, pp. 308 - 322.

[29] Vijayaraghavan, R., and Pavithra, A. (2022a). Selection of Life Test Sampling Inspection Plans for Continuous Production, *Reliability: Theory & Applications*, **17**, pp. 371 - 381.

[30] Vijayaraghavan, R., and Pavithra, A. (2022b). Construction of Life Test Sampling Inspection Plans by Attributes Based on Marshall - Olkin Extended Exponential Distribution, *Reliability: Theory & Applications*, **17**, pp. 421 - 429.

[31] Lomax, K. S. (1954). Business Failures: Another Example of the Analysis of Failure Data, *Journal of the American Statistical Association*, **49**, 847 - 852.

[32] Abdul, M. I. B. (2012). Recurrence Relations for Moments of Lower Generalized Order Statistics from Exponentiated Lomax Distribution and Its Characterization, *International Journal of Mathematical Archive* **3**, pp. 2144 - 2150.

[33] Lemonte, A. J. and Cordeiro, G. M. (2013). An Extended Lomax Distribution, *Statistics*, **47**, 800–816.

[34] Ramos, M. W. A., Marinho, P. R. D., *de* Silva R. V., and Cordeiro, G. M. (2013). The Exponentiated Lomax Poisson Distribution with an Application to Lifetime Data, *Advances and Applications in Statistics*, **34**, 107 - 135.

[35] Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., Adejumo, A. O., and Odetunmibi, O. A. (2017). A New Generalization of the Lomax Distribution with Increasing, Decreasing, and Constant Failure Rate, *Modelling and Simulation in Engineering*, Article ID 6043169, **2017**, pp. 1- 6.

[36] United States Department of Defense. (1961). Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Mean Life Criterion), *Quality Control and Reliability Technical Report (TR 3)*, Office of the Assistant Secretary of Defense Installation and Logistics), U. S. Government Printing Office, Washington D. C.

[37} United States Department of Defense. (1962). Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Hazard Rate Criterion), *Quality Control and Reliability Technical Report (TR 4)*, Office of the Assistant Secretary of Defense Installation and Logistics), U. S. Government Printing Office, Washington D. C.

[38] United States Department of Defense. (1963). Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Reliability Life Criterion), *Quality Control and Reliability Technical Report (TR 6)*, Office of the Assistant Secretary of Defense Installation and Logistics), U. S. Government Printing Office, Washington D. C.
[39] United States Department of Defense. (1965). Factors and Procedures for Applying MIL - STD-105D Sampling Plans to Life and Reliability Testing, *Quality Control and Reliability Assurance Technical Report (TR 7),* Office of the Assistant Secretary of Defense (Installations and Logistics), Washington, DC.
[40] Guenther, W. C. (1969). Use of the Binomial, Hypergeometric and Poisson Tables to Obtain Sampling Plans, *Journal of Quality Technology*, **1**, pp. 105 - 109.
[41] Cameron, J. M. (1952). Tables for Constructing and for Computing the Operating Characteristics of Single Sampling Plans, *Industrial Quality Control*, **9**, pp. 37 - 39.
[42] United States Department of Defense (1989). *Military Standard, Sampling Procedures and Tables for Inspection by Attributes (MIL-STD-105E)*, U.S. Government Printing Office, Washington, DC.

A NEW RANKING IN HEPTAGONAL FUZZY NUMBER AND ITS APPLICATION IN PROJECT SCHEDULING

Adilakshmi Siripurapu[1], Ravi Shankar Nowpada[2]

•

Dept. of Basic Science and Humanities, Vignan's Institute of Information Technology(A), Duvvada, Visakhapatnam, AP, India[1]
slakshmijagarapu@gmail.com
Dept. of Mathematics, Institute of Science, GITAM (Deemed to be University), Visakhapatnam, AP, India[2]
drravi68@gmail.com

Abstract

Ranking fuzzy numbers is significant in optimization approaches such as assignment challenges, transportation problems, project schedules, artificial intelligence, data analysis, network flow analysis, an uncertain environment in organizational economics etc. This paper introduces a new fuzzy ranking in Heptagonal fuzzy numbers and arithmetic operations of Heptagonal fuzzy numbers defined. In the network, every activity duration is viewed by a Heptagonal fuzzy number. Every Heptagonal fuzzy number is transformed into a crisp number using the ranking function. By applying the traditional method, we calculate the fuzzy critical path. These procedures are illustrated with numerical examples and compared with existing ranking functions.

Keywords: Activity duration, centroid, fuzzy ranking, fuzzy critical path, heptagonal fuzzy number.

I. Introduction

One of the most significant concepts in network analysis is the critical path approach. It is utilized to resolve project problems by preparing the networks and determining the earliest date an activity may begin and be completed. It is also an algorithm for scheduling a collection of project networks. It is also frequently used in connection with the Program Evaluation and Review Technique (PERT).

Zadeh [10] introduced the 'fuzzy logic' concept considering inaccuracies and inconsistencies. Several academics have utilized various forms of fuzzy numbers to develop mathematical models over the last few decades. Examples for fuzzy numbers include Triangular fuzzy numbers, Trapezoidal fuzzy numbers, Pentagonal fuzzy numbers, etc.

In many practical situations, the variables that define information uncertainty or vagueness are usually Triangular or Trapezoidal fuzzy numbers. Chandrasekaran et al. [1] developed a new arithmetic operation in Heptagonal fuzzy numbers and solved the transportation problem in 2013. Rathi et al. [8] defined a new non-normal fuzzy number called the Heptagonal fuzzy number and arithmetic computations, suggested a parametric ranking strategy for ordering Heptagonal fuzzy numbers and employed the fuzzy assignment problem in 2014. The Heptagonal fuzzy number ranking is derived from the centroid of centroids and incentre of Heptagonal fuzzy numbers by Namarta et al. in 2017[7]. Developed a new ranking in Heptagonal fuzzy numbers and adapted it to transportation problems by Sahaya et al. [9]. Karthik et al. [4] in 2019 proposed linear and non-linear Heptagonal fuzzy numbers under uncertain environments and derived the Haar ranking technique for the Heptagonal fuzzy number. Malini [6] suggested a new ranking in Heptagonal fuzzy numbers and applied transportation problems. Hamildon et al. [3] determine the fuzzy critical path with normalized Heptagonal fuzzy data.

II. Preliminaries

In this section, we will look at a few key definitions.

I. Fuzzy Set [10]

As stated in Zadeh's paper, the formalization of a fuzzy set is:
Let X be a space of points (objects), with a generic element of X denoted by x. Thus, $X = \{x\}$. A fuzzy set (class) A in X is characterized by a membership (characteristic function) function $\mu_A(x)$, which associates with each point in X a real number in the interval [0,1], with the value of $\mu_A(x)$ at x representing the "grade of membership" of x in A. When A set in the ordinary sense of the term, its membership function can take on only two values, 0 and 1, $\mu_A(x) = 1 \; or \; 0$ according to x does or does not belong to A.

II. Fuzzy Number [5]

It is a Fuzzy set of the following conditions:

- Convex fuzzy set
- Normalized fuzzy set.
- Its membership function is piece-wise continuous.
- It is defined in the real number.

Fuzzy numbers should be normalized and convex. Here the condition of normalization implies that the maximum membership value is 1.

III. Heptagonal fuzzy number (HFN) [3]

A fuzzy number $\tilde{A} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7)$ is a normal Heptagonal fuzzy number, and its membership function is expressed as;

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{1}{2}\frac{(x-a_1)}{(a_2-a_1)}, & for \;\; a_1 \le x \le a_2 \\ 0.5, & for \;\; a_2 \le x \le a_3 \\ \frac{1}{2}+\frac{1}{2}\frac{(x-a_3)}{(a_4-a_3)}, & for \;\; a_3 \le x \le a_4 \\ \frac{1}{2}+\frac{1}{2}\frac{(a_5-x)}{(a_5-a_4)}, & for \; a_4 \le x \le a_5 \\ 0.5, & for \;\; a_5 \le x \le a_6 \\ \frac{1}{2}\frac{(a_7-x)}{(a_7-a_6)}, & for \;\; a_6 \le x \le a_7 \\ 0, & otherwise \end{cases}$$

The graphical depiction of normalized Heptagonal fuzzy number is represented in Figure 1.

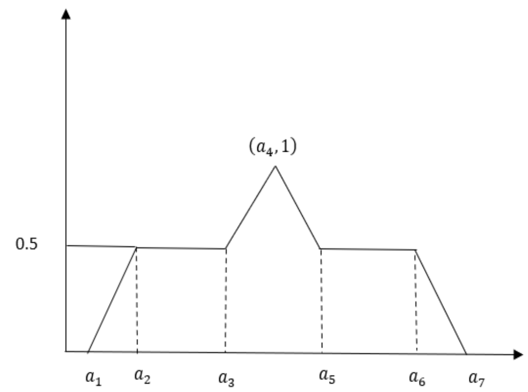


Figure 1: *Graphical representation of Heptagonal fuzzy number*

IV. Generalized Heptagonal fuzzy number (GHFN) [3]

A generalized Heptagonal fuzzy number is denoted by $\tilde{A}_{\tilde{H}} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, \omega)$ and its membership function is expressed as;

$$\mu_{\tilde{A}_{\tilde{H}}}(x) = \begin{cases} \frac{\omega}{2}\frac{(x-a_1)}{(a_2-a_1)}, & for \ a_1 \leq x \leq a_2 \\ \frac{\omega}{2}, & for \ a_2 \leq x \leq a_3 \\ \frac{\omega}{2}+\frac{\omega}{2}\frac{(x-a_3)}{(a_4-a_3)}, & for \ a_3 \leq x \leq a_4 \\ \frac{\omega}{2}+\frac{\omega}{2}\frac{(a_5-x)}{(a_5-a_4)}, & for \ a_4 \leq x \leq a_5 \\ \frac{\omega}{2}, & for \ a_5 \leq x \leq a_6 \\ \frac{\omega}{2}\frac{(a_7-x)}{(a_7-a_6)}, & for \ a_6 \leq x \leq a_7 \\ Otherwise, \ 0 & \end{cases}$$

The graphical depiction of generalized Heptagonal fuzzy number is represented in Figure 2.

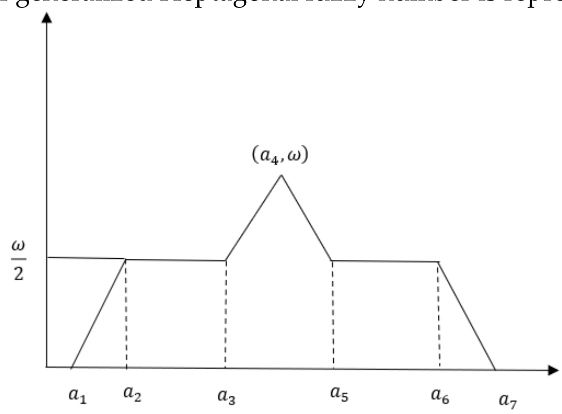


Figure 2: *Graphical representation of Generalized Heptagonal fuzzy number*

V. Arithmetic Operations of Heptagonal fuzzy number

According to Dubois [2], defined the arithmetic operation of Heptagonal fuzzy number.
Let $\tilde{A} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7)$ and $\tilde{B} = (b_1, b_2, b_3, b_4, b_5, b_6, b_7)$ be two Heptagonal fuzzy numbers then;

$k\tilde{A} = k(a_1, a_2, a_3, a_4, a_5, a_6, a_7) = (ka_1, ka_2, ka_3, ka_4, ka_5, ka_6, ka_7)$
$\tilde{A} \oplus \tilde{B} = (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4, a_5 + b_5, a_6 + b_6, a_7 + b_7)$
$\tilde{A} \ominus \tilde{B} = (a_1 - b_1, a_2 - b_2, a_3 - b_3, a_4 - b_4, a_5 - b_5, a_6 - b_6, a_7 - b_7)$
$\tilde{A} \otimes \tilde{B} = (a_1 * b_1, a_2 * b_2, a_3 * b_3, a_4 * b_4, a_5 * b_5, a_6 * b_6, a_7 * b_7)$
$\tilde{A} \oslash \tilde{B} = \left(\frac{a_1}{a_2}, \frac{a_2}{b_2}, \frac{a_3}{b_3}, \frac{a_4}{b_4}, \frac{a_5}{b_5}, \frac{a_6}{b_6}, \frac{a_7}{b_7}\right)$

Example:
Let $\tilde{A} = (3,6,9,12,15,18,21)$ and $\tilde{B} = (2,4,6,8,10,12,14)$ then
$\tilde{A} \oplus \tilde{B} = (5,10,15,20,25,30,35)$
$\tilde{A} \ominus \tilde{B} = (1,2,3,4,5,6,7)$
$\tilde{A} \otimes \tilde{B} = (6,24,54,96,150,216,294)$
$\tilde{A} \oslash \tilde{B} = (1.5, 1.5, 1.5, 1.5, 1.5, 1.5, 1.5)$
Remark: Some authors defined $\tilde{A} \ominus \tilde{B} = (a_1 - b_7, a_2 - b_6, a_3 - b_5, a_4 - b_4, a_5 - b_3, a_6 - b_2, a_7 - b_1)$.
How is it possible?
Here I consider one example.
Let $\tilde{A} = (2,4,6,8,10,12,14)$ & $\tilde{B} = (2,4,6,8,10,12,14)$, Here both $\tilde{A}$ *and* $\tilde{B}$ are same HFNs.
Now $\tilde{A} \ominus \tilde{B} = (2 - 14, 4 - 12, 6 - 10, 8 - 8, 10 - 6, 12 - 4, 14 - 2)$
$= (-12, -8, -4, 0, 4, 8, 12)$.
It is a completely wrong output since both $\tilde{A}$ *and* $\tilde{B}$ are the same HFNs.
According to my definition;
$\tilde{A} \ominus \tilde{B} = (2 - 2, 4 - 4, 6 - 6, 8 - 8, 10 - 10, 12 - 12, 14 - 14) = (0,0,0,0,0,0,0)$

III. Existing Rankings

In this section, we explained existing ranking functions.

I. Existing Ranking1[7]

Namarta et al. suggested a ranking process for HFN prediction on the centroid of centroids and incentre of centroids. Their suggested order is as follows:

$$G(x_0, y_0) = \left(\frac{2a_1 + 7a_2 + 7a_3 + 22a_4 + 7a_5 + 7a_6 + 2a_7}{54}, \frac{11\omega}{54}\right)$$

II. Existing Ranking 2 [3]

Hamildon et al. identify the critical path of a project network with normal HFNs. In his approach,

$$\Re(\tilde{A}_{\tilde{H}}) = \frac{a_1 + a_2 + a_3 + 2a_4 + a_5 + a_6 + a_7}{8}$$

IV Proposal Ranking Function

We suggest an effective tool for calculating the rank of HFN. The proposal ranking in a HFN diagram is represented in Figure 3.

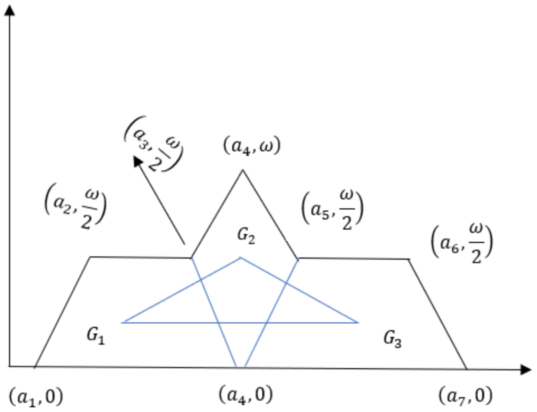


Figure 3: *Proposal ranking in GHFN*

In Figure 3, the heptagonal is divided into two trapezoidal and one rhombus. By applying the centroid formula of trapezoidal and rhombus, calculate the centroid of trapezoidal and rhombus, respectively. The circumcentre of the centroids of the HFN is taken into a balancing point of the Heptagon in Figure 7.3. The distance from the origin to the circumcentre of the centroids of this three-plane figure consider as a generalized HFN ranking function. Let the centroid of the three planar figures be G_1, G_2, G_3.

G_1 gives the centroid of the trapezoidal with vertices

$$(a_1,0),\left(a_2,\frac{\omega}{2}\right),\left(a_3,\frac{\omega}{2}\right)(a_4,0)$$

G_2 gives the centroid of the rhombus with vertices

$$(a_{4,}0),\left(a_3,\frac{\omega}{2}\right),(a_4,\omega),\left(a_5,\frac{\omega}{2}\right)$$

G_3 gives the centroid of the trapezoidal with vertices

$$(a_4,0),\left(a_5,\frac{\omega}{2}\right),\left(a_6,\frac{\omega}{2}\right),(a_7,0)$$

The centroid of these three planes is;

$$G_1=\left(\frac{a_1+a_2+a_3+a_4}{4},\frac{\omega}{4}\right),G_2=\left(\frac{a_3+a_4+a_5+a_4}{4},\frac{\omega}{2}\right),G_3=\left(\frac{a_4+a_5+a_6+a_7}{4},\frac{\omega}{4}\right)\text{ respectively.}$$

The circumcentre of G_1, G_2, and G_3 is

$$G_{\tilde{A}_{\tilde{H}}}(x_0,y_0)=\left(\frac{a_1+a_2+2a_3+4a_4+2a_5+a_6+a_7}{12},\frac{\omega}{3}\right).$$

The generalized Heptagonal fuzzy number $\tilde{A}_{\tilde{H}}=(a_1,a_2,a_3,a_4,a_5,a_6,a_{7,}\omega)$ new ranking function is;

$$\Re(\tilde{A}_{\tilde{H}})=\sqrt{x_0^2+y_0^2}$$

I. Ordering of a Heptagonal fuzzy number

Comparing fuzzy numbers using the ranking function $\Re: F(\mathcal{R})\longrightarrow\mathcal{R}$ is a successful approach**.** The ordering of two HFNs is described as follows;

- If $\Re(\tilde{A}_{\tilde{H}})>\Re(\tilde{B}_{\tilde{H}})\Rightarrow\tilde{A}_{\tilde{H}}>\tilde{B}_{\tilde{H}}$
- If $\Re(\tilde{A}_{\tilde{H}})<\Re(\tilde{B}_{\tilde{H}})\Rightarrow\tilde{A}_{\tilde{H}}<\tilde{B}_{\tilde{H}}$
- If $\Re(\tilde{A}_{\tilde{H}})=\Re(\tilde{B}_{\tilde{H}})\Rightarrow\tilde{A}_{\tilde{H}}=\tilde{B}_{\tilde{H}}$

Here, we utilize three sets of Heptagonal fuzzy numbers. Analyze the ranking of 3 sets by proposal ranking function and existing ranking functions. The sets and the outcome obtained by the proposal and existing rankings are given in Table 1. Here I consider ω=1.

Table 1: *Analysis of ranking order by proposal ranking function*

HFN	The rank of HFN	Conclusion

	Set-1	
$\tilde{\boldsymbol{A}}$ =(1,2,3,4,5,6,7)	4.0138	$\mathcal{R}(\tilde{C}) > \mathcal{R}(\tilde{B}) > \mathcal{R}(\tilde{D}) > \mathcal{R}(\tilde{A})$ $\Rightarrow \tilde{C} > \tilde{B} > \tilde{D} > \tilde{A}$
$\tilde{\boldsymbol{B}}$ =(3,4,5,7,9,10,11)	7.0079	
$\tilde{\boldsymbol{C}}$ =(2,3,4,8,13,14,15)	8.3399	
$\tilde{\boldsymbol{D}}$ =(1,2,3,8,9,11,13)	6.9246	
	Set-2	
$\tilde{\boldsymbol{A}}$ =(6,8,9,10,11,12,13)	9.9334	$\mathcal{R}(\tilde{C}) > \mathcal{R}(\tilde{A}) > \mathcal{R}(\tilde{D}) > \mathcal{R}(\tilde{B})$ $\Rightarrow \tilde{C} > \tilde{A} > \tilde{D} > \tilde{B}$
$\tilde{\boldsymbol{B}} = (\mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9})$	6.0277	
$\tilde{\boldsymbol{C}} = (\mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13}, \mathbf{15})$	11.0983	
$\tilde{\boldsymbol{D}} = (\mathbf{5}, \mathbf{7}, \mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12})$	8.9353	
	Set-3	
$\tilde{\boldsymbol{A}} = (\mathbf{5}, \mathbf{10}, \mathbf{15}, \mathbf{22}, \mathbf{23}, \mathbf{24}, \mathbf{25})$	19.0087	$\mathcal{R}(\tilde{A}) > \mathcal{R}(\tilde{B}) > \mathcal{R}(\tilde{C}) > \mathcal{R}(\tilde{D})$ $\Rightarrow \tilde{A} > \tilde{B} > \tilde{C} > \tilde{D}$
$\tilde{\boldsymbol{B}} = (\mathbf{4}, \mathbf{10}, \mathbf{12}, \mathbf{17}, \mathbf{18}, \mathbf{19}, \mathbf{21}$	15.1776	
$\tilde{\boldsymbol{C}} = (\mathbf{3}, \mathbf{10}, \mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{16}, \mathbf{17})$	12.5133	
$\tilde{\boldsymbol{D}} = (\mathbf{3}, \mathbf{6}, \mathbf{8}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13})$	9.3511	

The order of a HFN with the proposal and existing rankings are presented in Table 2.

Table 2: *Order of a HFN with the proposal and existing rankings*

Ranking function	Set-1	Set-2	Set-3
Namarta (2017)	$\tilde{C} > \tilde{B} > \tilde{D} > \tilde{A}$	$\tilde{C} > \tilde{A} > \tilde{D} > \tilde{B}$	$\tilde{A} > \tilde{B} > \tilde{C} > \tilde{D}$
Hamildon (2021)	$\tilde{C} > \tilde{B} > \tilde{D} > \tilde{A}$	$\tilde{C} > \tilde{A} > \tilde{D} > \tilde{B}$	$\tilde{A} > \tilde{B} > \tilde{C} > \tilde{D}$
Proposal ranking	$\tilde{C} > \tilde{B} > \tilde{D} > \tilde{A}$	$\tilde{C} > \tilde{A} > \tilde{D} > \tilde{B}$	$\tilde{A} > \tilde{B} > \tilde{C} > \tilde{D}$

V Application

This section performs an analytical example of the proposed fuzzy set CPM-based approach on an activity network. Think of a plant with 14 vertices and needs 21 primary activities, each activity connected by a direct link, like in the following graph (Figure 1.6). The fuzzy activity time is represented as a Heptagonal fuzzy number for every activity in Table 3 (All durations in days).

Table 3: *Project network with Heptagonal fuzzy number*

Activity	Heptagonal Fuzzy Number
1→2	(2,3,4,8,13,14,15)

1→3	(1,2,4,5,11,12,13)
1→4	(0,1,3,4,9,10,11)
1→5	(1,3,6,13,15,16,17)
1→6	(1,2,5,10,12,14,17)
2→7	(1,2,11,13,14,15,16)
3→7	(2,3,4,8,14,15,16)
3→10	(1,5,7,11,12,13,14)
4→8	(0,1,2,3,5,6,8)
4→9	(1,2,4,5,6,7,9)
5→9	(3,7,9,13,15,16,18)
5→13	(5,10,15,22,23,24,25)
6→14	(1,3,4,5,7,8,10)
7→11	(10,12,15,20,21,22,23)
7→12	(3,5,6,7,8,9,10)
8→12	(7,8,9,10,11,12,13)
9→13	(6,6,6,6,7,7,7)
10→14	(3,4,5,6,7,8,9)
11→14	(4,6,7,9,10,11,12)
12→14	(2,3,4,5,6,7,8)
13→14	(5,7,8,9,10,11,12)

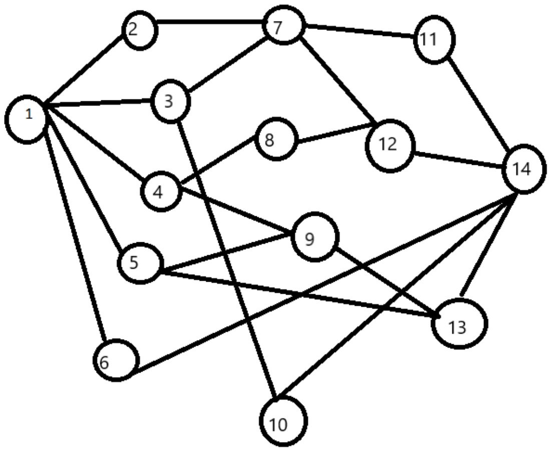


Figure 4: *Fuzzy project network*

I. Expected time of activities

Heptagonal fuzzy number transformed into an activity duration by proposal ranking function. This activity period is taken as the time within the nodes, and the fuzzy critical path is calculated by applying the conventional process. The activity period of the project network is represented in

Table 4, and Figure 5 represents the project network with defuzzified values of Heptagonal fuzzy numbers.

Table 4: *Activity duration with defuzzified value of HFN*

Activity	HFN	Activity period ($\tilde{t}_{ij}$)
1→2	(2,3,4,8,13,14,15)	8.3399
1→3	(1,2,4,5,11,12,13)	6.5085
1→4	(0,1,3,4,9,10,11)	5.1774
1→5	(1,3,6,13,15,16,17)	10.9217
1→6	(1,2,5,10,12,14,17)	9.0061
2→7	(1,2,11,13,14,15,16)	11.3382
3→7	(2,3,4,8,14,15,16)	8.6730
3→10	(1,5,7,11,12,13,14)	9.5891
4→8	(0,1,2,3,5,6,8)	3.4328
4→9	(1,2,4,5,6,7,9)	4.9279
5→9	(3,7,9,13,15,16,18)	12.0046
5→13	(5,10,15,22,23,24,25)	19.0029
6→14	(1,3,4,5,7,8,10)	5.3437
7→11	(10,12,15,20,21,22,23)	18.2530
7→12	(3,5,6,7,8,9,10)	6.9246
8→12	(7,8,9,10,11,12,13)	10.0055
9→13	(6,6,6,6,7,7,7)	6.3420
10→14	(3,4,5,6,7,8,9)	6.0092
11→14	(4,6,7,9,10,11,12)	8.5898
12→14	(2,3,4,5,6,7,8)	5.0110
13→14	(5,7,8,9,10,11,12)	8.9228

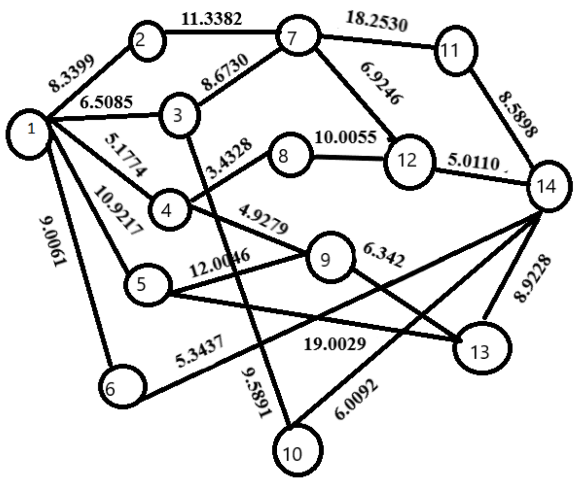


Figure 5: *Project network with defuzzified values of HFN*

II. Procedure for Fuzzy critical path method Based on Heptagonal fuzzy numbers

Step 1: Construct the network diagram of a given project.
Step 2: Represent every fuzzy activity time as a defuzzified value of the Heptagonal fuzzy number.
Step 3: Let $\tilde{E}_1 = (0,0,0,0,0,0)$ and calculate $\tilde{E}_i, i = 2,3, \dots \dots, n$ by using;

$$\tilde{E}_i = maximum\ end\ time\ of\ immediate\ predecessor + activity\ duration.$$

Step 4: Calculate the earliest finish time of an activity $i \to j$. That is earliest finish time is; $E\tilde{F}_{ij} = E\tilde{S}_{ij} + Fuzzy\ activity\ time$, where $E\tilde{S}_{ij} = \tilde{E}_i$.
Step 5: Let $\tilde{E}_n = \tilde{L}_n$ and calculate $\tilde{L}_i$, where $i = n-1, n-2, \dots \dots, 2,1$.

$$\tilde{L}_i = minimum\ end\ time\ of\ immediate\ successor -\ activity\ duration$$

Step 6: Calculate the latest start time of an activity $i \to j$. That is latest start time is;

$L\tilde{S}_{ij} = L\tilde{F}_{ij} - Fuzzy\ activity\ time$=, where $L\tilde{F}_{ij} = \tilde{L}_j$
Step 7: Calculate total float $T\tilde{F}_{ij} = L\tilde{F}_{ij} - E\tilde{F}_{ij}\ \ or\ \ L\tilde{S}_{ij} - E\tilde{S}_{ij}$.
Step 8: If $T\tilde{F}_{ij} = 0$, consider those activities as critical activities of a given project network.

III. Calculation of Earliest times

Node 1 is the starting node in the above network, and node 14 is the end node. Let $\tilde{E}_1 = 0$, and label node one as 0.
Iteration 1:
Node1 is the predecessor of node2.

$$\therefore \tilde{E}_2 = \tilde{E}_1 + \tilde{t}_{12} = 0 + 8.3399 = 8.3399$$

Iteration 2:
Node1 is the predecessor of node3.

$$\therefore \tilde{E}_3 = \tilde{E}_1 + \tilde{t}_{13} = 0 + 6.5085 = 6.5085$$

Iteration 3:
Node1 is the predecessor of node4.

$$\therefore \tilde{E}_4 = \tilde{E}_1 + \tilde{t}_{14} = 0 + 5.1774 = 5.1774$$

Iteration 4:
Node1 is the predecessor of node5.

$$\therefore \tilde{E}_5 = \tilde{E}_1 + \tilde{t}_{15} = 0 + 10.9217 = 10.9217$$

Iteration 5:
Node1 is the predecessor of node6.

$$\therefore \tilde{E}_6 = \tilde{E}_1 + \tilde{t}_{16} = 0 + 9.0061 = 9.0061$$

Iteration 6:
Node2 and node3 are predecessor of node 7.

$$\therefore \tilde{E}_7 = max\{\tilde{E}_2 + \tilde{t}_{27}, \tilde{E}_3 + \tilde{t}_{37}\} = \tilde{E}_2 + \tilde{t}_{27} = 8.3399 + 11.3382 = 19.6781$$

Iteration 7:
Node4 is the predecessor of node8.

$$\therefore \tilde{E}_8 = \tilde{E}_4 + \tilde{t}_{48} = 5.1774 + 3.4328 = 8.6102$$

Iteration 8:
Node4 and node5 are the predecessors of node9.

$$\therefore \tilde{E}_9 = max\{\tilde{E}_4 + \tilde{t}_{49}, \tilde{E}_5 + \tilde{t}_{59}\} = \tilde{E}_5 + \tilde{t}_{59} = 10.9217 + 12.0046 = 22.6343$$

Iteration 9:
Node3 is the predecessor of node10.

$$\therefore \tilde{E}_{10} = \tilde{E}_3 + \tilde{t}_{310} = 6.5085 + 9.5891 = 16.0976$$

Iteration 10:
Node7 is the predecessor of node11.

$$\therefore \tilde{E}_{11} = \tilde{E}_7 + \tilde{t}_{711} = 19.6781 + 18.2530 = 37.9311$$

Iteration 11:

Node7 and node8 are the predecessors of node 12.

$$\tilde{E}_{12} = max\{\tilde{E}_7 + \tilde{t}_{712}, \tilde{E}_8 + \tilde{t}_{812}\} = \tilde{E}_7 + \tilde{t}_{712} = 19.6781 + 6.9246 = 26.6027$$

Iteration 12:

Node5 and node 9 are predecessor of node 13.

$$\therefore \tilde{E}_{13} = max\{\tilde{E}_5 + \tilde{t}_{513}, \tilde{E}_9 + \tilde{t}_{913}\} = \tilde{E}_9 + \tilde{t}_{913} = 22.6343 + 6.3420 = 28.9763$$

Iteration 13:

Node6, node10, node11, node12, node13 are the predecessor of node14.

$$\therefore \tilde{E}_{14} = max\{\tilde{E}_6 + \tilde{t}_{614}, \tilde{E}_{10} + \tilde{t}_{1014}, \tilde{E}_{11} + \tilde{t}_{1114}, \tilde{E}_{12} + \tilde{t}_{1214}, \tilde{E}_{13} + \tilde{t}_{1314}\}$$
$$= \tilde{E}_{11} + \tilde{t}_{1114} = 37.9311 + 8.5898 = 46.5209$$

IV. Calculation of Latest times

In the above network, the end node is node14.

Let $\tilde{E}_{14} = \tilde{L}_{14}$ and label node14 as 46.5209.

Iteration1:

The successor of node 13 is node14.

$$\therefore \tilde{L}_{13} = \tilde{L}_{14} - \tilde{t}_{1314} = 46.5209 - 8.9228 = 37.5981$$

Iteration 2:

The successor of 12 is node14.

$$\therefore \tilde{L}_{12} = \tilde{L}_{14} - \tilde{t}_{1214} = 46.5209 - 5.0110 = 41.5099$$

Iteration 3:

The successor of 11 is node 14.

$$\therefore \tilde{L}_{11} = \tilde{L}_{14} - \tilde{t}_{1114} = 46.5209 - 8.5898 = 37.9311$$

Iteration 4:

The successor of 10 is node14.

$$\therefore \tilde{L}_{10} = \tilde{L}_{14} - \tilde{t}_{1014} = 46.5209 - 6.0092 = 40.5117$$

Iteration 5:

The successor of 9 is node13.

$$\therefore \tilde{L}_9 = \tilde{L}_{13} - \tilde{t}_{913} = 37.5981 - 6.3420 = 31.2561$$

Iteration 6:

The successor of node 8 is node 12.

$$\therefore \tilde{L}_8 = \tilde{L}_{12} - \tilde{t}_{812} = 41.5099 - 10.0085 = 31.5014$$

Iteration 7:

The successor of node 7 is node11 and node 12.

$$\therefore \tilde{L}_7 = min\{\tilde{L}_{11} - \tilde{t}_{711}, \tilde{L}_{12} - \tilde{t}_{712}\} = \tilde{L}_{11} - \tilde{t}_{711} = 37.9311 - 18.2530 = 19.6781$$

Iteration 8:

The successor of node 6 is node 14.

$$\therefore \tilde{L}_6 = \tilde{L}_{14} - \tilde{t}_{614} = 46.5209 - 5.3437 = 41.1772$$

Iteration 9:

The successor of node5 is node 9 and node 13.

$$\therefore \tilde{L}_5 = min\{\tilde{L}_9 - \tilde{t}_{59}, \tilde{L}_{13} - \tilde{t}_{513}\} = \tilde{L}_9 - \tilde{t}_{59} = 31.2561 - 12.0046 = 19.2515$$

Iteration 10:

The successor of node 4 is node8 and node9.

$$\therefore \tilde{L}_4 = min\{\tilde{L}_8 - \tilde{t}_{48}, \tilde{L}_9 - \tilde{t}_{49}\} = \tilde{L}_9 - \tilde{t}_{49} = 31.2561 - 4.9277 = 26.3284$$

Iteration 11:

The successor of node 3 is node7 and node 10.

$$\therefore \tilde{L}_3 = min\{\tilde{L}_7 - \tilde{t}_{37}, \tilde{L}_{10} - \tilde{t}_{310}\} = \tilde{L}_7 - \tilde{t}_{37} = 19.6781 - 8.6730 = 11.0051$$

Iteration 12:

The successor of node 2 is node7.

$$\therefore \tilde{L}_2 = \tilde{L}_7 - \tilde{t}_{27} = 19.6781 - 11.3382 = 8.3399$$

Iteration 13:

The successor of node1 is node2, node3, node4, node5 and node6.

$$\therefore \tilde{L}_1 = min\{\tilde{L}_2 - \tilde{t}_{12}, \tilde{L}_3 - \tilde{t}_{13}, \tilde{L}_4 - \tilde{t}_{14}, \tilde{L}_5 - \tilde{t}_{15}, \tilde{L}_6 - \tilde{t}_{16}\} = \tilde{L}_2 - \tilde{t}_{12} = 8.3399 - 8.3399 = 0$$

V. Calculation of Total float

Computed Earliest finish time, Latest start time and total float using formulas mentioned in procedure step 4, step 6, and step 7, respectively.
The earliest start and finish times, the latest start and finish times, total float of fuzzy activities are depicted in Table 5.

Table 5: ***The earliest, latest times and a total float of activities with defuzzified values***

Activity	$\tilde{t}_{ij}$	$E\tilde{S}_{ij}$	$E\tilde{F}_{ij}$	$L\tilde{S}_{ij}$	$L\tilde{F}_{ij}$	$T\tilde{F}_{ij}$
1→2	8.3399	0	8.3399	0	8.3399	0*
1→3	6.5085	0	6.5085	4.4966	11.0051	4.4966
1→4	5.1774	0	5.1774	21.151	26.3284	5.1774
1→5	10.9217	0	10.9217	8.3298	19.2515	7.6742
1→6	9.0061	0	9.0061	32.1711	41.1772	32.1711
2→7	11.3382	8.3399	19.6781	8.3399	19.6781	0*
3→7	8.6730	6.5085	15.1815	11.0051	19.6781	4.4966
3→10	9.5891	6.5085	16.0976	30.9226	40.5117	24.4141
4→8	3.4328	5.1774	8.6102	28.0686	31.5014	22.8912
4→9	4.9279	5.1774	10.1053	26.3282	31.2561	21.1508
5→9	12.0046	10.9217	22.9263	19.2515	31.2561	8.3298
5→13	19.0029	10.9217	29.9246	18.5952	37.5981	7.6735
6→14	5.3437	9.0061	14.3498	41.1772	46.5209	32.1711
7→11	18.2530	19.6781	37.9311	19.6781	37.9311	0*
7→12	6.9246	19.6781	26.6027	34.5853	41.5099	14.9072
8→12	10.0085	8.6102	18.6187	31.5014	41.5099	22.8912
9→13	6.342	22.6343	28.9763	31.2561	37.5981	8.6218
10→14	6.0092	16.0976	22.1068	40.5307	46.5399	24.4331
11→14	8.5898	37.9311	46.5209	37.9311	46.5209	0*
12→14	5.0110	26.6027	31.6137	41.5099	46.5209	14.9072
13→14	8.9228	28.9763	37.8991	37.5981	46.5209	8.6218

From the above table, the critical activities are 1→2, 2→7, 7→11, 11→14.
Therefore, the project critical path is 1→2→7→11→14.
As a result, the project will be completed in 46.5209≅ 46.5 days.

VI Results

Table 6 represents the fuzzy critical path and project completion time by proposal method and existing methods. The result graph is presented in Figure 6.

Table 6: ***Proposal method correlated with existing methods***

Ranking Method	Critical Path	Project completion time
Namarta (2017)	1→2→7→11	45.375
Hamildon (2021)	1→2→7→11	46.7853
Proposal method	1→2→7→11	46.5399

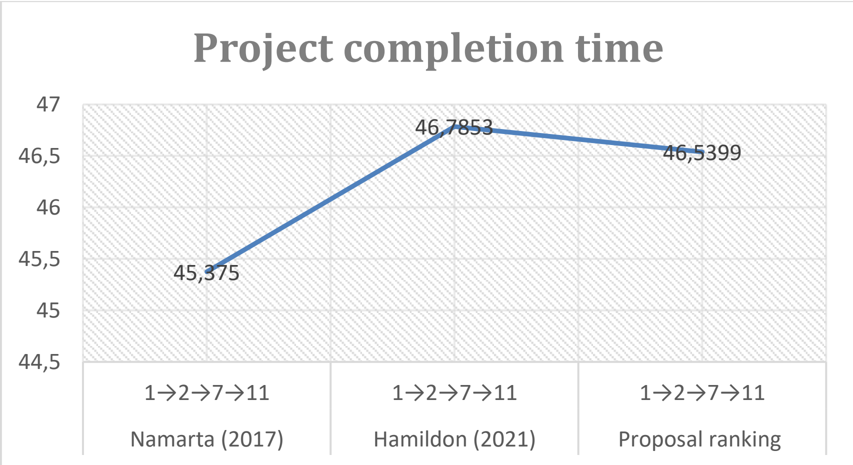


Figure 6: ***Proposal ranking results correlated with existing ranking results***

VII Conclusion

This paper introduced a new ranking function in Heptagonal fuzzy number. The proposed ranking function is derived from the centroid of HFN. In the network, every activity period is expressed by an HFN. The duration of every activity is transformed into the normal number or crisp number by a new ranking function. This normal number is considered as the expected time of activity. A conventional procedure identified the fuzzy critical path and project completion time. Numerous experiments have been conducted, and the results are correlated with some of the available ranking formulas. The attained results are similar to existing ranking results and the same critical path in all the methods. The proposal ranking can also be applied to more complex project networks in the real world. We can apply the ranking function of HFN to solve game problems and transportation problems.

Acknowledgement

There is no conflict of interest from co-author.

References

[1] Chandrasekaran, G. Kokila, Junu Saju, 2015. Ranking of Heptagon Number using Zero Suffix Method. *International Journal of Science and Research*, Volume 4 Issue 5, 2256 – 2257.

[2] Dubois. D, Henri Prade (1987), The mean value of a fuzzy number, *Fuzzy Sets and Systems*, Volume 24, Issue 3, Pages 279-300, ISSN 0165-0114.

[3] Hamildon Abd El-Wahed Khalifa, Majed G. Alharbi, and Pavan Kumar, (2021). On Determining the Critical Path of Activity Network with Normalized Heptagonal Fuzzy Data. *Wireless Communication and Mobile Computing*, Vol 2021.

[4] Karthik, S.N.P. & Dash. S. (2019). Haar Ranking of Linear and Non-linear Heptagonal fuzzy number and its application. *International Journal of Innovative Technology and Exploring Engineering*, 8(6): 1212-1220.

[5] Kaufmann, A. and Madan M. Gupta (1986). Introduction to fuzzy arithmetic: theory and applications. *Van Nostrand Reinhold*, New York.

[6] Malini. P, 2019. A new ranking technique on heptagonal fuzzy numbers to solve the fuzzy transportation problem. *International Journal of Mathematics in Operational Research*. Vol. 15(3), pages 364-371.

[7] Namarta, Thakur, D.N., & Gupta, D.U. (2017). Ranking of Heptagonal Fuzzy Numbers Using Incentre of Centroids. *International Journal of Advanced Technology in Engineering and Science*. 5(7), 248-255.

[8] Rathi. K & S. Balamohan (2014). Representation and ranking of fuzzy numbers with heptagonal membership function using value and ambiguity Index. *Applied Mathematical Sciences*. 8. 4309-4321.

[9] Sahaya Sudha, D.A., & Karunambigai, S. (2017). Solving a Transportation Problem using a Heptagonal Fuzzy Number. *International Journal of Advanced Research in Science, Engineering and Technology*. 4(1), 3118-3125.

[10] Zadeh L.A. (1965), Fuzzy sets, *Information and Control*, Volume 8, Issue 3, Pages 338-353, ISSN 0019-9958.

.

Weibull Inverse Power Rayleigh Distribution with Applications Related to Distinct Fields of Science

Muzamil Jallal[1], Aijaz Ahmad[2], Rajnee Tripathi[3]
1,2,3 Department of Mathematics, Bhagwant University, Ajmer
Email: muzamiljallal@gmail.com Email: ahmadaijaz4488@gmail.com
Email: rajneetripathi8@gmail.com

Abstract

In this paper an extension of Weibull Power Rayleigh Distribution has been introduced, and named it is as Weibull Inverse Power Rayleigh Distribution. This distribution is obtained by adopting T-X family technique. Various Structural properties, Reliability measures and Characteristics have been calculated and discussed. The behaviour of Probability density function, Cumulative distribution function, Survival function, Hazard rate function and mean residual function are illustrated through different graphs. Various parameters are estimated through the technique of MLE. The versatility and flexibility of the new distribution is done by using real life data sets. To evaluate and compare the out effectiveness of estimators, a simulation analysis has also been carried out.

Keywords:- *Weibull distribution, Inverse Rayleigh distribution, Renyi entropy, maximum likelihood estimation, Order statistics.*

I. INTRODUCTION

Weibull distribution, although being first identified by Frechet [8] and first applied by Rosin and Rammler [9] to describe particle size distribution, was named after Swedish Mathematician Waloddi Weibull, who in 1951 described the distribution in detail in his paper "A Statistical Distribution Function of Wide Applicability". This distribution is now commonly used to assess product reliability, analyze life data and model failure times, see inverse Weibull-G by Amal S. Hassan et al [3]. This distribution has found its importance in the fields of biology, economics, engineering sciences and hydrology.

The pdf of Weibull distribution is given by

$$f(t;\alpha,\beta)=\alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}} \quad \alpha,\beta>0, t\geq 0 \tag{1}$$

Where α,β and t are Location, Shape and Scale parameters respectively.

Rayleigh being a one-parameter continuous and simplest velocity probability distribution has a diverse range of applications. Various researchers have developed several extensions and modifications of the distribution, resulting in some flexible and more effective distributions by adding some more parameters or by compounding and thus showing its importance in various fields like

Social, Medical and Engineering Sciences. For instance transmuted Rayleigh distribution by Faton Merovci [6], Weibull-Rayleigh distribution by Faton Merovci et al [7], odd generalized exponential Rayleigh distribution by Albert Luguterah [1], Topp-Leone Rayleigh distribution by Fatoki Olayode [5], odd Lindley-Rayleigh distribution by Terna Godfrey Ieren [10], new generalisation of Rayleigh distribution by A.A Bhat et al [4].

The Inverse Rayleigh distribution finds its application in the field of reliability studies. Voda [11] worked out that the lifetime distributions of various types of experimental units can be approximated by Inverse Rayleigh distribution. In this article we use Weibull and Inverse Power Rayleigh Distributions to define a new model which generalizes the Inverse power Rayleigh distribution.

The cdf of Power inverse Rayleigh Distribution is given by

$$G(x,\theta,\lambda)=e^{-\frac{\theta}{2x^{2\lambda}}}\quad x>0,\theta,\lambda>0 \tag{2}$$

$$\overline{G}(x,\theta,\lambda)=1-e^{-\frac{\theta}{2x^{2\lambda}}} \tag{3}$$

And its associated pdf is given by

$$g(x,\theta,\lambda)=\frac{\theta\lambda}{x^{2\lambda+1}}e^{-\frac{\theta}{2x^{2\lambda}}}\qquad x>0,\theta,\lambda>0 \tag{4}$$

II. Materials and Methods

Transformed-transformer (T-X) family of distributions (Alzaatreh et al [2]) is given by

$$\mathrm{F(x)} = \int_0^{W[G(x)]} f(t)dt \tag{5}$$

Where f(t) is the probability density function of a random variable T and $W[G(x)]$ be a function of cumulative density function of random variable X .

Suppose $[G,\varsigma]$denotes the baseline cumulative distribution function, which depends on parameter vector ς. Now using T-X approach, the cumulative distribution function of Weibull Inverse Power Rayleigh distribution can be derived by replacing f(t) in equation(5) by equation (1) and $W[G(x)]=\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}$,where $\overline{G}(x,\varsigma)=1-G(x,\varsigma)$ which follows

$$\mathrm{F}(x,\alpha,\beta,\varsigma) = \int_0^{\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}} \alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}\,dt$$

$$F(x,\alpha,\beta,\varsigma)=1-e^{-\alpha\left[\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}\right]^{\beta}} \tag{6}$$

The corresponding pdf of (6) becomes

$$f(x,\alpha,\beta,\varsigma)=\alpha\beta g(x,\varsigma)\frac{[G(x,\varsigma)]^{\beta-1}}{[\overline{G}(x,\varsigma)]^{\beta-1}}e^{-\alpha\left[\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}\right]^{\beta}} \tag{7}$$

Where $G(x,\theta,\lambda)=e^{-\frac{\theta}{2x^{2\lambda}}}$ is known as Power Inverse Rayleigh Distribution. (8)

$$\overline{G}(x,\theta,\lambda)=1-G(x,\theta,\lambda)=1-e^{-\frac{\theta}{2x^{2\lambda}}} \tag{9}$$

And $g(x,\theta,\lambda)=\frac{\theta\lambda}{x^{2\lambda+1}}e^{-\frac{\theta}{2x^{2\lambda}}}\ ,\theta>0,\lambda>0$ (10)

III. Linear Transformation

Apply Taylor series expansion to the pdf in equation (7) we have

$$e^{-\alpha\left[\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}\right]^{\beta}}=\sum_{i=0}^{\infty}\frac{(-1)^i}{i!}\alpha^i\left[\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}\right]^{\beta i} \tag{11}$$

On substituting equation (11) in (7), we get the following expression

$$f(x,\varsigma)=\alpha\beta g(x,\varsigma)\frac{[G(x,\varsigma)]^{\beta-1}}{[\overline{G}(x,\varsigma)]^{\beta+1}}\sum_{i=0}^{\infty}\frac{(-1)^i}{i!}\alpha^i\left[\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}\right]^{\beta i}$$

$$f(x,\varsigma)=\sum_{i=0}^{\infty}\frac{(-1)^i}{i!}\alpha^{(i+1)}\beta\, g(x,\varsigma)[G(x,\varsigma)]^{\beta(i+1)-1}[\overline{G}(x,\varsigma)]^{-[\beta(i+1)+1]} \tag{12}$$

Using generalized binomial expansion, we have

$$(1-x)^{-a}=\sum_{j=0}^{\infty}\binom{a+j-1}{j}x^j$$

$$f(x,\varsigma)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta\, g(x,\varsigma)[G(x,\varsigma)]^{j+\beta(i+1)-1} \tag{13}$$

Using (8) and (9) in (13) we get

$$f(x,\varsigma)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta\frac{\theta\lambda}{x^{2\lambda+1}}e^{\frac{-\theta}{2x^{2\lambda}}}\left[e^{\frac{-\theta}{2x^{2\lambda}}}\right]^{j+\beta(i+1)-1} \tag{14}$$

$$f(x,\varsigma)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\delta_{ij}\frac{\theta\lambda}{x^{2\lambda+1}}e^{\frac{-\theta}{2x^{2\lambda}}}\left[e^{\frac{-\theta}{2x^{2\lambda}}}\right]^{j+\beta(i+1)-1} \tag{15}$$

Where $\delta_{ij} = \frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta$

IV. Weibull Inverse Power Rayleigh Distribution

Let X be a random variable_following Weibull Inverse Power Rayleigh Distribution with its cdf given by

$$F(x;\alpha,\beta,\theta,\lambda) = 1 - e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}} \quad , x>0,\alpha,\beta,\theta,\lambda>0 \tag{16}$$

And its pdf is given by

$$f(x;\alpha,\beta,\theta,\lambda) = \frac{\alpha\beta\theta\lambda}{x^{2\lambda+1}} e^{\frac{\theta}{2x^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1} e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}} \quad , x>0,\alpha,\beta,\theta,\lambda>0 \tag{17}$$

V. Reliability measures

The Survival function, Hazard function, Cumulative Hazard function, Reverse Hazard function is given by:

$$s_x(x) = e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}$$

$$h_x(x) = \frac{\alpha\beta\theta\lambda}{x^{2\lambda+1}} e^{\frac{\theta}{2x^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1}$$

$$H_x(x) = \alpha\left[e^{\frac{\theta}{2x^{2\lambda}}}-1\right]^{-\beta}$$

$$\tau_x = \frac{\frac{\alpha\beta\theta\lambda}{x^{2\lambda+1}} e^{\frac{\theta}{2x^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1} e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}}{1-e^{-\alpha\left(e\frac{\theta}{2x^{2\lambda}}-1\right)^{-\beta}}}$$

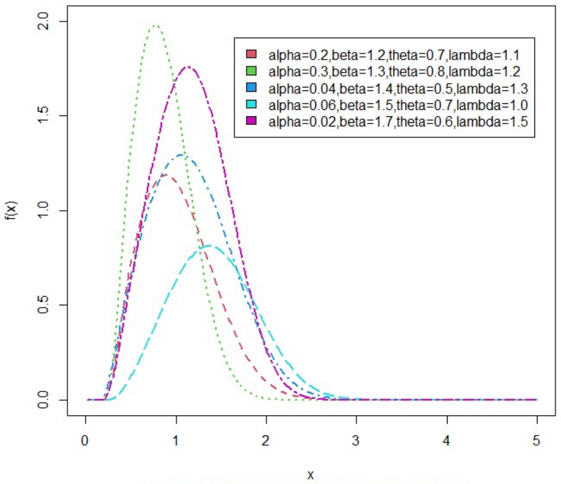


Fig a pdf of WPIRD under different values to parameters

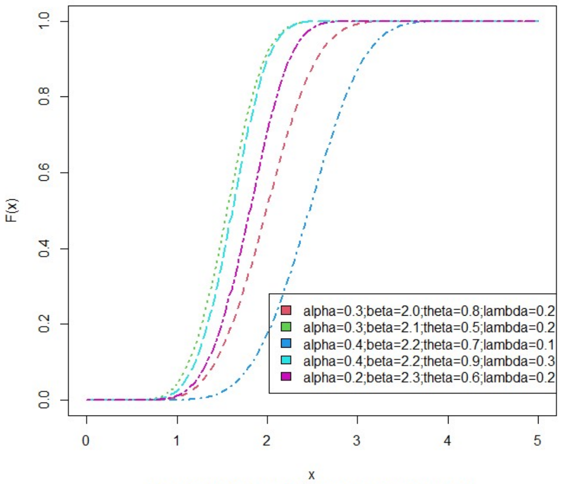


Fig b cdf of WIPRD under different values to parameters

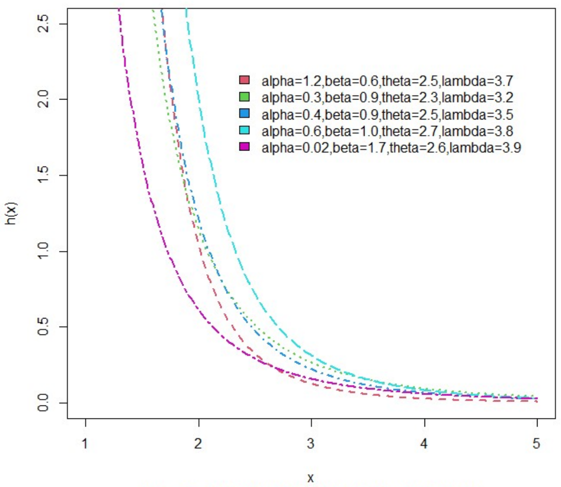


Fig c HRF of WPIRD under different values to parameters

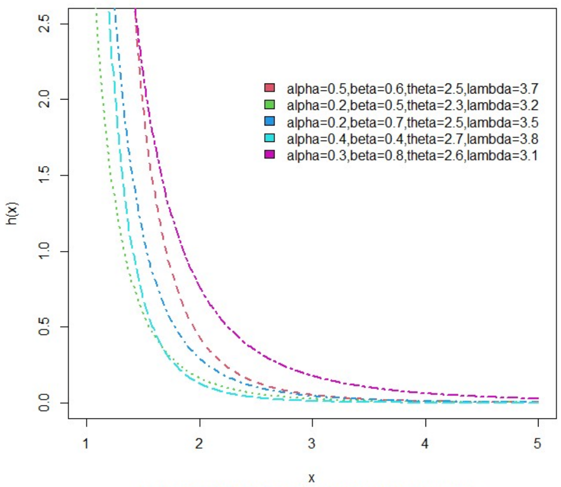


Fig d HRF of WPIRD under different values to parameters

Theorem 1: Show that the Quantile Function of Weibull Inverse Power Rayleigh Distribution is given by

$$x(u)=\left[\frac{\theta}{2\log\left[\left(\frac{-1}{\alpha}\log(1-u)\right)^{-\frac{1}{\beta}}+1\right]}\right]^{\frac{1}{2\lambda}}$$

Proof: Let x be a random variable following Weibull Inverse Power Rayleigh Distribution with parameters α,β,θ and λ, then we derive its Quantile Function from the corresponding cdf as given below

We know that cdf of WIPRD is given by

$$1-e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}$$

$$\text{Put } 1-e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}=u$$

Apply log on both sides

$$-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)=\log(1-u)$$

After solving we get

$$x(u)=\left[\frac{\theta}{2\log\left[\left(\frac{-1}{\alpha}\log(1-u)\right)^{-\frac{1}{\beta}}+1\right]}\right]^{\frac{1}{2\lambda}} \tag{18}$$

Median:

The median for the new WIPRD can be derived from the quantile function in equation (18) by putting u=0.5 as below

$$M_e=Q_x(0.5)=\left[\frac{\theta}{2\log\left[\left(\frac{-1}{\alpha}\log(\frac{1}{2})\right)^{-\frac{1}{\beta}}+1\right]}\right]^{\frac{1}{2\lambda}}$$

Theorem 2: If $X \sim WIPRD(\alpha,\beta,\theta,\lambda)$ then its rth moment is given by.

$$E(X)^r=\mu_r{}'=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta\left(\frac{\theta}{2}\right)^{\frac{r}{2\lambda}}\frac{\Gamma\left(1-\frac{r}{2\lambda}\right)}{[j+\beta(i+1)]^{1-\frac{r}{2\lambda}}}$$

Proof: We know that rth moment about origin is given by

$$E(X)^r = \mu_r' = \int_0^\infty x^r f(x;\alpha,\beta,\theta,\lambda)dx$$

Using (14) we get the following equation

$$\mu_r' = \int_0^\infty x^r \sum_{i=0}^\infty \sum_{j=0}^\infty \frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta \frac{\theta\lambda}{x^{2\lambda+1}} e^{\frac{-\theta}{2x^{2\lambda}}}\left[e^{\frac{-\theta}{2x^{2\lambda}}}\right]^{j+\beta(i+1)-1} dx$$

$$\mu_r' = \sum_{i=0}^\infty \sum_{j=0}^\infty \delta_{ij} \int_0^\infty x^r \frac{\theta\lambda}{x^{2\lambda+1}} e^{\frac{-[j+\beta(i+1)]\theta}{2x^{2\lambda}}} dx$$

Where $\delta_{ij} = \frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta$

$$\mu_r' = \sum_{i=0}^\infty \sum_{j=0}^\infty \delta_{ij}\, \theta\lambda \int_0^\infty x^{r-2\lambda-1} e^{\frac{-[j+\beta(i+1)]\theta}{2x^{2\lambda}}} dx$$

Put $\frac{-[j+\beta(i+1)]\theta}{2x^{2\lambda}} = z,$ we get

$$E(X) = \mu_r' = \sum_{i=0}^\infty \sum_{j=0}^\infty \delta_{ij}\left(\frac{\theta}{2}\right)^{\frac{r}{2\lambda}} \frac{\Gamma\left(1-\frac{r}{2\lambda}\right)}{[j+\beta(i+1)]^{1-\frac{r}{2\lambda}}}$$

On substituting r=1,2,3,4 we get first four moments about origin.

Theorem 3. Show that the Moment generating function of Weibull Inverse Power Rayleigh Distribution is given by

$$M_x(t) = \sum_{r=0}^\infty \frac{t^r}{r!} \sum_{i=0}^\infty \sum_{j=0}^\infty \frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta\left(\frac{\theta}{2}\right)^{\frac{r}{2\lambda}} \frac{\Gamma\left(1-\frac{r}{2\lambda}\right)}{[j+\beta(i+1)]^{1-\frac{r}{2\lambda}}}$$

Proof: We know that Moment generating function is given by

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x;\alpha,\beta,\theta,\lambda)dx$$

$$M_x(t) = E(e^{tx}) = \int_0^\infty \left\{1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \dots\right\} f(x;\alpha,\beta,\theta,\lambda)dx$$

$$M_x(t) = E(e^{tx}) = \int_0^\infty \sum_{r=0}^\infty \frac{(tx)^r}{r!} f(x;\alpha,\beta,\theta,\lambda)dx$$

$$M_x(t) = E(e^{tx}) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \int_0^{\infty} x^r f(x;\alpha,\beta,\theta,\lambda)dx$$

After solving the above equation we get,

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \delta_{ij} \left(\frac{\theta}{2}\right)^{\frac{r}{2\lambda}} \frac{\Gamma\left(1-\frac{r}{2\lambda}\right)}{[j+\beta(i+1)]^{1-\frac{r}{2\lambda}}}$$

Theorem 4. Show that the Characteristic function of Weibull Inverse Power Rayleigh Distribution is given by

$$\phi_x(it) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^i}{i!} \binom{\beta(i+1)+j}{j} \alpha^{(i+1)} \beta \left(\frac{\theta}{2}\right)^{\frac{r}{2\lambda}} \frac{\Gamma\left(1-\frac{r}{2\lambda}\right)}{[j+\beta(i+1)]^{1-\frac{r}{2\lambda}}}$$

Proof: We know that Characteristic function is given by

$$\phi_x(it) = E(e^{itx}) = \int_0^{\infty} e^{itx} f(x;\alpha,\beta,\theta,\lambda)dx$$

$$\phi_x(it) = E(e^{itx}) = \int_0^{\infty} \sum_{r=0}^{\infty} \frac{(itx)^r}{r!} f(x;\alpha,\beta,\theta,\lambda)dx$$

$$\phi_x(it) = E(e^{itx}) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \int_0^{\infty} x^r f(x;\alpha,\beta,\theta,\lambda)dx$$

After solving the above equation we get

$$\phi_x(it) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \delta_{ij} \left(\frac{\theta}{2}\right)^{\frac{r}{2\lambda}} \frac{\Gamma\left(1-\frac{r}{2\lambda}\right)}{[j+\beta(i+1)]^{1-\frac{r}{2\lambda}}}$$

VI. Incomplete Moments

We know that

$$T_q(S) = \int_0^{s} x^{s} f(x;\alpha,\beta,\theta,\lambda)dx$$

Put (14) in the above equation, we get

$$T_q(S) = \int_0^{s} x^s \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^i}{i!} \binom{\beta(i+1)+j}{j} \alpha^{(i+1)} \beta \frac{\theta\lambda}{x^{2\lambda+1}} e^{\frac{-\theta}{2x^{2\lambda}}} \left[e^{\frac{-\theta}{2x^{2\lambda}}} \right]^{j+\beta(i+1)-1} dx$$

$$T_q(S)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\delta_{ij}\theta\lambda\int_0^s \frac{x^s}{x^{2\lambda+1}}\,e^{\frac{-(j+\beta(i+1))\theta}{2x^{2\lambda}}}dx$$

$$T_q(S)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}2\delta_{ij}\left[\frac{(j+\beta(i+1))\theta}{2}\right]^{\frac{s-2\lambda}{2\lambda}}\Gamma\left(1-\frac{s}{2\lambda},M\right)$$

Where $M=\dfrac{(j+\beta(i+1))\theta}{2x^{2\lambda}}$

VII. Renyi Entropy

If X is a continuous random variable following WIPRD with pdf $f(x;\alpha,\beta,\theta,\lambda)$, then

$$T_R(\rho)=\frac{1}{1-\rho}\log\left\{\int_0^{\infty}f^{\rho}(x,\varsigma)dx\right\}$$

$$T_R(\rho)=\frac{1}{1-\rho}\log\left\{\int_0^{\infty}\left(\alpha\beta g(x,\varsigma)\frac{[G(x,\varsigma)]^{\beta-1}}{[\bar{G}(x,\varsigma)]^{\beta+1}}e^{-\alpha\left[\frac{G(x,\varsigma)}{\bar{G}(x,\varsigma)}\right]^{\beta}}\right)^{\rho}dx\right\}$$

$$T_R(\rho)=\frac{1}{1-\rho}\log\left\{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\upsilon_{ij}\int_0^{\infty}[g(x,\varsigma)]^{\rho}[G(x,\varsigma)]^{j+(\rho+i)\beta-\rho}dx\right\} \tag{19}$$

Where $\upsilon_{ij}=\dfrac{(-1)^i}{i!}(\rho\alpha)^i(\alpha\beta)^{\rho}\begin{pmatrix}\beta(\rho+1)+\rho+j-1\\ j\end{pmatrix}$

Now put (10) and (8) in equation (19) we get

$$T_R(\rho)=\frac{1}{1-\rho}\log\left\{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\upsilon_{ij}(\theta\lambda)^{\rho}\int_0^{\infty}x^{-\rho(2\lambda+1)}e^{\frac{-[j+(\rho+i)\beta]\theta}{2x^{2\lambda}}}dx\right\}$$

$$Put\ \frac{[j+(\rho+i)\beta]\theta}{2x^{2\lambda}}=z \quad\Rightarrow\quad dx=\frac{\left\{\frac{[j+(\rho+i)\beta]\theta}{2z}\right\}^{\frac{2\lambda+1}{2\lambda}}}{[j+(\rho+i)\beta]\theta\lambda}\quad dz$$

$$T_R(\rho)=\frac{1}{1-\rho}\log\left\{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\upsilon_{ij}\frac{(\theta\lambda)^{\rho-1}}{[j+(\rho+i)\beta]}\left[\frac{j+(\rho+i)\beta\theta}{2}\right]^{\frac{-\rho(2\lambda+1)+2\lambda+1}{2\lambda}}\Gamma\left(\frac{\rho(2\lambda+1)-1}{2\lambda}\right)\right\}$$

VIII. Tsallis Entropy of WIPRD

We know that Tsallis Entropy is defined as

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

$$S_\rho = \frac{1}{\rho-1}\left\{1-\int_0^\infty\left(\alpha\beta g(x,\varsigma)\frac{[G(x,\varsigma)]^{\beta-1}}{[\overline{G}(x,\varsigma)]^{\beta+1}}e^{-\alpha\left[\frac{G(x,\varsigma)}{\overline{G}(x,\varsigma)}\right]^\beta}\right)^\rho dx\right\}$$

$$S_\rho = \frac{1}{\rho-1}\left\{1-\left(\sum_{i=0}^\infty\sum_{j=0}^\infty \upsilon_{ij}\int_0^\infty [g(x,\varsigma)]^\rho[G(x,\varsigma)]^{j+(\rho+i)\beta-\rho}dx\right)\right\} \tag{20}$$

Now put (10) and (8) in equation (20) we get

$$S_\rho = \frac{1}{\rho-1}\left\{1-\left(\sum_{i=0}^\infty\sum_{j=0}^\infty \upsilon_{ij}\int_0^\infty\left[\frac{\theta\lambda}{x^{2\lambda+1}}e^{\frac{-\theta}{2x^{2\lambda}}}\right]^\rho\left(e^{\frac{-\theta}{2x^{2\lambda}}}\right)^{j+(\rho+i)\beta-\rho}dx\right)\right\}$$

After solving the above equation we get

$$S_\rho = \frac{1}{\rho-1}\left\{1-\left[\sum_{i=0}^\infty\sum_{j=0}^\infty \upsilon_{ij}\frac{(\theta\lambda)^{\rho-1}}{[j+(\rho+i)\beta]}\left[\frac{j+(\rho+i)\beta\theta}{2}\right]^{\frac{-\rho(2\lambda+1)+2\lambda+1}{2\lambda}}\Gamma\left(\frac{\rho(2\lambda+1)-1}{2\lambda}\right)\right]\right\}$$

IX. Mean Deviation from Mean

We know that

$$D(\mu) = E(|x-\mu|)$$

$$D(\mu) = \int_0^\infty |x-\mu| f(x)dx$$

$$D(\mu) = 2\mu F(\mu) - 2\int_0^\mu xf(x)dx \tag{21}$$

$$\text{Now we have}\int_0^\mu xf(x)dx = \int_0^\mu x\sum_{i=0}^\infty\sum_{j=0}^\infty\frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta\frac{\theta\lambda}{x^{2\lambda+1}}e^{\frac{-\theta}{2x^{2\lambda}}}\left[e^{\frac{-\theta}{2x^{2\lambda}}}\right]^{j+\beta(i+1)-1}dx$$

$$\int_0^\mu xf(x)dx = \sum_{i=0}^\infty\sum_{j=0}^\infty \delta_{ij}\left(\frac{\theta}{2}\right)^{\frac{1}{2\lambda}}\Gamma\left\{\left((1-\frac{1}{2\lambda}\right),\frac{(j+\beta(i+1))\theta}{2\mu^{2\lambda}}\right\} \tag{22}$$

Substituting (22) in (21) we get

$$D(\mu)=2\mu\left[1-e^{-\alpha\left(e^{\frac{\theta}{2\mu^{2\lambda}}}-1\right)^{-\beta}}\right]-2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\delta_{ij}\left(\frac{\theta}{2}\right)^{\frac{1}{2\lambda}}\Gamma\left\{\left(1-\frac{1}{2\lambda}\right),\frac{(j+\beta(i+1))\theta}{2\mu^{2\lambda}}\right\}$$

X. Mean Deviation from Median

We know that

$$D(M)=E(|x-M|)$$

$$D(M)=\int_{0}^{\infty}|x-M|\,f(x)dx$$

$$D(M)=\mu-2\int_{0}^{M}xf(x)dx \tag{23}$$

Now we have

$$\int_{0}^{M}xf(x)dx=\int_{0}^{M}x\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^i}{i!}\binom{\beta(i+1)+j}{j}\alpha^{(i+1)}\beta\frac{\theta\lambda}{x^{2\lambda+1}}\,e^{\frac{-\theta}{2x^{2\lambda}}}\left[e^{\frac{-\theta}{2x^{2\lambda}}}\right]^{j+\beta(i+1)-1}dx$$

$$\int_{0}^{M}xf(x)dx=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\delta_{ij}\left(\frac{\theta}{2}\right)^{\frac{1}{2\lambda}}\Gamma\left\{\left(1-\frac{1}{2\lambda}\right),\frac{(j+\beta(i+1))\theta}{2M^{2\lambda}}\right\} \tag{24}$$

Substituting (24) in (23) we get

$$D(M)=\mu-2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\delta_{ij}\left(\frac{\theta}{2}\right)^{\frac{1}{2\lambda}}\Gamma\left\{\left(1-\frac{1}{2\lambda}\right),\frac{(j+\beta(i+1))\theta}{2M^{2\lambda}}\right\}$$

XI. Maximum Likelihood Estimation

Let $x_1, x_2, x_3, \ldots, x_n$ be n random samples from Weibull Inverse Power Rayleigh Distribution, and then its likelihood function is given by

$$l=\prod_{i=1}^{n}f(x;\alpha,\beta,\theta,\lambda)$$

$$l=(\alpha\beta\theta\lambda)^n\prod_{i=1}^{n}x_i^{-(2\lambda+1)}e^{\frac{\theta}{2}\sum_{i=1}^{n}x_i^{-2\lambda}}\left(e^{\frac{\theta}{2x^{2x}}}-1\right)^{-\beta-1}e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}$$

Its Log Likelihood function is given by

$$\log l = n\log\alpha + n\log\beta + n\log\theta + n\log\lambda - (2\lambda+1)\sum_{i=1}^{n}\log x_i + \frac{\theta}{2}\sum_{i=1}^{n} x_i^{-2\lambda}$$

$$-(\beta+1)\sum_{i=1}^{n}\log\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right) - \alpha\sum_{i=1}^{n}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}$$

Differentiating the above w.r.t $\alpha, \beta, \theta \, and \, \lambda$, we will get the following equations

$$\frac{\partial \log l}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}$$

$$\frac{\partial \log l}{\partial\beta} = \frac{n}{\beta} - \sum_{i=1}^{n}\log\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right) + \alpha\sum_{i=1}^{n}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}\log\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)$$

$$\frac{\partial \log l}{\partial\theta} = \frac{n}{\theta} + \frac{1}{2}\sum_{i=1}^{n} x_i^{-2\lambda} - (\beta+1)\sum_{i=1}^{n}\frac{1}{\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)}\frac{1}{2x_i^{2\lambda}}e^{\frac{\theta}{2x_i^{2\lambda}}} + \alpha\beta\sum_{i=1}^{n}\frac{1}{2x_i^{2\lambda}}e^{\frac{\theta}{2x_i^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1}$$

$$\frac{\partial \log l}{\partial\lambda} = \frac{n}{\lambda} - 2\sum_{i=1}^{n}\log x_i - \theta\sum_{i=1}^{n} x_i^{-2\lambda}\log x_i + (\beta+1)\theta\sum_{i=1}^{n}\frac{1}{\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)}e^{\frac{\theta}{2x_i^{2\lambda}}} x_i^{-2\lambda}\log x_i$$

$$-\alpha\beta\theta\sum_{i=1}^{n} e^{\frac{\theta}{2x_i^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1} x_i^{-2\lambda}\log x_i$$

The above equations are non-linear equations which cannot be expressed in compact form and it is difficult to solve these equations explicitly for α, β, θ and λ . By applying the iterative methods such as Newton–Raphson method, secant method, Regula-Falsi method etc. the MLE of the parameters denoted as $\hat{\eta}(\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\lambda})$ of $\eta(\alpha, \beta, \theta, \lambda)$ can be obtained by using the above methods.

Since the MLE of $\hat{\eta}$ follows asymptotically normal distribution as given as follows $\sqrt{n}(\hat{\eta}-\eta) \rightarrow N(0, I(\eta))$

Where $I^{-1}(\eta)$ is the limiting variance covariance matrix $\hat{\eta}$ and $I(\eta)$ is a 4x4 Fisher Information matrix i.e

$$I(\eta)=-\frac{1}{n}\begin{bmatrix} E\left(\frac{\partial^2 \log l}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha \partial \beta}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha \partial \theta}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha \partial \lambda}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \beta \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial \beta^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \beta \partial \theta}\right) & E\left(\frac{\partial^2 \log l}{\partial \beta \partial \lambda}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \theta \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial \theta \partial \beta}\right) & E\left(\frac{\partial^2 \log l}{\partial \theta^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \theta \partial \lambda}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \lambda \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial \lambda \partial \beta}\right) & E\left(\frac{\partial^2 \log l}{\partial \lambda \partial \theta}\right) & E\left(\frac{\partial^2 \log l}{\partial \lambda^2}\right) \end{bmatrix}$$

Hence the approximate $100(1-\psi)\%$ confidence interval for α, β, θ and λ are respectively given by

$$\hat{\alpha} \pm Z_{\frac{\psi}{2}} \sqrt{I_{\alpha\alpha}^{-1}(\hat{\eta})} \qquad \hat{\beta} \pm Z_{\frac{\psi}{2}} \sqrt{I_{\beta\beta}^{-1}(\hat{\eta})} \qquad \theta \pm Z_{\frac{\psi}{2}} \sqrt{I_{\theta\theta}^{-1}(\hat{\eta})} \qquad \lambda \pm Z_{\frac{\psi}{2}} \sqrt{I_{\lambda\lambda}^{-1}(\hat{\eta})}$$

Where $Z_{\frac{\psi}{2}}$ is the ψ^{th} percentile of the standard distribution.

XII. Order Statistics

Let $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(n)}$ denotes the order statistics of n random samples drawn from Weibull Inverse Power Rayleigh Distribution, then the pdf of $X_{(k)}$ is given by

$$f_{x(k)}(x;\theta)=\frac{n!}{(k-1)!(n-k)!} f_X(x)\left[F_X(x)\right]^{k-1}\left[1-F_X(x)\right]^{n-k}$$

$$f_{x(k)}(x;\theta)=\frac{n!}{(k-1)!(n-k)!}\frac{\alpha\beta\theta\lambda}{x^{2\lambda+1}} e^{\frac{\theta}{2x^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1} e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\left\{1-e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\right\}^{k-1}\left(e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\right)^{n-k}$$

Then the pdf of first order $X_{(1)}$ Weibull Inverse Power Rayleigh Distribution is given by

$$f_{x(1)}(x;\theta)=\frac{n\alpha\beta\theta\lambda}{x^{2\lambda+1}} e^{\frac{\theta}{2x^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1} e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\left(e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\right)^{n-1}$$

and the pdf of nth order $X_{(n)}$ Weibull Inverse Power Rayleigh Distribution is given by

$$f_{x(n)}(x;\theta)=\frac{n\alpha\beta\theta\lambda}{x^{2\lambda+1}} e^{\frac{\theta}{2x^{2\lambda}}}\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta-1} e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\left\{1-e^{-\alpha\left(e^{\frac{\theta}{2x^{2\lambda}}}-1\right)^{-\beta}}\right\}^{n-1}$$

XIII. Simulation Study

In this section, we study the performance of ML estimators for different sample sizes (n=, 50,150, 250,500). We have employed the inverse CDF technique for data simulation for WIPRD distribution using R software. The process was repeated 1000 times for calculation of bias, variance and MSE.

Table 1: *The Mean values, Average bias and MSEs of 1,000 simulations of WIPRD for parameter values*

Sample Size n	parameters	$\alpha = 1.0$ $\beta = 0.9$ $\theta = 2.3$ and $\lambda = 0.3$			
		Average	Bias	Variance	MSE
50	α	0.9493	-0.050	0.0001	0.0027
	β	0.7717	-0.128	0.0001	0.0166
	θ	2.0815	-0.218	0.0246	0.0723
	λ	2.0510	1.7510	0.0768	3.1429
150	α	0.9537	-0.046	8.41e-05	0.0022
	β	0.7760	-0.123	8.11e-05	0.0154
	θ	2.0854	-0.214	7.07e-03	0.0531
	λ	1.9675	1.6675	2.55e-02	2.8063
250	α	0.9545	-0.045	6.23e-05	0.0021
	β	0.7767	-0.123	5.81e-05	0.0152
	θ	2.0866	-0.213	4.24e-05	0.0497
	λ	1.9479	1.6479	1.56e-05	2.7313
500	α	0.9556	-0.044	3.43e-05	0.0020
	β	0.7778	-0.122	3.11e-05	0.0149
	θ	2.0905	-0.209	2.15e-05	0.0460
	λ	1.9261	1.6261	8.36e-05	2.6528
		$\alpha = 1.1$ $\beta = 0.8$ $\theta = 2.0$ and $\lambda = 0.5$			
50	α	0.9465	-0.153	0.0001	0.0236
	β	0.7688	-0.031	0.0002	0.0011
	θ	1.7951	-0.204	0.0216	0.0635
	λ	1.1839	0.6839	0.0248	0.4926
150	α	0.9496	-0.150	6.99e-o5	0.0226
	β	0.7716	-0.028	7.04e-05	0.0008
	θ	1.8038	-0.196	6.59e-03	0.0450
	λ	1.1476	0.6476	8.11e-03	0.4275
250	α	0.9508	-0.149	5.19e-05	0.0223
	β	0.7728	-0.027	5.04e-05	0.0007
	θ	1.8116	-0.188	4.33e-03	0.0398
	λ	1.1355	0.6355	4.83e-03	0.4087
500	α	0.9513	-0.148	2.83e-05	0.0221
	β	0.7733	-0.026	2.69e-05	0.0007
	θ	1.8128	-0.187	2.08e-03	0.0371
	λ	1.1277	0.6277	2.59e-03	0.3967

As is clear from table 1, decreasing trend is being observed in average bias, variance and MSE as we increase the sample size. Hence, the performance of ML estimators is quite well, consistent in case of Weibull Inverse Power Rayleigh Distribution.

XIV. Application

In this segment, the efficacy of the developed distribution has been assessed using two realistic sets of data. As the new distribution is compared to New Modified Weibull distribution (NMWD), Additive Weibull distribution (AWD), Power Gompertz distribution (PGD), Inverse power Rayleigh distribution (IPRD), Weibull distribution (WD), Lindley distribution (LD) and Hamza distribution (HD). It is revealed that the new developed distribution offers an appropriate fit.

Various criterion including the AIC (Akaike information criterion), CAIC (Consistent Akaike information criterion), BIC (Bayesian information criterion), HQIC (Hannan-Quinn information criteria) and KS (Kolmogorov-Smirnov) are used to compare the fitted models. The p-value of each model is also recorded. A distribution having lesser AIC, CAIC, BIC, HQIC and KS values and with large p-value is considered better one.

$$AIC = 2k - 2\ln l \qquad CAIC = \frac{2k}{n-k-1} - 2\ln l$$

$$BIC = k\ln n - 2\ln l \qquad HQIC = 2k\ln(\ln(n)) - 2\ln l$$

Data set 1:- The following represents the dataset of 63 0bservations of the tensile strength measurements on 1000 carbon fiber-impregnated tows at four different gauge lengths. The data reported by Bader and Priest (1982) as follows:

1.901, 2.132, 2.203, 2.228, 2.257,2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886 ,3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

Data set 2: The second data represents COVID-19 mortality rates data belongs to Italy of 59 days that is recorded from 27 February to 27 April 2020. The data is as follows:

4.571, 7.201, 3.606, 8.479, 11.410, 8.961, 10.919, 10.908, 6.503, 18.474, 11.010 ,17.337, 16.561, 13.226, 15.137, 8.697, 15.787,13.333, 11.822, 14.242, 11.273, 14.330, 16.046, 11.950, 10.282, 11.775, 10.138, 9.037, 12.396, 10.644, 8.646, 8.905, 8.906, 7.407, 7.445 ,7.214, 6.194, 4.640, 5.452, 5.073 ,4.416 ,4.859 ,4.408, 4.639, 3.148, 4.040, 4.253, 4.011, 3.564, 3.827, 3.134, 2.780, 2.881, 3.341, 2.686, 2.814, 2.508, 2.450, 1.518.

The fitted models are compared using empirical goodness of fit measures such as the AIC (Akaike information criterion), CAIC (Consistent Akaike information criterion), BIC (Bayesian information criterion), HQIC (Hannan-Quinn information criteria), and KS (Kolmogorov- Smirnov). Each model's p-value is also displayed. A distribution with a lower AIC, CAIC, BIC, and HQIC together with a higher p value is rated as the top distribution

Table 2 and 5 shows the descriptive statistics for data set 1 and data set 2 respectively. Table 3 and 6 displays the parameter estimates for the data set 1 and data set 2 respectively. Table 4 and 7 displays the log-likelihood, Akaike information criteria (AIC), BIC (Bayesian information criterion) etc. details and some other statistics for the data set 1 and data set 2.

Table 2. *Descriptive Statistics for data set 1*

Mean	Min.	Max.	Q_1	Q_3	Median	S.D	Skew.	Kurt.
3.059	1.901	5.020	2.554	3.421	2.996	0.620	0.632	3.286

Table 3. *The ML Estimates and standard error of the unknown parameters*

Model		WIPRD	NMWD	AWD	PGD	IPRD	WD	LD	HD
$\hat{\alpha}$		11.4370	0.01469	0.00292	0.00438	---------	0.003775	---------	---------
$\hat{\beta}$		0.7586	0.00100	0.00100	2.72725	---------	4.69909	---------	0.00100
$\hat{\theta}$		60.8587	3.14576	4.68182	0.71363	15.1809	-----------	0.53923	2.28724
$\hat{\lambda}$		0.91032	0.43441	4.54990	---------	0.91031	-----------	---------	---------
Stan-	$\hat{\alpha}$	5.880	0.0063	0.00081	0.00165	---------	0.00105	---------	---------
dard	$\hat{\beta}$	0.5707	0.00034	0.00058	0.51536	---------	0.22866	---------	0.62231
Error	$\hat{\theta}$	53.0940	0.29681	0.21350	0.09294	2.34246	-----------	0.04958	0.05084
	$\hat{\lambda}$	0.0810	0.26257	0.26577	---------	0.08109	-----------	---------	---------

Table 4. *Performance of distributions*

Model	WIPRD	NMWD	AWD	PGD	IPRD	WD	LD	HD
-2logl	111.612	138.816	124.704	134.466	185.528	124.546	242.714	150.514
AIC	119.612	146.816	132.705	140.466	189.528	128.546	244.715	154.515
AICC	120.301	147.505	133.395	146.896	193.814	128.746	244.780	154.715
HQIC	122.983	150.187	136.077	140.873	189.728	130.232	245.558	156.201
BIC	128.184	155.388	141.278	142.995	191.214	132.832	246.858	158.802
K-S Value	0.08215	0.1448	0.11151	0 .143	0.35439	0.11139	0.4308	0.21545
P Value	0.7888	0.1424	0.4137	0.152	2.68e-07	0.4151	1.39e-10	0.00576

Table 5. *Descriptive Statistics for data set 2*

Mean	Min.	Max.	Q_1	Q_3	Median	S.D	Skew.	Kurt.
8.156	1.518	18.474	4.146	11.341	7.445	4.5267	0.4523	2.1281

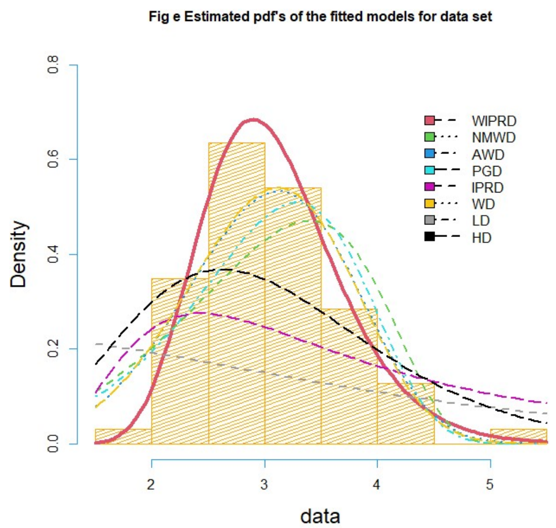
Table 6. *The ML Estimates and standard error of the unknown parameters*

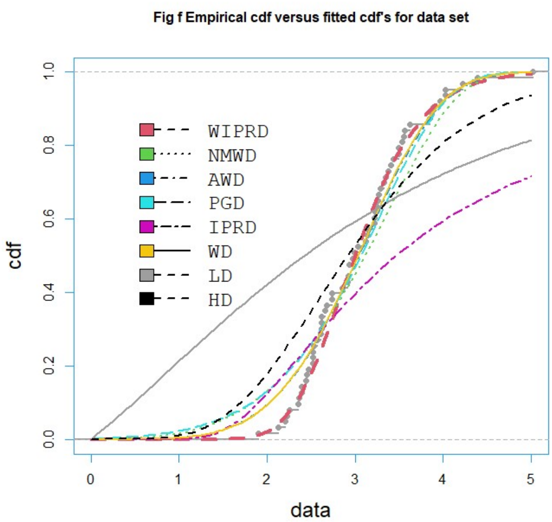
Model		WIPRD	NMWD	AWD	PGD	IPRD	WD	LD	HD
$\hat{\alpha}$		0.2065	0.02612	0.02612	0.01746	---------	0.014105	---------	---------
$\hat{\beta}$		0.82124	0.00100	0.00100	0.00387	---------	1.918049	---------	21.5900
$\hat{\theta}$		11.26308	1.434910	1.43491	1.75839	22.8479	-----------	0.222869	0.81384
$\hat{\lambda}$		0.51936	0.293231	0.29323	---------	0.51936	-----------	---------	---------
Standard Error	$\hat{\alpha}$	0.16217	0.01111	0.01111	0.01006	---------	0.006603	---------	---------
	$\hat{\beta}$	0.12048	0.013344	0.01334	0.00864	---------	0.18601	---------	28.7503
	$\hat{\theta}$	7.0582	0.15826	0.15826	0.31366	3.6430	-----------	0.02068	0.05698
	$\hat{\lambda}$	0.04781	0.23093	0.230931	---------	0.04781	-----------	---------	---------

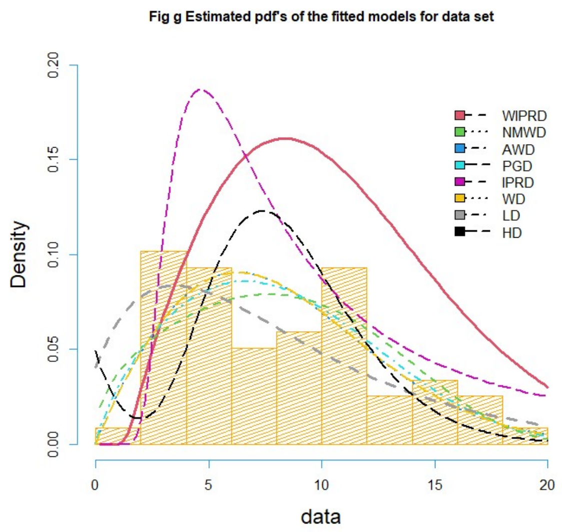
Table 7. *Performance of distributions*

Model	WIPRD	NMWD	AWD	PGD	IPRD	WD	LD	HD
-2logl	267.932	336.232	336.232	334.882	289.302	335.404	346.598	362.508
AIC	275.932	344.232	344.232	340.882	293.302	339.404	348.599	366.508
AICC	276.673	344.973	344.973	341.319	293.516	339.618	348.669	366.722
HQIC	279.176	347.476	347.476	343.315	294.924	341.026	349.410	368.129
BIC	284.242	352.543	352.543	347.115	297.457	343.559	350.677	370.663
K-S Value	0.10235	0.12345	0.21187	0.12154	0.22795	0.121	0.14516	0.21409
P Value	0.5331	0.3041	0.00839	0.3216	0.00354	0.3268	0.1506	0.00748

As it is obvious from table 4 and table 7 that the Weibull inverse power Rayleigh distribution has smaller values for AIC, AICC, BIC, HQIC and K-S statistics as compared with its sub models. Accordingly we arrive at the conclusion that Weibull inverse power Rayleigh distribution provides an adequate fit than the compared ones.







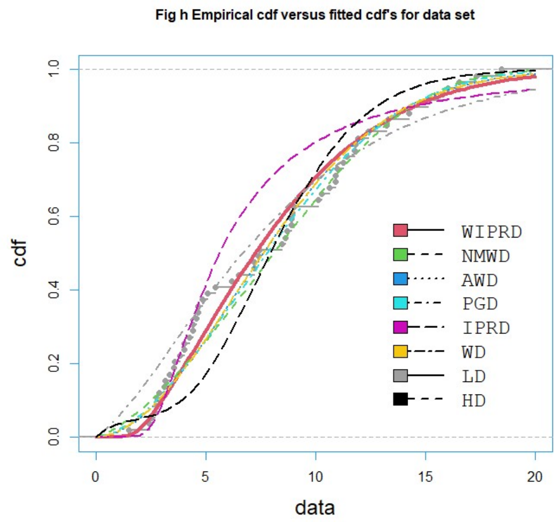


Figure e, f, g and h represents the estimated densities and cdfs of the fitted distributions to data set 1st and 2nd.

XV. Conclusion

This newly introduced distribution "Weibull inverse power Rayleigh distribution" which is obtained by T-X method. Several mathematical quantities for the newly developed distribution are derived including moments, moment generating function, incomplete moments, order statistics, different measure of entropies etc. To show the behavior of p.d.f, c.d.f and other related measures different plots have been drawn. The parameters are obtained by MLE technique. Lastly by carrying out through two real life data sets to show that the formulated distribution leads an improved fit than the compared ones.

References

[1] Albert L (2016).Odd Generalized Exponential Rayleigh Distribution. Advances and Applications in Statistics, Vol 48(1) , 33-48.

[2] Alzaatreh A, Lee C, Famoye F (2013).A New Method for Generating Families of Distributions. Metron, 71, 63-79.

[3] Amal H and Said G.N (2018).The Inverse Weibull-G Family. Journal of Data Science, 723-742.

[4] Bhat A.A and Ahmad S.P (2020).A New Generalization of Rayleigh Distribution: Properties and Applications. Pakistan Journal of Statistics, 36(3), 225-250.

[5] Fatoki O (2019).The Topp-Leone Rayleigh Distribution with Application. American Journal of Mathematics and Statistics, 9(6), 215-220.

[6] Faton M (2013).Transmuted Rayleigh Distribution. Austrian Journal of Statistics, 42(1), 21-31.

[7] Faton M and Ibrahim E (2015).Weibull-Rayleigh Distribution Theory and Applications. Applied Mathematics and Information Sciences, an International Journal, Vol 9(5), 1-11.

[8] Frechet M (1927).Sur La Loi De Probabilite De Lecart Maximum. Ann. Soc. Polon. De Math. Cracovie. 6:93-116.

[9] Rosin, P; Rammler, E (1933).“The laws Governing the Fineness of Powdered Coal”. Journal of the Institute of Fuel, 7:29-36.
[10] Terna G.I, Sauta A, Issa A.A (2020).Odd Lindley-Rayleigh Distribution, Its Properties and Applications to Simulated and Real Life Datasets. Journal of Advances in Mathematics and Computer Science, 35(1), 63-88.
[11] Voda, R (1972). “On the Inverse Rayleigh Variable”, Rep. Stat. Res. Ju, Vol. 19(4), 15-21.

RT&A, No 2 (68)
Volume 17, June 2022

Estimation and Prediction for Exponentiated Exponential Distribution under Generalised Progressive Hybrid Censoring

Aakriti Pandey[1], A. Kaushik[*2], & S. K. Singh[3]

•

[1,2,3] Department of Statistics, Institute of Science, Banaras Hindu University,
Varanasi, India-221005.
Email: [2]arundevkauhsik@gmail.com, [1]akritibhu@gmail.com ,
[3]singhsk64@gmail.com
[*] Corresponding author

Abstract

In this article, we propose the estimators for the parameters of exponentiated exponential distribution under generalized progressive hybrid censoring scheme obtained through different methods of estimations like maximum likelihood, Maximum product spacing, Bootstrap and Bayesian. Asymptotic confidence, Bootstrap and HPD intervals have also been computed. Moreover, Stress Strength reliability estimation is also discussed. The performance of the estimators have been studied in terms of their MSEs. Bayesian prediction of future observations has also been attempted. For illustrating the proposed methodology, a real data set is taken into account.

Keywords: Exponentiated Exponential distribution, generalized progressive hybrid censoring scheme, Stress-Strength Reliability, Bayes estimates, Bayesian prediction.

1. Introduction

The progressive hybrid censoring schemes has gained considerable attention in past few years. [1] and [2] have considered progressive hybrid type-I censoring scheme. In PHT-I, n units are put on test with progressive censoring scheme $(R_1, R_2, \cdots, R_m)$ and the time of termination of the experiment is fixed as $T^* = min\{X_{m:m:n}, T\}$($X_{m:m:n}$ denotes mth order failure time), $T \in (0, \infty)$ and $1 \leq m \leq n$ are prefixed constants. In PHT-I,test can never exceed T which facilitates reduction in time and cost of the experimentation. The problem arises when the unknown average life time is higher than the stopping time that results to a status with less than m failures to be observed. It ultimately reduces the efficiency of the inference based on the censored data.

[2] further proposed progressive hybrid type-II censoring scheme. In PHT-II, the experiment is terminated at time $T^* = max\{X_{m:m:n}, T\}$ which assures at least m number of failures (see [2]). When $X_{m:m:n} > T$, the experiment is terminated at mth failure with withdrawals occurring at each failure according to pre-specified progressive scheme $(R_1, R_2, \cdots, R_m)$ that may lead to significant increase in the time of termination. On the other hand, when $X_{m:m:n} < T$, we observe failure upto time T. The termination time in this censoring scheme is unknown to the experimenter. From the above discussion, we note that PHT-I censoring keeps the termination time of the experiment below prefixed value by compromising the efficiency where as PHT-II censoring ensures efficiency more than the prefixed level but compromises the termination time. Therefore, need for a censoring scheme controlling termination time and efficiency simultaneously was felt. Keeping this point in mind, [3] introduced a censoring scheme called generalized progressive hybrid (GPH) censoring scheme.

The paper envisages exponentiated exponential distribution(EED) as a lifetime distribution which was introduced by [4]. Many author studied the EED, see [5], [6] and [7].

The probability density function of EED is given as

$$f(x|\alpha,\beta) = \alpha\beta e^{-\beta x}(1-e^{-\beta x})^{\alpha-1}; x \geq 0, \alpha, \beta > 0. \tag{1}$$

where α and β are the shape and scale parameters and the cumulative distribution function is given by

$$F(x|\alpha,\beta) = (1-e^{-\beta x})^{\alpha}. \tag{2}$$

The requisition for the prediction of future sample on the basis of current information started burgeoning demand since it has found application in a wide range of activities including science, engineering, statistics, social sciences and other applied areas. Predictive inference facilitates to infer about future lifetimes using observed data. This impetus is welcomed by many authors and they made successful effort to draft the problem of Bayesian prediction of future observations based on various types of censored data from different lifetime models. A lot of research paper is available in literature on the prediction problem in case of censored as well as complete data (see [8], [9],[10], [11],[12]).

The paper is organized as follows. Section 2 describes the censoring scheme in detail. Section 3 provides estimates through different methods of estimation i.e. Maximum Likelihood estimates, estimates through Maximum product spacing method, Bootstrap estimates and Bayesian Estimates. Section 4 is about Bayesian Prediction of future observations based on GPH censored data in which both one sample prediction and two sample prediction has been attempted. It also contains stress strength reliability estimates. In section 5, a real data set is considered to demonstrate the applicability of the methodology. In section 6, a simulation study is conducted. Finally, conclusions are summarized in section 7.

2. The Censoring Scheme

Let the experiment begins with n units. The lifetime of the sample units $X_1, X_2, \cdots, X_n$ are supposed to be independent and identically distributed random variables from a distribution with cumulative density function (cdf) $F(\cdot)$ and probability density function (pdf) $f(\cdot)$. we have prefixed integers $k, m \in \{1, 2, \cdots, n\}$ such that $k < m$. Let R_i denotes the number of units which are randomly removed from the experiment at i^{th} failure obeying the condition that $\sum_{i=1}^{m} R_i + m = n$. The test is terminated at the stopping time $T^* = max\{X_{k:m:n}, min\{X_{m:m:n}, T\}\}$. It may be noted that this scheme guarantees a bare minimum number of k failures. Let D be the number of observed failures up to time T. Then for observed observations, following three cases arise under this scheme:

$$\begin{aligned}
&\textbf{Case-I:} && X_{1:m:n}, \cdots, X_{2:m:n}, \cdots, X_{k:m:n}, && \text{if } T < X_{k:m:n},\\
&\textbf{Case-II:} && X_{1:m:n}, \cdots, X_{k:m:n}, \cdots, X_{D:m:n}, && \text{if } X_{k:m:n} < T < X_{m:m:n},\\
&\textbf{Case-III:} && X_{1:m:n}, \cdots, X_{k:m:n}, \cdots, X_{m:m:n}, && \text{if } X_{k:m:n} < X_{m:m:n} < T.
\end{aligned}$$

A schematic representation of this censoring scheme is given in figure 1. Note that for the Case-I, $X_{k+1:m:n}, \cdots, X_{m:m:n}$ are not observed; likewise for the Case-II, $X_{D+1:m:n}, \cdots, X_{m:n:n}$ are not observed. Given a generalised progressive censored sample, the likelihood function for Case-I, Case-II and Case-III denoted by $L_I(\alpha,\beta)$, $L_{II}(\alpha,\beta)$and $L_{III}(\alpha,\beta)$ is given below:

$$\begin{aligned}
&\textbf{Case-I:}\ L_I(\alpha,\beta) = K_1 \prod_{j=1}^{k-1} f(x_{j:m:n})[1-F(x_{j:m:n})]^{R_j} f(x_{k:m:n})[1-F(x_{k:m:n})]^{R_k^*},\\
&\textbf{Case-II:}\ L_{II}(\alpha,\beta) = K_2 \prod_{j=1}^{D} f(x_{j:m:n})[1-F(x_{j:m:n})]^{R_j}[1-F(T)]^{R_{D+1}^*},\\
&\textbf{Case-III:}\ L_{III}(\alpha,\beta) = K_3 \prod_{j=1}^{m} f(x_{j:m:n})[1-F(x_{j:m:n})]^{R_j},
\end{aligned} \tag{3}$$

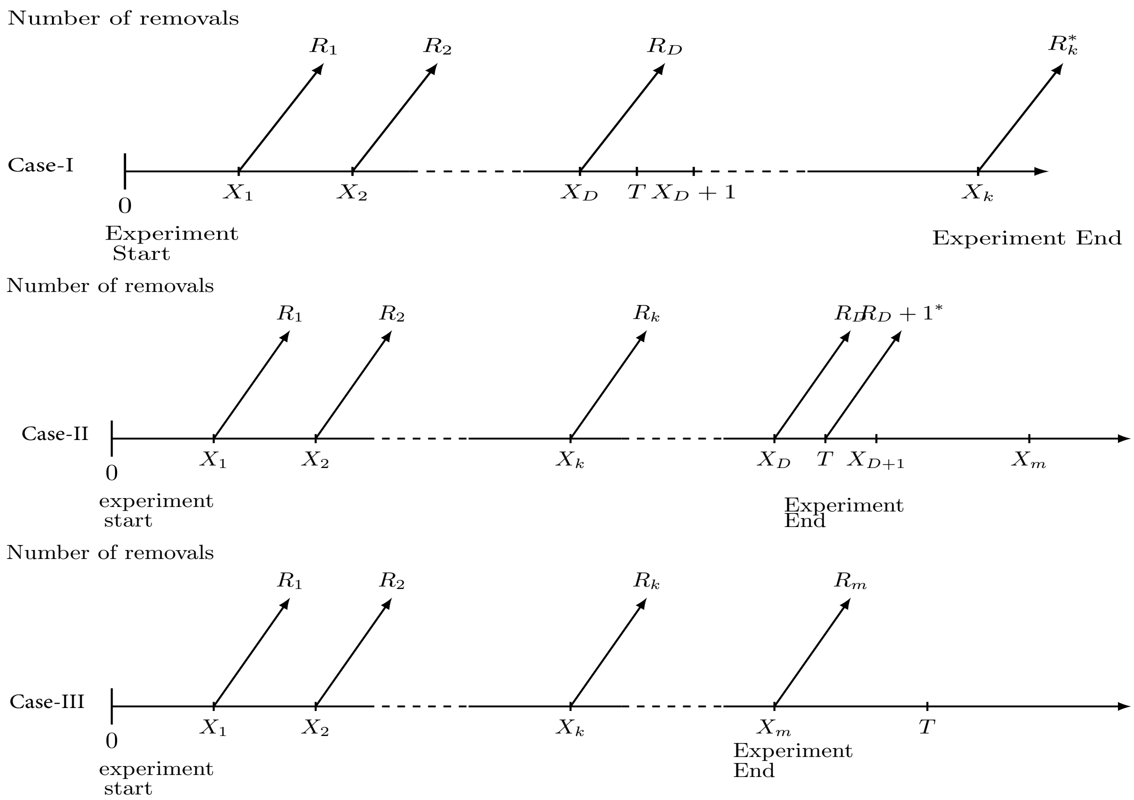


Figure 1: *Schematic representation of generalised progressive hybrid censoring Scheme*

where $K_1 = [\prod_{j=1}^{k}\sum_{k=j}^{m}(R_k+1)], K_2 = [\prod_{j=1}^{D}\sum_{k=j}^{m}(R_k+1)], K_3 = [\prod_{j=1}^{m}\sum_{k=j}^{m}(R_k+1)], R_k^* = [n-k-\sum_{i=1}^{k-1}R_i]$ and $R_{D+1}^*=[n-D-\sum_{i=1}^{D}R_i]$.

3. Estimation

3.1. Classical Estimation

3.1.1 Maximum Likelihood Estimation

Maximum likelihood estimation is one of the most felicitous method in classical paradigm to obtain the estimates of the parameters of proposed distribution. In this section, we will find the MLEs of α and β of the considered distribution. The MLEs $\hat{\alpha}$ and $\hat{\beta}$ of α and β, respectively can be obtained by maximising the likelihood function. Using the equations (1) and (2), the likelihood can be written as :

$$L(\alpha,\beta) \propto (\alpha\beta)^J \prod_{j=1}^{J}(1-e^{-\beta x_j})^{\alpha-1}e^{-\beta x_j}[1-(1-e^{-\beta x_j})^{\alpha}]^{R_j} \times W(\alpha,\beta), \tag{4}$$

where $W(\alpha,\beta) = \begin{cases} 1, & \text{if } J = k, m \\ \left[1-(1-e^{-\beta T})^{\alpha}\right]^{R_{D+1}^*}, & \text{if } J = D. \end{cases}$
and hence log-likelihood equation will be

$$\begin{aligned} l(\alpha,\beta) \propto & J \times \ln(\alpha\beta) + (\alpha-1)\sum_{j=1}^{J}\ln(1-e^{-\beta x_j}) \\ & + \sum_{j=1}^{J} R_j \ln[1-(1-e^{-\beta x_j})^{\alpha}] - \beta\sum_{j=1}^{J} x_j + \ln W(\alpha,\beta), \end{aligned} \tag{5}$$

Differentiating it with respect to the parameters α and β respectively, we get :

$$\frac{\partial l(\alpha,\beta)}{\partial \alpha} = \frac{J}{\alpha} + \sum_{j=1}^{J} \ln(1-e^{-\beta x_j}) - \sum_{j=1}^{J} \frac{R_j(1-e^{-\beta x_j})^{\alpha} \ln(1-e^{-\beta x_j})}{1-(1-e^{-\beta x_j})^{\alpha}} + \frac{\partial \ln W(\alpha,\beta)}{\partial \alpha} \tag{6}$$

$$\frac{\partial l(\alpha,\beta)}{\partial \beta} = \frac{J}{\beta} + (\alpha-1)\sum_{j=1}^{J} \frac{x_j e^{-\beta x_j}}{(1-e^{-\beta x_j})} - \sum_{j=1}^{J} \frac{R_j \alpha (1-e^{-\beta x_j})^{\alpha-1} x_j e^{-\beta x_j}}{1-(1-e^{-\beta x_j})^{\alpha}} - \sum_{j=1}^{J} x_j + \frac{\partial \ln W(\alpha,\beta)}{\partial \beta} \tag{7}$$

where $\frac{\partial \ln W(\alpha,\beta)}{\partial \alpha} = -\frac{R^*_{D+1}(1-e^{-\beta T})^{\alpha}}{1-(1-e^{-\beta T})^{\alpha}} \times \ln(1-e^{-\beta T})$ and $\frac{\partial \ln W(\alpha,\beta)}{\partial \beta} = -\frac{R^*_{D+1}\alpha(1-e^{-\beta T})^{\alpha-1}}{1-(1-e^{-\beta T})^{\alpha}} \times T e^{-\beta T}$, if $J = D$ and zero otherwise. The MLE of α and β can be obtained by solving likelihood equations (6) and (7) simultaneously. But it may be noted that explicit solutions of the above equations are difficult to find. Therefore, we propose the use of numerical method to obtain the solution for the above two non linear equations.

3.1.2 Maximum Product Spacing Method

It has underscored that for small samples MPS often perform better than MLE. The asymptotic behaviour of both MPS and MLE methods is same. In this section the method of product of spacings is proposed for point estimation of parameters of EED under GPH censoring scheme. The product spacing, denoted as G, is defined as product of the probabilities of an observation lying in the intervals induced by the sample observations and MPS estimates are values of the parameters that maximizes G . The expressions for G and corresponding equations whose solutions provide the MPS estimates are given below:
For Case I and Case III:

$$\begin{aligned} G &\propto \prod_{i=1}^{J+1}[F(x_i) - F(x_{i-1})] \prod_{i=1}^{J}[1-F(x_i)]^{R_i} \\ &\propto \prod_{i=1}^{J+1}[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}] \prod_{i=1}^{J}[1-(1-e^{-\beta x_i})^{\alpha}]^{R_i} \end{aligned}$$

The logarithm of G is

$$logG \propto \sum_{i=1}^{J+1} log[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}] + \sum_{i=1}^{J} R_i log[1-(1-e^{-\beta x_i})^{\alpha}] \tag{8}$$

The partial derivatives with respect to the parameters when equated to zero gives the following:

$$\begin{aligned} \frac{dlogG}{d\alpha} =& \sum_{i=1}^{J+1} \frac{(1-e^{-\beta x_i})^{\alpha} log(1-e^{-\beta x_i}) - (1-e^{-\beta x_{i-1}})^{\alpha} log(1-e^{-\beta x_{i-1}})}{[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}]} \\ &- \sum_{i=1}^{J} R_i \frac{(1-e^{-\beta x_i})^{\alpha} log(1-e^{-\beta x_i})}{[1-(1-e^{-\beta x_i})^{\alpha}]} = 0 \end{aligned} \tag{9}$$

$$\begin{aligned} \frac{dlogG}{d\beta} =& \sum_{i=1}^{J+1} \frac{\alpha(1-e^{-\beta x_i})^{\alpha-1} e^{-\beta x_i} x_i - \alpha(1-e^{-\beta x_{i-1}})^{\alpha-1} e^{-\beta x_{i-1}} x_{i-1}}{[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}]} \\ &- \sum_{i=1}^{J} R_i \frac{\alpha(1-e^{-\beta x_i})^{\alpha-1} e^{-\beta x_i} x_i}{[1-(1-e^{-\beta x_i})^{\alpha}]} = 0 \end{aligned} \tag{10}$$

For Case II:

$$\begin{aligned} G &\propto \prod_{i=1}^{D}[F(x_i) - F(x_{i-1})](F(T) - F(D))(1-F(T)) \prod_{i=1}^{D}[1-F(x_i)]^{R_i}[1-F(T)]^{R^*_{D+1}} \\ &\propto \prod_{i=1}^{D}[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}][(1-e^{-\beta T})^{\alpha} - (1-e^{-\beta D})^{\alpha}][1-(1-e^{-\beta T})^{\alpha}] \\ &\prod_{i=1}^{D}[1-(1-e^{-\beta x_i})^{\alpha}]^{R_i}[1-(1-e^{-\beta T})^{\alpha}]^{R^*_{D+1}} \end{aligned}$$

and

$$logG \propto \sum_{i=1}^{D} log[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}] + log[(1-e^{-\beta T})^{\alpha} - (1-e^{-\beta D})^{\alpha}]$$
$$+ log[1-(1-e^{-\beta T})^{\alpha}] + \sum_{i=1}^{D} R_i log[1-(1-e^{-\beta x_i})^{\alpha}] + R^*_{D+1} log[1-(1-e^{-\beta T})^{\alpha}]$$

The resulting equations are

$$\begin{aligned} \frac{dlogG}{d\alpha} =& \sum_{i=1}^{D} \frac{(1-e^{-\beta x_i})^{\alpha} log(1-e^{-\beta x_i}) - (1-e^{-\beta x_{i-1}})^{\alpha} log(1-e^{-\beta x_{i-1}})}{[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}]} \\ &+ \frac{(1-e^{-\beta T})^{\alpha} log(1-e^{-\beta T}) - (1-e^{-\beta D})^{\alpha} log(1-e^{-\beta D})}{[(1-e^{-\beta T})^{\alpha} - (1-e^{-\beta D})^{\alpha}]} \\ &- \frac{(1-e^{-\beta T})^{\alpha} log(1-e^{-\beta T})}{[1-(1-e^{-\beta T})^{\alpha}]} - \sum_{i=1}^{D} R_i \frac{(1-e^{-\beta x_i})^{\alpha} log(1-e^{-\beta x_i})}{[1-(1-e^{-\beta x_i})^{\alpha}]} \\ &- R^*_{D+1} \frac{(1-e^{-\beta T})^{\alpha} log(1-e^{-\beta T})}{[1-(1-e^{-\beta T})^{\alpha}]} = 0 \end{aligned} \tag{11}$$

$$\begin{aligned} \frac{dlogG}{d\beta} =& \sum_{i=1}^{D} \frac{\alpha(1-e^{-\beta x_i})^{\alpha-1} e^{-\beta x_i} x_i - \alpha(1-e^{-\beta x_{i-1}})^{\alpha-1} e^{-\beta x_{i-1}} x_{i-1}}{[(1-e^{-\beta x_i})^{\alpha} - (1-e^{-\beta x_{i-1}})^{\alpha}]} \\ &+ \frac{\alpha(1-e^{-\beta T})^{\alpha-1} e^{-\beta T} T - \alpha(1-e^{-\beta D})^{\alpha-1} e^{-\beta D} D}{[(1-e^{-\beta T})^{\alpha} - (1-e^{-\beta D})^{\alpha}]} \\ &- \frac{\alpha(1-e^{-\beta T})^{\alpha-1} e^{-\beta T} T}{[1-(1-e^{-\beta T})^{\alpha}]} - \sum_{i=1}^{D} R_i \frac{\alpha(1-e^{-\beta x_i})^{\alpha-1} e^{-\beta x_i} x_i}{[1-(1-e^{-\beta x_i})^{\alpha}]} \\ &- R^*_{D+1} \frac{\alpha(1-e^{-\beta T})^{\alpha-1} e^{-\beta T} T}{[1-(1-e^{-\beta T})^{\alpha}]} = 0 \end{aligned} \tag{12}$$

Similar to the likelihood equations, MPS equations are also non-linear equations, thus an approach similar to that proposed for MLE is to used here also.

3.1.3 Bootstrap Estimates

The asymptotic confidence intervals are generally expected not to perform well when the effective sample size is small. The key provision that seeks to address such problems is re-sampling technique such as bootstrap. This tool provides a tactical way as it would give more accurate approximate confidence interval. Basically a couple of methods are available in literature to find bootstrap confidence intervals for the parameter of interest. One is percentile bootstrap (Boot-p) confidence interval given by [13] and another one is student's t bootstrap (Boot-t) confidence interval suggested by [14]. The generation algorithm for GPH censored sample is discussed in section 6. We propose to use the following algorithm to generate parametric bootstrap samples, suggested by [15] as given below.
Algorithm:

Step 1. Compute $\hat{\alpha}$ and $\hat{\beta}$ which is nothing but the ML estimate of the parameter α and β respectively based on GPH censored sample $x = (x_{1:J:n}, x_{2:J:n}, \cdots, x_{J:J:n})$.

Step 2. Generate the bootstrap GPH censored samples $x^* = (x^*_{1:J:n}, x^*_{2:J:n}, \cdots, x^*_{J:J:n})$ from EED with parameters $\hat{\alpha}$ and $\hat{\beta}$ by using algorithm given in section 6. From these data, we compute bootstrap estimates say, $\hat{\alpha}^*$ and $\hat{\beta}^*$.

Step 3. Repeat step 2, B times to obtain a set of bootstrap GPH censored samples of α and β as $(\hat{\alpha}^*_1, \hat{\alpha}^*_2, \cdots, \hat{\alpha}^*_B)$ and $(\hat{\beta}^*_1, \hat{\beta}^*_2, \cdots, \hat{\beta}^*_B)$.

From the bootstrap samples generated from the above algorithm, the bootstrap confidence intervals for the parameters α and β can be obtained by Boot-p method which is as follows: Let $G(x) = Prob(\hat{\alpha}^* \leq x)$ be the cdf of $\hat{\alpha}^*$. Define $\hat{\alpha}_{Boot}(x)$=$G^{-1}(x)$ for given x. The $100(1-\zeta)\%$ bootstrap percentile interval for α is given by $\left(\hat{\alpha}_{Boot}\left(\frac{\zeta}{2}\right), \hat{\alpha}_{Boot}\left(1-\frac{\zeta}{2}\right)\right)$, which is $\frac{\zeta}{2}$ and $(1-\frac{\zeta}{2})$ quantiles of bootstrap sample $\hat{\alpha}_1^*, \hat{\alpha}_2^*, \cdots, \hat{\alpha}_B^*$. In the similar fashion we can obtain bootstrap percentile interval for β.

3.2. Bayesian Estimation

This section contains the method of obtaining Bayes estimates of the parameters α and β based on GPH censored data with prefixed removals. In order to obtain Bayes estimators, we assume that the parameters α and β are random variables which are independently distributed having prior distribution as:

$$g_1(\alpha) = \frac{\lambda_1^{\nu_1}}{\Gamma\nu_1} e^{(-\lambda_1\alpha)}\alpha^{\nu_1-1}; 0 < \alpha < \infty, \lambda_1 > 0, \nu_1 > 0 \tag{13}$$

$$g_2(\beta) = \frac{\lambda_2^{\nu_2}}{\Gamma\nu_2} e^{(-\lambda_2\beta)}\beta^{\nu_2-1}; 0 < \beta < \infty, \lambda_2 > 0, \nu_2 > 0 \tag{14}$$

respectively. Combining the priors given by (13) and (14) with likelihood function in (4), we obtain the joint posterior density function of α and β given as $\pi(\alpha, \beta|x, R) = \dfrac{J_1}{J_0}$, where

$$J_1 = \frac{\lambda_1^{\nu_1}}{\Gamma\nu_1}\frac{\lambda_2^{\nu_2}}{\Gamma\nu_2} e^{-(\lambda_1\alpha+\lambda_2\beta)}\alpha^{(J+\nu_1-1)}\beta^{(J+\nu_2-1)}\prod_{j=1}^{J}(1-e^{-\beta x_j})^{(\alpha-1)}[1-(1-e^{-\beta x_j})^{\alpha}]^{R_j}e^{-\beta x_j} \tag{15}$$

and $J_0 = \int\int J_1 d\alpha d\beta$. Bayes estimators $\hat{\alpha}_E$ and $\hat{\beta}_E$ of α and β under SELF(squared error loss function) can be obtained as

$$\hat{\alpha}_B = \int_0^\infty \int_0^\infty \alpha\pi_1(\alpha|x, R)d\alpha d\beta, \qquad \hat{\beta}_B = \int_0^\infty \int_0^\infty \beta\pi_2(\beta|x, R)d\alpha d\beta \tag{16}$$

The integrals involved in equation (16) can not be simplified in any standard form. So we use MCMC numerical technique to obtain the estimates. We have used Metropolis-Hastings within Gibbs sampling. The full conditional posterior distribution of parameters α and β are

$$\pi_1(\alpha|x, R) = \frac{\lambda_1^{\nu_1}}{\Gamma\nu_1}\frac{\lambda_2^{\nu_2}}{\Gamma\nu_2} e^{-(\lambda_1\alpha)}\alpha^{(J+\nu_1-1)}\prod_{j=1}^{J}(1-e^{-\beta x_j})^{(\alpha-1)}[1-(1-e^{-\beta x_j})^{\alpha}]^{R_j}W(\alpha, \beta) \tag{17}$$

$$\pi_2(\beta|x, R) = \frac{\lambda_1^{\nu_1}}{\Gamma\nu_1}\frac{\lambda_2^{\nu_2}}{\Gamma\nu_2} e^{-(\lambda_2\beta)}\beta^{(J+\nu_2-1)}\prod_{j=1}^{J}(1-e^{-\beta x_j})^{(\alpha-1)}[1-(1-e^{-\beta x_j})^{\alpha}]^{R_j}e^{-\beta x_j}W(\alpha, \beta) \tag{18}$$

The algorithm consists of following steps:

1. Set i=1 and initial guesses of α and β say the are α_0 and β_0.
2. Using Metropolis algorithm, Generate (α_i, β_i) from $\pi(\alpha, \beta, x, R)$.
3. Repeat steps 2-4, N number of times and obtain (α_1, β_1),$(\alpha_2, \beta_2)\cdots$ (α_N, β_N).
4. Obtain the Bayes estimates of α and β under SELF as $[E(\alpha|data)] = \left[\dfrac{1}{N-N_0}\sum_{i=N_0+1}^{N}\alpha_i\right]$ and $[E(\beta|data)] = \left[\dfrac{1}{N-N_0}\sum_{i=N_0+1}^{N}\beta_i\right]$. Where N_0 is the burn in period.
5. The HPD credible intervals for α and β can be obtained by using algorithm given by [16]. Let the ordered MCMC sample be $(\alpha_{[1]}, \alpha_{[2]}, \cdots, \alpha_{[N]})$ and $(\beta_{[1]}, \beta_{[2]}, \cdots, \beta_{[N]})$. Futher $100(1-\zeta)\%$ credible intervals of the parameters are constructed say $(\alpha_{\lfloor 1\rfloor}, \alpha_{\lfloor N(1-\zeta)\rfloor}), \cdots, (\alpha_{\lfloor N\zeta\rfloor}, \alpha_{\lfloor N\rfloor})$ and $(\beta_{\lfloor 1\rfloor}, \beta_{\lfloor N(1-\zeta)\rfloor}), \cdots, (\beta_{\lfloor N\zeta\rfloor}, \beta_{\lfloor N\rfloor})$. Where$\lfloor x\rfloor$ denotes the greatest integer less than or equal to x. The HPD credible interval of α and β is that interval which has shortest length.

4. Bayesian Prediction

4.1. One Sample Prediction

The Bayes prediction of an unknown observation which belongs to the future sample based on the current information available to us, would be a great tool to have an idea about lifetime of unobserved data. The one sample facilitates us to predict the lifetime of future ordered observation which may not be observed due to censoring. In the current section, we have derived the predictive posteriors for the future observations from EED using the informative sample that has been observed under GPH censoring scheme.

Let $x_{(1)}, x_{(2)}, \cdots, x_{(i)}$ be the ordered observed sample and $y_{(1:r_1)}, y_{(2:r_2)}, \cdots, y_{(s:r_i)}$ be future ordered sample from same parent population where $s = 1, 2, \cdots, r_i$ and $i = 1, 2, \cdots, J$ ($J = k$ for case-II, $J = D$ for case-I, and $J = m$ for case-III) . From [17], the conditional PDF of $y_{(s:r_i)}$ given $x_{(i)}$ is obtained as

$$f(y_{(s:r_i)}|x_{(i)}) = \begin{cases} \frac{(n-i)!}{(s-1)!(n-i-s)!} \frac{[1-F(y_{(s:r_i)})]^{n-i-s}}{[1-F(x_{(i)})]^{n-i}} [F(y_{(s:r_i)}) - F(x_{(i)})]^{s-1} f(y_{(s:r_i)}); \\ \text{for case-I and case-III} \\ \frac{(n-i)!}{(s-1)!(n-i-s)!} \frac{[1-F(y_{(s:r_i)})]^{n-i-s}}{[1-F(T)]^{n-i}} [F(y_{(s:r_i)}) - F(T)]^{s-1} f(y_{(s:r_i)}); \\ \text{for case-II.} \end{cases} \tag{19}$$

After putting the pdf and cdf from equations (1) and (2), we get

$$f(y_{(s:r_i)}|x_{(i)}) = \begin{cases} \frac{(n-i)!}{(s-1)!(n-i-s)!} \frac{[1-(1-exp(-\beta y_{(s:r_i)}))^{\alpha}]^{n-i-s}}{[1-(1-exp(-\beta x_{(i)}))^{\alpha}]^{n-i}} \\ [(1-exp(-\beta y_{(s:r_i)}))^{\alpha} - (1-exp(-\beta x_{(i)}))^{\alpha}]^{s-1} \\ \alpha\beta exp(-\beta y_{(s:r_i)})(1-exp(-\beta y_{(s:r_i)}))^{\alpha-1}; \text{for case I and case-III} \\ \frac{(n-i)!}{(s-1)!(n-i-s)!} \frac{[1-(1-exp(-\beta y_{(s:r_i)}))^{\alpha}]^{n-i-s}}{[1-(1-exp(-\beta T))^{\alpha}]^{n-i}} \\ [(1-exp(-\beta y_{(s:r_i)}))^{\alpha} - (1-exp(-\beta T))^{\alpha}]^{s-1} \\ \alpha\beta exp(-\beta y_{(s:r_i)})(1-exp(-\beta y_{(s:r_i)}))^{\alpha-1}; \text{for case-II.} \end{cases} \tag{20}$$

Then, the predictive posterior density of future observation under GPH censoring scheme can be obtained as

$$f_1(y_{(s:r_i)}|\tilde{x}) = \int_0^{\infty}\int_0^{\infty} f(y_{(s:r_i)}|\alpha,\beta,\tilde{x})\pi(\alpha,\beta|\tilde{x})d\alpha d\beta. \tag{21}$$

Since the integrals involved in the expression can not be simplified to closed form suggest to use the numerical methods.

The MCMC sample $\{(\alpha_i, \beta_i), i = 1, 2, \cdots, M\}$ obtained from $\pi(\alpha, \beta|\tilde{x})$ using Gibbs algorithm can be utilized to obtain consistent estimate of $f_1(y_{(s:r_i)}|\tilde{x})$ as

$$f_1^*(y_{(s:r_i)}|\tilde{x}) = \frac{1}{M-M_0}\sum_{i=1}^{M-M_0} f(y_{(s:r_i)}|\alpha_i,\beta_i,\tilde{x}) \tag{22}$$

Where M_0 denotes the burn-in-period of Markov Chain. Furthermore, the Survival Function of the future sample can be obtained as

$$\begin{aligned} S_{y_{(s:r_i)}}(T) &= 1 - \int_{y_{(s:r_i)}=x_{(J)}}^{T} f_1(y_{(s:r_i)}|\tilde{x})dy_{(s:r_i)} \\ &= 1 - \int_{y_{(s)}=x_{(J)}}^{T}\int_0^{\infty}\int_0^{\infty} f(y_{(s:r_i)}|\alpha,\beta,\tilde{x})\pi(\alpha,\beta|\tilde{x})d\alpha d\beta dy_{(s:r_i)} \end{aligned} \tag{23}$$

Moreover, the two sided $100(1-\delta)\%$ prediction intervals $(L_{s:r_i}, U_{s:r_i})$ for $y_{(s:r_i)}$ can be obtained by solving the following two equations $P(y_{(s:r_i)} > U_{s:r_i}|\tilde{x}) = \frac{\delta}{2}$ and $P(y_{(s:r_i)} > L_{s:r_i}|\tilde{x}) = 1 - \frac{\delta}{2}$. The confidence interval can be obtained by any iterative method as the above equations can not be solved analytically.

4.2. Two Sample Prediction

There may occur a situation, where the distribution of k^{th} order statistics is independent of the informative sample i.e. $f(y_{(k)}|\alpha,\beta,\tilde{x})$ is same as $f(y_{(k)}|\alpha,\beta)$. This is the case of two sample prediction problem. The experimenter is interested in k^{th} failure time of a future sample of size N following the same lifetime distribution. The PDF of k^{th} order statistics is given by: (see [17])

$$p(y_{(k)}|\alpha,\beta) = \frac{N!}{(k-1)!(N-k)!}[F(y_{(k)})]^{k-1}[1-F(y_{(k)})]^{N-k}f(y_{(k)}). \tag{24}$$

after putting the pdf and cdf from equations (1) and (2), we get

$$\begin{aligned} p(y_{(k)}|\alpha,\beta) =& \frac{N!}{(k-1)!(N-k)!}(1-exp(-\beta y_{(k)}))^{\alpha(k-1)}[1-(1-exp(-\beta y_{(k)}))^{\alpha}]^{N-k} \\ & \alpha\beta exp(-\beta y_{(k)})(1-exp(-\beta y_{(k)}))^{\alpha-1} \end{aligned} \tag{25}$$

The predictive posterior density of future observations under GPH censoring scheme is given by

$$p_1(y_{(k)}|\tilde{x}) = \int_0^\infty \int_0^\infty p(y_{(k)}|\alpha,\beta,\tilde{x})\pi(\alpha,\beta|\tilde{x})d\alpha d\beta. \tag{26}$$

Since the above equation can not be solved analytically. Therefore we use MCMC method along with Gibbs algorithm to obtain the consistent estimator of $p_1(y_{(k)}|\tilde{x})$ is given by

$$p_1^*(y_{(k)}|\tilde{x}) = \frac{1}{M-M_0}\sum_{i=1}^{M-M_0} p(y_{(k)}|\alpha_i,\beta_i), \tag{27}$$

where M_0 is the burn-in-period. The survival function of future sample can be defined as

$$\begin{aligned} S_{y_{(k)}}(T) =& 1-\int_{y_{(k)}=0}^{T} p_1(y_{(k)}|\tilde{x})dy_{(k)} \\ =& 1-\int_{y_{(k)}=0}^{T}\int_0^\infty\int_0^\infty p(y_{(k)}|\alpha,\beta,\tilde{x})\pi(\alpha,\beta|\tilde{x})d\alpha d\beta dy_{(k)}. \end{aligned} \tag{28}$$

The two sided $100(1-\delta)\%$ prediction interval (L_k, U_k) for $y_{(k)}$ can be obtained by solving the equations $P(y_{(k)} > U_k|\tilde{x}) = \frac{\delta}{2}$ and $P(y_{(k)} > L_k|\tilde{x}) = 1-\frac{\delta}{2}$. The confidence interval can be obtained by any iterative method as the above equations can not be solved analytically.

4.3. Stress Strength Reliability

The problem of stress strength reliability estimation is not new rather it has a long history. The term stress strength was first introduced by [18]. Since then, a lot of works have been done in this direction including parametric and non-parametric in nature. One may navigate through some recent works as [19] and [20]. This section discusses the inferential procedure of stress strength reliability $R = P(X < Y)$, when X and Y are independent and following EED with parameters (α_1,β_1) and (α_2,β_2) respectively. In a reliability study, let X denotes the strength of the unit and Y denotes the magnitude of stress applied to the unit by operating environment. A unit will function well if its strength is greater than the stress imposed on it. The stress-strength reliability R of the system is defined as

$$\begin{aligned} R =& Pr[X<Y] = \int_0^\infty \int_y^\infty f_X(x)f_Y(y)dxdy \\ =& \int_0^\infty f_Y(y)(1-F_X(y))dy \\ =& \int_0^\infty \alpha_2\beta_2 e^{-\beta_2 y}(1-e^{-\beta_2 y})^{\alpha_2-1}[1-(1-e^{-\beta_1 y})^{\alpha_1}]dy \\ =& 1-\alpha_2\sum_{i=0}^{\infty}\binom{\alpha_1}{i}(-1)^i B\left(\alpha_2, \frac{i\beta_1}{\beta_2}+1\right) \end{aligned} \tag{29}$$

In a special case, when $\beta_1 = \beta_2$, this reduced to $R = \dfrac{\alpha_1}{\alpha_1 + \alpha_2}$. Further, one can compute the estimate of R by using the method discussed in section 3.

5. Real Data Illustration

This section deals with real life applicability of proposed methodology. The data set comprises survival times of two groups of patients suffering from head neck cancer disease which was reported by [21]. First group of patients are treated with radiotherapy whereas second group of patients are treated with both radiotherapy and chemotherapy. The data set is as follows:
Data-1(X): *6.53, 7, 10.42, 14.48, 16.10, 22.70, 34, 41.55, 42, 45.28, 49.40, 53.62, 63, 64, 83, 84, 91, 108, 112, 129, 133,133, 139, 140, 140, 146, 149, 154, 157, 160, 160, 165, 146, 149, 154, 157, 160, 160, 165, 173, 176, 218, 225, 241, 248, 273, 277, 297, 405, 417, 420, 440, 523, 583, 594, 1101, 1146, 1417.*
Data-2(Y): *12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81, 43, 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, 1776.*
Before advancing further, we first check the validity of EED for the above data sets. The summary of data fit is quoted here:

Data Set	p-value	KS-distance	LogL	AIC	BIC	$\hat{\alpha}_{ML}$	SE($\hat{\alpha}_{ML}$)	$\hat{\beta}_{ML}$	SE($\hat{\beta}_{ML}$)
RT	0.0646	0.1720	-372.3767	748.7535	752.8743	1.0636	0.1851	0.0046	0.0007
RT+CT	0.2505	0.1498	-281.9551	567.9101	571.4785	1.0730	0.2178	0.0047	0.0009

So, it is clear that EED fits to the above two data sets. For illustrating the proposed methodology, we have generated censored data for a prefixed m, k, T and number of removals. We have considered different removal patterns by fixing values of $R_1, R_2, \cdots, R_m$ for a set of values of m, k and T. The schemes those have been considered are as follows:

- ${S_{m:n}}^{(1)}$: All the removals are at the last failure, i.e. $R_m = n - m$.
- ${S_{m:n}}^{(2)}$: All the removals are at the first failure, i.e. $R_1 = n - m$.
- ${S_{m:n}}^{(3)}$: The removals are at the first and last failure, i.e. $R_1 = R_m = (n - m)/2$.
- ${S_{m:n}}^{(4)}$: The removals are at middle failure, i.e. $R_{m/2} = R_{m/2+1} = (n - m)/2$.

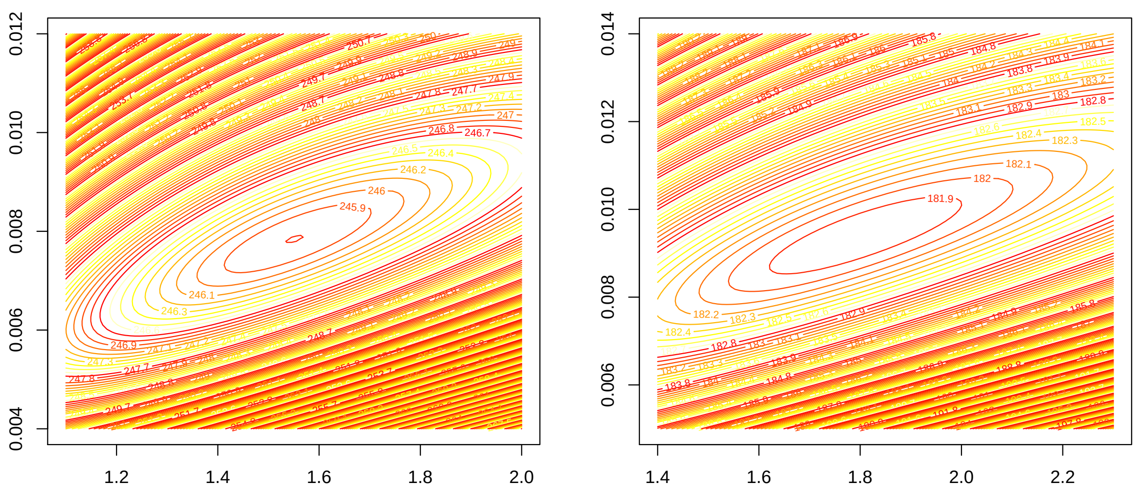


Figure 2: *Contour plot for parameters α(Left) and β(Right) based on generated censored dataset-2.4.*

Thus, generated censored datasets are given in Table 1. As mentioned earlier, the likelihood, product spacing equations and posterior integral do not have explicit solutions, therefore, numerical approach coupled with R software is maneuvered. Basically $optim(\cdot)$ function is used here to find the ML and MPS estimates of the parameters. We have used contour plot which is shown in figure 2 to provide initial guess to the $optim(\cdot)$ function. For more details reader may see [22]. Using the concept of large sample theory, the asymptotic confidence interval for α and β are also computed and variance of the estimates evaluated by inverse of estimated Fisher Information matrix. The point and asymptotic confidence interval estimates, thus obtained, are summarized in table 2.

Table 1: *The censored datasets generated from real datasets by considering different choices of k, m, T and removal patterns*

k	m	T	Scheme	Generated data-points	Data Name
				Data-1(RT)	
25	40	600	$S^{(4)}$	(6.53, 7, 10.42, 14.48, 16.1, 22.7, 34, 41.55, 42, 45.28, 49.4, 53.62, 63, 64, 83, 84, 91, 108, 112, 129, 133, 140, 140, 146, 154, 160, 160, 160, 165, 218, 225, 241, 248, 273, 417, 523, 594)	1.1
25	40	600	$S^{(3)}$	(6.53, 7, 10.42, 14.48, 16.1, 22.7, 41.55, 42, 45.28, 53.62, 63, 64, 84, 91, 108, 112, 129, 133, 133, 140, 146, 146, 149, 154, 154, 157, 157, 160, 160, 160, 160, 165, 165, 173, 176, 218, 225, 241, 248, 273)	1.2
25	40	600	$S^{(2)}$	(6.53, 10.42, 14.48, 16.1, 22.7, 41.55, 42, 45.28, 49.4, 63, 64, 84, 91, 112, 133, 139, 140, 146, 146, 149, 157, 157, 160, 160, 160, 165, 173, 218, 225, 241, 248, 273, 277, 297, 405, 420, 583, 594)	1.3
25	40	600	$S^{(1)}$	(6.53, 7, 10.42, 14.48, 16.1, 22.7, 34, 41.55, 42, 45.28, 49.4, 53.62, 63, 64, 83, 84, 91, 108, 112, 129, 133, 133, 139, 140, 140, 146, 146, 149, 149, 154, 154, 157, 157, 160, 160, 160, 160, 165, 165, 173)	1.4
				Data-2(RT+CT)	
20	30	600	$S^{(4)}$	(12.2, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 43, 47.38, 55.46, 58.36, 63.47, 68.46, 74.47, 78.26, 84, 92, 119, 127, 133, 140, 146, 173, 179, 194, 195, 209, 281, 339, 519)	2.1
20	30	600	$S^{(3)}$	(12.2, 23.56, 23.74, 25.87, 31.98, 37, 43, 47.38, 55.46, 58.36, 63.47, 68.46, 74.47, 81, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 159, 179, 194, 195, 209, 249)	2.2
20	30	600	$S^{(2)}$	(12.2, 23.56, 23.74, 25.87, 37, 43, 63.47, 68.46, 74.47, 78.26, 81, 84, 92, 110, 112, 119, 127, 133, 146, 155, 159, 173, 179, 194, 209, 249, 319, 469, 519)	2.3
20	30	600	$S^{(1)}$	(12.2, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 43, 47.38, 55.46, 58.36, 63.47, 68.46, 74.47, 78.26, 81, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173)	2.4

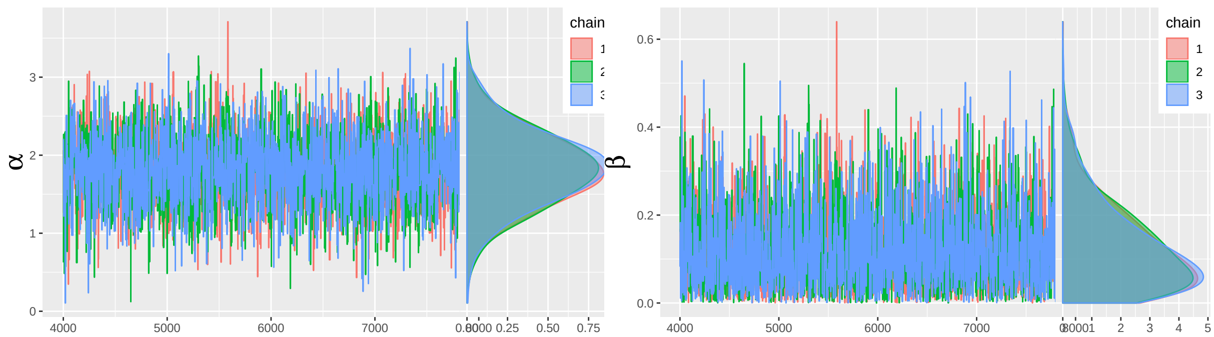


Figure 3: *Iteration and density plot of MCMC sample for parameters α (Left) and β (Right) for dataset-2.4.*

To compute Bayes estimates for considered dataset, we have used MCMC technique discussed in Section 3.2. Following [23], we ran three MCMC chains with initial values selected as MLE, MLE - (asymptotic standard deviation) and MLE + (asymptotic standard deviation), respectively.

Table 2: *ML, MPS, Bayes and Bootstrap estimates with their respective 95% CI of parameters (within bracket) for generated censored datasets*

Datasets	$\hat{\alpha}_{ML}$	$\hat{\alpha}_{MP}$	$\hat{\alpha}_{B}$	$\hat{\alpha}_{Boot}$
1.1	0.9116(0.5719,1.2443)	0.8971(0.5889,1.2346)	0.8955(0.5912,1.2341)	0.9003(0.5892,1.2800)
1.2	1.3696(0.8197,1.9074)	1.2563(0.8332,1.9063)	1.2674(0.8572,1.9057)	1.3119(0.8329,2.0743)
1.3	1.0873(0.6322,1.5072)	1.0305(0.6688,1.5061)	1.2249(0.6933,1.4851)	1.0344(0.6688,1.5021)
1.4	1.5528(0.8724,2.1812)	1.4249(0.9307,2.1751)	1.4378(0.9512,2.1749)	1.4689(0.9355,2.2880)
2.1	1.6421(0.8167,2.4140)	1.5305(0.8729,2.4106)	1.4925(0.8728,2.1817)	1.5491(0.7875,2.4105)
2.2	2.3914(1.1093,3.5832)	2.2012(1.2013,3.5812)	2.4258(1.2016,3.2355)	2.1641(1.1022,3.5816)
2.3	1.3544(0.6903,1.9820)	1.3471(0.7367,1.9713)	1.4143(0.7366,1.8731)	1.3384(0.6989,1.9716)
2.4	1.8013(0.8838,2.7044)	1.6441(0.9086,2.6934)	1.7765(0.9084,2.6250)	1.9375(0.9082,2.6942)
Datasets	$\hat{\beta}_{ML}$	$\hat{\beta}_{MP}$	$\hat{\beta}_{B}$	$\hat{\beta}_{Boot}$
1.1	0.0032(0.0001,0.0123)	0.0032(0.0010,0.0044)	0.0033(0.0013,0.0045)	0.0031(0.0001,0.0115)
1.2	0.0069(0.0034,0.0121)	0.0068(0.0040,0.0094)	0.0067(0.0036,0.0093)	0.0069(0.0033,0.0116)
1.3	0.0047(0.0017,0.0163)	0.0047(0.0017,0.0065)	0.0049(0.0021,0.0065)	0.0049(0.0018,0.0155)
1.4	0.0078(0.0042,0.0148)	0.0078(0.0048,0.0108)	0.0077(0.0041,0.0107)	0.0081(0.0038,0.0145)
2.1	0.0084(0.0047,0.0176)	0.0081(0.0046,0.0120)	0.0091(0.0051,0.0120)	0.0082(0.0045,0.0169)
2.2	0.0124(0.0072,0.0195)	0.0116(0.0064,0.0171)	0.0114(0.0067,0.0171)	0.0115(0.0056,0.0181)
2.3	0.0065(0.0031,0.0134)	0.0065(0.0031,0.0093)	0.0068(0.0032,0.0093)	0.0071(0.0021,0.0126)
2.4	0.0094(0.0040,0.0210)	0.0093(0.0048,0.0134)	0.0101(0.0043,0.0135)	0.0095(0.0042,0.0219)

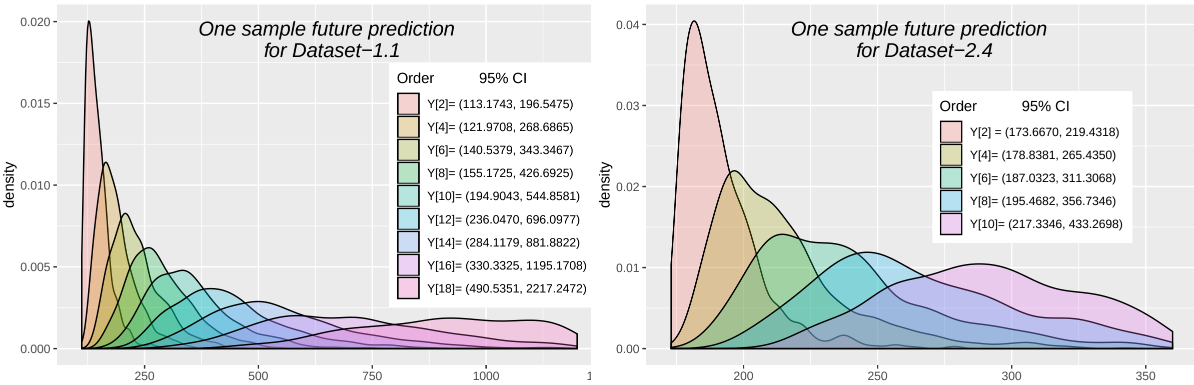


Figure 4: *Density plot of one sample predicted order statistics with their respective 95% confidence interval.*

Figure 3 shows the iterations and density plot of samples generated from the posterior distribution using the MCMC technique. From this figure, we see that all the three chains have converged and are well mixed. It is further noted that the posterior of α is approximately symmetric, but the posterior of β is left skewed. Utilizing these MCMC samples, we computed the Bayes estimates, following the method discussed in Section 3.2. The ML, MPS, Bayes and bootstrap estimates of α are denoted by $\hat{\alpha}_{ML}$, $\hat{\alpha}_{MP}$, $\hat{\alpha}_{B}$ and $\hat{\alpha}_{Boot}$ respectively. Similarly, the ML, MPS, Bayes and bootstrap estimates of β are denoted by $\hat{\beta}_{ML}$, $\hat{\beta}_{MP}$, $\hat{\beta}_{B}$ and $\hat{\beta}_{Boot}$ respectively. The point and HPD interval estimates, thus obtained, are summarized in table 2. In the table 3, we have provided the ML, MPS, Bayes and bootstrap estimates of stress-strength reliability for various combination of censored real dataset.

One and two sample predictive densities along with prediction interval for the future observations are presented in figures 4 and 5, respectively. From figure 4, it is observed that the proposed predictive interval for ordered observations contain the observed sample observations and it verifies the applicability of the prediction techniques for real problems.

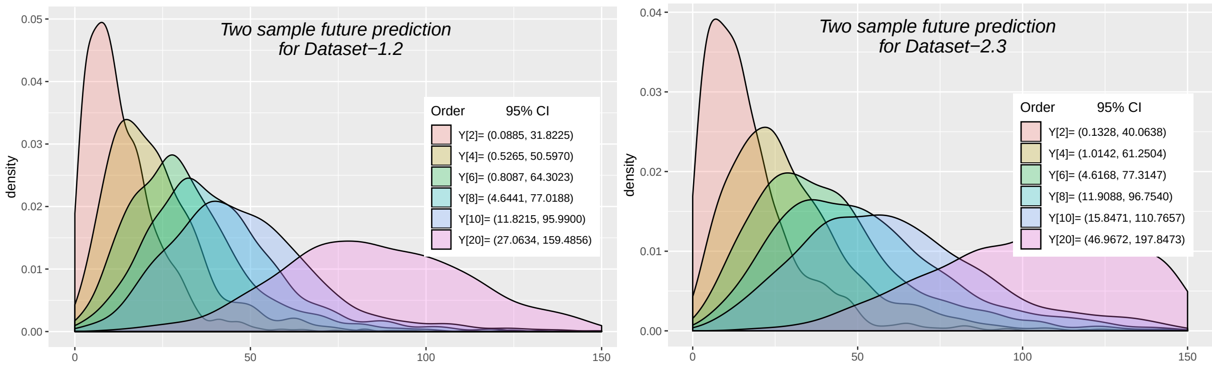


Figure 5: *Density plot of two sample predicted order statistics with their respective 95% confidence interval.*

Table 3: *ML, MPS, Bayes and Bootstrap estimates of reliability i.e. $P[X < Y]$ and their respective 95% CI (within bracket) for generated censored datasets*

X-data	Y-data	MLE	MPS	Bayes	Bootstrap
1.1	2.1	0.3878(0.1787,0.5288)	0.3654(0.1924,0.4988)	0.4060(0.1824,0.5277)	0.3813(0.1705,0.5355)
	2.2	0.3526(0.1606,0.5766)	0.3301(0.1708,0.5361)	0.3328(0.1736,0.5652)	0.3377(0.1530,0.5715)
	2.3	0.4149(0.2216,0.5831)	0.4046(0.2226,0.5489)	0.4084(0.2640,0.5470)	0.4071(0.2203,0.5587)
	2.4	0.3765(0.2414,0.5330)	0.3479(0.2480,0.5074)	0.3707(0.2725,0.5292)	0.3613(0.2388,0.5508)
1.2	2.1	0.4814(0.2659,0.6414)	0.4356(0.2925,0.6119)	0.4864(0.2869,0.5833)	0.5050(0.2615,0.6333)
	2.2	0.4389(0.2341,0.6136)	0.3978(0.2566,0.6091)	0.4169(0.2526,0.5616)	0.4512(0.2279,0.6179)
	2.3	0.5116(0.3238,0.6443)	0.4713(0.3539,0.6079)	0.4821(0.3366,0.6337)	0.5153(0.3193,0.6693)
	2.4	0.4683(0.3082,0.6499)	0.4289(0.3259,0.6055)	0.4426(0.3256,0.5865)	0.4846(0.3013,0.6371)
1.3	2.1	0.4359(0.2384,0.5971)	0.4134(0.2535,0.5661)	0.4201(0.2575,0.5614)	0.4424(0.2502,0.5871)
	2.2	0.3963(0.1921,0.6534)	0.3776(0.2043,0.6376)	0.4115(0.2257,0.6047)	0.3915(0.1931,0.6517)
	2.3	0.4646(0.2676,0.6356)	0.4570(0.2764,0.5951)	0.5066(0.3044,0.6284)	0.4526(0.2723,0.6223)
	2.4	0.4234(0.2299,0.6457)	0.3895(0.2320,0.5831)	0.4316(0.2562,0.6262)	0.4115(0.2274,0.6610)
1.4	2.1	0.4913(0.2661,0.6446)	0.4572(0.2873,0.5917)	0.4515(0.2743,0.6207)	0.4774(0.2763,0.6586)
	2.2	0.4477(0.2258,0.6491)	0.4205(0.2358,0.6083)	0.4763(0.2534,0.6061)	0.4638(0.2280,0.6809)
	2.3	0.5221(0.3798,0.6829)	0.4964(0.3917,0.6572)	0.4962(0.4234,0.6533)	0.4970(0.3979,0.6771)
	2.4	0.4778(0.2264,0.7030)	0.4654(0.2479,0.6595)	0.4398(0.2321,0.7011)	0.4642(0.2253,0.7258)

6. Simulation Study

A simulation study is conducted here to study the performance of the estimates of the parameters α and β on the basis of MSEs under considered censoring scheme. It is to be mentioned here that the exact expressions for the MSEs cannot be obtained because the estimators are not found in explicit form. Therefore, the MSEs of the estimators are estimated on the basis of simulation study of $5,000$ samples. It may be noted here that the MSEs of the estimators will depend on values of n, k, m, T, α and β, and hence various choices have been made to study the effect thereof. To generate a GPH censored sample from the considered distribution, see [24].

Here, we considered a number of values for sample size n. For an informative prior, the hyper-parameters are chosen on the basis of information possessed by the experimenter. In most cases, the experimenter can have a notion of what are the expected value of the parameter and can always associate a degree of belief to this value. In other words, the experimenter can specify the prior mean and prior variance for the parameters. The prior mean reflects the experimenter's belief about the parameter in the form of its expected value, and the prior variance reflects his confidence in this expected value. Keeping this point in mind, we have chosen hyper-parameters in such a way that the prior mean is equal to the true value of the parameter, and the belief in the prior mean is either strong or weak, i.e. the prior variance is small or large, respectively; for details see [25]. The average ML, MPS, Bayes and bootstrap estimates of parameter with corresponding MSEs are given in table 4 and table 5. The bias and MSEs of ML, MPS, Bayes and bootstrap estimates of reliability function are given in table 6.

Table 4: *Average estimate and MSE (within bracket) of different estimators of parameters for varying n and m*

n	m	Scheme	$\hat{\alpha}_{ML}$	$\hat{\alpha}_{MP}$	$\hat{\alpha}_{B}$	$\hat{\beta}_{ML}$	$\hat{\beta}_{MP}$	$\hat{\beta}_{B}$
90	80	$S^{(1)}$	2.0683(0.1029)	2.0159(0.1025)	2.0686(0.0930)	2.0430(0.0583)	1.9429(0.0577)	2.0516(0.0522)
		$S^{(2)}$	2.0659(0.1029)	1.9642(0.1026)	2.0753(0.0928)	2.0514(0.0581)	1.9653(0.0575)	2.0457(0.0521)
		$S^{(3)}$	2.0615(0.1013)	2.0304(0.1008)	2.0572(0.0911)	2.0229(0.0562)	1.9410(0.0562)	2.0213(0.0506)
		$S^{(4)}$	2.0572(0.1147)	2.0421(0.1132)	2.0529(0.1027)	2.0499(0.0586)	1.9735(0.0582)	2.0524(0.0527)
90	60	$S^{(1)}$	2.0716(0.1157)	2.0028(0.1146)	2.0665(0.1041)	2.0496(0.0599)	2.0306(0.0591)	2.0474(0.0536)
		$S^{(2)}$	2.0533(0.1136)	2.0070(0.1139)	2.0597(0.1025)	2.0423(0.0578)	1.9675(0.0580)	2.0427(0.0522)
		$S^{(3)}$	2.0493(0.1116)	1.9896(0.1120)	2.0566(0.1007)	2.0215(0.0579)	1.9366(0.0577)	2.0391(0.0520)
		$S^{(4)}$	2.0437(0.1205)	1.9697(0.1197)	2.0608(0.1081)	2.0347(0.0620)	1.9715(0.0616)	2.0378(0.0557)
90	30	$S^{(1)}$	2.0427(0.1213)	2.0220(0.1212)	2.0542(0.1094)	2.0419(0.0633)	2.0297(0.0633)	2.0290(0.0571)
		$S^{(2)}$	2.0526(0.1190)	2.0442(0.1184)	2.0556(0.1066)	2.0381(0.0620)	1.9808(0.0615)	2.0414(0.0556)
		$S^{(3)}$	2.0597(0.1253)	1.9608(0.1238)	2.0517(0.1124)	2.0275(0.0592)	1.9911(0.0592)	2.0227(0.0534)
		$S^{(4)}$	2.0652(0.1288)	2.0557(0.1279)	2.0645(0.1159)	2.0332(0.0621)	1.9379(0.0619)	2.0346(0.0561)
70	60	$S^{(1)}$	2.0788(0.1344)	2.0575(0.1334)	2.0653(0.1205)	2.0415(0.0862)	1.9671(0.0861)	2.0381(0.0779)
		$S^{(2)}$	2.0941(0.1319)	2.0183(0.1311)	2.0906(0.1187)	2.0525(0.0852)	1.9892(0.0846)	2.0499(0.0766)
		$S^{(3)}$	2.0459(0.1311)	1.9820(0.1299)	2.0457(0.1176)	2.0458(0.0855)	1.9620(0.0848)	2.0502(0.0768)
		$S^{(4)}$	2.0779(0.1334)	2.0621(0.1332)	2.0878(0.1202)	2.0640(0.0856)	1.9870(0.0852)	2.0706(0.0773)
70	36	$S^{(1)}$	2.0863(0.1379)	2.0374(0.1373)	2.0739(0.1238)	2.0470(0.0885)	1.9673(0.0886)	2.0562(0.0798)
		$S^{(2)}$	2.0689(0.1368)	2.0609(0.1361)	2.0622(0.1233)	2.0355(0.0874)	1.9803(0.0870)	2.0365(0.0790)
		$S^{(3)}$	2.0585(0.1399)	1.9804(0.1399)	2.0544(0.1260)	2.0425(0.0868)	1.9478(0.0870)	2.0445(0.0783)
		$S^{(4)}$	2.0668(0.1344)	2.0540(0.1332)	2.0738(0.1205)	2.0424(0.0864)	2.0032(0.0863)	2.0434(0.0777)
70	26	$S^{(1)}$	2.0771(0.1433)	1.9734(0.1414)	2.0694(0.1284)	2.0490(0.0895)	1.9615(0.0889)	2.0585(0.0806)
		$S^{(2)}$	2.0553(0.1459)	2.0013(0.1454)	2.0697(0.1308)	2.0335(0.0889)	1.9611(0.0885)	2.0396(0.0803)
		$S^{(3)}$	2.0675(0.1478)	2.0608(0.1462)	2.0672(0.1324)	2.0371(0.0880)	1.9804(0.0875)	2.0302(0.0792)
		$S^{(4)}$	2.0822(0.1544)	2.0206(0.1539)	2.0721(0.1391)	2.0393(0.0901)	1.9362(0.0892)	2.0475(0.0809)
60	50	$S^{(1)}$	2.1001(0.1649)	2.0472(0.1631)	2.0961(0.1479)	2.0675(0.0938)	2.0571(0.0941)	2.0819(0.0846)
		$S^{(2)}$	2.0895(0.1657)	2.0075(0.1641)	2.0825(0.1488)	2.0349(0.0946)	1.9453(0.0941)	2.0524(0.0847)
		$S^{(3)}$	2.1037(0.1668)	2.0851(0.1670)	2.1047(0.1505)	2.0471(0.0941)	1.9920(0.0936)	2.0618(0.0845)
		$S^{(4)}$	2.1348(0.1621)	2.0864(0.1602)	2.1254(0.1452)	2.0286(0.0944)	1.9679(0.0939)	2.0310(0.0846)
60	30	$S^{(1)}$	2.1312(0.1714)	2.1120(0.1699)	2.1313(0.1542)	2.0852(0.0970)	2.0561(0.0963)	2.0740(0.0868)
		$S^{(2)}$	2.1003(0.1721)	2.0910(0.1718)	2.0911(0.1550)	2.0643(0.1087)	1.9843(0.1084)	2.0733(0.0976)
		$S^{(3)}$	2.0848(0.1877)	1.9918(0.1862)	2.0852(0.1691)	2.0417(0.0969)	2.0095(0.0963)	2.0417(0.0869)
		$S^{(4)}$	2.0903(0.1707)	2.0684(0.1693)	2.0889(0.1531)	2.0698(0.0964)	2.0126(0.0965)	2.0588(0.0868)
60	20	$S^{(1)}$	2.0772(0.1917)	1.9855(0.1908)	2.0894(0.1728)	2.0512(0.1004)	1.9888(0.1001)	2.0654(0.0907)
		$S^{(2)}$	2.0738(0.1928)	2.0258(0.1910)	2.0671(0.1735)	2.0656(0.1136)	2.0601(0.1128)	2.0578(0.1018)
		$S^{(3)}$	2.0822(0.1991)	2.0537(0.1981)	2.0764(0.1792)	2.0573(0.1026)	1.9807(0.1023)	2.0581(0.0927)
		$S^{(4)}$	2.0578(0.1989)	1.9652(0.1975)	2.0492(0.1783)	2.0851(0.1135)	2.0743(0.1128)	2.0853(0.1019)

In the table 4, we have computed average ML, MPS and Bayes estimates of the considered parameters and their corresponding MSEs for different choices of n, m and censoring schemes with fix values of parameters $\alpha = 0.5$ and $\beta = 0.5$ and $T = 10$. Here, three choices of n i.e. n=60(small), 70(moderate), 90(large) are considered. The value of k are set to be 50% of n respectively. From this table, we can note that in general, the MSEs decrease as n or m increases in all the considered cases. It can also be seen that the MSE of the MLE is more than that of the corresponding MPS and Bayes estimate in all cases; but the difference between the MSEs of the Bayes and ML estimates decreases for increases in the value of n. Further, MSE of the Bayes estimate is least among all the considered estimators. For small number of removals i.e. for large m, the MSEs of both the parameters is less for the removal pattern ${S_{m:n}}^{(2)}$ in comparison to ${S_{m:n}}^{(1)}$ and MSEs under ${S_{m:n}}^{(3)}$ is observed to be lesser than that for ${S_{m:n}}^{(4)}$. For large number of removals; the MSEs of both the parameters under removal pattern ${S_{m:n}}^{(1)}$ are lesser than those under ${S_{m:n}}^{(2)}$ and MSEs under ${S_{m:n}}^{(4)}$ are observed to be less than those for ${S_{m:n}}^{(3)}$ i.e. the trend shows a reversal from small number of removals.

Table 5 shows the performances of estimates(ML, MPS, Bayes and bootstrap) in terms of MSEs for varying parameters α and β under GPH censoring scheme for various choices of n and fixed

Table 5: *Average estimate and MSE (within bracket) of different estimators of parameters for varying α and β.*

CensoringScheme	α	β	$\hat{\alpha}_{ML}$	$\hat{\alpha}_{MP}$	$\hat{\alpha}_{B}$	$\hat{\beta}_{ML}$	$\hat{\beta}_{MP}$	$\hat{\beta}_{B}$
	0.5	0.5	0.5403(0.0107)	0.5348(0.0103)	0.5377(0.0095)	0.5576(0.0278)	0.5511(0.0261)	0.5542(0.0256)
		1	0.5421(0.0108)	0.5365(0.0098)	0.5370(0.0100)	1.1394(0.1274)	1.1275(0.1232)	1.1365(0.1209)
		2	0.5391(0.0110)	0.5336(0.0103)	0.5363(0.0099)	2.2996(0.5353)	2.2757(0.5192)	2.3057(0.5080)
	1	0.5	1.0861(0.0558)	1.0748(0.0532)	1.0935(0.0526)	0.5476(0.0180)	0.5416(0.0171)	0.5493(0.0171)
n=30,k = 15, m = 22, T = 10, R = (0*10,4*2,0*10)		1	1.0786(0.0572)	1.0673(0.0547)	1.0867(0.0540)	1.0979(0.0736)	1.0859(0.0705)	1.1050(0.0691)
		2	1.1054(0.0637)	1.0940(0.0618)	1.1054(0.0597)	2.1994(0.3117)	2.1769(0.3018)	2.1887(0.2957)
	2	0.5	2.2099(0.2964)	2.1874(0.2868)	2.2053(0.2812)	0.5402(0.0125)	0.5339(0.0119)	0.5384(0.0116)
		1	2.1867(0.3161)	2.1644(0.3061)	2.1925(0.2993)	1.0775(0.0509)	1.0666(0.0489)	1.0775(0.0474)
		2	2.2146(0.3599)	2.1921(0.3485)	2.2202(0.3413)	2.1547(0.2041)	2.1330(0.1976)	2.1725(0.1934)
	0.5	0.5	0.5151(0.0058)	0.5092(0.0048)	0.5151(0.0049)	0.5404(0.0214)	0.5344(0.0203)	0.5416(0.0196)
		1	0.5247(0.0067)	0.5189(0.0060)	0.5251(0.0061)	1.0995(0.0801)	1.0878(0.0774)	1.1063(0.0754)
		2	0.5267(0.0068)	0.5213(0.0061)	0.5243(0.0055)	2.1800(0.3354)	2.1573(0.3245)	2.1724(0.3184)
	1	0.5	1.0615(0.0358)	1.0504(0.0346)	1.0690(0.0337)	0.5350(0.0116)	0.5292(0.0110)	0.5326(0.0101)
n=50, k = 25, m = 36,T = 10, R =(0*17,7*2,0*17)		1	1.0605(0.0362)	1.0493(0.0342)	1.0659(0.0337)	1.0738(0.0474)	1.0629(0.0455)	1.0715(0.0450)
		2	1.0651(0.0379)	1.0539(0.0362)	1.0750(0.0357)	2.1568(0.1843)	2.1344(0.178)	2.1626(0.1743)
	2	0.5	2.1991(0.1615)	2.1763(0.1560)	2.1869(0.1529)	0.5335(0.0082)	0.5280(0.0076)	0.5341(0.0068)
		1	2.1376(0.1696)	2.1159(0.1641)	2.1168(0.1608)	1.0350(0.0269)	1.0240(0.0259)	1.0341(0.0251)
		2	2.1527(0.1754)	2.1310(0.1697)	2.1387(0.1663)	2.1274(0.1095)	2.1058(0.1061)	2.1468(0.1033)
	0.5	0.5	0.5094(0.0027)	0.5040(0.0025)	0.5081(0.0020)	0.5183(0.0081)	0.5129(0.0077)	0.5176(0.0069)
		1	0.5058(0.0028)	0.4998(0.0025)	0.508(0.0021)	1.0371(0.0337)	1.0263(0.0319)	1.0391(0.0320)
		2	0.5134(0.0028)	0.5075(0.0026)	0.5107(0.0019)	2.0918(0.1205)	2.0706(0.1164)	2.0924(0.1139)
	1	0.5	1.0293(0.0145)	1.0186(0.0137)	1.0383(0.0131)	0.5131(0.0045)	0.5075(0.0042)	0.5126(0.0038)
n=100, k = 50, m = 72, T = 10, R = (0*35,14*2,0*35)		1	1.0324(0.0151)	1.0215(0.0141)	1.0365(0.0141)	1.0438(0.0203)	1.0332(0.0197)	1.0369(0.0191)
		2	1.0242(0.0159)	1.0131(0.0153)	1.0159(0.0150)	2.0526(0.0775)	2.0316(0.0746)	2.0599(0.0735)
	2	0.5	2.0463(0.0779)	2.0255(0.0750)	2.0663(0.0738)	0.5101(0.0032)	0.5041(0.0027)	0.5111(0.0023)
		1	2.0972(0.0798)	2.0755(0.0771)	2.0777(0.0752)	1.0324(0.0143)	1.0221(0.0137)	1.0413(0.0136)
		2	2.0750(0.0808)	2.0542(0.0778)	2.0663(0.0767)	2.0466(0.0525)	2.0256(0.0507)	2.0666(0.0499)

Table 6: *Bias and MSE (within bracket) of different estimators of $R = P[X < Y]$ when $\alpha = 0.5$, $\beta = 0.5$, $T = 10$, k_1 and k_2 are half of sample size n_1 and n_2, and m_1 and m_2 are 80% of sample size n_1 and n_2 . All entries of the table are multiplied by 10^3.*

Scheme	n_1	n_2	MLE	MPS	Bootstrap	Bayes		
						Large Variance	Non-Informative	Small Variance
$S^{(2)}$	30	30	6.9337(4.8475)	6.2974(4.2886)	6.8986(4.7014)	5.8717(4.0551)	5.9223(4.1511)	5.6398(3.0416)
	30	50	6.0457(4.2175)	5.4149(3.8425)	6.0143(4.3102)	5.2663(3.7421)	5.2601(3.7564)	5.0293(2.6965)
	30	100	5.5791(4.0310)	4.9611(3.5106)	5.6192(3.9752)	4.9102(3.3407)	5.0089(3.3857)	4.6493(2.4892)
	50	50	5.7328(4.1779)	5.1995(3.6288)	5.9919(4.1521)	5.0901(3.4534)	4.9623(3.6513)	4.7896(2.5583)
	50	100	5.4295(3.6835)	4.8097(3.3036)	5.2663(3.6921)	4.6123(3.1861)	4.5537(3.2653)	4.4485(2.3702)
	100	100	4.8737(3.3677)	4.3531(3.0344)	4.7554(3.4227)	4.1374(2.8558)	4.2813(2.8962)	4.0495(2.1783)
$S^{(3)}$	30	30	6.9518(4.8499)	6.0947(4.2810)	6.7646(4.8703)	5.7705(4.1187)	5.9781(4.0731)	5.6565(2.9414)
	30	50	6.2685(4.3231)	5.6613(3.9404)	6.1842(4.2264)	5.2618(3.5897)	5.2718(3.6786)	4.9798(2.6695)
	30	100	5.6802(3.9142)	5.0137(3.6531)	5.5496(3.9760)	4.8688(3.2799)	4.8871(3.3660)	4.5926(2.5324)
	50	50	5.7935(4.1735)	5.2874(3.7631)	5.7932(4.1630)	4.9381(3.5414)	5.1437(3.5796)	4.8483(2.6166)
	50	100	5.4147(3.6682)	4.9194(3.3341)	5.3210(3.7371)	4.5276(3.2698)	4.6319(3.2168)	4.3927(2.3942)
	100	100	4.9886(3.4919)	4.4981(3.0521)	4.9136(3.4366)	4.2458(2.8857)	4.2665(2.9377)	3.9843(2.1317)
$S^{(1)}$	30	30	8.3766(5.8594)	7.3233(5.2685)	8.3944(5.8504)	6.9325(4.8481)	7.0932(5.0225)	6.5129(3.6627)
	30	50	7.3513(5.2099)	6.7839(4.6143)	7.3860(5.0897)	6.2203(4.2930)	6.4552(4.3801)	5.9409(3.2971)
	30	100	6.7621(4.8604)	6.0695(4.3071)	6.7758(4.8412)	5.7347(3.9563)	5.8965(4.0652)	5.4157(2.9415)
	50	50	6.9321(5.0217)	6.4680(4.5044)	7.0883(4.9558)	6.0779(4.1896)	6.0609(4.2478)	5.6684(3.1288)
	50	100	6.2776(4.5562)	5.9264(4.1362)	6.5143(4.5732)	5.4045(3.8858)	5.6245(3.8524)	5.1871(2.8698)
	100	100	5.8052(4.1824)	5.1675(3.6788)	5.8722(4.0758)	4.9535(3.5410)	5.0076(3.5632)	4.6606(2.5188)
$S^{(4)}$	30	30	7.6202(5.3718)	6.8561(4.8040)	7.3301(5.3458)	6.5334(4.4716)	6.5969(4.5147)	6.1735(3.3876)
	30	50	6.8518(4.7686)	6.0057(4.2925)	6.8758(4.7493)	5.6725(4.0401)	5.8644(4.1988)	5.5385(2.9678)
	30	100	6.1996(4.2686)	5.6710(3.8474)	6.2351(4.2745)	5.3898(3.7433)	5.4337(3.7927)	5.1184(2.7952)
	50	50	6.3107(4.4244)	5.6776(4.0887)	6.4848(4.4322)	5.3416(3.8225)	5.6016(3.8429)	5.3026(2.8149)
	50	100	6.0107(4.1180)	5.2660(3.6801)	5.8027(4.0877)	5.1367(3.5053)	5.1041(3.5395)	4.7381(2.5381)
	100	100	5.3390(3.6663)	4.8599(3.3559)	5.4634(3.7267)	4.6548(3.1825)	4.5542(3.1891)	4.3599(2.4226)

$T = 10$. The value of k and m are set to be 50% and 72% of n respectively. The removal pattern is taken as ${S_{m:n}}^{(4)}$ i.e. $R_{m/2}$ = $R_{m/2+1}$ = (n - m)/2. From this table, we can conclude that for fix α, as β increases, MSE of β increases. Similarly, for fix β, as α increases, the MSE of α increases. For fix α and β as n (for fixed proportions k, m and fixed T) increases, the MSEs of both the parameters decrease.

In the table 6, we have presented the bias and MSE of different estimators(ML, MPS, Bayes and bootstrap) of $R = P[X < Y]$ when $\alpha = 0.5$, $\beta = 0.5$, $T = 10$, k_1 and k_2 are half of sample size n_1 and n_2, and m_1 and m_2 are set to be 80% of sample size n_1 and n_2 respectively. From this table, it can be easily seen that as n_1 or n_2 increases the bias and MSE decrease for all the estimators and for all the considered censoring schemes. Further, MSE of the Bayes estimate is least among all the considered estimators.

7. Conclusion

The article considers the problem of estimation and prediction for exponentiated exponential distribution from a generalised progressive hybrid censored sample. It is clear from above discussions that the proposed estimation procedures under GPH censoring scheme can be easily implemented with specific choice of T and m. The MPS procedure provides more precise estimates than those obtained from maximum likelihood and bootstrap procedures. The Bayesian procedure delivers more accurate and precise estimates of the parameters even if we consider the vague prior. The HPD intervals for the parameters are also obtained and it is verified that the width of the HPD interval is smaller than asymptotic and bootstrap confidence intervals. Therefore, we may conclude that the use of HPD interval under considered situation can safely be recommended. Moreover, Bayesian prediction of unknown future observation has far flung applicability in different areas of applied statistics. The Bayesian approach using MCMC method can be effectively used to solve prediction problems. Finally, we can conclude that the discussed methodology can be extensively used in various disciplines of scientific areas where such life-tests are needed.

References

[1] Kundu, D. and Joarder, A. (2006). Analysis of type-ii progressively hybrid censored data. *Computational Statistics & Data Analysis*, 50(10):2509–2528.

[2] Childs, A., Chandrasekar, B., and Balakrishnan, N. (2008). Exact likelihood inference for an exponential parameter under progressive hybrid censoring schemes. In *Statistical models and methods for biomedical and technical systems*, pages 319–330. Springer.

[3] Cho, Y., Sun, H., and Lee, K. (2015). Exact likelihood inference for an exponential parameter under generalized progressive hybrid censoring scheme. *Statistical Methodology*, 23:18–34.

[4] Gupta, R. D. and Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

[5] Gupta, R. D. and Kundu, D. (2001a). Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical journal*, 43(1):117–130.

[6] Gupta, R. D. and Kundu, D. (2001b). Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, 69(4):315–337.

[7] Singh, S. K. (2011). Estimation of parameters and reliability function of exponentiated exponential distribution: Bayesian approach under general entropy loss function. *Pakistan journal of statistics and operation research*, 7(2).

[8] Al-Hussaini, E. K. (1999). Predicting observables from a general class of distributions. *Journal of Statistical Planning and Inference*, 79(1):79–91.

[9] Al-Hussaini, E. K. (2001). On bayes prediction of future median. *Communications in Statistics-Theory and Methods*, 30(7):1395–1410.

[10] Pradhan, B. and Kundu, D. (2011). Bayes estimation and prediction of the two-parameter gamma distribution. *Journal of Statistical Computation and Simulation*, 81(9):1187–1198.

[11] Kundu, D. and Raqab, M. Z. (2012). Bayesian inference and prediction of order statistics for a type-ii censored weibull distribution. *Journal of statistical planning and inference*, 142(1):41–47.

[12] Kundu, D. and Howlader, H. (2010). Bayesian inference and prediction of the inverse weibull distribution for type-ii censored data. *Computational Statistics & Data Analysis*, 54(6):1547–1558.

[13] Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*, volume 38. Siam.

[14] Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *Annals of Statistics*, 16:927–953.

[15] Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

[16] Chen, M.-H. and Shao, Q.-M. (1999). Monte carlo estimation of bayesian credible and hpd intervals. *Journal of Computational and Graphical Statistics*, 8(1):69–92.

[17] David, H. A. and Nagaraja, H. (2003). Order statistics: Wiley series in probability and statistics.

[18] Church, J. D. and Harris, B. (1970). The estimation of reliability from stress-strength relationships. *Technometrics*, 12(1):49–54.

[19] Sharma, V. K., Singh, S. K., Singh, U., and Agiwal, V. (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.

[20] Al-Mutairi, D. K., Ghitany, M. E., and Kundu, D. (2013). Inferences on stress-strength reliability from lindley distributions. *Communications in Statistics-Theory and Methods*, 42(8):1443–1463.

[21] Efron, B. (1988). Logistic regression, survival analysis, and the kaplan-meier curve. *Journal of the American statistical Association*, 83(402):414–425.

[22] Kaushik, A., Pandey, A., Singh, U., and Singh, S. K. (2017). Bayesian estimation of the parameters of exponentiated exponential distribution under progressive interval type-i censoring scheme with binomial removals. *Austrian Journal of Statistics*, 46(2):43–47.

[23] Robert, C. P. (2015). The metropolis-hastings algorithm. *arXiv:1504.01896v1 [stat.CO]*.

[24] Pandey, A., Kaushik, A., Singh, S. K., and Singh, U. (2021). On the estimation problems for exponentiated exponential distribution under generalized progressive hybrid censoring: On the generalised progressive hybrid censoring. *Austrian Journal of Statistics*, 50(1):24–40.

[25] Singh, S. K., Singh, U., and Kumar, D. (2011). Estimation of parameters and reliability function of exponentiated exponential distribution: Bayesian approach under general entropy loss function. *Pakistan Journal of Statistics and Operational Research*, VII(2):199–216.

RELEASE TIME ANALYSIS OF OPEN SOURCE SOFTWARE USING ENTROPY AND RELIABILITY

Vishal Pradhan, Gunjan Tripathi, Ajay Kumar AND Joydip Dhar

•

Department of Applied Sciences, ABV-IIITM Gwalior,
Gwalior-474015, MP, India
vishal.iiitmg@gmail.com, gunjantripathi7@gmail.com, ajayfma@iiitm.ac.in, jdhar@iiitm.ac.in

Abstract

Any software system, however securely written or precise the code is, is always susceptible to failure. These factors, such as the number of errors in the program or the mean-time for software failure, measure the program's reliability. In order to meet more customer needs, current OSS products must be reliable. To measure these parameters, like the reliability of the software, we use different growth models called Software Reliability Growth Models. These models help us in determining the different reliability measures. Faults occur due to several reasons in software- sometimes, it is the environmental factors. It can also be because of casual human behavior. Faults may also occur during the process of removal of previous faults. Whenever the code is changed, randomness in the software increases. We can calculate the optimal release time of a software product based on the calculated reliability measures, which have entropy also been considered. Finally, the user's satisfaction level can also be considered.

Keywords: Entropy, Debugging, Feature improvement, Feature addition, Optimal release time, Software repositories

1. Introduction

The whole idea behind development in the open-source domain is that we initially developed the software's basic model. It can be done by a single individual or a group of developers. After developing the basic software, the code is made publicly available so that people can make changes to it to add new features or make improvements in the current features. People also help in making the software famous amongst the developer community. Code changes can be done by anyone residing remotely in any part of the world. The person can simply request the owner of the software to suggest some new features or modify existing features of the software.

Source code is then again redistributed among the community of developers to cross-check and verify the changes made in the code. Open source software (OSS) is released based on two different methods which decide when the next release is to be scheduled:

1. **Time-based strategy:** In the time-based strategy, specific dates are predicted for the releases and different versions of the software are released on those dates only.
2. **Feature-based strategy:** In the feature-based strategy, new versions are released after implementing certain set of features. New features may be added to the software, or previous features may be modified to increase the satisfaction of the user community.

OSS has a number of characteristics that distinguish it from typical closed source software development.

1.1. Entropy in OSS

Changes are made to code to improve the functionality of the system, add new features, or make modifications to features suggested by the users and developers. When these changes are included, the uncertainty of the software increases. To calculate the increased randomness, information theory can be used as suggested in [10]. This idea of entropy calculation given in the information theory helps us in determining randomness increased in the software.

In this work, the effort is made to calculate the randomness in terms of the entropy in a system by assuming that the randomness increases when the source code is changed for modifications and improvements as discussed in [16]. There are various reasons for making changes to the code. Sometimes the user is not satisfied with the performance due to performance bugs. In the open-source community, since innovators are involved in the development of the software, it is important to keep their interest alive in the product. For this reason, releases are done more often in the open-source products so that contributors can relate to the product and keep making changes to it. The product remains relevant in the market.

Any software goes through a process of evolution in which bugs are fixed, new features are introduced and some features are modified according to the user's satisfaction level. Using this calculated randomness, we can also determine when the next version of the software can come into the market. In this work, non-homogeneous Poisson process (NHPP) based models are discussed. In the software testing process, it is commonly assumed that the cumulative number of failures follows the NHPP and these models are believed to be a reliable model in performance analysis for software [2, 5, 13]. Various NHPP based reliability models for OSS have been developed in recent years [10, 12].

There are two types of environment discussed in the debugging domain [7]. No new bugs are introduced in the perfect debugging model while eradicating the previous bugs. However, in the imperfect debugging model, new errors may get introduced while removing previous ones. Any software goes through a process of evolution in which bugs are fixed, new features are introduced and some features are modified according to the user's satisfaction level.

Whenever bugs are fixed, new features are introduced and modifications are made to the software, it gets better and that is what the motive is. Any organization's goal is to make its software relevant and updated as new changes keep coming into the market. Evolution is a part of any growth process. It is the same for the software industry also. An effort has been made to develop a mathematical model through which we can predict the total number of bugs to be fixed in the future and new features to be added and features to be modified. Since it is not possible to fix all the bugs in a single release, this process continues over the product's whole life cycle.

The objectives are to calculate the entropy of the software using code changes done in the different files of the system and propose a model to determine the total issues based on the calculated entropy and consider the user's satisfaction level. Also, make the prediction of release time of the next version based on entropy which the product manager can use to reduce the overall cost of the product.

2. Literature review

For any software to remain relevant in the market, it needs to grow with the user's expectations and new technologies coming into the market. Changes are a necessity for the growth and development of the product as in [19]. When new changes are made in the code, the randomness in the software increases.

Previous research has been done in this direction in considering this entropy as a measure in the calculation of release date of the software [6, 11, 15, 17, 18]. When the entropy increases, the software becomes more complex with time. This research has been made to quantify these changes and calculate the randomness that occurred because of those changes. After the entropy is calculated through file changes, the total number of issues to be fixed, including the bugs, newly introduced features, and feature improvements, are predicted. In the end, an effort is also put to predict the next release time of the version based on the entropy calculation.

In the past, numerous efforts have been made in the direction of measuring the reliability of the software and thus decreasing the overall software assessment budget of the product [10]. In most of the earlier models, only two factors are mostly considered in the software design process-reliability and the cost of the product. These are inversely proportional to each other. If we wish to increase the reliability of the software, we need to gice more time to it which increases the cost or if we try to decrease the cost, we compromise with the reliability of the software [3]. Many of the models used for prediction in the closed source projects give too optimistic results in the open source domain [8, 9].

In the models used for prediction in the open source domain, testing efforts are considered and time taken to correct the faults are also measured [3]. Dai et al. [4] put an effort to find out the randomness in the reliability modelling of a single component within a large software and attributes are also assumed to be correlated. Authors have also taken into consideration, the views of the subject expert and the historical change data of the software.

Kamavaram and Goseva-Popstojanova [8] performed the research based on including entropy for the calculation of software reliability engineering by using it in the Markov model which is used for software specifications. Randomness of the operational profile is calculated and the model proposed is architecture based. effort is also made to introduce the concept of conditional entropy. Kerzazi and Khomh [9] in this research considers one of the most important thing in release engineering, the time for a software code to reach from the development environment to the production environment considering the quality analysis at each stage. Only data from the actual industrial organizations are taken over a long period of time (15 months) and around 250 releases are taken into consideration.

Li et al. [10] investigated one of the left out factors in open source software- reliability. Two important parameters which are stressed upon are the fast release of software and the reliability. Though these are contradicting in nature but they are most important factors to b considered. Release planning model is presented through this research. Multi attribute theory is also considered. Michlmayr et al.[11] focused on the time-based release model of OSS. In the time-based model, new releases come in the market after a certain pre-decided interval of time. They have taken interviews of members from seven open source organizations with volunteer workers and have analyzed the benefits of time-based release models.

Ruhe [14] researched thoroughly on the tools and techniques through which products can be made in a manner such that resources are utilized efficiently and users expectations can also be met by the product. Methods to build successful product are discussed. Release planning problem is also discussed.

3. Objectives

Any software goes through a process of evolution in which bugs are fixed, new features are introduced and some features are modified according to the user's satisfaction level.

Also, certain issues are left unnoticed, which the innovators in future development processes remove. This can happen due to correction delay. When the life cycle process of the development starts, the issues which were not fixed in previous releases are also taken into consideration and are added to the issue content of the current work. The objectives are:

- Calculate the entropy of the software using code changes done in the different files of the system.
- Propose a model to determine the total issues based on the calculated entropy and consider the user's satisfaction level.
- Prediction of release time of the next version is based on entropy which the product manager can reduce the product's overall cost.

4. Research Methodology

4.1. Dataset

The dataset collection process involved collecting issues from the issues directory of different Apache products. Products have multiple releases and issues are fixed before each release.

Multiple open-source projects have the details of their issue available on multiple platforms like Bugzilla. Apache projects are big projects and they involve a large number of contributors involved in the development of the products. Different products of Apache open source are considered, for example, Avro, jUDDI and Hive. All the fixed issues - either bugs, new features, or improvised features are downloaded from the issue tracking repository available on the Bugzilla website.

4.2. Methodology

After collecting data of issues from the issues directory, release dates are also noted from the git-hub repository. Then the date of fixing issues is mapped with the release date of products. Then, monthly entropy can be calculated according to the number of changes made in different projects and modules of the product. After the calculation of entropy, this is mapped with the release date noted from the git-hub repository. In the calculation of entropy, the code change process is termed as an event and noted down. Here, an event is defined as the process during which the code of a software is changed.

First, all the issues, including the bugs, new features introduced or suggested by the users of the software and feature improvements, are downloaded from the issue tracking repository. All product details with issues are available on the issues.apache.org website. Then for each product, their bugs, feature additions and feature improvements are calculated on a month-to-month basis. We also made a note of the date when these issues were fixed.

The release date of the products can be extracted from the git-hub using the git-hub tool. Whenever changes are made in the source code and committed, they are being done to fix some issues or add some new features. Git-hub tool is easily available on the git-hub site.

4.3. Implementation

Issues are defined as the bugs present in the software, new features to be added and the feature modifications to be performed. There are two types of contributors, innovators and imitators. Innovators fix the issues at rate 'p' and imitators fix issues at rate 'q'. We assume that the total number of issues identified as the target issues to be fixed in a particular release is constant and represented by 'a'. Initially, since no issues are fixed, we can represent this condition with a differential equation as follows:

$$\frac{d(X(t))}{dt} = p(a - X(t)) + q\frac{X(t)}{a}(a - X(t)) \tag{1}$$

where X(t) represents the value of fixed bugs cumulatively.

Assuming the initial conditions at t = 0 , no bugs are fixed, that is X(0) = 0, we have:

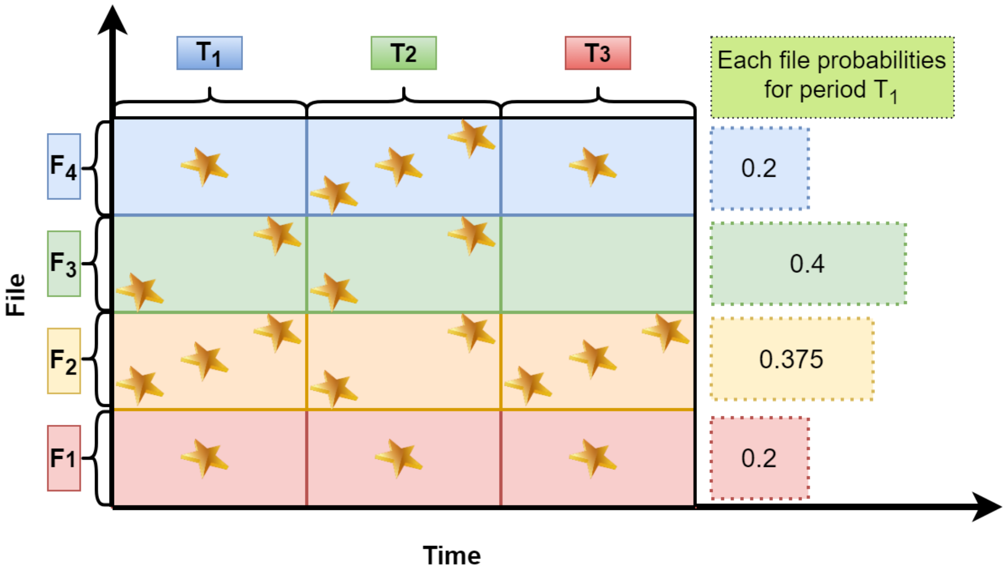
$$X(t) = a\left[\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{(-(p+q)t)}}\right] \tag{2}$$

Here, q/p remains constant and p+q represent the rate at which issues are fixed per remaining issue.

5. Entropy calculation

To calculate entropy, we must first find out the amount of information in the software code. Information theory can be used to find that. For our model, we define data as the event in which a file is changed for modification in the code. This can be used to calculate the amount of randomness associated with the software system.

We have considered four files and the time interval during which changes were made in those files is noted down. Where p gives the probability that the file will be changed during that period, we calculated it by dividing the number of times this specific file is changed in that period by the total number of changes made in all the files. This is shown in the figure.



There are multiple reasons for changing the code of the software. When users suggest bugs and new user requirements come up, developers make changes in the code to modify modules or add new modules. Sometimes there are logical errors present that also need to be rectified. Using the Cobb-Douglas equation, we can include both time and entropy for the calculation of output as follows, where time is represented by 's' and entropy is represented by 'u'.

$$t = s^{\alpha} u^{(1-\alpha)} \tag{3}$$

where $0 <= \alpha <= 1$

Using the Cobb-Douglas function, we can integrate time and entropy in the calculation of X(t) to observe the effect of both time and entropy. The model to calculate issues fixed can thus be represented as having both entropy and time-integrated. The model can thus be represented as:

$$X(s,u) = a \left[\frac{1 - e^{-b(s^{\alpha} u^{(1-\alpha)})}}{1 + \beta e^{-b(s^{\alpha} u^{(1-\alpha)})}} \right] \tag{4}$$

In most of the OSSs, all issues are not fixed in the same release. Some are passed on to the next release. These issues are then added to the content of issues for a particular version and then fixed by the developer team.

5.1. Code History Metric

Code History Metric(CHM) is a method with which we calculate the complexity of code changes in the software [1]. The whole concept of code changes is used for the calculation of CHM. We also measure one more factor, which is Code History Period Factor(CHPF), for a file j during a time period of i as:

$$CHPF(i,j) = C_{ij}H_i where j \in F_i \tag{5}$$

Here, H_i is the entropy calculated for the changes done in the file during a period of time interval i and C_{ij} represents the contribution of entropy for the given file j. We are considering several cases here:

Case 1 : If C_{ij} = 1, we assume full complexity and all files having the same weight during this interval.

Case 2 : If C_{ij} = P_j, we assume contribution is equal to the probability of file j being changed in the time interval.

5.2. Releasing Versions of Software

Steps for calculating the matrix are as follows:

1. First a note of the release date is taken for every version of the software.
2. Date on which each file was changed is also noted.
3. Classification of bugs fixed in the files as a new feature, feature improvement and feature modification.
4. Calculate the total number of bugs fixed.
5. Record the issues in each released version of the software.
6. Arrange the changes according to the month in which they were made and the matrix is calculated.
7. For every release, the time at which the software was released is found.

Though making new changes to the software is a requirement for every organization, it brings new faults into the system. The code becomes more complex over time and it becomes challenging to keep it relevant and reliable at the same time. Predicting the modules that have more number of bugs than the others is beneficial in the sense that then the product manager can decide how to allocate resources among the parts of the software such that an efficient version is released with more reliability and user satisfaction level.

Predicting bugs also helps in reducing the testing time of the team and thus, the overall expenditure can be reduced.

Based on the change of code of the software, bugs and release time are predicted. As a part of the previous work, a few previous models are considered and their results are compared with the proposed model in the next section.

6. Experiments and results

6.0.1 Experiments

Assuming that initially no issues are there and issues will occur in further releases only after suggestions from the users come, at t = 0, X(t) = 0. This gives the following equation :

$$X(t) = a\left[\frac{1-e^{-bt}}{1+\beta e^{-bt}}\right] \tag{6}$$

Here, value of q/p does not change with time and p+q is the rate at which issues are fixed per remaining issues.

Rate at which cumulative number of issues are fixed is given by:

$$\frac{dX(t)}{dt} = \frac{f(t)}{1 - F(t)}(a - X(t)) \tag{7}$$

Here, F(t) is the distribution function representing the iussues to be fixed anf f(t) is the density function, relation between them is f(t) = d F(t)/dt.

At any given time t, (a - X(t)) represents the number of issues yet to be fixed in the code. As time increases, more issues are fixed and (a - X(t)), i.e, remaining issues in the software decreases with time. This in turn increases the reliability of the software[4].

For a particular software product, 'Ai' is the real number of issues to be fixed, including the bugs, new features to be added and the modifications to be performed in the software. Where 'ai' is the potential issues that need to be fixed. 'ai-Ai' represents the issues left in the previous release and are now in the issue content of the current release. Fixing efficiency is denoted by 'bi'.

Comparing the results of our model with the JM and S-shaped mode, we get the results as shown in the table 1.

Table 1: *Comparison Table*

Release Number	Model	a	b	Ai	ai-Ai
1	JM Model	360	0.121		28
1	S-shaped Model	621	0.173		289
1	Proposed Model	364	0.527	332	32
2	JM Model	197	0.13		14
2	S-shaped Model	221	0.268		38
2	Proposed Model	507	0.076	183	324
3	JM Model	264	0.078		55
3	S-shaped Model	300	0.174		91
3	Proposed Model	233	0.3	209	24
4	JM Model	219	0.063		77
4	S-shaped Model	181	0.211		38
4	Proposed Model	210	0.173	142	9
5	JM Model	85	0.116		10
5	S-shaped Model	104	0.233		9
5	Proposed Model	75	0.743	75	0

The release time of the software is calculated using linear regression in multiple stages. Assuming that the time, p_0 is a dependent variable and the Code History Matrix (x_0) and the number of bugs ($x1$) are independent variables, time can be calculated as:

$$p_o = a_0 + a_1 x_0 + a_2 x_1 \tag{8}$$

Here, a_0 and a_1 are regression coefficients whose value can be found out by linear regression method.

Calculated values of time to release from the Code History Matrix are also shown in the table 2.

Table 2: *CHM Table*

Total Changes	CHM(1)	CHM(2)	Time
673	7.288	2.228	2.248
286	2.335	0.499	2.723
1626	8.931	3.405	2.726
774	10.07	3.056	3.934
738	4.034	0.901	2.467
911	6.545	1.698	3.303
1524	8.361	2.311	2.812
302	3.296	0.804	3.421
2414	10.19	2.522	5.074
2331	3.535	0.783	3.058
943	5.795	1.374	3.317
576	6.463	1.593	2.562
413	3.456	0.886	1.143
382	5.178	1.32	3.335
1883	5.302	1.422	1.895
234	1.757	0.309	0.663
500	2.324	1.273	1.086
898	5.361	1.256	4.918
1254	5.252	1.978	3.337
827	7.04	0.256	2.443
342	1.629	1.304	0.661
734	5.377	1.543	2.803
891	5.464	1.064	3.007
1353	4.795	1.476	1.412
850	6.447	0.143	2.461
1135	1.578	0.214	1.128
334	1.645	0.583	1.143

7. Results and discussion

The whole idea behind OSS development is that we initially developed the software's basic model. It can be done by a single individual or a group of developers. After developing the basic software, the code is made publicly available so that people can make changes to it to add new features or make improvements in the current features.

The aptitude of the testing team also determines the number of faults introduced into a software. Environmental factors such as different operating system specifications can also contribute to introducing bugs into the software. Whenever developers try to remove a bug from the code, they make changes to the software. These changes can be quantified in terms of entropy. The time of the next release is helpful both to the product manager and the user of the software.

7.1. Ease of Testing

One of the main stages of the software development process is the testing phase. This model helps to ease the process in the testing phase by providing a direction to the testers so that they can focus their energy and resources in one direction and get better and more efficient results. Once the cost of the testing process is minimized, the whole development process cost is reduced.

7.2. Managerial implications

Product managers can use this approach to find out the best time to release the software in the market. Rapid release strategy and reliability are conflicting parameters and it is crucial to balance between them. Product managers can look at the future trend of bugs and plan their releases accordingly.

8. Conclusion

There is always scope for improvement. Although the proposed model helps determine the important parameters of modelling, the manager's decision is still subjective. There are still many more different factors in different projects that matter when product release is considered. So our model can be used to make a general idea about the situation, but the final decision needs to be taken considering many other factors also.

Other factors that need to be considered are past data of the product and historical experiences of the team. Only then a more trusted decision can be taken. Efforts should also be made to study more closed source projects and compare them to the calculation of different parameters.

One more limitation is that calculation of entropy based on the files changed is done manually. This process can also be made more precise and accurate by using other methods of entropy calculation if available. More research can be done in this direction.

References

[1] Anand, A., Bharmoria, S. and Ram, M., 2019, Characterizing the Complexity of Code Changes in Open Source Software, *In Recent Advancements in Software Reliability Assurance*, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742: CRC Press, pp. 1–14.

[2] Anand, A., Kaur, J., Singh, O. and Ram, M., 2021, Optimal Resource Allocation for Software Development under Agile Framework, *Reliability: Theory & Applications*, (SI 2 (64)), pp. 48–58.

[3] Cathedral, R.: 2001, the bazaar raymond es the cathedral and the bazaar: Musings on linux and open source by an accidental revolutionary.

[4] Dai, Y.-S., Xie, M., Long, Q. and Ng, S.-H.: 2007, Uncertainty analysis in software reliability modeling by bayesian analysis with maximum-entropy principle, *IEEE Transactions on Software Engineering* **33**(11), 781–795.

[5] Dhaka, R., Pachauri, B. and Jain, A., 2021, Two-Dimensional SRGM with Delay in Debugging by Considering the Uncertainty Factor and Predictive Analysis, *Reliability: Theory & Applications*, (SI 2 (64)), pp .82–94.

[6] Gacek, C. and Arief, B.: 2004, The many meanings of open source, *IEEE software* **21**(1), 34–40.

[7] Hassan, A. E.: 2009, Predicting faults using the complexity of code changes, *2009 IEEE 31st international conference on software engineering*, IEEE, pp. 78–88.

[8] Kamavaram, S. and Goseva-Popstojanova, K.: 2002, Entropy as a measure of uncertainty in software reliability, *13th Int'l Symp. Software Reliability Engineering*, pp. 209–210.

[9] Kerzazi, N. and Khomh, F.: 2014, Factors impacting software release engineering: A longitudinal study, *Proc. 2nd Workshop Release Eng*, pp. 1–5.

[10] Li, X., Li, Y. F., Xie, M. and Ng, S. H.: 2011, Reliability analysis and optimal version-updating for open source software, *Information and Software Technology* **53**(9), 929–936.

[11] Michlmayr, M., Fitzgerald, B. and Stol, K.-J.: 2015, Why and how should open source projects adopt time-based releases?, *IEEE Software* **32**(2), 55–63.

[12] Pradhan, V., Kumar, A. and Dhar, J., 2022, Modeling Multi-Release Open Source Software Reliability Growth Process with Generalized Modified Weibull Distribution, *Evolving Software Processes: Trends and Future Directions*, pp. 123–133.

[13] Pradhan, V., Kumar, A. and Dhar, J. 2022, Modelling software reliability growth through generalized inflection S-shaped fault reduction factor and optimal release time, *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 236(1), pp. 18–36.

[14] Ruhe, G.: 2010, *Product release planning: methods, tools and applications*, Auerbach Publications.

[15] Saxena, P., Kumar, V. and Ram, M., 2021, Ranking of Software Reliability Growth Models: A Entropy-ELECTRE Hybrid Approach, *Reliability: Theory & Applications*, (SI 2 (64)), pp. 95–113.

[16] Singh, V., Sharma, M. and Pham, H.: 2017, Entropy based software reliability analysis of multi-version open source software, *IEEE Transactions on Software Engineering* **44**(12), 1207–1223.

[17] Spinellis, D.: 2015, The strategic importance of release engineering, *IEEE software* **32**(2), 3–5.

[18] Svahnberg, M., Gorschek, T., Feldt, R., Torkar, R., Saleem, S. B. and Shafique, M. U.: 2010, A systematic review on strategic release planning models, *Information and software technology* **52**(3), 237–248.

[19] Wright, H. K. and Perry, D. E.: 2012, Release engineering practices and pitfalls, *Proceedings of the 34th International Conference on Software Engineering*, IEEE Press, pp. 1281–1284.

Performance Analysis of System where Service Type for Boiler Depends Upon Major or Minor Failures

Upasana Sharma[1] and Rajveer Kaur[2*]

•

[1]Department of Statistics, Punjabi University, Patiala-India, usharma@pbi.ac.in
[2*]Department of Statistics, Punjabi University, Patiala-India,
Corresponding author: rajveersidhu700@gmail.com

Abstract

In industries, the type of failure sensitively affects the system. So, it is essential to Categorize these failures into different categories to enhance the system performance. In this research, concentration made in differentiating the failure type into major/minor categories with repair/replacement facility for the service by single repairman. Currently, we studied the boiler system of the steam generation plant to perform the task of repair/replacement with a single repairman. A reliability model constructs to compute MTSF(mean time to system failure), availability, Busy period for repair/replacement, and profit evaluation. The above measures were estimated numerically and plotted graphically using semi-Markov processes and regenerative point technique. Various effectiveness measures show how system performance gets affected by major/minor failures & the type of service provided.

Keywords: Regenerative point technique, major/minor failure, repair/replacement, Reliability modeling, semi-Markov processes.

I. Introduction

Proper functioning of any system is a prerequisite of any industrial process, as an interruption in the operation of a system causes not only deterioration in the quality of manufactured products but damages to the plant itself. Thus the reliability of the system becomes much more essential. Many contributors pay their efforts in the literature of reliability. [5], [8], [2] have worked on cost-effectiveness and reliability analysis on different situations. Various situations on repair, replacement, & inspection have been investigated by the authors [6], [4], [9], [3], [1]. However, the distinction between major and minor faults has not been the topmost research topic in reliability. [7] have discussed the concept of major/minor failures subjected to the ordinary and expert repairman. [10] revealed the possibility of immediate repair on minor failure and waiting time for repair on major failure at night hours. But the concept of repair/replacements depending upon the minor/major faults has not been seen in the literature of reliability modeling. In addition to the above idea, the boiler of the Steam Generation Plant studied, in which the type of service provided for a boiler depends upon major and minor faults. Minor faults are repaired easily while major faults are replaceable.

Initially, a three-unit system with a boiler and two FD fans has considered for the study. When a boiler fails system stops working immediately, but if any one of the FD fans fails system goes on reduced capacity. For continuous functioning, a boiler and 2 FD fans should function. On boiler failure, two possibilities arise major and minor faults. Repair facility provided for a small crack in the outer chamber of the boiler that occurs due to overheating & erosion, whereas the repairman performed replacement for the major equipment failures for a boiler. Priority of repair is given to the boiler over fans so that the system can operate for a long time. For the FD fans, repair priority is on an FCFS basis.

The following reliability measures were computed numerically using semi-Markov processes and regenerative point techniques and also plotted graphically based on the information gathered from the industry :

- Mean time to system failure (MTSF).
- Availability analysis at full capacity.
- Availability analysis at reduced capacity.
- Busy period for repair time only.
- Busy period for replacement time only.
- Expected no. of repairs.
- Expected no. of replacements.
- Cost-benefit analysis.

II. Model Description

I. State Transition Diagram

Figure 1, shows the state transitions diagram of the steam generation plant consisting of one boiler and two FD fans.

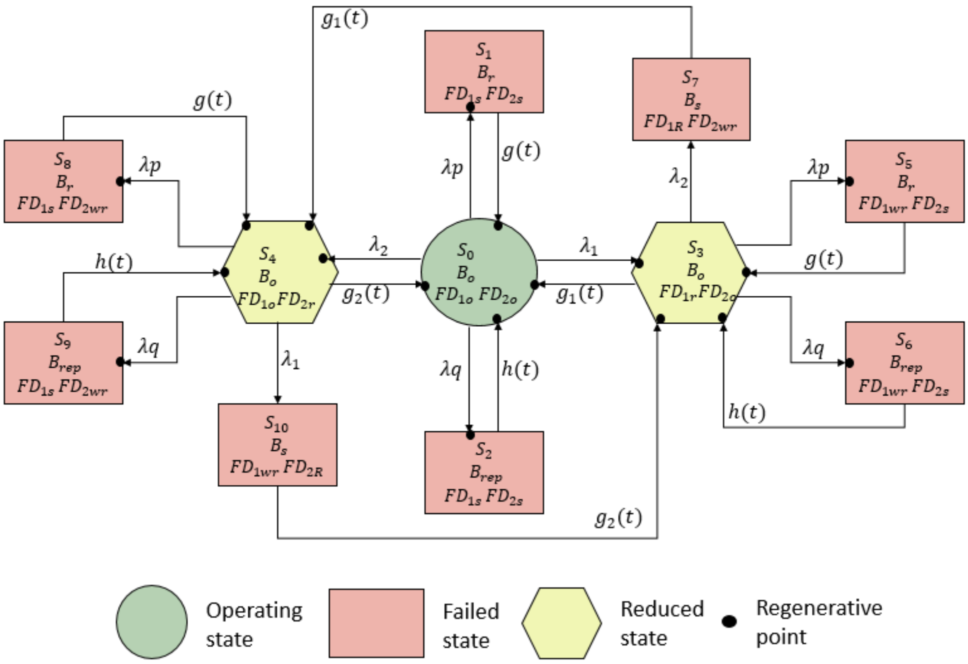


Figure 1: *State Transition Diagram*

Table 1: *State Description*

States	Discription
S_0	This is the operating state of the system
$S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_8, S_9$	The epoch of entry into these states are regenerative points thus, these states are called regenerative states
S_3, S_4	These are reduced capacity states
$S_1, S_2, S_5, S_6, S_7, S_8, S_9, S_{10}$	These are failed states of the system.

II. Assumptions

The stated model follows these assumptions.

- All the random variables are independent.
- Failure times distribution is exponential, whereas repair times have distributed arbitrarily.
- System works as well as new after every repair.
- when the failure occurs repairman will come immediately.

III. Nomenclature

Table 2: *Notations & symbols for the states of system*

Notations	
Notations	Discription
λ	Constant failure rate of Boiler.
λ_1	Constant failure rate of FD fan 1.
λ_2	Constant failure rate of FD fan 2.
α	Repair rate of Boiler.
α_1	Repair rate of FD fan 1.
α_2	Repair rate of FD fan 2.
γ	Replacement rate of Boiler.
p	probability of minor failure in Boiler.
q	probability of major failure in Boiler.
$G(t)$, $g(t)$	c.d.f. & p.d.f of repair time of Boiler.
$H(t)$, $h(t)$	c.d.f. & p.d.f of replacement time of Boiler.
$G_1(t)$, $g_1(t)$	c.d.f. & p.d.f of repair time of FD fan 1.
$G_2(t)$, $g_2(t)$	c.d.f. & p.d.f of repair time of FD fan 2.
Symbols for the states of the system	
Symbols	Discription
S_i	states of the system, $i = 1, 2, 3, ..., 10$
B_o, FD_{1o}, FD_{2o}	Boiler, FD fans 1 and 2 are in operating state respectively.
B_r	Boiler under repair.
B_{rep}	Boiler under replacement.
B_s, FD_{1s}, FD_{2s}	Boiler, FD fans 1 and 2 in standby state.
FD_{1r}, FD_{2r}	FD fans 1 and 2 under repair.
FD_{1wr}, FD_{2wr}	FD fans 1 and 2 are waiting for repair.
FD_{1R}, FD_{2R}	FD fans 1 and 2 are under repair from previous state.

IV. Transition Probabilities & Mean Sojourn Times

The p_{ij} represents non-zero elements which are given below

$$p_{01} = \frac{\lambda p}{\lambda + \lambda_1 + \lambda_2}, \qquad p_{02} = \frac{\lambda q}{\lambda + \lambda_1 + \lambda_2},$$

$$p_{03} = \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2}, \qquad p_{04} = \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2},$$

$$p_{10} = p_{20} = 1, \qquad p_{30} = g_1{}^*(\lambda + \lambda_2),$$

$$p_{35} = \frac{\lambda p}{\lambda + \lambda_2}[1 - g_1{}^*(\lambda + \lambda_2)], \qquad p_{36} = \frac{\lambda q}{\lambda + \lambda_2}[1 - g_1{}^*(\lambda + \lambda_2)],$$

$$p_{37} = \frac{\lambda_2}{\lambda+\lambda_2}[1 - g_1{}^*(\lambda+\lambda_2)], \qquad p_{34}^{(7)} = \frac{\lambda_2}{\lambda+\lambda_2}[1 - g_1{}^*(\lambda+\lambda_2)],$$
$$p_{40} = g_2{}^*(\lambda+\lambda_1), \qquad p_{48} = \frac{\lambda p}{\lambda+\lambda_1}[1 - g_2{}^*(\lambda+\lambda_1)],$$
$$p_{49} = \frac{\lambda q}{\lambda+\lambda_1}[1 - g_2{}^*(\lambda+\lambda_1)], \qquad p_{4,10} = \frac{\lambda_1}{\lambda+\lambda_1}[1 - g_2{}^*(\lambda+\lambda_1)],$$
$$p_{4,3}^{(10)} = \frac{\lambda_1}{\lambda+\lambda_1}[1 - g_2{}^*(\lambda+\lambda_1)], \qquad p_{53} = 1,$$
$$p_{63} = 1, \qquad p_{74} = 1,$$
$$p_{84} = 1, \qquad p_{94} = 1,$$
$$p_{10,3} = 1$$

It can be verified by these probabilities that

$$p_{01} + p_{02} + p_{03} + p_{04} = 1, \qquad p_{10} = 1, \qquad p_{20} = 1,$$
$$p_{30} + p_{35} + p_{36} + p_{37} = 1, \qquad p_{30} + p_{35} + p_{36} + p_{34}^{(7)} = 1, \qquad p_{40} + p_{48} + p_{49} + p_{4,10} = 1,$$
$$p_{40} + p_{48} + p_{49} + p_{43}^{(10)} = 1, \qquad p_{53} = p_{63} = p_{74} = 1, \qquad p_{84} = p_{94} = p_{10,3} = 1$$

Also μ_i, the mean sojourn times in state S_j are

$$\mu_0 = \frac{1}{\lambda+\lambda_1+\lambda_2}, \qquad \mu_1 = -g^{*\prime}(0), \quad \mu_2 = -h^{*\prime}(0), \quad \mu_3 = \frac{1}{\lambda+\lambda_2}[1 - g_1{}^*(\lambda+\lambda_2)],$$
$$\mu_4 = \frac{1}{\lambda+\lambda_1}[1 - g_2{}^*(\lambda+\lambda_1)], \quad \mu_5 = -g^{*\prime}(0), \quad \mu_6 = -h^{*\prime}(0), \quad \mu_7 = -g_1{}^{*\prime}(0),$$
$$\mu_8 = -g^{*\prime}(0), \qquad \mu_9 = -h^{*\prime}(0), \quad \mu_{10} = -g_2{}^{*\prime}(0),$$

The unconditional mean time taken by the system to transit for any regenerative state $'j'$ when it (time) is counted from the epoch of entrance into state $'i'$ is mathematically represented as

$$m_{ij} = \int_0^\infty t dQ_{ij}(t) = -q_{ij}{}^{*\prime}(0)$$

$$m_{01} + m_{02} + m_{03} + m_{04} = \mu_0, \qquad m_{10} = \mu_1,$$
$$m_{20} = \mu_2, \qquad m_{30} + m_{35} + m_{36} + m_{37} = \mu_3,$$
$$m_{30} + m_{35} + m_{36} + m_{34}^{(7)} = \mu_3 + K_1, \qquad m_{40} + m_{48} + m_{49} + m_{4,10} = \mu_4,$$
$$m_{40} + m_{48} + m_{49} + m_{43}^{(10)} = \mu_4 + K_2, \qquad m_{53} = \mu_5,$$
$$m_{63} = \mu_6, \qquad m_{74} = \mu_7,$$
$$m_{84} = \mu_8, \qquad m_{94} = \mu_9,$$
$$m_{10,3} = \mu_{10}$$

where

$$K_1 = \frac{\lambda_2}{\lambda}\int_0^\infty t g_1(t) dt, \qquad K_2 = \frac{\lambda_1}{\lambda}\int_0^\infty t g_2(t) dt, \tag{1}$$

III. Reliability Measures for System Effectiveness

I. Mean Time to System Failure (MTSF)

When the system starts from the initial state S_0, Mean time to system failure (MTSF) of the system is determined by considering failed state as absorbing state as given below

$$MTSF = T_0 = \lim_{s\to 0} \frac{1 - \phi_0{}^{**}(s)}{s} \tag{2}$$

Using L' Hospital Rule & putting the value of $\phi_0{}^{**}(s)$, we have

$$T_0 = \frac{N}{D} \tag{3}$$

where
$N = \mu_0 + \mu_3 p_{03} + \mu_4 p_{04}$
$D = 1 - p_{03}p_{30} - p_{04}p_{40}$

II. Availability Analysis at Full Capacity

Using the theory of regenerative processes, the availability AF_0 of the system at full capacity is given by

$$A_{F0} = \lim_{s\to 0}(s\,A_{F0}{}^*(s)) = \frac{N_1}{D_1}$$

where

$$N_1 = \mu_0[(1 - p_{48} - p_{49})(1 - p_{35} - p_{36}) - p_{34}^{(7)} p_{43}^{(10)}] \tag{4}$$

$$\begin{aligned} \& \quad D_1 = {} & \mu_0[p_{40} - p_{30}p_{43} - p_{40}p_{35}p_{53} - p_{40}p_{63}p_{36}] + (\mu_1 p_{01} + \mu_2 p_{02})[(1 - p_{48}p_{84} - p_{49}p_{94}) \\ & (1 - p_{36}p_{63} - p_{35}p_{53}) - p_{43}p_{34}] + [\mu_3 + K_1][p_{03} + p_{04}p_{43} - p_{03}p_{48}p_{84} - p_{03}p_{49}p_{94}] \\ & + [\mu_4 + K_2][p_{04} + p_{03}p_{34} - p_{04}p_{35}p_{53} - p_{04}p_{36}p_{63}] \\ & + (\mu_5 p_{35} + \mu_6 p_{36})[(1 - p_{01}p_{10} - p_{02}p_{20})(1 - p_{48}p_{84} - p_{49}p_{94}) - p_{04}p_{40}] \\ & + (\mu_8 p_{48} + \mu_9 p_{49})[(1 - p_{01}p_{10} - p_{02}p_{20})(1 - p_{35}p_{53} - p_{63}p_{36}) - p_{03}p_{30}] \end{aligned} \tag{5}$$

III. Availability Analysis at Reduced Capacity

Using the theory of regenrative processes, the availability AR_0 of the system at reduced capacity is given by

$$A_{R0} = \lim_{s\to 0}(s\,A_{R0}{}^*(s)) = \frac{N_2}{D_1}$$

where

$$N_2 = \mu_3[p_{03}(1 - p_{48} - p_{49}) + p_{04}p_{43}^{(10)}] + \mu_4[p_{04}(1 - p_{35} - p_{36}) + p_{03}p_{34}^{(7)}] \tag{6}$$

and D_1 is already specified in equation (5).

IV. Busy Period for Repair Time only

In steady state, busy period for repair time is defined as the time for which system is under repair by repairman and is given by

$$B_{R0} = \lim_{s\to 0}(sB_{R0}{}^*(s)) = \frac{N_3}{D_1} \tag{7}$$

Where

$$\begin{aligned} N_3 = {} & \mu_1 p_{01}[(1 - p_{48} - p_{49})(1 - p_{35}p_{36}) - p_{43}{}^{(10)} p_{34}{}^{(7)}] + \mu_3[p_{03}(1 - p_{48} - p_{49}) + p_{04}p_{43}{}^{(10)}] \\ & + \mu_4[p_{04}(1 - p_{35} - p_{36}) + p_{03}p_{34}{}^{(7)}] + \mu_5 p_{35}[p_{03}p_{40} + (p_{03} + p_{04})p_{43}{}^{(10)}] \\ & + \mu_8 p_{48}[p_{04}(1 - p_{35} - p_{36}) + p_{03}p_{34}{}^{(7)}] \end{aligned} \tag{8}$$

& D_1 is already specified in equation (5).

V. Busy Period for Replacement Time only

In steady state, busy period for replacement time is defined as the time for which system is busy under replacements and is given by

$$B_{RP0} = \lim_{s\to 0}(sB_{RP0}{}^{*}(s)) = \frac{N_4}{D_1} \tag{9}$$

where

$$N_4 = \mu_2 p_{02}[(1 - p_{48} - p_{49})(1 - p_{35}p_{36}) - p_{43}{}^{(10)}p_{34}{}^{(7)}] + \mu_6 p_{36}[p_{03}(1 - p_{48} - p_{49}) + p_{43}{}^{(10)}] + \mu_9 p_{49}[p_{04}(1 - p_{35} - p_{36}) + p_{34}{}^{(7)}] \tag{10}$$

& D_1 is already specified in equation (5).

VI. Expected No. of Repairs

Expected number of repairs per unit time for the system is given by

$$V_{R0} = \lim_{s\to 0}[sV_{R0}{}^{**}(s)] = \frac{N_5}{D_1} \tag{11}$$

where

$$N_5 = (1 - p_{02})(1 - p_{35} - p_{36})(1 - p_{48} - p_{49}) - (1 - p_{02})p_{34}{}^{(7)}p_{43}^{(10)} + p_{03}p_{48}p_{34}^{(7)} + p_{04}p_{35}p_{43}^{(10)} + p_{04}p_{48}(1 - p_{35} - p_{36}) + p_{03}p_{35}(1 - p_{48} - p_{49}) \tag{12}$$

& D_1 is already specified in equation (5).

VII. Expected No. of Replacements

Expected number of replacements per unit time for the system is given by

$$V_{RP0} = \lim_{s\to 0}[sV_{RP0}{}^{**}(s)] = \frac{N_6}{D_1} \tag{13}$$

where

$$N_6 = p_{02}((1 - p_{35} - p_{36})(1 - p_{48} - p_{49}) - p_{34}{}^{(7)}p_{43}^{(10)}) + p_{03}(p_{36}(1 - p_{48} - p_{49}) + p_{49}p_{34}^{(7)}) + p_{04}(p_{49}(1 - p_{35} - p_{36}) + p_{36}p_{43}^{(10)}) \tag{14}$$

& D_1 is already specified in equation (5).

IV. Cost-Benefit Analysis

The expected total profit incurred to the system is

$$P_0 = C_0A_{F0} + C_1A_{R0} - C_2B_{R0} - C_3B_{RP0} - C_4V_{R0} - C_5V_{RP0}$$

where
C_0 = revenue per unit up time at full capacity
C_1 = revenue per unit time at reduced capacity
C_2 = cost per unit time when repairman is busy in doing repair
C_3 = cost per unit time when repairman is busy in doing replacement
C_4 = cost per repair.
C_5 = cost per replacement.

V. Particular Cases

For the numerical evaluation and graphical plotting for various reliability measures, the following particular cases are considered.
Let us assume that $g(t) = \alpha e^{-\alpha t}, h(t) = \gamma e^{-\gamma t}, g_1(t) = \alpha_1 e^{-\alpha_1 t}, g_2(t) = \alpha_2 e^{-\alpha_2 t}$, and the remaining distributions are the same as in the general case. Therefore, we have

$$p_{01} = \frac{\lambda p}{\lambda + \lambda_1 + \lambda_2}, \quad p_{02} = \frac{\lambda q}{\lambda + \lambda_1 + \lambda_2}, \quad p_{03} = \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2},$$

$$p_{04} = \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2}, \quad p_{10} = 1, \quad p_{20} = 1,$$

$$p_{30} = \frac{\alpha_1}{\lambda + \alpha_1 + \lambda_2}, \quad p_{35} = \frac{\lambda p}{\lambda + \alpha_1 + \lambda_2}, \quad p_{36} = \frac{\lambda q}{\lambda + \alpha_1 + \lambda_2},$$

$$p_{37} = \frac{\lambda_2}{\lambda + \alpha_1 + \lambda_2} = p_{34}{}^{(7)}, \quad p_{40} = \frac{\alpha_2}{\lambda + \lambda_1 + \alpha_2}, \quad p_{48} = \frac{\lambda p}{\lambda + \alpha_2 + \lambda_1},$$

$$p_{49} = \frac{\lambda q}{\lambda + \alpha_2 + \lambda_1}, \quad p_{4,10} = \frac{\lambda_1}{\lambda + \alpha_2 + \lambda_1} = p_{43}{}^{(10)}, \quad p_{53} = p_{63} = p_{74} = 1$$

$$p_{84} = p_{94} = p_{10,3} = 1, \quad \mu_0 = \frac{1}{\lambda + \lambda_1 + \lambda_2}, \quad \mu_1 = \frac{1}{\alpha},$$

$$\mu_2 = \frac{1}{\gamma}, \quad \mu_3 = \frac{1}{\lambda + \alpha_1 + \lambda_2}, \quad \mu_4 = \frac{1}{\lambda + \alpha_2 + \lambda_1},$$

$$\mu_5 = \frac{1}{\alpha}, \quad \mu_6 = \frac{1}{\gamma}, \quad \mu_7 = \frac{1}{\alpha_1},$$

$$\mu_8 = \frac{1}{\alpha}, \quad \mu_9 = \frac{1}{\gamma}, \quad \mu_{10} = \frac{1}{\alpha_2},$$

Table 3: *Computation of various rates/costs on the basis of actual data collected from industry*

Various rates/ cost associated	corresponding values
Failure rate of Boiler (λ)	0.0001186/hr
Failure rate of FD fan 1 (λ_1)	0.0001171/hr
Failure rate of FD fan 2 (λ_2)	0.000101295/hr
Repair rate of Boiler (α)	0.00738/hr
Replacement rate of Boiler (γ)	0.0008733 /hr
Repair rate of FD fan 1 (α_1)	0.024272/hr
Repair rate of FD fan 2 (α_2)	0.048544/hr
Expected cost per repair (C_4)	Rs. 14282
Expected cost per replacement (C_5)	Rs. 1579627

Hypothetical values have been taken for remaining rates/costs. arious reliability measures for system performance have been computed in table 4 by putting the values given in the table 3 based on particular cases.

Table 4: *Computation of various measures of system effectiveness*

Mean time to system failure	08376.82/hr
Availability of the system at full capacity	0.344986/hr
Availability of the system at reduced capacity	0.639788/hr
Busy period of repairman for repair time only	0.644142/hr
Busy period of repairman for replacement time only	0.010379/hr
Expected no. of repairs	0.000107/hr
Expected no. of replacements	0.000009/hr

VI. Results and Discussion

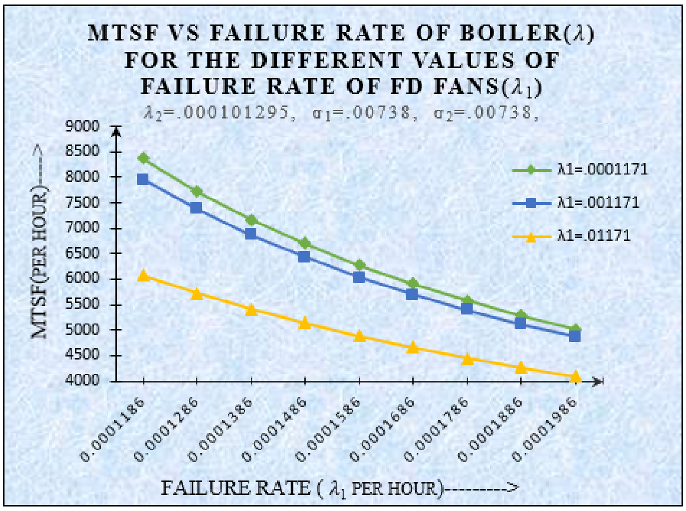


Figure 2: *MTSF vs Failure rate of Boiler*

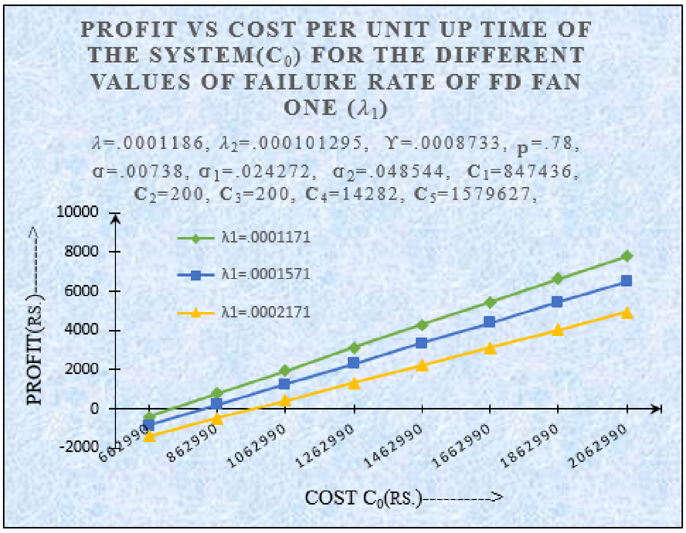


Figure 3: *Profit vs Revenue up-time for the system*

Table 5: *Cut Point for profit w.r.t. Revenue Up-time of the system.*

Failure rate of FD fan one (/hr)	Revenue per unit up time (Rs.)	Profit (Rs.)
$\lambda_1 = .0001171$	$C_0 < or = or > 731316.24$	negative or zero or positive
$\lambda_1 = .0001571$	$C_0 < or = or > 826502.352$	negative or zero or positive
$\lambda_1 = .0002171$	$C_0 < or = or > 969279.171.$	negative or zero or positive

In figure 2, the effect of the failure rate of Boiler(λ) on MTSF has shown for the different values of the failure rate of FD fan one (λ_1). As the failure rate (λ) increases, the MTSF of the system decreases. Also, as the failure rate (λ_1) increases MTSF of the system decreases. In figure 3, the effect of cost per unit up time of the system (C_0) w.r.t. profit has shown for the different values of failure rate of FD fan one (λ_1). As the cost C_0 increases, profit of the system increases. Also, as the failure rate of FD fan one (λ_1) increases profit decreases. Various cut point formed from the graph of profit w.r.t. revenue Up-time of the system as shown in table 5.

VII. Conclusion

Reliability modeling established for steam generation plant that shows the effect of service type & type of failures on system performance. The study reveals that the busy period of repairman for repair time is more as compared to replacement time. Also, cut-off points formed from profit helps the industrialist to maintain the economy of their system. In addition, any industry can consider the stated model to enhance the performance of their system using the different rates for repair/replacement.

References

[1] Agrawal, A. et. al (2021), Performance Analysis of the Water Treatment Reverse Osmosis Plant, *Reliability: Theory & Applications*, 16(3(63)): 16-25.

[2] Al Balushi, N. (2021), A Review of the Reliability Analysis of the Complex Industrial Systems, *Advances in Dynamical Systems and Applications*, 16(1):257–297.

[3] Al Rahbi, Y., & Rizwan, S. M. (2020), A Comparative Analysis between the Models of a Single Component with Single Repairman & Multiple Repairmen of an Aluminium Industry, *In 2020 International Conference on Computational Performance Evaluation (ComPE), IEEE*, 132–135.

[4] Bhardwj, R. K., & Malik, S. C. (2010). MTSF and Cost effectiveness of 2-out-of-3 cold standby system with probability of repair and inspection, *Int. J. of Eng. Sci. and Tech*, 2(1):5882–5889.

[5] Carrasco, J. A. (2003), Solving dependability/performability irreducible Markov models using regenerative randomization, *IEEE Transactions on Reliability*, 52(3):319–329.

[6] Li, W. et. al (1998), Stochastic analysis of a repairable system with three units and two repair facilities, *Microelectronics Reliability*, 38(4):585–595.

[7] Parashar, B., & Taneja, G., (2007), Reliability and profit evaluation of a PLC hot standby system based on a master-slave concept and two types of repair facilities, *IEEE Transactions on reliability*, 56(3):534–539.

[8] Rani, P., & Duhan, M. (2020), Reliability Modeling, Cost-Benefit Analysis with Inspection, Repair and Replacement Facility for Optical Fiber Network in DCRUST, *IUP Journal of Telecommunications*, 12(2):51-61.

[9] Sattiraju, R. et. al (2014), Reliability analysis of a wireless transmission as a repairable system, *In 2014 IEEE Globecom Workshops (GC Wkshps), IEEE*, 1397–1401.

[10] Sheetal, D. S., & Taneja, G. (2018), Reliability Analysis of a System Working in High Temperature Zones with Fault-Dependent Repair during Night hours. *International Journal of Applied Engineering Research*, 13(20):14650-14656.

Estimation of Average Degree of Social Network Using Clique, Shortest Path and Cluster Sampling to monitor Network Reliability

Vivek Kumar Gupta[1], Diwakar Shukla[2]

•

Department of Mathematics and Statistics
Dr. Harisingh Gour Vishwavidyalaya
Sagar, M.P., 470003, India
[1]v.vivekgupta@yahoo.com; [2]diwakarshukla@rediffmail.com

Abstract

In recent past, Online Social Networks (OSN) has emerged as a platform for sharing information, thoughts, and activities. In the real-world network, method of considering the appropriate samples is most frequently used for network analysis. Graph sampling is a procedure used for computing unknown parameters. Many sampling algorithms exist in literature such as Random node, Random edge sampling, Rank degree, etc. can be used for estimation. This paper presents a comparison of clique based procedure (CBP) and shortest path based procedure (SPP) to estimate the average degree of a vertex in a social network using an overlapping cluster sampling. A comparative procedure is used to obtain the lower and upper limit of confidence intervals with the help of multiple samples. Ogive based simulation is also used for single value computation of limits of CI. The results, obtained from simulation, show that clique based sampling algorithm (CBP) is more efficient than the shortest path based sampling algorithm (SPP). The estimated confidence intervals can be used for monitoring the reliability of a social network in terms of control over average network degree.

Keywords: Graph, Sampling, Social network, Overlapping cluster, Confidence interval (CI), Shortest path procedure (SPP), Clique based procedure (CBP), Reliability, Percentage relative gain(PRG)

1. Introduction

Online Social Networks (OSN) are used by large numbers of people around the world interacting with each other by forming like minded groups, based on the commonness of characters. Many real-world complex systems can be represented as a collection of vertices and edges — for example, information networks, communication networks, biological networks, etc. Recently evolved a surge of interest for exploring the characteristics of these networks, modeling their structure, develop algorithms for them, and examining systems that govern networks [8]. However, many of the real-world networks are too large to acquire, store or analyze, e.g. 3 billion emails per day worldwide from multiple sources to multiple destinations. The scientific community focuses on developing scalable analytic methods for different size datasets. In order to facilitate the development and testing of systems for network domains, it is often necessary to take a sample (smaller subgraphs) from a large network structure. A sampled subgraph can be used to drive realistic simulations and experimentation. Just to have a precise assessment of the performance of such systems, it is suggested by many scientists to use appropriate sampling methods that can select a good representative of networks. Graph sampling [4] is used to study small subsets of networks along with preserving the main features of the original network [6] [7].

Physical distances are utilized to get interaction between the different system variables. For example, the distance between two atoms or between two galaxies in the universe to evaluate the intensity of force of attraction.
A good sampling [1] algorithm for estimating a parameter must have:

- Cost effectiveness.
- Sample size suitability for unbiased parameter estimation.
- Practical and effective ways of accessing the graph.
- Lesser amount of time and reduction of computational efforts.

In networks, distance is a kind of path linkage used in a different manner. Distance between two web pages or between two unknown individual physical distances [9] [12] is not relevant. A path is a link in the network and distance in network represents the number of links the path contains.

In this paper, a method of cluster sampling for networks is presented using the concept of the shortest path and cliques. The approach has focus to find the shortest path and cliques between several pairs of vertices by selecting random pairs. The degree sequence of vertices in these shortest paths is taken for construction of overlapping clusters [18, 2]. The sampled pair of vertices of the social network contains only a fraction of all possible pairs of vertices.

The aim is to obtain an estimate of average degree which is a valid parameter of real network. Paper is organized as Section 2 contains definitions, overlapping sampling, motivation, and related work described in brief. Section 3 describes a sampling scheme with properties like bias and variance estimate. The performance of the proposed procedures is examined through ogive based simulation whose results are reported and comparision of efficiency are in Section 4-6. Other sections 7 to 9 reveal reliability, discussion, efficiency comparison, and conclusion.

2. Definition and Related Work

A Social network(graph) G(V, E) is represented as a pair of a vertices set V(G) and an edge set E(G), the number of vertices in G is N. Simple Graph(Network) G(V, E) contains undirected, unweighted edges, neither loops nor multiple edges. The neighborhood of u is N(u) = {v: (u, v) $\in$ E(G)}. It forms a set of edges connected to u. The degree is the number of connections that a vertex has, degree (v) = $|N(v)|$.

Average degree of a vertex is the average number of edges per vertex in the graph. It is defined as:

$$\frac{\textbf{Total number of Edges}}{\textbf{Total number of vertices}} = \textbf{Average Degree}$$

2.0.1 Clique

A clique is a subset of a network in which the vertices are more closely and intensely tied to one another than they are to other vertices of the network. The term "Dyad" is the smallest clique composed of two adjacent vertices. The chain of adjacent cliques is used as a tool for forming the community. Community detection allows professionals like election planners, community specialist physicians to understand the characteristics and role within the network and outside the network [16]. Concerned literature of methodologies of community detection [11] [17] have been developed a lot and several methods are in picture. Every algorithm has advantages, disadvantages, and working limitations over others. Many of them fail while dealing with overlapping communities [20] [19]. For example, the community of soldiers and community of drinkers may be overlapping where unique identification is a difficult procedure. One can find out new ways and means to generate community detection in networks. Mathematically clique is a subset of vertices all adjacent to each other. It can be used for community structure detection for large scale networks. The community identification in these methods is defined as a chain of adjacent cliques. Some methods can find the community structure for very large-scale networks. A method proposed in [19] is also useful for such cases.

Note 2.1 : If symbol e_{ij} is used as an edge from vertex vx_i to vx_j then shortest path from one vertex to another is a path sequence of vertices $(vx_1, vx_2, \ldots, vx_n)$ so that overall possible n minimizes the $\sum_{i=1}^{n-1} f(e_{i,i+1})$.

Note 2.2 : If any two vertices are selected in a graph there may exist a shortest path. Also, there may multiple shortest paths of same length d_{ij} between vertices vx_i and vx_j. Note that the shortest path does not consider any loop or any intersect itself. Further, in an undirected graphical network, these lengths of shortest path $d_{ij} = d_{ji}$ holds between any two vertices vx_i and vx_j. But in a directed graph network, it may happen that $d_{ij} \neq d_{ji}$.

2.1. General Computational Algorithm

- Take a network G(V, E), where V = set of vertices $(vx_1, vx_2, \ldots, vx_N)$, E = set of edges $\{e_1, e_2, \ldots, e_m\}$.
- Using random sampling, select K pair of vertices or set of vertices from G(V, E) as the case may be.
- Apply an appropriate procedure of overlapping cluster formation.
- Create a degree sequence of vertices (clusters).
- Estimate average degree of network using overlapping cluster sampling mean estimation method.

2.2. Computational Procedure for Creating Clusters

2.2.1 Shortest Path Procedure(SPP)

In this, non-adjacent pair of vertices are selected in a graphical network, and using Dijkstra's algorithm [10] one can find the shortest path whose degree sequence can be obtained.

2.2.2 Clique Based Procedure (CBP)

In this, the K vertices are selected as source vertices and one can find the clique, where a clique is a complete subgraph whose degree sequence can be calculated.

2.3. Motivation

The Clique Based Procedure (CBP) was used by [16] for computing the average edge length for community detection. This procedure provides the construction of overlapping clusters. The shortest path procedure (SPP) also provides the construction of overlapping clusters. In sampling theory, there exist methodologies to estimate average value of a parameter in the setup of overlapping clusters. This paper presents a comparison of SPP and CBP using the mathematical approach of cluster sampling techniques for network mean degree estimation. Newman and Milgram [11, 14, 13] suggested with evidence of why using the concept of shortest path for sampling social networks. Newman's [11] experiment on scientific collaboration shows that on average 64% scientists collaborator shortest path pass through one's top-ranked collaborator and 17% pass through the second-ranked one. Milgram's [14] [13] experiment of small-world phenomena concludes that delivering a message from one person to another by using shortest path based on local information exist in large social networks and that by using only local information. In general, in social networks, information [5] propagates along the shortest paths of users as a direct and simple way to communicate. For example, smart advertisement of products with minimum cost by maximum influence path. Above discussion motivates to take the cliques and shortest paths [3] as the building blocks to sampled network. By using it one can estimate different network parameters and at the same time can preserve the network functionalities. A clique may be an alternative of shortest path and need to be examined.

3. Estimation of Average Degree using Overlapping Cluster Sampling

Let there are total K clusters, many of them are having overlapping vertices of different degrees formed by appropriate computational method. The i^{th} cluster (i = 1, 2, 3, ..., K) contains N_i units. Suppose the term Y_{ij} denotes degree of j^{th} vertex belonging to i^{th} clusters and F_i be the frequency of j^{th} vertex occurring in K clusters. Total distinct vertices in the network graph are N.

Step 1: Let k out of K (k< K) clusters are selected randomly who are formed either by method SPP or by CBP.

Step 2: From the i^{th} cluster of size N_i, the $n_i(n_i < N_i)$ vertices are selected by SRSWR.

Define $D_{ij} = \frac{MY_{ij}}{NF_j}$, i = 1,2, ..., K and j = 1, 2, ...,N_i,
where, M denotes total vertices in all K cluster (including overlapping, M > N). The D_{ij} indicates that degree values Y_{ij} at K clusters are normalized and converted. Overall average unknown network parameter is:

$$\bar{D} = \frac{1}{K}\sum_{i=1}^{K}\frac{1}{N_i}\sum_{j=1}^{N_i} D_{ij} \tag{1}$$

Theorem 1. A biased estimator of average $\bar{D}$ is given by [2]

$$\bar{d} = \frac{1}{k}\sum_{i=1}^{k}\frac{1}{n_i}\sum_{j=1}^{n_i} d_{ij} \tag{2}$$

where, d_{ij} represents D_{ij} units who are present in sample n_i.

Proof. Let us consider (see [2])
E_2 = The conditional expectation over a given sample of cluster
E_1 = The expectations for over all such sample,

$$\begin{aligned} E(\bar{d}) &= \left(\tfrac{1}{k}\textstyle\sum_{i=1}^{k}\tfrac{1}{n_i}\sum_{j=1}^{n_i} d_{ij}\right) \\ &= E_1E_2\left(\tfrac{1}{k}\textstyle\sum_{i=1}^{k}\tfrac{1}{n_i}\sum_{j=1}^{n_i} d_{ij}\right) \\ &= E_1\left(\tfrac{1}{k}\textstyle\sum_{i=1}^{k} E_2(\bar{d}_{i\bullet})\right) \\ &\text{where, } \bar{d}_{i\bullet} = \text{sample within clusters} \\ &= E_1\left(\tfrac{1}{k}\textstyle\sum_{i=1}^{k} E_2(\bar{D}_{i\bullet})\right) = \bar{D} \neq \bar{Y} \end{aligned}$$

Hence the theorem. ■

Note 3.1 The $\bar{d}$ is a biased estimator of $\bar{Y}$ and its bias is given by $Bias(\bar{d}) = E(\bar{d}) - \bar{Y}$
This bias can be estimated by [2]

$$\widehat{Bias}(\bar{d}) = \frac{K-1}{K\bar{N}(k-1)}\sum_{i=1}^{k}(N_i - \bar{n})(\bar{d}_{i\bullet} - \bar{d}) \tag{3}$$

3.0.1 Estimation of Variance:

Consider average square between cluster averages in the sample is

$$s_b^2 = \frac{1}{k-1}\sum_{i=1}^{k}\left(\bar{d}_{i\bullet} - \bar{d}\right)^2 \tag{4}$$

It can be shown that

$$E(s_b^2) = S_b^2 + \frac{1}{K}\sum_{i=1}^{K}\left(\frac{1}{n_i} - \frac{1}{N_i}\right)S_i^2 \tag{5}$$

Also, one can define

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (d_{ij} - \bar{d}_{i\bullet})^2 \tag{6}$$

Further,

$$E(s_i^2) = S_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (d_{ij} - \bar{D}_i)^2 \tag{7}$$

So,

$$E\left[\frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) s_i^2\right] = \frac{1}{K}\sum_{i=1}^{K}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) S_i^2 \tag{8}$$

Thus, one can express

$$E(s_b^2) = S_b^2 + E\left[\frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) s_i^2\right] \tag{9}$$

and an unbiased estimator of S_b^2 is

$$\widehat{S_b^2} = s_b^2 - \frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right) s_i^2 \tag{10}$$

Also, an estimator of the variance can be obtained by replacing S_b^2 and S_i^2 by their unbiased estimators as:

$$\widehat{Var}(\bar{d}) = \left(\frac{1}{k} - \frac{1}{K}\right)\widehat{S_b^2} + \frac{1}{kK}\sum_{i=1}^{k}\left(\frac{1}{n_i} - \frac{1}{N_i}\right)\widehat{S_i^2} \tag{11}$$

3.0.2 Confidence Interval (CI)

Let a and b are the two real numbers and P(A) denotes the probability of an event A. The 95% confidence interval is defined as $P[a < \theta < b] = 0.95$, where θ is an unknown parameter. As per theory of normal distribution the best choice of a and b is

a = Estimated average -1.96$\sqrt{\text{Estimated variance}}$,
b = Estimated average + 1.96$\sqrt{\text{Estimated variance}}$.

4. Proposed Sampling Scheme And Dataset

To evaluate amd compare the two methods CBP and SPP sampling the well known Zachary's Karate Club [15] network datasets have taken into account. Zachary network are widely used to study the efficiency of different graph sampling techniques.

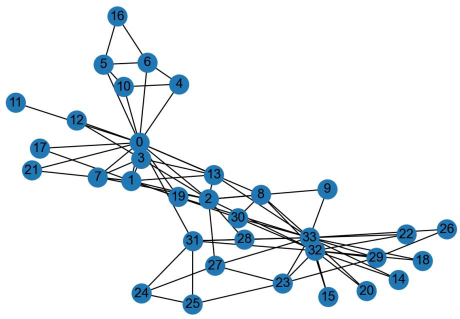


Figure 1: *Karate Club Network [15].*

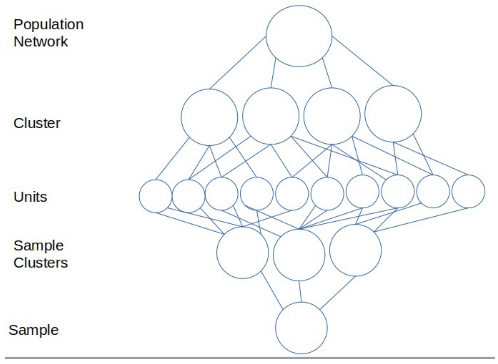


Figure 2: *Overlapping cluster sampling scheme diagram.*

Table 1: *Dataset Description.*

Description of Dataset(network)			
Network	vertex	Edge	Description
Karate	34	78	Zachary Karate Club Network [15]

4.1. Clique Based Procedure (CBP)

The computational procedure(CBP) prposed by authors is as under:
Step 1: Choose randomly K non-adjacent vertices vx_s vertices.
Step 2: Find cliques using vx_s as source vertex.
Step 3: Take degree of vertices of cliques in overlapping cluster.
Step 4: By SRSWR rule, choose k overlapping clusters from K clusters.
Step 5: By SRSWR rule, select sample of n_i vertices from N_i vertices among k clusters.

Table 2: *Clique of random vertices in Karate Club graph*

Cliques of vertices in Karate Club graph			
Serial No.	Vertices	Cliques	Degree sequence
S_1	vx_1	$[vx_0, vx_1, vx_2, vx_3, vx_7]$	(16, 9, 10, 6, 4)
S_2	vx_3	$[vx_0, vx_1, vx_2, vx_3, vx_{13}]$	(16, 9, 10, 6, 5)
S_3	vx_{15}	$[vx_{33}, vx_{32}, vx_{15}]$	(17, 12, 2)
S_4	vx_{28}	$[vx_{33}, vx_{28}, vx_{31}]$	(17, 3, 6)
S_5	vx_{22}	$[vx_{33}, vx_{32}, vx_{22}]$	(17, 12, 2)
S_6	vx_9	$[vx_2, vx_9]$	(10, 2)
S_7	vx_5	$[vx_5, vx_{16}, vx_6]$	(4, 2, 4)
S_8	vx_{10}	$[vx_0, vx_4, vx_{10}]$	(16, 3, 3)
S_9	vx_{12}	$[vx_0, vx_{12}, vx_3]$	(16, 2, 6)
S_{10}	vx_{30}	$[vx_{33}, vx_{32}, vx_8, vx_{30}]$	(17, 12, 5, 4)
S_{11}	vx_{14}	$[vx_{33}, vx_{32}, vx_{14}]$	(3, 3, 6)
S_{12}	vx_{23}	$[vx_{33}, vx_{27}, vx_{23}]$	(17, 4, 5)
S_{13}	vx_{26}	$[vx_{33}, vx_{26}, vx_{29}]$	(17, 2, 4)
S_{14}	vx_{20}	$[vx_{33}, vx_{32}, vx_{20}]$	(17, 12, 2)
S_{15}	vx_{11}	$[vx_0, vx_{11}]$	(16, 1)
S_{16}	vx_{21}	$[vx_0, vx_1, vx_{21}]$	(16, 9, 2)
S_{17}	vx_{19}	$[vx_0, vx_1, vx_{19}]$	(16, 9, 3)
S_{18}	vx_{31}	$[vx_{24}, vx_{25}, vx_{31}]$	(3, 3, 6)
S_{19}	vx_{17}	$[vx_0, vx_1, vx_{17}]$	(16, 9, 2)
S_{20}	vx_{18}	$[vx_{33}, vx_{32}, vx_{18}]$	(17, 12, 2)

4.2. Shortest Path Based Procedure (SPP)

Computaional procedure(SPP) existing in literature due to Dijkstra's algorithm [10] is as under:

Step 1: Choose randomly K pairs of non-adjacent vertices vx_s as source vertex and vx_d as destination vertex.
Step 2: Find shortest path between K pairs of vertices using shortest path algorithm through Dijkstra's algorithm [10].
Step 3: Degree sequence is formed to each vertex appearing in the computed shortest path.
Step 4: Take degree sequence as overlapping clusters which divide the graph vertices.

Step 5: By SRSWR rule, choose k overlapping clusters from K clusters.
Step 6: By SRSWR rule, choose sample n_i vertices from N_i among k clusters.

Table 3: *Shortest path of random pair of vertices in Karate Club graph*

Shortest path of random pair of vertices in Karate Club graph			
Serial No.	Pairs of vertices	Shortest Path	Degree sequence
S_1	(vx_{16}, vx_{26})	$[vx_{16}, vx_5, vx_0, vx_8, vx_{33}, vx_{26}]$	(2, 4, 16, 5, 17, 2)
S_2	(vx_{16}, vx_{26})	$[vx_{16}, vx_6, vx_0, vx_{19}, vx_{33}, vx_{26}]$	(2, 4, 16, 3, 17, 2)
S_3	(vx_{29}, vx_{12})	$[vx_{29}, vx_{32}, vx_2, vx_0, vx_{12}]$	(4,12,10,16,2)
S_4	(vx_{28}, vx_{16})	$[vx_{28}, vx_2, vx_0, vx_5, vx_{16}]$	(3, 10, 16, 4, 2)
S_5	(vx_{10}, vx_{15})	$[vx_{10}, vx_0, vx_{19}, vx_{33}, vx_{15}]$	(3,16, 3, 17,2)
S_6	(vx_{26}, vx_2)	$[vx_{26}, vx_{29}, vx_{32}, vx_2]$	(2, 4, 12, 10)
S_7	(vx_{25}, vx_7)	$[vx_{25}, vx_{31}, vx_0, vx_7]$	(3, 6, 16, 4)
S_8	(vx_{23}, vx_4)	$[vx_{23}, vx_{25}, vx_{31}, vx_0, vx_4]$	(5, 3, 6, 16, 3)
S_9	(vx_9, vx_{24})	$[vx_9, vx_2, vx_{27}, vx_{24}]$	(2,10, 4, 3)
S_{10}	(vx_{22}, vx_{24})	$[vx_{22}, vx_{32}, vx_{31}, vx_{24}]$	(2, 12, 6, 3)
S_{11}	(vx_{21}, vx_{27})	$[vx_{21}, vx_0, vx_2, vx_{27}]$	(2, 16, 10, 4)
S_{12}	(vx_{18}, vx_{25})	$[vx_{18}, vx_{32}, vx_{23}, vx_{25}]$	(2, 12, 5, 3)
S_{13}	(vx_{18}, vx_4)	$[vx_{18}, vx_{32}, vx_2, vx_0, vx_4]$	(2, 12, 10, 16, 3)
S_{14}	(vx_{17}, vx_5)	$[vx_{17}, vx_0, vx_2, vx_{32}, vx_1, vx_5]$	(2, 16, 10, 12, 9, 4)
S_{15}	(vx_{30}, vx_{25})	$[vx_{30}, vx_{32}, vx_{23}, vx_{25}]$	(4,12, 5, 3)
S_{16}	(vx_2, vx_{26})	$[vx_2, vx_8, vx_{33}, vx_{26}]$	(10, 5, 17, 2)
S_{17}	(vx_1, vx_{20})	$[vx_1, vx_2, vx_{32}, vx_{20}]$	(9, 10, 12, 2)
S_{18}	(vx_4, vx_8)	$[vx_4, vx_0, vx_2, vx_{32}, vx_1, vx_8]$	(3, 16, 10, 12, 9, 5)
S_{19}	(vx_3, vx_{26})	$[vx_3, vx_{13}, vx_{33}, vx_{26}]$	(6, 5, 17, 2)
S_{20}	(vx_{14}, vx_{11})	$[vx_{14}, vx_{32}, vx_{31}, vx_0, vx_{11}]$	(2,12, 6, 16, 1)

4.3. Frequency table of vertices in overlapping clusters

Table 4: *Frequency of vertices occuring in K-clusters*

Frequency of vertices in K-clusters, N = 34, $M_{sp} = 94, M_{cl} = 63$.											
vx	Y_{ij}	F_{ij}	F'_{ij}	D_{ij}	D'_{ij}	**vx**	Y_{ij}	F_{ij}	F'_{ij}	D_{ij}	D'_{ij}
vx_0	16	12	8	3.69	3.70	vx_{17}	2	1	1	5.53	3.70
vx_1	9	3	5	8.29	3.33	vx_{18}	2	2	1	2.76	3.70
vx_2	10	10	3	2.76	6.18	vx_{19}	3	2	1	4.15	5.56
vx_3	6	1	3	16.59	3.70	vx_{20}	2	1	1	5.53	3.70
vx_4	3	3	1	2.76	5.56	vx_{21}	2	1	1	5.53	3.70
vx_5	4	3	1	3.69	7.41	vx_{22}	2	1	1	5.53	3.70
vx_6	4	1	1	11.06	7.41	vx_{23}	5	3	1	4.61	9.26
vx_7	4	1	1	11.06	7.41	vx_{24}	3	2	1	4.15	5.56
vx_8	5	3	1	4.61	9.26	vx_{25}	3	4	1	2.07	5.56
vx_9	2	1	1	5.53	3.71	vx_{26}	2	5	1	1.11	3.70
vx_{10}	3	1	1	8.29	5.56	vx_{27}	4	2	1	5.53	7.41
vx_{11}	1	1	1	2.76	1.85	vx_{28}	3	1	1	8.29	5.56
vx_{12}	2	1	1	5.53	3.71	vx_{29}	4	2	1	5.53	7.41
vx_{13}	5	1	1	13.82	9.26	vx_{30}	4	1	1	11.06	7.41
vx_{14}	2	1	1	5.53	3.70	vx_{31}	6	4	2	4.15	5.56
vx_{15}	2	1	1	5.53	3.70	vx_{32}	12	10	6	3.32	3.70
vx_{16}	2	3	1	1.84	3.70	vx_{33}	17	5	9	9.4	3.5

In Table-2 and Table-3, overlapping clusters were collected using CBP and SPP. In which many vertices lie in clusters repeatedly. To use overlapping cluster sampling, degree values of vertices are normalized using its frequency by

$$D_{ij} = \frac{MY_{ij}}{NF_j}, \text{ i = 1, 2,..., K and j = 1, 2, ..., } N_i$$

where,
N = Total number of distinct vertices in network.
M = Total number of vertices in overlapping clusters.
M_{sp} = Total number of vertices in overlapping clusters obtained by SPP.
M_{cl} = Total number of vertices in overlapping clusters obtained by CBP.
Y_{ij} = Degree of vertices in network.
F_{ij} = Frequency of vertices in clusters formed by SPP.
F'_{ij} = Frequency of vertices in clusters formed by CBP.
D_{ij} = Normalised degree of vertices in clusters created by SPP.
D'_{ij} = Normalised degree of vertices in clusters created by CBP.

5. Experimental Results

5.1. Ogive Based Simulation Procedure

Step 1: Draw sample of k clusters by SRSWR from K clusters(k < K).

Step 2: Draw sample of n_i vertices of second stage units from N_i among k clusters.

Step 3: Calculate lower limit and upper limit of confidence interval(CI).

Step 4: Repeat step I, II and III for P times(P is positive integer).

Step 5: Draw two ogive curves separately for lower limit and upper limit of confidence intervals.

Step 6: Draw a perpendicular from point of intersections of two ogive curves to find lower and upper limits of CI.

5.2. Numerical Illustration

Consider the Karate Club Network datasets (figure-1) which has N=34 identifiable distinct units(vertices). In Table-2 and Table-3 using method CBP and SPP overlapping clusters based on cliques and shortest path are obtained which contain each unit of the network. The objective is to estimate average degree of network and relative efficiency of estimate using confidence interval size. For numerical evaluation, one can take sample in two stages (figure-2). In the first stage sample of size k = 15 clusters are taken from K = 20 clusters. In the second stage sample of vertices from each clusters are taken randomly. Further, one can use ogive simulation for P times.

5.3. Shortest Path Based Procedure for Parameter Estimation [10]

A sample of cluster of size k = 15 of size K = 20 (Table-3) is taken by SRSWR and in each overlapping sampled cluster, a percentage of sample vertices are chosen randomly to calculate confidence intervals and average degree.

Table 5: *Sample cluster units (by SPP)*

Sampled vertices and degree sequence using SPP			
Serial No.	Sample pair of vertices	Shortest Path	Normalised Degree sequence
S_1	(vx_{16}, vx_{26})	$[vx_5, vx_0, vx_8, vx_{26}]$	(3.69, 3.69, 4.61, 1.11)
S_2	(vx_{16}, vx_{26})	$[vx_{16}, vx_6, vx_{19}, vx_{33}]$	(1.84, 11.06, 4.15, 9.4)
S_3	(vx_{29}, vx_{12})	$[vx_{29}, vx_2, vx_0, vx_{12}]$	(5.53, 2.76, 3.69, 5.53)
S_4	(vx_{10}, vx_{15})	$[vx_{10}, vx_{19}, vx_{33}, vx_{15}]$	(8.29, 4.15, 9.4, 5.53)
S_5	(vx_{26}, vx_2)	$[vx_{29}, vx_{32}, vx_2]$	(5.53, 3.32, 2.76)
S_6	(vx_{25}, vx_7)	$[vx_{25}, vx_{31}, vx_7]$	(2.07, 4.15, 11.06)
S_7	(vx_{23}, vx_4)	$[vx_{23}, vx_{25}, vx_0, vx_4]$	(4.61, 2.07, 3.69, 2.76)
S_8	(vx_9, vx_{24})	$[vx_9, vx_{27}, vx_{24}]$	(5.53, 5.53, 4.15)
S_9	(vx_{22}, vx_{24})	$[vx_{22}, vx_{32}, vx_{24}]$	(5.53, 3.32, 4.15)
S_{10}	(vx_{21}, vx_{27})	$[vx_{21}, vx_2, vx_{27}]$	(5.53, 2.76, 5.53)
S_{11}	(vx_{18}, vx_{25})	$[vx_{18}, vx_{32}, vx_{25}]$	(2.76, 3.32, 2.07)
S_{12}	(vx_{17}, vx_5)	$[vx_{17}, vx_0, vx_2, vx_5]$	(5.53, 3.69, 2.76, 3.69)
S_{13}	(vx_1, vx_{20})	$[vx_2, vx_{32}, vx_{20}]$	(2.76, 3.32, 5.53)
S_{14}	(vx_4, vx_8)	$[vx_4, vx_{32}, vx_1, vx_8]$	(2.76, 3.32, 8.29, 4.61)
S_{15}	(vx_{14}, vx_{11})	$[vx_{14}, vx_{32}, vx_0, vx_{11}]$	(5.53, 3.32, 3.69, 2.76)

Table 6: *Sample based Computation (for SPP)*

Sample based computation for confidence interval(using SPP)						
S. No.	Degree sequence	$\bar{d}_{(sp)i\bullet}$	$(\bar{d}_{i\bullet} - \bar{d})^2$	$s^2_{(sp)i}$	95% C.I.	CI size
S_1	(3.69, 3.69, 4.61, 1.11)	3.275	1.437	2.2713	[1.798, 4.752]	2.954
S_2	(1.84, 11.06, 4.15, 9.4)	6.6125	4.575	18.797	[2.364, 10.861]	8.497
S_3	(5.53, 2.76, 3.69, 5.53)	4.3775	0.009	1.915	[3.021, 5.734]	2.713
S_4	(8.29, 4.15, 9.4, 5.53)	6.8425	5.612	5.87	[4.468, 9.217]	4.749
S_5	(5.53, 3.32, 2.76)	3.87	0.364	2.145	[2.213, 5.527]	3.314
S_6	(2.07, 4.15, 11.06)	5.76	1.654	22.15	[0.434, 11.086]	10.652
S_7	(4.61, 2.07, 3.69, 2.76)	3.2825	1.419	1.224	[0.434, 11.086]	10.652
S_8	(5.53, 5.53, 4.15)	5.07	0.356	0.6348	[4.168, 5.972]	1.804
S_9	(5.53, 3.32, 4.15)	4.33	0.021	1.246	[3.067 ,5.593]	2.526
S_{10}	(5.53, 2.76, 5.53)	4.607	0.018	2.557	[2.798, 6.416]	3.618
S_{11}	(2.76, 3.32, 2.07)	2.717	3.086	0.392	[2.009, 3.425]	1.416
S_{12}	(5.53, 3.69, 2.76, 3.69)	3.9175	0.309	1.348	[2.780, 5.055]	2.275
S_{13}	(2.76, 3.32, 5.53)	3.87	0.364	2.145	[2.213, 5.527]	3.314
S_{14}	(2.76, 3.32, 8.29, 4.61)	4.7475	0.075	6.185	[2.308, 7.182]	4.874
S_{15}	(5.53, 3.32, 3.69, 2.76)	3.825	0.421	1.438	[2.650, 5.000]	2.35
	Average Value	$\bar{d}_{sp} = 4.4736$	$s^2_{sp} = 1.408$	$\hat{s}^2_i = 4.688$	[2.4483,6.8288]	4.3805

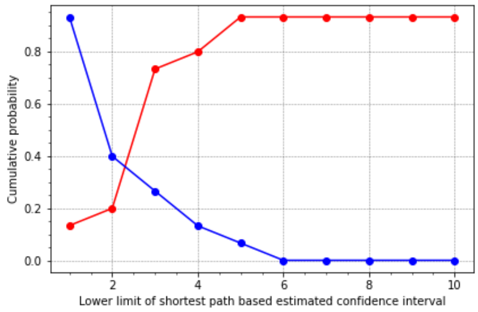


Figure 3: *Ogive for lower limit of CI for SPP.*

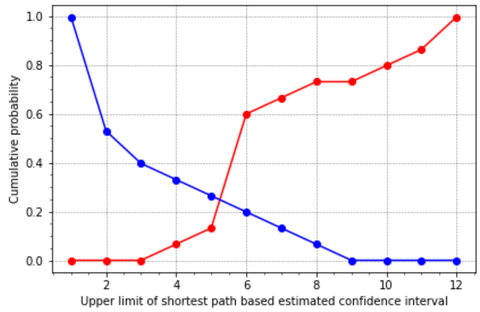


Figure 4: *Ogive for upper limit of CI for SPP.*

$$\widehat{Var}(\bar{d}_{(sp)}) = \left(\frac{1}{k} - \frac{1}{K}\right) \hat{S}_{sp}^2 + \frac{1}{kK} \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i}\right) \hat{S}_i^2$$

Estimated average degree = $\mathbf{\bar{d}_{sp}} = \mathbf{4.47}$
Estimated variance for average degree = $\widehat{Var}(\bar{d}_{(sp)})$ = 1.06
Average CI size = 4.38
A 95% confidence interval estimate using SPP for average degree is $[\mathbf{2.4483}, \mathbf{6.8288}]$.
Through ogive based simulation(figure 3 & 4) for average degree the confidence interval is $[\mathbf{2.38}, \mathbf{5.42}]$

5.4. Clique Based Procedure (CBP) for Parameter Estimation

Consider sample of cluster clique having size k = 15 of size K = 20(Table-2) by SRSWR in each overlapping sampled cluster. Herein a percentage is used to calculate confidence intervals and average degree of network.

Table 7: *Clique of random vertex in Karate Club Graph (by CBP)*

Sampled vertices of clique using CBP			
Serial No.	Sample vertices	Sample cliques	Normalised degree sequence
S_1	vx_3	$[vx_0, vx_1, vx_3, vx_{13}]$	(3.70, 3.33, 3.70, 9.26)
S_2	vx_{15}	$[vx_{32}, vx_{15}]$	(3.70, 3.70)
S_3	vx_{28}	$[vx_{33}, vx_{28}]$	(3.5, 5.56)
S_4	vx_9	$[vx_2, vx_9]$	(6.18, 3.71)
S_5	vx_5	$[vx_5, vx_{16}]$	(7.41, 3.70)
S_6	vx_{10}	$[vx_0, vx_4]$	(3.70, 5.56)
S_7	vx_{12}	$[vx_{12}, vx_3]$	(3.71, 3.70)
S_8	vx_{30}	$[vx_{33}, vx_{32}, vx_{30}]$	(3.5, 3.70, 7.41)
S_9	vx_{23}	$[vx_{33}, vx_{27}]$	(3.5, 7.41)
S_{10}	vx_{26}	$[vx_{33}, vx_{26}]$	(3.5, 3.70)
S_{11}	vx_{20}	$[vx_{33}, vx_{20}]$	(3.5, 3.70)
S_{12}	vx_{11}	$[vx_0, vx_{11}]$	(3.70, 1.85)
S_{13}	vx_{21}	$[vx_0, vx_1]$	(3.70, 3.33)
S_{14}	vx_{31}	$[vx_{24}, vx_{31}]$	(5.56, 5.56)
S_{15}	vx_{18}	$[vx_{33}, vx_{18}]$	(3.5, 3.70)

Table 8: *Sample based Computation (for CBP)*

Sample based computation for confidence interval(Using CBP)						
S. No.	Degree sequence	$\bar{d}_{(cl)i\bullet}$	$(\bar{d}_{i\bullet} - \bar{d})^2$	$s^2_{(cl)i}$	95% C.I.	CI size
S_1	(3.70, 3.33, 3.70, 9.26)	4.99	0.436	2.2713	[2.207, 7.788]	5.581
S_2	(3.70, 3.70)	3.7	0.397	0	[3.700, 3.700]	0
S_3	(3.5, 5.56)	4.53	0.04	2.122	[2.511, 6.549]	4.038
S_4	(6.18, 3.71)	4.94	0.325	3.05	[2.525, 7.365]	4.84
S_5	(7.41, 3.70)	5.55	1.488	6.882	[1.914, 9.186]	7.272
S_6	(3.70, 5.56)	4.63	0.09	1.73	[2.808, 6.452]	3.644
S_7	(3.71, 3.70)	3.70	0.397	0.0001	[3.695, 3.715]	0.02
S_8	(3.5, 3.70, 7.41)	4.87	0.292	4.849	[2.378, 7.362]	4.984
S_9	(3.5, 7.41)	5.45	1.254	7.644	[1.623, 9.287]	7.664
S_{10}	(3.5, 3.70)	3.6	0.533	0.02	[3.404, 3.796]	0.392
S_{11}	(3.5, 3.70)	3.6	0.533	0.02	[3.404, 3.796]	0.392
S_{12}	(3.70, 1.85)	2.77	2.434	1.7113	[0.962, 4.588]	3.626
S_{13}	(3.70, 3.33)	3.51	0.672	0.0685	[3.152, 3.878]	0.726
S_{14}	(5.56, 5.56)	5.56	1.513	0	[5.560, 5.560]	0
S_{15}	(3.5, 3.70)	3.6	0.533	0.02	[3.404, 3.796]	0.392
	Average value	$\bar{d}_{cl} = 4.33$	$s^2_{cl} = 0.781$	$\hat{s}^2_i = 0.4876$	[2.883, 5.787]	2.9047

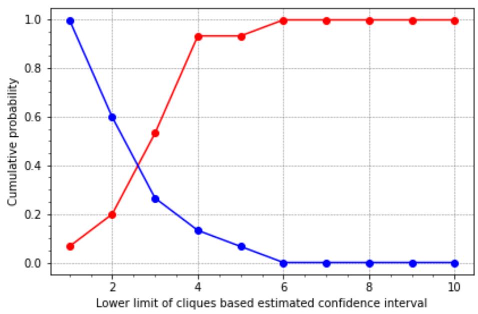


Figure 5: *Ogive for lower limit of CI by CBP.*

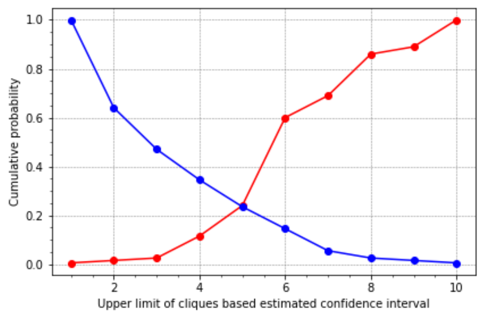


Figure 6: *Ogive for upper limit of CI by CBP.*

$$\widehat{Var}(\bar{d}) = \left(\frac{1}{k} - \frac{1}{K}\right) \hat{S^2_{cl}} + \frac{1}{kK} \sum_{i=1}^{k} \left(\frac{1}{n_i} - \frac{1}{N_i}\right) \hat{S}^2_i$$

Estimated average degree = $\bar{d}_{cl} = 4.33$
Estimated variance for average degree = $\widehat{Var}(\bar{d}_{cl}) = \mathbf{0.022813}$
Average confidence interval(CI) size = $[5.787 - 2.883] = \mathbf{2.90}$
The 95% confidence interval estimate using CBP is $[\mathbf{2.8831}, \mathbf{5.7878}]$.
Through ogive based simulation(figure 5 & 6) the confidence interval(CI) is $[\mathbf{2.63}, \mathbf{5.01}]$.

6. Comparision

The Percentage Relative Efficiency (PRE) of estimators $\bar{d}_{cl}$, $\bar{d}_{sp}$ is defined as under:

$$\textbf{PRE} = \frac{\widehat{Var}(\bar{d}_{sp}) - \widehat{Var}(\bar{d}_{cl})}{\widehat{Var}(\bar{d}_{sp})} \times 100 = \frac{1.06 - 0.0228}{1.06} \times 100 = 97.84\%$$

The Percentage Relative Gain(PRG) over the length of confidence intervals is defined as:

$$\textbf{PRG} = \frac{(\text{length of CI})_{SPP} - (\text{length of CI})_{CBP}}{(\text{length of CI})_{SPP}} \times 100 = \frac{4.3805 - 2.9047}{4.3805} \times 100 = 33.69\%$$

Using ogive based simulation, the Percentage Relative Gain is:

$$(\textbf{PRG})_{ogive} = \frac{[(\text{length of CI})_{SPP}]_{ogive} - [(\text{length of CI})_{CBP}]_{ogive}}{[(\text{length of CI})_{SPP}]_{ogive}} \times 100 = \frac{3.04 - 2.38}{3.04} \times 100 = 21.71\%$$

7. Reliability of Social Networks as an Application

As considered, the average degree estimation of a social network leads to monitoring the reliability of the network. People join the social network at any point of time and leave it at any other instant. Addition and deletion in a social network is a common continuous process. A network is said to be reliable if the average degree of social network remains controlled over the time framework. The upper limit and lower limit of confidence intervals are useful measures to make a benchmark for checking of growth or decay of social networks over the time domain.

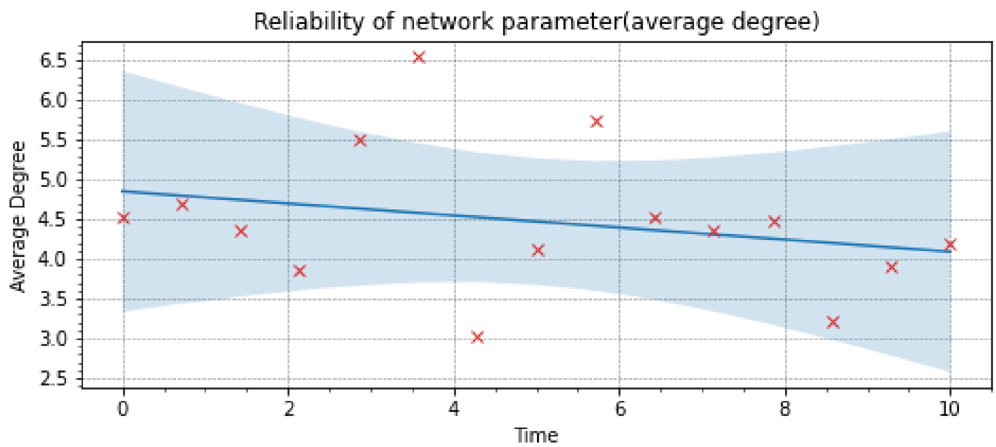


Figure 7: *Network reliabiity based on vertex degree estimate.*

8. Discussion

The two methods SPP and CBP are compared in a common setup of a social network for the objective of computation of average degree. A social network in general can be represented as a graph of vertices and edges. Clusters of vertices are formed by using both methods SPP and CBP. After comparison of percentage relative efficiency, the CPP found efficient by 97.8% over SPP. The simulated confidence interval for clique procedure (CBP) is [2.8-5.7] which is catching the true value of average degree 4.58 of vertices, which is also supported by figures 5 and 6. The same calculation for simulated confidence interval using shortest path procedure(SPP) is [2.4-6.8] which is longer than earlier (see figure 3 & 4). Ogive based simulation procedure also supports for better efficiency of [(2.38, 5.42) for SPP, (2.63, 5.01) for CBP] the proposed for network degree evaluation and the proposed is useful for network reliability (figure-7).

9. Conclusion

This paper contains an overlapping cluster sampling based comparative approach using the shortest path and cliques method over created clusters. A graphical structure has been taken under consideration representing the social network. In order to estimate the unknown parameter(like average degree), the proposed sampling method takes into account the cliques and compares with shortest paths between several pairs of vertices in a setup of the overlapping cluster of degree sequence. The proposed method is examined by conducting an experiment on a well-known real network keeping in view that the average degree is an important property of network. To evaluate the comparative statistical significance of proposed procedure CBP, the 95% confidence intervals were computed for both methods. It has been found as an outcome of the study that 95% confidence intervals contain the true value. The Ogive based simulation procedure has been implemented which shows cluster based method using clique (CBP) provides a better estimate of the parameter(average degree) than the cluster based method using shortest path (SPP). The network reliability could be monitored over the long time domain by the bench-mark values of confidence intervals. This contribution opens up new avenues and opportunities for network degree parameter estimation. One can think of the inclusion of the additional network measures for future studies that will help to bring up new insights to the development of graph sampling cluster methods. In order to have more a comprehensive evaluation of the existing social networking, the sampling methods could be considered by involving the other kinds of parametric network measures and properties.

References

[1] Cochran, W. G. (2005). Sampling Techniques, *John Willey and Sons, New York*.

[2] Singh S. (1988). Estimation in overlapping clusters, *Communications in Statistics: Theory and Methods*, 17:613-621.

[3] Rezvanian, A. and Meybodi, M. R. (2015). Sampling social networks using shortest paths, *Physica A: Statistical Mechanics and its Applications*, 424: 254-268.

[4] Zhang, LC. and Patone, M. (2017). Graph sampling. *Metron*, 75: 277-299.

[5] Alim, A. and Shukla, D. (2021). Double sampling based parameter estimation in big data and application in control charts. *Reliability: Theory & Applications*, 16(2 (62)): 72-114.

[6] Katzir, L., Liberty, E., Somekh, O. & Cosma, I. A. (2014). Estimating sizes of social networks via biased sampling, *Internet Mathematics*, 10:3-4, 335-359.

[7] Kurant, M.,Butts, C. T. and Markopoulou, A. (2012). Graph size estimation. *CoRR, abs/1210.0460*.

[8] McCormick, T. H., Moussa, A., Ruf, J., DiPrete, T. A., Gelman, A., Teitler J., and Zheng, T. (2013). A practical guide to measuring social structure using indirectly observed network data. *Journal of Statistical Theory and Practice*, 7(1):120–132.

[9] McCormick, H., Salganik, M.J. and Zheng, T. (2010). How many people do you know?: Efficiently estimating personal network size. *JASA*, 105(489):59–70.

[10] Dijkstra, E. W. (1959). A note on two problems in connexion with graphs, *Numerische Mathematik*, 1(1):269–271.

[11] Newman, M.E. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks and centrality, *Physical Review E Stat Nonlin Soft Matter Phys.*, 64(1):016132.

[12] Chen, W., Wang, C. and Wang, Y. (2010). Scalable influence maximization for prevalent viral marketing in large-scale social networks, *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovxery and Data Mining*, pp. 1029–1038.

[13] Milgram, S. (1967). The small world problem, *Psychology Today*, 2(1): 60–67

[14] Travxers, J. and Milgram, S. (1969). An experimental study of the small world problem, *Sociometry* ,32(4):425–443.

[15] Zachary, W. W. (1977). An information flow model for conflict and fission in small groups, *Journal of Anthropological Research* 33(4): 452–473.

[16] Milan, R. and Shukla, D. (2021). Kernel sampling based parameter estimation in detected community in weighted graph in big data. *Reliability: Theory* & *Applications*, 16(4 (65)), 105-120.

[17] Pandey, K. K. and Shukla, D. (2021). Stratified linear systematic sampling based clustering approach for detection of financial risk group by mining of big data. *International Journal of System Assurance Engineering and Management*, 1-15.

[18] Lee, C., Reid, F., McDaid, A., Hurley, N. (2010). Detecting highly overlapping community structure by greedy clique expansion. *SNA-KDD: Social Network Mining and Analysis* pp. 33–42.

[19] Shang, R., Luo, S., Li, Y., Jiao, L. and Stolkin, R. (2015). Large-scale community detection based on node membership grade and sub-communities integration. *Physica A: Statistical Mechanics and its Applications*, 428: 279-294.

[20] Shen, H., Cheng, X., Cai, K. and Hu, M. B. (2009). Overlapping and hierarchical community structure in networks. *Physica A: Statistical Mechanics and its Applications*, 388(8): 1706-1712.

Inverse Weibull-Burr III Distribution with Properties and Application Related to Survival Rates in Animals

Aijaz Ahmad*

•

Department of Mathematics, Bhagwant University, Ajmer, India
aijazahmad4488@gmail.com

Mujamil Jallal

•

Department of Mathematics, Bhagwant University, Ajmer, India
muzamiljallal@gmail.com

I. H. Dar

•

Department of Statistics, University of Kashmir, Srinagar, India
ishfaqqh@gmail.com

Rajnee Tripathi

•

Department of Mathematics, Bhagwant University, Ajmer, India
rajneetripathi@hotmail,com

Abstract

The objective of this study is to develop an extension of the Burr-III distribution which is achieved by adopting the inverse Weibull-G family of distribution and is referred as inverse Weibull-Burr III distribution (IWB-III) to evaluate complicated data. Different structural characteristics of the suggested distribution have been determined and analysed. Distinct plots depict the behaviour of the probability density function (pdf) and the cumulative distribution function (cdf). The maximum likelihood estimation method is applied to estimate the stated distribution parameters. To assess and investigate the efficacy of estimators in terms of bias, variance, and mean square error (MSE), a simulation study was conducted. Lastly, the effectiveness of the stated distribution is proven by an actual data set relevant to survival rates in animals.

Keywords: Inverse Weibull-G family; Burr-III distribution; moments, Renyi entropy; simulation; maximum likelihood estimation.

1. Introduction

Over numerous decades, academics have been attempting to develop a number of novel distributions to satisfy certain realistic demands. The rationale is that conventional distributions have generally been shown to lack fit in actual applications, such as medicinal research, engineering, hydrology, environmental science, and many more. In particular, the objective of creating novel distributions or generalisations is to construct adaptable statistical models effective at dealing with complicated real-world data. This adaptability may be obtained in a straightforward manner by introducing new parameters to the standard distribution.

The Weibull distribution has been utilised in a variety of disciplines and applications. The hazard function of the Weibull distribution can only be monotonic in nature. As a result, it can not be employed to simulate lifespan data with a bathtub-shaped hazard function.

Let X be a random variable that follows the Weibull distribution with parameters α and β. Then its probability density function (pdf) is defined as

$$\psi(x,\alpha,\beta) = \alpha^{\beta}\beta x^{\beta-1}e^{-\alpha^{\beta}x^{\beta}}; \quad x>0,\alpha,\beta>0$$

The transformation $T = \frac{1}{X}$, yields the inverse of the Weibull distribution. As a result, the probability density function (pdf) of the inverse Weibull distribution assumes the following structure.

$$h(t,\alpha,\beta) = \alpha^{\beta}\beta t^{-(\beta+1)}e^{-\alpha^{\beta}t^{-\beta}}; \quad t>0,\alpha,\beta>0 \tag{1}$$

In this work, we construct the inverse Weibull-Burr III distribution, which is an expansion of the Burr-III distribution. Burr, I.W [4] advocated a family of twelve cumulative distribution functions for simulating lifespan data. Burr-type III and Burr-type XII distributions were two prevalent members of the family. The Burr-III distribution has been studied thoroughly and employed in a range of aspects of research. Daniyal et al [9], Al-Dayian et al [10] and B.A. para et al [6] provide further information on the characteristics of the Burr-III distribution. The probability density function (pdf) of Burr-III distribution is stated as.

$$g(y,\theta,\lambda) = \theta\lambda y^{-\theta-1}\left(1+y^{-\theta}\right)^{-\lambda-1}; \quad y>0,\theta,\lambda>0 \tag{2}$$

The associated cumulative distribution function (cdf) of equation (1.2) is given as

$$G(y,\theta,\lambda) = \left(1+y^{-\theta}\right)^{-\lambda}; \quad y>0,\theta,\lambda>0 \tag{3}$$

In recent decades, researchers have concentrated on discovering novel generators from continuous conventional distributions. As an outcome, the resulting distribution enhances the efficacy and adaptability of data analysis. The following are some generated families of distribution: the beta-G family of distribution investigated by Eugene et al [11], the gamma-G family by Zagrofos and Balakrishana [13], the kumaraswamy-G family by Cordeiro et al [8], the transformedtransformer(T-X) by Alzaatrh et al [1], the Weibull-G by Bourguignon et al [3],Brito et al. [5] created the Topp-Leone odd log-logistic family of distributions, Morad Alizadeh et al. [12] constructed the Gompertz-G distribution family, and Amal S. Hassan et al. [2] established the inverse Weibull-G distribution.
T-X family of distributions defined by Alzaatreh et al [1] is given by

$$F(y) = \int_0^{W[G(y)]} r(t)dt \tag{4}$$

Where $r(t)$ be the probability density function of a random variable $T$and $W[G(y)]$ be a function of cumulative density function of random variable Y.
Suppose $G(y,\phi)$,denotes the baseline cumulative distribution function, which depends on parameter vector ϕ. Now using T-X approach, the cumulative distribution function $F(y)$ of inverse Weibull

generator (IWG) can be derived by replacing $r(t)$ in equation (1.4) with (1.1) and $W[G(y)] = \frac{G(y,\phi)}{\bar{G}(y,\phi)}$, where $\bar{G}(y,\phi) = 1 - G(y,\phi)$ which follows

$$\begin{aligned} F(y,\phi) &= \int_0^{\frac{G(y,\phi)}{\bar{G}(y,\phi)}} \alpha^\beta \beta t^{-(\beta+1)} e^{-\alpha^\beta t^{-\beta}} dy \\ &= e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}} \quad ; y > 0, \alpha, \beta, \phi > 0 \end{aligned} \tag{5}$$

where $\bar{G}(y,\phi) = 1 - G(y,\phi)$
The associated pdf of (1.5) is given as

$$f(y,\alpha,\phi) = \alpha^\beta \beta g(y,\phi) \frac{(G(y,\phi))^{-\beta-1}}{(\bar{G}(y,\phi))^{-\beta+1}} e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}} \quad ; y > 0, \alpha, \beta, \phi > 0 \tag{6}$$

In addition, Reliability function denoted as $\bar{F}(y,\phi)$, hazard rate function denoted as $h(y,\phi)$ and cumulative hazard rate function denoted as $H(y,\phi)$ are respectively given as

$$\begin{aligned} \bar{F}(y,\phi) &= 1 - e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}} \\ h(y,\phi) &= \frac{\alpha^\beta \beta g(y,\phi) \frac{(G(y,\phi))^{-\beta-1}}{(\bar{G}(y,\phi))^{-\beta+1}} e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}}}{1 - e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}}} \\ H(y,\phi) &= -\, ln[\bar{F}(y,\phi)] = -ln \left\{ 1 - e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}} \right\} \end{aligned}$$

1.1. Usefull Expansions

Applying Taylor series expansion to the exponential function of the pdf in equation (1.6) we have

$$e^{-\alpha^\beta \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}} = \sum_{s=0}^{\infty} \frac{(-1)^s}{s!} \alpha^{s\beta} \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta s} \tag{7}$$

substituting equation (1.7) in equation (1.6), we have

$$f(y,\phi) = \sum_{s=0}^{\infty} \frac{(-1)^s}{s!} \alpha^{\beta(s+1)} \beta g(y,\phi) \frac{(G(y,\phi))^{-\beta(s+1)-1}}{(\bar{G}(y,\phi))^{-\beta(s+1)+1}} \tag{8}$$

$$= \sum_{s=0}^{\infty} \frac{(-1)^s}{s!} \alpha^{\beta(s+1)} \beta g(y,\phi) \left(G(y,\phi)\right)^{-\beta(s+1)-1} \left(1 - G(y,\phi)\right)^{\beta(s+1)-1} \tag{9}$$

using generalised binomial theorem, we have

$$\begin{aligned} (1-z)^{a-1} &= \sum_{p=0}^{\infty} (-1)^p \binom{a-1}{p} z^p \\ (\bar{G}(y,\phi))^{-\beta(s+1)+1} &= (1 - G(y,\phi))^{\beta(p+1)-1} = \sum_{p=0}^{\infty} (-1)^p \binom{\beta(p+1)-1}{p} (G(y,\phi))^p \\ &= \sum_{s=0}^{\infty} \sum_{p=0}^{\infty} \frac{(-1)^{s+p}}{s!} \binom{\beta(s+1)-1}{p} \alpha^{\beta(s+1)} \beta g(y,\phi) (G(y,\phi))^{p-\beta(s+1)-1} \\ &= \sum_{s=0}^{\infty} \sum_{p=0}^{\infty} \zeta_{s,p} g(y,\phi) (G(y,\phi))^{p-\beta(s+1)-1} \end{aligned} \tag{10}$$

where

$$\zeta_{s,p} = \frac{(-1)^{s+p}}{s!} \binom{\beta(s+1)-1}{p} \alpha^{\beta(s+1)} \beta$$

Now using equation (2) and (3) in equation (10), we obtain pdf of formulated distribution in mixture form, as follows

$$f(y,\alpha,\beta,\theta,\lambda) = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty} \zeta_{s,p}\theta\lambda y^{-\theta-1}\left(1+y^{-\theta}\right)^{-\lambda-1}\left[(1+y^{-\theta})^{-\lambda}\right]^{p-\beta(s+1)-1}$$
$$;y>0,\alpha,\beta,\theta,\lambda>0 \qquad (11)$$

2. Inverse Weibull-Burr III Distribution

In this part, we'll investigate the inverse Weibull-Burr III distribution and look at aspects of its statistical characteristics. We derive the cumulative distribution function (cdf) of the given distribution using equation (3) in equation (5) as follows.

$$F(y,\alpha,\beta,\theta,\lambda) = e^{-\alpha^{\beta}((1+y^{-\theta})^{\lambda}-1)^{\beta}} \qquad (12)$$

Figure 1: Expounds some of possible layouts of the cdf of IWB-III distribution for distinct choice of parameters

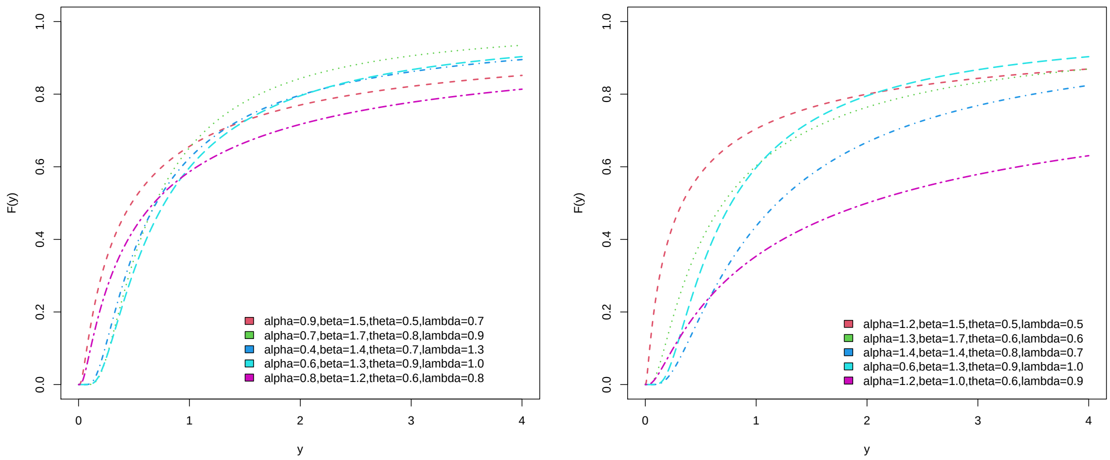


Figure 1: *plots of cdf for IWB-III distribution*

The associated pdf of (12) is given as

$$f(y,\alpha,\beta,\theta,\lambda) = \alpha^{\beta}\beta\theta\lambda y^{-\theta-1}(1+y^{-\theta})^{\lambda-1}((1+y^{-\theta})^{\lambda}-1)^{\beta-1}e^{-\alpha^{\beta}((1+y^{-\theta})^{\lambda}-1)^{\beta}} \qquad (13)$$
$$;\quad \alpha>0,\beta>0,\theta>0,\lambda>0$$

Figure 2: Expounds some of possible layouts of the pdf of IWB-III distribution for distinct choice of parameters

3. Reliability Measures of (IWB-III) Distribution

This section is focused on researching and developing distinct ageing indicators for the formulated distribution.

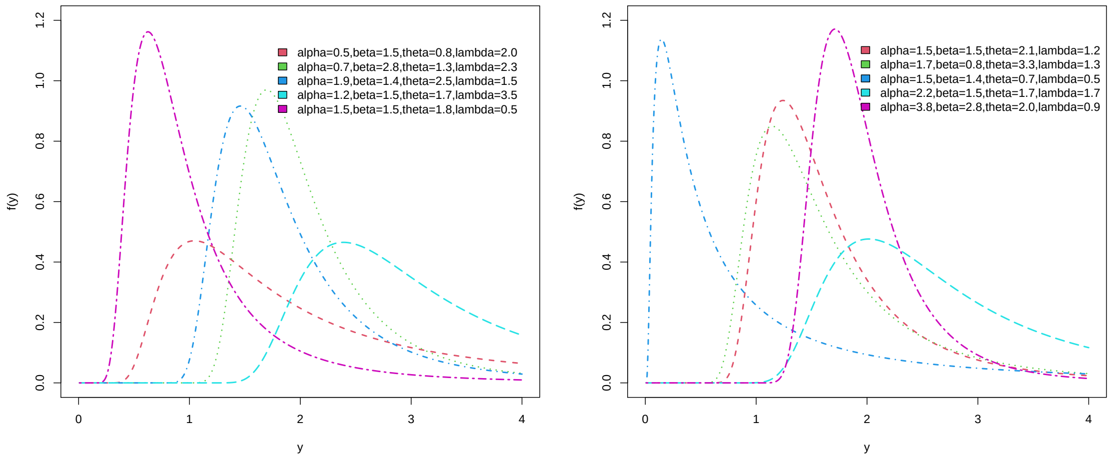


Figure 2: *plots of pdf for IWB-III distribution*

3.1. Survival function

Suppose Y be a continuous random variable with cdf $F(y)$.Then its Survival function which is also called reliability function is defined as

$$S(y) = p_r(Y > y) = \int_y^\infty f(y)dy = 1 - F(y)$$

Therefore, the survival function for IWB-III distribution is given as

$$\begin{aligned} S(y, \alpha, \beta, \theta, \lambda) =& 1 - F(y, \alpha, \beta, \theta, \lambda) \\ =& 1 - e^{-\alpha^\beta((1+y^{-\theta})^\lambda - 1)^\beta} \end{aligned} \tag{14}$$

3.2. Hazard rate function

The hazard rate function of a random variable y is denoted as

$$h(y, \alpha, \beta, \theta, \lambda) = \frac{f(y, \alpha, \beta, \theta\lambda)}{F(y, \alpha, \beta, \theta\lambda)} \tag{15}$$

using equation (12) and (13) in equation (15), then the hazard rate function of IWB-III distribution is given as

$$h(y, \alpha, \beta, \theta\lambda) = \frac{\alpha^\beta \beta\theta\lambda y^{-\theta-1}(1 + y^{-\theta})^{\lambda-1}((1 + y^{-\theta})^\lambda - 1)^{\beta-1} e^{-\alpha^\beta((1+y^{-\theta})^\lambda - 1)^\beta}}{1 - e^{-\alpha^\beta((1+y^{-\theta})^\lambda - 1)^\beta}} \tag{16}$$

Figure 3: Expounds some of possible layouts of the hazard function of IWB-III distribution for distinct choice of parameters

3.3. Cumulative hazard rate function

The cumulative hazard rate function of a random variable y is given as

$$H(y, \alpha, \beta, \theta, \lambda) = -\,ln[\bar{F}(y, \alpha, \beta, \theta, \lambda)] \tag{17}$$

using equation (12) in equation (17), then we obtain cumulative hazard rate function of IWB-III distribution

$$H(y, \alpha, \beta, \theta, \lambda) = -\,ln\left\{1 - e^{-\alpha^\beta((1+y^{-\theta})^\lambda - 1)^\beta}\right\} \tag{18}$$

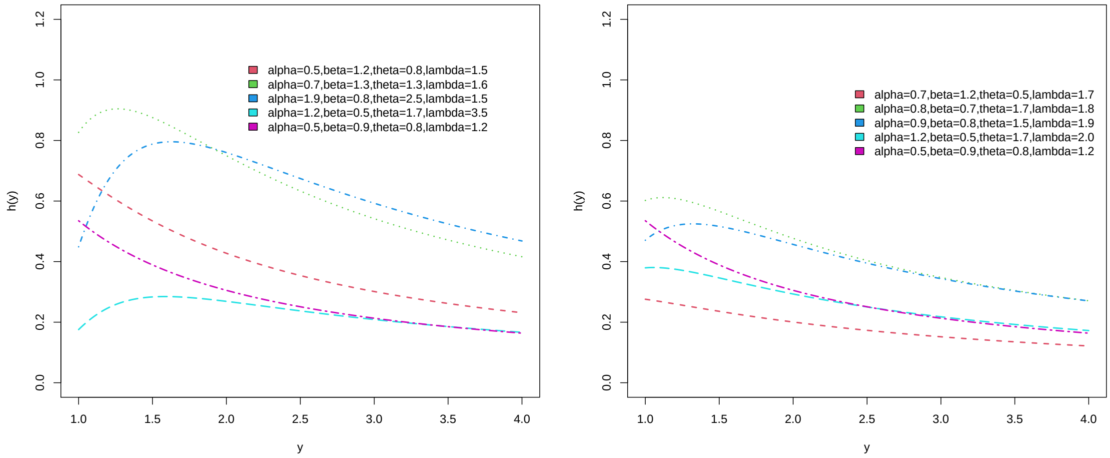


Figure 3: *plots of hazard function for IWB-III distribution*

3.4. Mean residual function

The mean residual lifetime is the predicted residual life or the average completion period of the constituent after it has exceeded a certain duration y. It is extremely significant in reliability investigations.
Mean residual function of random y variable can be obtained as

$$\begin{aligned} m(y,\alpha,\beta,\theta,\lambda) =& \frac{1}{S(y,\alpha,\beta,\theta,\lambda)} \int_y^\infty tf(t,\alpha,\beta,\theta,\lambda)dt - y \\ =& \frac{1}{\left\{1 - e^{-\alpha^\beta((1+y^{-\theta})^\lambda - 1)^\beta}\right\}} \sum_{s=0}^\infty \sum_{p=0}^\infty \zeta_{s,p}\theta\lambda \\ & \times \int_y^\infty t^{-\theta}\left(1+t^{-\theta}\right)^{-\lambda-1}\left[(1+t^{-\theta})^{-\lambda}\right]^{p-\beta(s+1)-1} dt - y \end{aligned}$$

Making substitution $(1+t^{-\theta})^{-\lambda} = z$, sothat $(1+y^{-\theta})^{-\lambda} \leq z \leq 1$, we have

$$m(y,\alpha,\beta,\theta,\lambda) = \int_{(1+y^{-\theta})^{-\lambda}}^1 z^{p-\beta(s+1)+\frac{1}{\lambda\theta}-1}(1-z^{\frac{1}{\lambda}})^{-\frac{1}{\theta}}dz$$

After solving the integral, we get

$$B\left(1-(1+y^{-\theta})^{-1},(p-\beta(s+1))\lambda + \frac{1}{\theta}, 1-\frac{1}{\theta}\right)$$

Where $B(x,a,b) = \int_0^x u^{a-1}(1-u)^{b-1}$denotes incomplete beta function

4. Statistical Properties Of (IWB-III) Distribution

This section is devoted to derive and examine disttinct properties of IWB-III

4.1. Moments

Let y denotes a random variable, then the r^{th} moment of IWB-III is denoted as μ_r' and is given by

$$\begin{aligned}\mu_r' =& E(y^r) = \int_0^\infty y^r f(y,\alpha,\beta,\theta,\lambda)dy \\ =& \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\theta\lambda\int_0^\infty y^r(1+y^{-\theta})^{-\lambda-1}\left[(1+y^{-\theta})^{-\lambda}\right]^{p-\beta(s+1)-1}dy\end{aligned}$$

Making substitution $(1+y^{-\theta})^{-\lambda}=z$, so that $0<z<1$, we have

$$\mu_r' = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\int_0^1 z^{p-\beta(s+1)+\frac{r}{\lambda\theta}-1}\left(1-z^{\frac{1}{\lambda}}\right)^{\frac{-r}{\theta}}dy$$

After solving the integral, we have

$$\mu_r' = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\lambda B\left((p-\beta(s+1))\lambda+\frac{r}{\theta},1-\frac{r}{\theta}\right)$$

Where $B(.)$ denotes incomplete beta function.

4.2. Moment generating function

suppose Y denotes a random variable follows IWB-III distribution. Then the moment generating function of the distribution denoted by $M_Y(t)$is given

$$\begin{aligned}M_Y(t) = E(e^{ty}) &= \int_0^\infty e^{ty} f(y,\alpha,\beta,\theta,\lambda)dy \\ &= \int_0^\infty\left(1+ty+\frac{(ty)^2}{2!}+\frac{(ty)^3}{3!}+....\right)f(y,\alpha,\beta,\theta,\lambda)dy \\ &= \sum_{r=0}^{\infty}\frac{t^r}{r!}\int_0^\infty y^r f(y,\alpha,\beta,\theta,\lambda)dy \\ &= \sum_{r=0}^{\infty}\frac{t^r}{r!}E(y^r) \\ &= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\frac{t^r}{r!}\zeta_{s,p}\lambda B\left((p-\beta(s+1))\lambda+\frac{r}{\theta},1-\frac{r}{\theta}\right)\end{aligned}$$

The characteristics function of the IWB-III distribution denoted as $\phi_Y(t)$ can be yeild by replacing $t=it$ where $i=\sqrt{-1}$

$$\phi_Y(t) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\frac{(it)^r}{r!}\zeta_{s,p}\lambda B\left((p-\beta(s+1))\lambda+\frac{r}{\theta},1-\frac{r}{\theta}\right)$$

4.3. Incomplete moments

The general expression for incomplete moments is given as

$$\begin{aligned}m(y) &= \int_0^y y^r f(y,\alpha,\beta,\theta,\lambda)dy \\ &= \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\theta\lambda\int_0^y y^{r-\theta-1}\left(1+y^{-\theta}\right)^{-\lambda-1}\left[(1+y^{-\theta})^{-\lambda}\right]^{p-\beta(s+1)-1}dy\end{aligned}$$

Making substitution $\left(1+y^{-\theta}\right)^{-\lambda}=z$, so that $0\leq z\leq\left(1+y^{-\theta}\right)^{-\lambda}$, we have

$$=\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\int_{0}^{(1+y^{-\theta})^{-\lambda}} z^{j-\beta(s+1)+\frac{r}{\theta\lambda}-1}\left(1-z^{\frac{1}{\lambda}}\right)^{-\frac{r}{\theta}}dy$$

After solving the integral, we get

$$m(y)=\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\lambda B\left(1-(1+y^{-\theta})^{-1};(p-\beta(s+1))\lambda+\frac{r}{\theta},1-\frac{r}{\theta}\right)$$

where $B(.)$ denotes the incomplete beta function.

4.4. Quantile function

The quantile function of a random variable Y, where $Y\sim IWB-III$ distribution can be obtained by inverting equation (12), we have

$$y_q=\left\{\left[1+\left(-\frac{1}{\alpha^{\beta}}log(q)\right)^{\frac{1}{\beta}}\right]^{\frac{1}{\lambda}}-1\right\}^{-\frac{1}{\theta}}$$

In particular, the median of the distribution can be obtained by setting $q=0.5$, we have

$$y_{0.5}=\left\{\left[1+\left(-\frac{1}{\alpha^{\beta}}log(0.5)\right)^{\frac{1}{\beta}}\right]^{\frac{1}{\lambda}}-1\right\}^{-\frac{1}{\theta}}$$

4.5. Random number generation

Suppose y denotes a random variable with pdf given in equation (2.1) . The random number of IWB-III distribution can be generated as

$$\begin{aligned} F(y)=&u \implies y=F^{-1}(u)\\ y=&\left\{\left[1+\left(-\frac{1}{\alpha^{\beta}}log(u)\right)^{\frac{1}{\beta}}\right]^{\frac{1}{\lambda}}-1\right\}^{-\frac{1}{\theta}}\end{aligned}$$

Where u is the uniform random variable defined in an open interval (0,1) .

4.6. Mean deviation about mean and median

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 IWB-III distribution with mean μ. Then the mean deviation from mean is defined as.

$$\begin{aligned} D(\mu)=&E\left(|Y-\mu|\right)\\ =&\int_0^{\infty}|Y-\mu|f(y)dy\\ =&\int_0^{\mu}(\mu-y)\,f(y)dy+\int_{\mu}^{\infty}(y-\mu)\,f(y)dy\\ =&2\mu F(\mu)-2\int_0^{\mu}yf(y)dy \end{aligned}\tag{19}$$

Now

$$\int_0^{\mu} yf(y)dy = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\theta\lambda \int_0^{\mu} y^{-\theta-1}\left(1+y^{-\theta}\right)^{-\lambda-1}\left[\left(1+y^{-\theta}\right)^{-\lambda-1}\right]^{p-\beta(s+1)-1} dy$$

Making substitution $\left(1+y^{-\theta}\right)^{-\lambda}=z$, so that $0\le z\le \left(1+\mu^{-\theta}\right)^{-\lambda}$, we have

$$\int_0^{\mu} yf(y)dy = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\int_0^{(1+\mu^{-\theta})^{-\lambda}} z^{p-\beta(s+1)+\frac{1}{\theta\lambda}-1}\left(1-z^{\frac{1}{\lambda}}\right)^{-\frac{1}{\theta}} dz$$

After solving the integral, we get

$$\int_0^{\mu} yf(y)dy = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\lambda B\left(1-(1+\mu^{-\theta})^{-1};(p-\beta(s+p))\lambda+\frac{1}{\theta},1-\frac{1}{\theta}\right) \tag{20}$$

Where $B(.)$ denotes incomplete beta function.
Substitute equation (20) in equation (19), we have

$$D(\mu) = \mu e^{-\alpha^{\beta}((1+\mu^{-\theta})^{\lambda}-1)^{\beta}} - \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\lambda B\left(1-(1+\mu^{-\theta})^{-1};(p-\beta(s+p))\lambda+\frac{1}{\theta},1-\frac{1}{\theta}\right)$$

Let Y be a random variable from IWB-III distribution with median M. Then the mean deviation from median is defined as.

$$\begin{aligned} D(M) &= E\left(|Y-M|\right) \\ &= \int_0^{\infty}|Y-M|f(y)dy \\ &= \int_0^{M}(M-y)f(y)dy + \int_M^{\infty}(y-M)f(y)dy \\ &= \mu - 2\int_0^{M} yf(y)dy \end{aligned} \tag{21}$$

NOW

$$\int_0^{M} yf(y)dy = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\theta\lambda \int_0^{M} y^{-\theta-1}\left(1+y^{-\theta}\right)^{-\lambda-1}\left[\left(1+y^{-\theta}\right)^{-\lambda-1}\right]^{p-\beta(s+1)-1} dy$$

Making substitution $\left(1+y^{-\theta}\right)^{-\lambda}=z$, so that $0\le z\le \left(1+M^{-\theta}\right)^{-\lambda}$, we have

$$\int_0^{M} yf(y)dy = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\int_0^{(1+M^{-\theta})^{-\lambda}} z^{p-\beta(s+1)+\frac{1}{\theta\lambda}-1}\left(1-z^{\frac{1}{\lambda}}\right)^{-\frac{1}{\theta}} dz$$

After solving the integral, we get

$$\int_0^{M} yf(y)dy = \sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\lambda B\left(1-(1+M^{-\theta})^{-1};(p-\beta(s+p))\lambda+\frac{1}{\theta},1-\frac{1}{\theta}\right) \tag{22}$$

Where $B(.)$ denotes incomplete beta function.
Substitute equation (22) in equation (21), we have

$$D(M) = \mu - 2\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\zeta_{s,p}\lambda B\left(1-(1+M^{-\theta})^{-1};(p-\beta(s+p))\lambda+\frac{1}{\theta},1-\frac{1}{\theta}\right)$$

5. Renyi Entropy

Let Y be a continuous random variable with probability density functionf $f(y)$. Then Renyi entropy is stated as

$$T_R(\rho) = \frac{1}{1-\rho} log\left[\int_0^\infty f^\rho(y)dy\right]$$

Where $\rho > 0$ and $\rho \neq 1$

$$\begin{aligned} T_R(\rho) =& \frac{1}{1-\rho} log\left[\int_0^\infty \left(\alpha^\beta \beta g(y,\phi)\frac{(G(y,\phi))^{-(\beta+1)}}{(\bar{G}(y,\phi))^{-(\beta-1)}} e^{-\alpha^\beta\left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}}\right)^\rho dy\right] \\ =& \frac{1}{1-\rho} log\left[\alpha^{\beta\rho}\beta^\rho \int_0^\infty (g(y,\phi))^\rho (G(y,\phi))^{-\rho(\beta+1)} (1-G(y,\phi))^{\rho(\beta+1)} e^{-\alpha^\beta\rho\left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta}} dy\right] \end{aligned} \quad (23)$$

Using the expansion $e^{-kx} = \sum_{s=0}^\infty \frac{(-1)^s}{s!}(kx)^s$ in equation (5.1), we have

$$=\frac{1}{1-\rho} log\left[\alpha^{\beta\rho}\beta^\rho \int_0^\infty (g(y,\phi))^\rho (G(y,\phi))^{-\rho(\beta+1)} (1-G(y,\phi))^{\rho(\beta+1)} \sum_{s=0}^\infty \frac{(-1)^s}{s!}(\alpha^{\beta\rho})^s \left(\frac{G(y,\phi)}{\bar{G}(y,\phi)}\right)^{-\beta s} dy\right]$$

$$=\frac{1}{1-\rho} log\left[\sum_{s=0}^\infty \frac{(-1)^s}{s!}\alpha^{\beta(\rho+s)}\beta^\rho\rho^s \int_0^\infty (g(y,\phi))^s (G(y,\phi))^{-\beta(\rho+s)}(1-G(y,\phi))^{\beta(\rho-s)+\rho} dy\right] \quad (24)$$

Using generalized bionomial expansion $(1-z)^{a-1} = \sum_{p=0}^\infty (-1)^p\binom{a-1}{p}z^p$ in equation (5.2), we have

$$T_R(\rho) = \frac{1}{1-\rho} log\left[\sum_{s=0}^\infty\sum_{p=0}^\infty \frac{(-1)^{s+p}}{s!}\binom{\beta(\rho-s)+\rho}{p}\alpha^{\beta(\rho+s)}(\beta s)^\rho \int_0^\infty (g(y,\phi))^\rho (G(y,\phi))^{p-\beta(\rho+s)} dy\right] \quad (25)$$

Using equation (12) and (13) in (25), we have

$$T_R(\rho) = \frac{1}{1-\rho} log\left[\sum_{s=0}^\infty\sum_{p=0}^\infty \zeta_{s,p}(\theta\lambda)^\rho \int_0^\infty y^{-\rho(\theta+1)}(1+y^{-\theta})^{-\rho(\lambda+1)} \left((1+y^{-\theta})^{-\lambda}\right)^{p-\beta(\rho+s)} dy\right]$$

Where

$$\zeta_{s,p} = \frac{(-1)^{s+p}}{s!}\binom{\beta(\rho-s)+\rho}{p}\alpha^{\beta(\rho+s)}(\beta s)^\rho$$

Making substitution $(1+y^{-\theta})^{-\lambda} = z, 0 < z < 1$, we have

$$T_R(\rho) = \frac{1}{1-\rho} log\left[\sum_{s=0}^\infty\sum_{p=0}^\infty \zeta_{s,p}(\theta\lambda)^{\rho-1} \int_0^1 z^{p-\beta(\rho+s)+\frac{\rho(\theta\lambda-1)+1}{\theta\lambda}-1}(1-z^{\frac{1}{\lambda}})^{\frac{\rho(\theta+1)}{\theta}-1} dy\right]$$

After solving the integral, we get

$$T_R(\rho) = \frac{1}{1-\rho} log\left[\sum_{s=0}^\infty\sum_{p=0}^\infty \zeta_{s,p}(\theta\lambda)^{\rho-1}\lambda B(m\lambda, n)\right]$$

Where $B(.)$ denotes beta function and $m = p-\beta(\rho+s)+\frac{\rho(\theta\lambda-1)+1}{\theta\lambda}, n = \frac{\rho(\theta+1)}{\theta}$

6. Order Statistics of (IWB-III) Distribution

Let us suppose $Y_1, Y_2, ..., Y_n$ be random samples of size n from IWB-III distribution with pdf $f(y)$ and cdf F(y). Then the probability density function of th k^{th} order statistics is given as

$$f_Y(k) = \frac{n!}{(k-1)!(n-1)!} f(y) \left[F(y)\right]^{k-1} \left[1 - F(y)\right]^{n-1} \tag{26}$$

Using equation (12) and (13) in equation (26), we have

$$\begin{aligned} f_Y(k) =& \frac{n!}{(k-1)!(n-1)!} \alpha^\beta \beta\theta\lambda y^{-\theta-1}(1+y^{-\theta})^{\lambda-1}((1+y^{-\theta})^\lambda - 1)^{\beta-1} e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta} \\ & \times \left[e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta}\right]^{k-1} \left[1 - e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta}\right]^{n-k} \end{aligned}$$

The pdf of the first order statistics Y_1 of IWB-III distribution is given by

$$\begin{aligned} f_Y(1) =& n\alpha^\beta \beta\theta\lambda y^{-\theta-1}(1+y^{-\theta})^{\lambda-1}((1+y^{-\theta})^\lambda - 1)^{\beta-1} e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta} \\ & \times \left[1 - e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta}\right]^{n-1} \end{aligned}$$

The pdf of the n^{th} order statistics Y_n of IWB-III distribution is given by

$$\begin{aligned} f_Y(1) = n\alpha^\beta \beta\theta\lambda y^{-\theta-1}(1+y^{-\theta})^{\lambda-1}((1+y^{-\theta})^\lambda - 1)^{\beta-1} e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta} \\ \left[e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta}\right]^{n-1} \end{aligned}$$

7. Maximum Likelihood Estimation of (IWB-III) Distribution

Let the random samples $y_1, y_2, y_3, ..., y_n$ are drawn from IWB-III distribution. The likelihood function of n observations is given as

$$L = \prod_{i=1}^{n} \left(\alpha^\beta \beta\theta\lambda y^{-\theta-1}(1+y^{-\theta})^{\lambda-1}((1+y^{-\theta})^\lambda - 1)^{\beta-1} e^{-\alpha^\beta((1+y^{-\theta})^\lambda-1)^\beta}\right)$$

The log-likelihood function is given as

$$\begin{aligned} l =& n\beta log(\alpha) + nlog(\beta) + nlog(\theta) + nlog(\lambda) - (\theta+1)\sum_{i=1}^{n} log y_i + (\lambda-1)\sum_{i=1}^{n} log(1+y_i^{-\theta}) \\ & + (\beta-1)\sum_{i=1}^{n} log\left((1+y_i^{-\theta})^\lambda - 1\right) - \alpha^\beta \sum_{i=1}^{n}\left((1+y_i^{-\theta})^\lambda - 1\right)^\beta \end{aligned} \tag{27}$$

The partial derivatives of the log-likelihood function with respect to α, β, θ and λ are given as

$$\frac{\partial l}{\partial \alpha} = \frac{n\beta}{\alpha} - \beta\alpha^{\beta-1}\sum_{i=1}^{n}\left((1+y_i^{-\theta})^{\lambda}-1\right)^{\beta} \tag{28}$$

$$\begin{aligned}\frac{\partial l}{\partial \beta} =& nlog(\alpha) + \frac{n}{\beta}\sum_{i=1}^{n}\left((1+y_i^{-\theta})^{\lambda}-1\right) - \alpha^{\beta}\sum_{i=1}^{n}\left((1+y_i^{-\theta})^{\lambda}-1\right)^{\beta} log\left((1+y_i^{-\theta})^{\lambda}-1\right)\\ &- \alpha^{\beta}log(\alpha)\sum_{i=1}^{n}\left((1+y_i^{-\theta})^{\lambda}-1\right)^{\beta}\end{aligned} \tag{29}$$

$$\begin{aligned}\frac{\partial l}{\partial \theta} =& \frac{n}{\theta} - \sum_{i=1}^{n}log(y_i) - (\lambda-1)\sum_{i=1}^{n}\frac{y_i^{-\theta}}{(1+y_i^{-\theta})}log(y_i) - (\beta-1)\lambda\sum_{i=1}^{n}\frac{\left(1+y_i^{-\theta}\right)^{\lambda-1}}{\left(1+y_i^{-\theta}\right)-1}y_i^{-\theta}log(y_i)\\ &+ \alpha^{\beta}\beta\lambda\sum_{i=1}^{n}\left((1+y_i^{-\theta})^{\lambda}-1\right)^{\beta-1}\left(1+y_i^{-\theta}\right)^{\lambda-1}y_i^{-\theta}log(y_i)\end{aligned} \tag{30}$$

$$\begin{aligned}\frac{\partial l}{\partial \lambda} =& \frac{n}{\lambda} + \sum_{i=1}^{n}log(1+y_i^{-\theta}) + (\beta+1)\sum_{i=1}^{n}\frac{\left(1+y_i^{-\theta}\right)^{\lambda}}{\left(1+y_i^{-\theta}\right)-1}log(1+y_i^{-\theta}) - \alpha^{\beta}\beta\sum_{i=1}^{n}\left((1+y_i^{-\theta})^{\lambda}-1\right)^{\beta}\\ &\times\left(1+y_i^{-\theta}\right)^{\lambda}log(1+y_i^{-\theta})\end{aligned} \tag{31}$$

Clearly the equations (28),(29),(30) and (31), are non-linear equations which cannot be expressed in compact form and it is difficult to solve them explicitly for α, β, θ and λ .By applying the iterative methods such as Newton–Raphson method, secant method, Regula-falsi method etc. The MLE of the parameters denoted as $\hat{\xi}(\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\lambda})$ of $\xi(\alpha, \beta, \theta, \lambda)$ can be obtained by using the above methods.
For interval estimation and hypothesis tests on the model parameters, an information matrix is required. The 3 by 3 observed matrix is

$$I(\xi) = \frac{-1}{n}\begin{bmatrix} E\left(\frac{\partial^2 logl}{\partial\alpha^2}\right) & E\left(\frac{\partial^2 logl}{\partial\alpha\partial\beta}\right) & E\left(\frac{\partial^2 logl}{\partial\alpha\partial\theta}\right) & E\left(\frac{\partial^2 logl}{\partial\alpha\partial\lambda}\right)\\ E\left(\frac{\partial^2 logl}{\partial\beta\partial\alpha}\right) & E\left(\frac{\partial^2 logl}{\partial\beta^2}\right) & E\left(\frac{\partial^2 logl}{\partial\beta\partial\theta}\right) & E\left(\frac{\partial^2 logl}{\partial\beta\partial\lambda}\right)\\ E\left(\frac{\partial^2 logl}{\partial\theta\partial\alpha}\right) & E\left(\frac{\partial^2 logl}{\partial\theta\partial\beta}\right) & E\left(\frac{\partial^2 logl}{\partial\theta^2}\right) & E\left(\frac{\partial^2 logl}{\partial\theta\partial\lambda}\right)\\ E\left(\frac{\partial^2 logl}{\partial\lambda\partial\alpha}\right) & E\left(\frac{\partial^2 logl}{\partial\lambda\partial\beta}\right) & E\left(\frac{\partial^2 logl}{\partial\lambda\partial\theta}\right) & E\left(\frac{\partial^2 logl}{\partial\lambda^2}\right)\end{bmatrix}$$

The elements of above information matrix can be obtain by differentiating equations (28),(29),(30) and (31) again partially. Under standard regularity conditions when $n \to \infty$ the distribution of $\hat{\xi}$ can be approximated by a multivariate normal $N(0, I(\hat{\xi})^{-1})$ distribution to construct approximate confidence interval for the parameters. Hence the approximate $100(1-\psi)\%$ confidence interval for α, β, θ and λ are respectively given by
$\hat{\alpha} \pm Z_{\frac{\psi}{2}}\sqrt{I_{\alpha\alpha}^{-1}(\hat{\xi})}$, $\hat{\beta} \pm Z_{\frac{\psi}{2}}\sqrt{I_{\beta\beta}^{-1}(\hat{\xi})}$, $\hat{\theta} \pm Z_{\frac{\psi}{2}}\sqrt{I_{\theta\theta}^{-1}(\hat{\xi})}$ and $\hat{\lambda} \pm Z_{\frac{\psi}{2}}\sqrt{I_{\lambda\lambda}^{-1}(\hat{\xi})}$

8. Simulation Analysis

The MLE'S, bias and mean quare error (MSE) were all addressed to simulation analysis. From IWB-III with N=1000, samples of size n=50,150,250,350 and 500 were obtained. The following expression has been used to produce random numbers.

$$y = \left\{\left[1+\left(-\frac{1}{\alpha^{\beta}}log(u)\right)^{\frac{1}{\beta}}\right]^{\frac{1}{\lambda}} - 1\right\}^{-\frac{1}{\theta}}$$

Where u is uniform random numbers with $u \in (0,1)$. For various parameter combinations, simulation results have been achieved. The MLE's bias, and MSE values are calculated and presented in table 1 and 2. As the sample size increases, this becomes apparent that these estimates are relatively consistent and approximate the actual values of parameters. Interestingly, with all parameter combinations, the bias and MSE reduce as the sample size increases.

Table 1: *Average values of MLEs their corresponding MSEs and Bias for different parameter values* $\alpha = 0.6, \beta = 1.8, \theta = 1.7, \lambda = 0.9$

Sample size	Parameters	MLEs	Bias	MSE
50	α	0.96617	0.36617	0.13433
	β	0.78938	-1.01061	1.02160
	θ	2.28698	0.58698	0.36917
	λ	0.97504	0.07504	0.03660
150	α	0.96446	0.36446	0.13040
	β	0.78690	-1.00309	1.00635
	θ	2.28292	0.58292	0.34793
	λ	0.90522	0.00522	0.01064
250	α	0.95675	0.35675	0.14210
	β	0.79891	-1.00108	1.00229
	θ	2.18506	0.48506	0.24773
	λ	0.89007	-0.00992	0.00780
350	α	0.94957	0.34957	0.14120
	β	0.70135	-1.09864	0.99739
	θ	2.18423	0.47623	0.24659
	λ	0.87259	-0.02740	0.00611
500	α	0.94070	0.34070	0.14103
	β	0.70123	-1.99776	0.99560
	θ	2.18660	0.46960	0.24483
	λ	0.86856	-0.03143	0.00485

Table 2: *Average values of MLEs their corresponding MSEs and Bias for different parameter values* $\alpha = 0.9, \beta = 1.8, \theta = 1.3, \lambda = 1.1$

Sample size	Parameters	MLEs	Bias	MSE
50	α	0.96470	0.06470	0.00444
	β	0.78783	-1.01216	1.02473
	θ	1.55140	0.25140	0.08037
	λ	0.85433	-0.24566	0.08322
150	α	0.96430	0.064308	0.00430
	β	0.78731	-1.01268	1.00554
	θ	1.50394	0.20394	0.08034
	λ	0.77932	-0.32067	0.01114
250	α	0.96324	0.06324	0.00421
	β	0.77744	-1.02255	1.00523
	θ	1.40394	0.20294	0.08009
	λ	0.77879	-0.32120	0.00351
350	α	0.95853	0.04853	0.00330
	β	0.60026	-1.99973	0.99958
	θ	1.31691	0.11691	0.00448
	λ	0.76227	-0.33772	0.00345
500	α	0.94868	0.03868	0.00245
	β	0.50124	-1.99975	0.99758
	θ	1.22246	0.10246	0.00370
	λ	0.75697	-0.34302	0.00341

9. Data Aanalysis

This subsection evaluates a real-world data set to demonstrate the IWB-III distribution's applicability and effectiveness. The IWB-III distribution's adaptability is determined by comparing its efficacy to that of other analogous distributions such as new modified Weibull distribution (NMWD), modified Weibull distribution (MWD),Topp-Leone Burr distribution (TLBD), inverse Weibull distribution (IWD) and Burr-III distribution (B-IIID),inverse Rayleigh distribution (IRD),inverse Lindley distributon (ILD).
To compare the versatility of the explored distribution, we consider the criteria like AIC (Akaike information criterion), CAIC (Consistent Akaike information criterion), BIC (Bayesian information criterion) and HQIC (Hannan-Quinn information criterion). Distribution having lesser AIC, CAIC, BIC and HQIC values is considered better.

$$AIC = -2l + 2p, \qquad AICC = -2l + 2pm/(m-p-1), \qquad BIC = -2l + p(log(m))$$
$$HQIC = -2l + 2plog(log(m)) \qquad K.S = max_{1\leq j\leq m}\left(F(x_j) - \frac{j-1}{m}, \frac{j}{m} - F(x_j)\right)$$

Where $'l'$ denotes the log-likelihood function,'p'is the number of parameters and'm'is the sample size.
Data set: Bjerkedel studied the survival rates (in days) of 72 guinea pigs treated with pathogenic turbercle bacteria [7]. The data are as follows
0.1, 0.33, 0.44, 0.56, 0.59, 0.59, 0.72, 0.74, 0.92, 0.93,0.96, 1, 1, 1.02, 1.05, 1.07, 1.07, 1.08, 1.08, 1.08, 1.09, 1.12, 1.13, 1.15, 1.16, 1.2, 1.21, 1.22, 1.22, 1.24, 1.3, 1.34, 1.36, 1.39, 1.44, 1.46,1.53, 1.59, 1.6, 1.63, 1.68, 1.71, 1.72, 1.76, 1.83, 1.95, 1.96, 1.97, 2.02, 2.13, 2.15, 2.16, 2.22, 2.3, 2.31, 2.4, 2.45, 2.51, 2.53, 2.54, 2.78,2.93, 3.27, 3.42, 3.47, 3.61, 4.02, 4.32, 4.58, 5.55, 2.54, 0.77.
The ML estimates with corresponding standard errors in parenthesis of the unknown parameters are presented in Table 4 and the comparison statistics, AIC, BIC, CAIC, HQIC and the goodness-of-fit statistic for the data set are displayed in Table 5.

Table 3: *Descriptive statistics for data set*

Min.	Max.	Ist Qu.	Med.	Mean	3rd Qu.	kurt.	Skew.
0.100	5.550	1.077	1.450	1.754	2.240	4.9139	1.3282

Table 4: *The ML Estimates (standard error in parenthesis) for data set*

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$
IWB-IIID	189.54	0.1577	15.511	0.1618	...
	(172.96)	(0.0159)	(0.0977)	(0.0584)	...
NMWD	0.0010	0.2922	1.7967	0.0010	1.7941
	(0.0035)	(0.0940)	(5.0481)	(0.0014)	(0.1570)
AWD	0.0010	0.2924	1.7961	1.7962	...
	(0.0205)	(0.0152)	(0.1563)	(0.1573)	...
TLBD	0.484	2.3688	1.8033	...	...
	(0.2556)	(0.8929)	(0.9392)	...	...
IWD	1.1753	1.0402	...	...	...
	(0.0849)	(0.1110)	...	...	...
B-IIID	...	...	2.3189	1.8576	...
	...	....	(0.2144)	(0.2192)	....
IRD	0.9112	...	...	...	...
	(0.1073)	...	...	...	...
ILD	1.5540	...	...	...	...
	(0.1434)	...	...	...	...

It is observed from table 5 that IWB-IIID provides best fit than other competative models based on the measures of statistics, AIC, BIC, AICC, HQIC and K-S statistic. Along with p-values of each model.

Table 5: *Comparison criterion and goodness-of-fit statistics for data set*

Model	-l	AIC	AICC	BIC	HQIC	K.S statistic	p-value
IWB-IIID	93.007	194.01	194.61	203.12	197.64	0.0765	0.7933
NMWD	96.03	202.07	202.98	213.45	206.60	0.0983	0.4897
AWD	96.02	200.05	200.64	209.15	203.67	0.0982	0.4902
TLBD	97.60	201.21	201.56	208.04	203.93	0.3966	2.907e-10
IWD	117.32	238.65	238.82	243.20	240.46	0.1899	0.01109
B-IIID	97.608	199.21	199.38	203.76	201.02	0.1092	0.3565
IRD	161.85	325.71	325.77	327.99	326.61	0.4674	4.352e-14
ILD	118.93	239.86	239.92	242.14	240.77	0.3219	2.2e-16

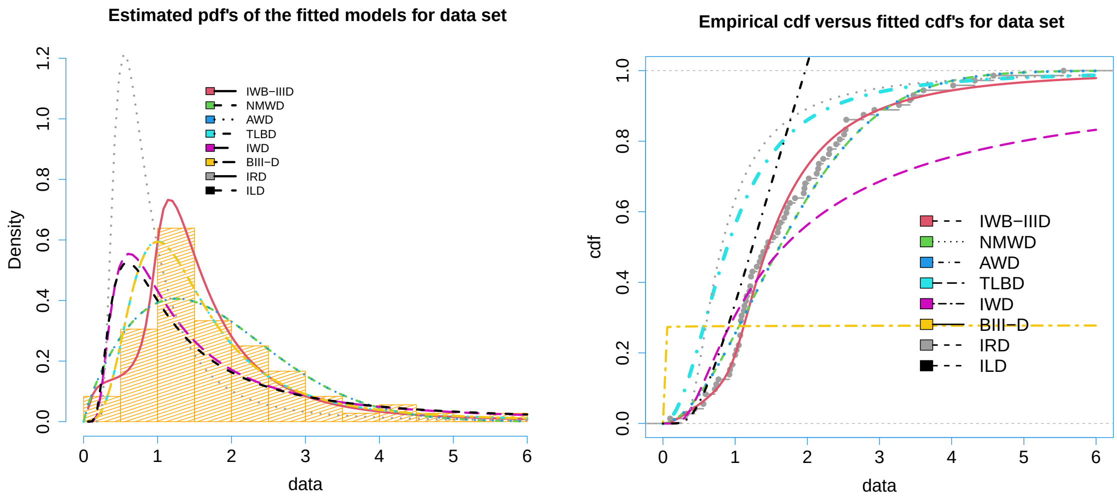


Figure 4: *fitted pdf's and cdf's for the data set*

10. Conclusion

In this work, we developed a novel flexible distribution known as the inverse Weibull-Burr III distribution. Numerous mathematical characteristics are determined for this distribution, including moments, moment generating functions, incomplete moments, order statistics, Renyi entropy, mean deviations, and reliability analysis. The maximum likelihood estimation approach was used to estimate the distribution's parameters. Ultimately, it has been demonstrated by employing a real-world data set that the stated distribution leads to a better fit than the comparable ones.

References

[1] A. Alzaatreh, C.Lee,F. Famoye . A new method for generating families of distributions. *Metron, 71 (2013),https://doi.org/10.1002/wics.1255*, 63-79.

[2] H. Amal and G.N. Said. The inverse Weibull-G familiy. *Journal of data science (2018)*, 723-742, DOI: 10.6339/JDS.201810-16(4).00004.

[3] M. Bourguingnon, R.B. Silva and G.M. Corderio. The Weibull-G family of probability distributions *Journal of data science,12 (2014), http://www.jds-online.com*, 53-68.

[4] I.W. Burr. Cummlative frequency function. *Annals of mathematical statistics,13 (1942),https://www.jstor.org/stable/2235756*, 215-232.

[5] E. Brito, G.M. Cordeiro, H.M Yousuf,M. Alizadeh,G.O Silva. The Topp-Leone odd log-logistic family of distributions *J stat comput simul, 87(15) (2017), https://doi.org/10.1080/00949655.2017.1351972*, 3040-3058.

[6] A.P. Bilal, J.Suriya and J. Tariq. Generalization of Burr-III distribution. *International journal of modern mathematical science, 13(3) (2015), https://www.ModernScientificPress.com/Journals/ijmms.aspx*,322-329.

[7] T. Bjerkedal. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *American journal of Epidemiology, 72(1) (1960)*,130-148.

[8] R. Corderio and G. Pulcini. A new family of generalized distributions. *Journal of statistical computation and simulation, 81 (2011),https://doi.org/10.1080/00949650903530745*, 883-893.

[9] M. Daniyal and M. Aleem. On Mixture of Burr XII and Weibull Distributions. *Journal of Statistics and Probability, 3(2) (2014)*, 251-267, DOI: 10.12785/jsap/030215.

[10] A. Danyian, G. Behairy, S.M. and A.A. EL-Helbawy. The Kumaraswamy-Burr type-III distribution: Properties and Estimation. *British journal of mathematics and computer science, 14(2) (2016)*, 1-21, DOI: 10.9734/BJMCS/2016/19958.

[11] N. Eugene, C. Lee and F. Famoye. Beta-normal distribution and its applications. *Communication in statistics- theory and methods,31 (2002),https://doi.org/10.108/STA-120003130*, 497-512.

[12] A. Morad, M.C. Gauss. The Gompertz-G family of distributions *Journal of statistical theory and practice,11(1)(2017),https:doi.org/10.108/15598608.2016.1267668*, 179-207.

[13] K. Zagrafos andN. Balakrishnan. On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology, 6 (2009),https://doi.org/10.1016/j.stamet.2008.12.003*, 344-362.

Reliability and Economic Analysis of Captive Power Plant With Reduced Capacity

Upasana Sharma[1] and Avtar Singh[2*]

•

[1]Department of Statistics, Punjabi University, Patiala-India, usharma@pbi.ac.in
[2*]Department of Statistics, Punjabi University, Patiala-India, avtardhillon.ad@gmail.com

Abstract

This paper reported the performance evaluation of Captive Power plant working in the fertilizer industry with possible production capacities. The idea of reduced capacity and load sharing to use the available system optimally is analyzed. The system works on two STG's (steam turbine generators) and one gridline. Gridline can bear the load of one or both STG's on failures. At the breakdown in gridline and STG, the system work at reduced capacity. Gridline repaired on a priority basis. The semi-Markov processses and regenerative point technique are used to evaluate the reliability and economic measures such as availability, busy period of repairman, and expected no. of repairs. The graphical study shows the relationships between these measures with the failure rates of STG and gridline.

Keywords: Steam Turbine Generators, Regenerative point technique, semi-Markov process, Reduced capacity, Reliability modeling.

I. Introduction

Nowadays, Captive power plants are a reliable and beneficial energy source for power-consuming production industries. Optimizing the operations of the power-producing units in these captive power plants can boost the industry's profit. A good number of researchers have worked on various reliability models with conditions of repair and maintenance. Gupta and Goel [3] studied a two-unit cold standby system working under abnormal weather conditions. Chandrashekhar et al. [1], Goyal et al. [2], Parashar et al. [4] have analyzed two and three-unit systems. Rizwan et al. [5] worked with the reliability of the hot standby industrial system.

Singh and Taneja [7], [8] analyzed power generating systems with various types of inspections. Rajesh et. al [10], [9] studied gas turbine power plants consisting of two and three units. These attempts to the literature create a motivation for the present study, to work with the economic benefits of the captive power plant. The captive power plants are auto producers of electricity, which operates off-grid or in parallel with gridline to make consistent and quality electricity supply for industries at reasonable costs. Availability of these power generating units in any possible way (full or reduced) can make a reliable electricity supply at less cost. Keeping an eye on the above fact economic analysis of Captive Power plant working in National Fertilizer limited, Bathinda, India has done.

The present system comprises two STG's (steam turbine generators) connected in parallel with the Gridline of PSPCL (Punjab state power corporation limited). These two STGs can fulfill the electrical load for the system. On failure of any one or both of STGs, the system operates with the help of gridline. The system will work at reduced capacity when only one STG is working (one STG and gridline failed). The failure of these three units leads to complete system failure. Repair of gridline is done on a priority basis among all units, whereas the FCFS repair pattern is applied on both STGs. The reliability measure MTSF (mean time to system failure) and economic measures such as availability, a busy period of the repairman, and expected no. of repairs have been derived using the semi-Markov processes and regenerative point techniques numerically. Also, graphical plotting was performed for these measures.

I. Assumptions for the model

- All failure time variables follow exponential distribution but repair times distributed generally..
- Every repaired unit works as new one.
- In the given model system initially started working from state S_0.

II. Nomenclature & Model Description

I. Notations & abbreviations

Notations		Discription
λ_1	:	Constant failure rate of STG 1.
λ_2	:	Constant failure rate of STG 2.
λ_3	:	Constant failure rate of Gridline.
α_1	:	Repair rate of STG 1.
α_2	:	Repair rate of STG 2.
α_3	:	Repair rate of Gridline.
$G_1(t)$, $g_1(t)$	:	c.d.f. & p.d.f of repair time of STG 1.
$G_2(t)$, $g_2(t)$	:	c.d.f. & p.d.f of repair time of STG 2.
$G_3(t)$, $g_3(t)$	:	c.d.f. & p.d.f of repair time of Gridline.
a	:	probability of transit from S_7 & S_8 to S_3 respectively after repair .
b	:	probability of transit from S_7 & S_8 to S_4 respectively after repair .
©	:	Laplace Convolution.
Ⓢ	:	Stieltjes Convolution.
$*/**$	:	Laplace Transformation/ Laplace Stieltjes Transformation.
$M_i(t)$	:	Probability that system is working at state S_i during the time interval $(0-t]$.
$W_i(t)$	:	Probability of repairman repairing at state S_i during the time interval $(0-t]$.

II. Symbols for States

Symbols for the states of the system:-
S_i : States of the system with number i, $i = 1, 2, 3, ...8$.
O_I, O_{II}, O_{III} : STG 1, STG 2, Gridline from PSPCL are in operating state.
CS_{III} : Gridline (PSPCL) in cold standby state.
Fr_I, Fr_{II}, Fr_{III} : STG 1, STG 2, Gridline under repair.
FR_I, FR_{II}, FR_{III} : STG 1, STG 2, Gridline under repair from previous state.
Fwr_I, Fwr_{II} :Failed Units STG 1, STG 2 waiting for repair.

III. State Transition Diagram

Figure 1, shows the state transitions diagram of the Captive power plant consisting of two STGs and one gridline from PSPCL. The states S_0, S_1, S_2, S_3, S_4 are operating states. The states $S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_8$ are regenerative states. The states S_5, S_6 are reduced capacity states. The states S_7, S_8 are failed states. Table 1 shows the description of every state of the system.

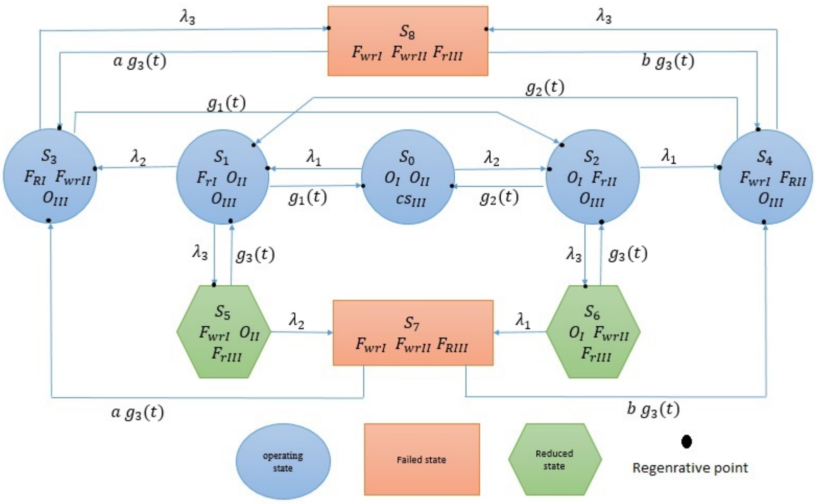


Figure 1: *State Transition Diagram*

Table 1: *State Discription*

State notation	States Discription
S_0	This is the initial full capacity working state where both STGs are working. Gridline is in a standby state.
S_1	System working at full capacity where STG 1 and gridline are working. STG 2 is in a failed state under repair.
S_2	System working full capacity where STG 2 and gridline are working. STG 1 is in failed state under repair.
S_3, S_4	System is operating at full capacity with gridline. Both STGs are in a failed state.
S_5	System operating at reduced capacity where only STG 2 is working. STG 1 and gridline are in the failed state.
S_6	System operating at reduced capacity where only STG 1 is working. STG 2 and gridline are in failed states.
S_7, S_8	These are failed states where all units are in a failed state.

IV. Transition Probabilities & Mean Sojourn Times

p_{ij} represents non-zero elements which are given below The non zero elements p_{ij}'s are given as:

$$p_{01} = \frac{\lambda_1}{\lambda_1 + \lambda_2}, \quad p_{02} = \frac{\lambda_2}{\lambda_1 + \lambda_2}, \quad p_{10} = g_1^*(\lambda_3 + \lambda_2),$$

$$p_{13} = \frac{\lambda_2}{\lambda_3 + \lambda_2}[1 - g_1^*(\lambda_3 + \lambda_2)], \quad p_{15} = \frac{\lambda_3}{\lambda_3 + \lambda_2}[1 - g_1^*(\lambda_3 + \lambda_2)], \quad p_{20} = g_2^*(\lambda_1 + \lambda_3),$$

$$p_{24} = \frac{\lambda_1}{\lambda_1 + \lambda_3}[1 - g_2^*(\lambda_1 + \lambda_3)], \quad p_{26} = \frac{\lambda_3}{\lambda_1 + \lambda_3}[1 - g_2^*(\lambda_1 + \lambda_3)], \quad p_{32} = g_1^*(\lambda_3),$$

$$p_{38} = [1 - g_1^*(\lambda_3)], \quad p_{41} = g_2^*(\lambda_3), \quad p_{48} = [1 - g_2^*(\lambda_3)]$$

$$p_{51} = g_3^*(\lambda_2), \quad p_{57} = [1 - g_3^*(\lambda_2)], \quad p_{53}^{(7)} = a[1 - g_3^*(\lambda_2)],$$

$$p_{54}^{(7)} = b[1 - g_3^*(\lambda_2)], \quad p_{62} = g_3^*(\lambda_1), \quad p_{67} = [1 - g_3^*(\lambda_1)],$$

$$p_{63}^{(7)} = a[1 - g_3^*(\lambda_1)], \quad p_{64}^{(7)} = b[1 - g_3^*(\lambda_1)], \quad p_{83} = a,$$

$$p_{84} = b$$

The mean sojourn time μ_i corresponding to regenerative state $'i'$ is given as:

$$\mu_0 = \frac{1}{\lambda_1+\lambda_2}, \quad \mu_1 = \frac{1}{\lambda_3+\lambda_2}[1-g_1^*(\lambda_3+\lambda_2)], \quad \mu_2 = \frac{1}{\lambda_1+\lambda_3}[1-g_2^*(\lambda_1+\lambda_3)],$$
$$\mu_3 = \frac{1}{\lambda_3}[1-g_1^*(\lambda_3)], \quad \mu_4 = \frac{1}{\lambda_3}[1-g_2^*(\lambda_3)], \quad \mu_5 = \frac{1}{\lambda_2}[1-g_3^*(\lambda_2)],$$
$$\mu_6 = \frac{1}{\lambda_1}[1-g_3^*(\lambda_1)], \quad \mu_8 = -g_3^{*'}(0)$$

The unconditional mean time m_{ij} required by the system to transit from state $'i'$ to any regenerative state $'j'$ when time is counted from the epoch of entrance into the state $'i'$ is mathematically stated as:

$$m_{ij} = \int_a^b t dQij(t) = -q_{ij}^{*'}(0) \tag{1}$$

So we have

$$m_{01}+m_{02} = \mu_0, \quad m_{10}+m_{13}+m_{15} = \mu_1, \quad m_{20}+m_{24}+m_{28} = \mu_2, \quad m_{32}+m_{38} = \mu_3,$$
$$m_{41}+m_{48} = \mu_4, \quad m_{51}+m_{53}^{(7)}+m_{54}^{(7)} = k_1, \quad m_{62}+m_{63}^{(7)}+m_{64}^{(7)} = k_1, \quad m_{83}+m_{84} = \mu_8$$

III. Reliability and Economic Measures for System Effectiveness

I. Mean Time to System Failure (MTSF)

Assume $\phi_i(t)$ as a distribution function of variable time (t) lapses during the system transition from a regenerative state S_i to any working or failed state where failed state act as an absorbing state. By probabilistic arguments, the following recursive relations are obtained:

$$\phi_0(t) = Q_{01}(t)Ⓢ\phi_1(t) + Q_{02}(t)Ⓢ\phi_2(t) \tag{2}$$
$$\phi_1(t) = Q_{10}(t)Ⓢ\phi_0(t) + Q_{13}(t)Ⓢ\phi_3(t) + Q_{15}(t)Ⓢ\phi_5(t) \tag{3}$$
$$\phi_2(t) = Q_{20}(t)Ⓢ\phi_0(t) + Q_{24}(t)Ⓢ\phi_4(t) + Q_{26}(t)Ⓢ\phi_6(t) \tag{4}$$
$$\phi_3(t) = Q_{32}(t)Ⓢ\phi_2(t) + Q_{38}(t) \tag{5}$$
$$\phi_4(t) = Q_{41}(t)Ⓢ\phi_1(t) + Q_{48}(t) \tag{6}$$
$$\phi_5(t) = Q_{51}(t)Ⓢ\phi_1(t) + Q_{57}(t) \tag{7}$$
$$\phi_6(t) = Q_{62}(t)Ⓢ\phi_2(t) + Q_{67}(t) \tag{8}$$

Transforming the equations(2-8) using Laplace Stieltjes Transformations to get $\phi_i^{**}(t)$. Mean Time to System Failure T_0 at steady state S_o is given by

$$T_0 = \lim_{s\to 0} \frac{1-\phi_0{}^{**}(s)}{s} \tag{9}$$

Using L' Hospital's rule here, we get

$$T_0 = N/D \tag{10}$$

where

$$\begin{aligned} N = {}& \mu_0(1 - p_{15}p_{51} - p_{26}p_{62} - p_{15}p_{26}p_{62}p_{57} + p_{13}p_{24}p_{32}p_{48} + p_{15}p_{26}p_{62} - p_{24}p_{13}p_{32}) \\ & + \mu_1(-p_{62}p_{26}p_{01} + p_{02}p_{24}p_{41} + p_{01}) + \mu_2(-p_{51}p_{02}p_{15} + p_{13}p_{32}p_{01} + p_{02}) + \mu_3(p_{13}p_{01} \\ & - p_{13}p_{01}p_{26}p_{62} + p_{24}p_{02}p_{13}p_{41}) + \mu_4(p_{24}p_{02} - p_{24}p_{02}p_{15}p_{51} + p_{13}p_{01}p_{24}p_{32}) + \mu_5(p_{15}p_{01} \\ & - p_{15}p_{01}p_{26}p_{62} + p_{24}p_{02}p_{15}p_{41}) + \mu_6(p_{26}p_{02} - p_{15}p_{51}p_{26}p_{02} + p_{13}p_{01}p_{26}p_{32}) \end{aligned} \tag{11}$$

and

$$\begin{aligned} D = {}& 1 + p_{51}p_{26}p_{15}p_{62} + p_{51}p_{02}p_{20}p_{15} - p_{15}p_{51} - p_{26}p_{62} + p_{62}p_{26}p_{01}p_{10} - p_{13}p_{32}p_{24}p_{41} \\ & - p_{13}p_{32}p_{01}p_{20} - p_{02}p_{10}p_{24}p_{41} - p_{02}p_{20} - p_{01}p_{10} \end{aligned} \tag{12}$$

II. Availability Analysis at Full & Reduced capacity

Let $A_i^F(t)$ notates the probability that system is available with full capacity to perform its intended task at a regenerative state S_i at time $t = 0$. The availability of system at successive regenerative state S_j $(j = 1, 2, ..6, 8)$ is independant from its previous transitions made. This phenomenon follows the theory of regenerative process techniques [6]. Thus following recursive relations are obtained:

$$A_0^F(t) = M_0(t) + q_{01}(t)©A_1^F(t) + q_{02}(t)©A_2^F(t) \tag{13}$$
$$A_1^F(t) = M_1(t) + q_{10}(t)©A_0^F(t) + q_{13}(t)©A_3^F(t) + q_{15}(t)©A_5(t) \tag{14}$$
$$A_2^F(t) = M_2(t) + q_{20}(t)©A_0^F(t) + q_{24}(t)©A_4^F(t) + q_{26}(t)©A_6(t) \tag{15}$$
$$A_3^F(t) = M_3(t) + q_{32}(t)©A_2^F(t) + q_{38}(t)©A_8^F(t) \tag{16}$$
$$A_4^F(t) = M_4(t) + q_{41}(t)©A_1^F(t) + q_{48}(t)©A_8^F(t) \tag{17}$$
$$A_5^F(t) = q_{51}(t)©A_1^F(t) + q_{53}^{(7)}(t)©A_3^F(t) + q_{54}^{(7)}(t)©A_4^F(t) \tag{18}$$
$$A_6^F(t) = q_{62}(t)©A_2^F(t) + q_{63}^{(7)}(t)©A_3^F(t) + q_{64}^{(7)}(t)©A_4^F(t) \tag{19}$$
$$A_8^F(t) = q_{83}(t)©A_3^F(t) + q_{84}(t)©A_4^F(t) \tag{20}$$

Where

$$M_0(t) = e^{-(\lambda_1+\lambda_2)t}, M_1(t) = e^{-(\lambda_1+\lambda_2)t}\bar{G_1}(t), M_2(t) = e^{-(\lambda_1+\lambda_2)t}\bar{G_2}(t),$$
$$M_3(t) = e^{-(\lambda_3)t}\bar{G_1}(t), M_4(t) = e^{-(\lambda_3)t}\bar{G_2}(t)$$

Transforming the equations(13-20) using Laplace transformations to get $A_0^{F*}(s)$. we have

$$A_0^F(t) = \lim_{s\to 0}(sA_0^{F*}(s)) \tag{21}$$

The steady state availability A_0^F of the system having full capacity is given by:

$$A_0^F = \lim_{t\to 0} A_0^F(t) = N_1/D_1 \tag{22}$$

where

$$\begin{aligned}
N_1 = \mu_0[&(p_{15}p_{51} - 1)(p_{26}p_{32}p_{63} + (1 - p_{62}p_{26})(p_{38}p_{83} + p_{84}p_{48}) + p_{26}p_{62}(1 + p_{15}p_{54}^{(7)}p_{41}p_{38}) \\
&- p_{83}p_{24}p_{32}p_{48} - 1) + (1 - p_{62}p_{26})((p_{83}p_{54}^{(7)}p_{15} - p_{84}p_{53}^{(7)}p_{15} - p_{84}p_{13})p_{41}p_{38}) \\
&- p_{84}p_{32}p_{15}p_{26}p_{51}p_{48}(1 - p_{62}(1 - p_{63}^{(7)})) - (1 - p_{26}p_{32})p_{15}p_{54}^{(7)}p_{41} - p_{64}^{(7)}p_{32}p_{26}p_{41}(p_{13} + p_{15}p_{54}^{(7)}) \\
&- p_{24}p_{32}P_{41}(p_{13} + p_{15}p_{53}^{(7)})] + \mu_1[p_{01}((1 - p_{84}p_{48} - p_{83}p_{38})(1 - p_{26}p_{62}) - p_{63}^{(7)}p_{32}p_{26}(1 - p_{84}p_{48}) \\
&- p_{83}p_{48}p_{32}(p_{24} + p_{64}^{(7)}p_{26})) + p_{02}((p_{64}^{(7)}p_{41}p_{26} + p_{24}p_{41})(1 - p_{83}p_{38}) + p_{84}p_{63}^{(7)}p_{41}p_{26}p_{38})] \\
&+ \mu_2[p_{02}((1 - p_{83}p_{38})(1 - p_{15}p_{51} - p_{15}p_{54}^{(7)}p_{41}) - p_{84}p_{48}(1 - p_{51}p_{15}) - p_{84}p_{38}p_{13}p_{41}) \\
&+ p_{01}p_{32}((p_{15}p_{53}^{(7)} + p_{13})(1 - p_{84}p_{48}) + p_{15}p_{54}^{(7)}p_{83}p_{48})] + \mu_3[(1 - p_{62}p_{26})(p_{15}p_{53}^{(7)}p_{01} + p_{01}p_{13} \\
&- p_{01}p_{84}p_{13}p_{48}) + (p_{84}p_{63}^{(7)} - p_{83}p_{64}^{(7)})(p_{15}p_{51}p_{02}p_{26}p_{48} - p_{02}p_{26}p_{48}) + p_{83}p_{24}p_{02}p_{48}(1 - p_{15}p_{51}) \\
&- p_{15}p_{26}p_{02}p_{41}(p_{54}^{(7)}p_{53}^{(7)} - p_{64}^{(7)}p_{53}^{(7)}) + p_{15}p_{01}p_{48}(p_{83}p_{54}^{(7)} - p_{84}p_{53}^{(7)}) + p_{32}p_{02}p_{26} - p_{15}p_{26}p_{63}^{(7)}p_{51}p_{02} \\
&+ p_{64}^{(7)}p_{02}p_{41}p_{13}p_{26} + p_{24}p_{15}p_{02}p_{41}p_{53}^{(7)} + p_{24}p_{02}p_{41}p_{13}] + \mu_4[(1 - p_{83}p_{38})(p_{24}p_{02} - p_{24}p_{15}p_{51}p_{02} \\
&- p_{15}p_{64}^{(7)}p_{51}p_{02}p_{26} + p_{64}^{(7)}p_{02}p_{26} + p_{15}p_{01}p_{54}^{(7)}) + (p_{13}p_{15}p_{53}^{(7)})(p_{01}p_{24}p_{32} + p_{01}p_{64}^{(7)}p_{32}p_{26}) \\
&+ p_{15}p_{84}p_{53}^{(7)}p_{38}(p_{01} - p_{02}p_{41}) + p_{84}p_{63}p_{02}p_{26}p_{38}(1 - p_{15}p_{51}) + p_{01}p_{54}^{(7)}p_{15}(p_{26}(p_{62}(p_{83}p_{38} \\
&- 1) - p_{63}^{(7)}p_{32}))]
\end{aligned} \tag{23}$$

and

$$\begin{aligned}
D_1 = {} & \mu_0((1 - p_{83}p_{38})(p_{24}p_{41}p_{10} + p_{41}p_{64}^{(7)}p_{26}p_{10} + p_{20} - p_{51}p_{20}p_{15}) + p_{41}p_{15}p_{38}p_{20}(p_{83}p_{54}^{(7)} - \\
& p_{84}p_{53}^{(7)}) - p_{84}p_{20}p_{48}(1 - p_{15}p_{51}) + p_{41}p_{84}p_{26}p_{63}^{(7)}p_{38}p_{10}) + \mu_1(p_{41}p_{84}p_{38}(1 - p_{26}p_{62}) + p_{32}p_{01}p_{20} \\
& (1 - p_{84}p_{48}) + p_{32}p_{24}p_{41} + p_{32}p_{41}p_{64}^{(7)}p_{26} - p_{41}p_{84}p_{38}p_{02}p_{20}+) + \mu_2(p_{32}p_{83}p_{48}(1 - p_{01}p_{10}) \\
& + p_{41}p_{02}p_{10}(1 - p_{83}p_{38}) + p_{32}p_{53}^{(7)}p_{41}p_{15} + p_{32}p_{13}p_{41} - p_{32}p_{15}p_{48}p_{51}p_{83}) + \mu_3((1 - p_{26}p_{62}) \\
& (p_{41}p_{84}p_{15}p_{53}^{(7)} + p_{41}p_{84}p_{13} - p_{01}p_{10}p_{83} + p_{83} - p_{83}p_{51}p_{15}) - p_{41}p_{83}p_{54}^{(7)}p_{15}(1 - p_{02}p_{02}) \\
& - p_{41}p_{84}p_{02}p_{20}(p_{13} + p_{53}^{(7)}p_{15}) - p_{24}p_{41}p_{83}p_{02}p_{10} - p_{41}p_{64}^{(7)}p_{83}p_{26}p_{02}p_{10} + p_{41}p_{83}p_{54}^{(7)}p_{15}p_{26}p_{62} \\
& + p_{83}p_{51}p_{15}p_{02}p_{20} + p_{41}p_{84}p_{26}p_{63}^{(7)}p_{02}p_{10}) + \mu_4(p_{32}p_{24}p_{83}(1 - p_{15}p_{51} - p_{01}p_{10}) + p_{84}(1 - p_{01}p_{10} \\
& - p_{26}p_{62} - p_{15}p_{51}) + p_{32}p_{84}p_{15}p_{26}p_{51}p_{63}^{(7)} + p_{32}p_{83}p_{54}^{(7)}p_{15}p_{01}p_{20} - p_{32}p_{64}^{(7)}p_{26}p_{15}p_{51}p_{83} \\
& + p_{84}p_{26}p_{62}(p_{15}p_{51} + p_{01}p_{10}) - p_{32}p_{84}p_{01}p_{20}(p_{53}^{(7)}p_{15} + p_{13}) - p_{84}p_{02}p_{20}(1 - p_{15}p_{51}) \\
& + (1 - p_{01}p_{10})(p_{32}p_{64}^{(7)}p_{83}p_{26} - p_{32}p_{84}p_{26}p_{63}^{(7)}) + k_1(p_{41}p_{84}p_{15}p_{38}(1 - p_{26}p_{62} - p_{02}p_{20}) \\
& + p_{32}p_{24}p_{41}p_{15} + p_{32}p_{01}p_{20}p_{15} + p_{32}p_{41}p_{64}^{(7)}p_{26}p_{15} - p_{32}p_{15}p_{84}p_{48}p_{01}p_{20}) + k_1(p_{32}p_{83}p_{26}p_{48}(1 \\
& - p_{01}p_{10}) + p_{41}p_{26}p_{02}p_{10}(1 - p_{83}p_{38}) + p_{32}p_{41}p_{53}^{(7)}p_{26}p_{15} + p_{32}p_{41}p_{26}p_{13}) + \mu_8((1 - p_{15}p_{51}) \\
& (p_{32}p_{24}p_{48} - p_{38}p_{02}p_{20}) + p_{32}p_{64}p_{26}p_{48}(1 - p_{01}p_{10}) - p_{41}p_{54}^{(7)}p_{15}p_{38}(1 - p_{02}p_{20} - p_{26}p_{63}^{(7)}) \\
& + p_{38}(-p_{26}p_{62} - p_{01}p_{10} - p_{15}p_{51}) + p_{26}p_{38}p_{62}(p_{15}p_{51} + p_{01}p_{10}) - p_{41}p_{38}p_{02}p_{10}(p_{24} + p_{64}^{(7)}p_{26}) \\
& - p_{32}p_{24}p_{48}p_{01}p_{10} - p_{32}p_{64}^{(7)}p_{26}p_{15}p_{48}p_{51})
\end{aligned} \tag{24}$$

Let $A_i^R(t)$ notates the probability that system is available with reduced capacity to work at a regenerative state S_i at time $t = 0$. The following recursive relations are obtained using the above described argument of regenrative process techniques:

$$A_0^R(t) = q_{01}(t)©A_1^R(t) + q_{02}(t)©A_2^R(t) \tag{25}$$

$$A_1^R(t) = q_{10}(t)©A_0^R(t) + q_{13}(t)©A_3^R(t) + q_{15}(t)©A_5^R(t) \tag{26}$$

$$A_2^R(t) = q_{20}(t)©A_0^R(t) + q_{24}(t)©A_4^R(t) + q_{26}(t)©A_6^R(t) \tag{27}$$

$$A_3^R(t) = q_{32}(t)©A_2^R(t) + q_{38}(t)©A_8^R(t) \tag{28}$$

$$A_4^R(t) = q_{41}(t)©A_1^R(t) + q_{48}(t)©A_8^R(t) \tag{29}$$

$$A_5^R(t) = M_5(t) + q_{51}(t)©A_1^R(t) + q_{53}^{(7)}(t)©A_3^R(t) + q_{54}^{(7)}(t)©A_4^R(t) \tag{30}$$

$$A_6^R(t) = M_6(t) + q_{62}(t)©A_2^R(t) + q_{63}^{(7)}(t)©A_3^R(t) + q_{64}^{(7)}(t)©A_4^R(t) \tag{31}$$

$$A_8^R(t) = q_{83}(t)©A_3^R(t) + q_{84}(t)©A_4^R(t) \tag{32}$$

where

$$M_5(t) = e^{-(\lambda_2)t}\bar{G_3}(t), M_6(t) = e^{-(\lambda_1)t}\bar{G_3}(t)$$

Transforming the equations(25-32) using Laplace transformations to get $A_0^{R*}(s)$. we have

$$A_0^R(t) = \lim_{s\to 0}(sA_0^{R*}(s)) \tag{33}$$

The steady state availability A_0^R of the system having reduced capacity is given by:

$$A_0^R = \lim_{s\to 0}(sA_0^{R*}(s)) = N_2/D_1 \tag{34}$$

where

$$
\begin{aligned}
N_2 = \mu_5(&p_{24}p_{02}p_{15}p_{41} - p_{15}p_{01}p_{26}p_{62} - p_{15}p_{01}p_{38}p_{83} - p_{84}p_{15}p_{01}p_{48} - p_{15}p_{01}p_{26}p_{32}p_{63} \\
&+ p_{64}^{(7)}p_{15}p_{41}p_{26}p_{02} + p_{15}p_{01} - p_{64}^{(7)}p_{15}p_{41}p_{26}p_{02}p_{38}p_{83} - p_{64}^{(7)}p_{15}p_{01}p_{26}p_{32}p_{83}p_{48} \\
&+ p_{84}p_{15}p_{41}p_{26}p_{02}p_{38}p_{63}^{(7)} + p_{84}p_{15}p_{02}p_{26}p_{32}p_{63}^{(7)}p_{48} + p_{15}p_{01}p_{26}p_{62}p_{38}p_{83} - p_{24}p_{02}p_{15}p_{41}p_{38}p_{83} \\
&+ p_{84}p_{15}p_{01}p_{26}p_{62}p_{48} - p_{24}p_{32}p_{83}p_{15}p_{01}p_{48}) + \mu_6(p_{26}p_{02} - p_{26}p_{02}p_{38}p_{83} + p_{26}p_{32}p_{13}p_{01} \\
&- p_{15}p_{51}p_{26}p_{02} - p_{84}p_{26}p_{02}p_{48} + p_{15}p_{51}p_{26}p_{02}p_{38}p_{83} + p_{15}p_{01}p_{26}p_{32}p_{53}^{(7)} + p_{54}^{(7)}p_{15}p_{41}p_{26}p_{02}p_{38}p_{83} \\
&+ p_{54}^{(7)}p_{15}p_{01}p_{26}p_{32}p_{83}p_{48} + p_{84}p_{51}p_{15}p_{26}p_{02}p_{48} - p_{84}p_{15}p_{41}p_{26}p_{02}p_{38}p_{53}^{(7)} - p_{84}p_{15}p_{01}p_{26}p_{32}p_{53}^{(7)}p_{48} \\
&- p_{84}p_{26}p_{02}p_{38}p_{13}p_{41} - p_{84}p_{26}p_{32}p_{13}p_{01}p_{48} - p_{54}^{(7)}p_{15}p_{41}p_{26}p_{02}) \quad (35)
\end{aligned}
$$

and D_1 is already specified in equation 24

III. Busy Period for Repairman

Let $B_i(t)$ notates the probability that the repairman is busy on the job when the system is at a regenerative state S_i at time $t = 0$. Using the probabilistic arguments as described above, The following recursive relations are obtained:

$$B_0(t) = q_{01}(t)©B_1(t) + q_{02}(t)©B_2(t) \quad (36)$$
$$B_1(t) = W_1(t) + q_{10}(t)©B_0(t) + q_{13}(t)©B_3(t) + q_{15}(t)©B_5(t) \quad (37)$$
$$B_2(t) = W_2(t) + q_{20}(t)©B_0(t) + q_{24}(t)©B_4(t) + q_{26}(t)©B_6(t) \quad (38)$$
$$B_3(t) = q_{32}(t)©B_2(t) + q_{38}(t)©B_8(t) \quad (39)$$
$$B_4(t) = q_{41}(t)©B_1(t) + q_{48}(t)©B_8(t) \quad (40)$$
$$B_5(t) = W_5(t) + q_{51}(t)©B_1(t) + q_{53}^{(7)}(t)©B_3(t) + q_{54}^{(7)}(t)©B_4(t) \quad (41)$$
$$B_6(t) = W_5(t) + q_{62}(t)©B_2(t) + q_{63}^{(7)}(t)©B_3(t) + q_{64}^{(7)}(t)©B_4(t) \quad (42)$$
$$B_8(t) = W_8(t) + q_{83}(t)©B_3(t) + q_{84}(t)©B_4(t) \quad (43)$$

Where

$$
\begin{aligned}
&W_0(t) = e^{-(\lambda_1+\lambda_2)t}, W_1(t) = e^{-(\lambda_1+\lambda_2)t}\bar{G_1}(t), W_2(t) = e^{-(\lambda_1+\lambda_2)t}\bar{G_2}(t), \\
&W_5(t) = e^{-(\lambda_2)t}\bar{G_3}(t), \\
&W_6(t) = e^{-(\lambda_1)t}\bar{G_3}(t), W_8(t) = (a+b)\bar{G_3}(t)
\end{aligned}
$$

Taking Laplace transform of equations (36-43) we get $B_0^*(s)$.
we have

$$B_0(t) = \lim_{s\to 0}(sB_0^*(s))$$

Expected busy period of a repairman is given by

$$B_0 = \lim_{t\to 0}(B_0(t)) = N_3/D_1 \quad (44)$$

where

$$
\begin{aligned}
N_3 = \mu_1[&(p_{26}p_{64}^{(7)} + p_{24})(p_{02}p_{41} - p_{41}p_{83}p_{38}p_{02} - p_{83}p_{32}p_{48}p_{01}) + (p_{84}p_{48} + p_{83}p_{38}) \\
&(p_{62}p_{26}p_{01} - p_{01}) - p_{63}^{(7)}p_{32}p_{26}p_{01}(1 - p_{84}p_{48}) - p_{01}(p_{62}p_{26} - 1)] + \mu_2[(1 - p_{84}p_{48})(p_{53}^{(7)}p_{32}p_{15}p_{01}
\end{aligned}
$$

$$+ p_{13}p_{32}p_{01} + p_{02} + p_{15}p_{51}p_{02}) - p_{83}p_{38}p_{02}(1 - p_{51}p_{15}) - p_{41}p_{15}p_{38}p_{02}(p_{53}^{(7)}p_{84} + p_{83}p_{54}^{(7)}) + p_{15}p_{54}^{(7)}(p_{83}p_{32}p_{48}p_{01} - p_{41}p_{02}) - p_{41}p_{13}p_{84}p_{38}p_{02}] + \mu_5[(1 - p_{62}p_{26})(p_{15}p_{01} - p_{84}p_{15}p_{48}p_{01} - p_{83}p_{15}p_{01}p_{38}) + (1 - p_{84}p_{48})(-p_{63}^{(7)}p_{32}p_{15}p_{26}p_{01}) + (p_{02}p_{41}p_{15}p_{26}p_{64}^{(7)} + p_{02}p_{41}p_{15}p_{24})(1 - p_{83}p_{38}) + p_{15}p_{26}(p_{41}p_{63}^{(7)}p_{84}p_{38}p_{02} - p_{83}p_{32}p_{64}^{(7)}p_{48}p_{01})] + \mu_6[(1 - p_{15}p_{51})(-p_{02}p_{84}p_{26}p_{48} - p_{83}p_{26}p_{38}p_{02} + p_{02}p_{26})(1 - p_{84}p_{48})(p_{53}^{(7)}p_{32}p_{15}p_{26}p_{01}p_{13}p_{32}p_{26}p_{01}) + p_{83}p_{32}p_{15}p_{54}^{(7)}p_{26}p_{48}p_{01}] + \mu_8[(1 - p_{62}p_{26})(p_{13}p_{01}p_{38} + p_{53}^{(7)}p_{15}p_{01}p_{38}) + (1 - p_{15}p_{51})(p_{02}p_{24}p_{48} + p_{02}p_{64}^{(7)}p_{26}p_{48} + p_{63}^{(7)}p_{26}p_{38}p_{02}) + (p_{24} + p_{26}p_{64}^{(7)})(p_{13}p_{32}p_{48}p_{01} + p_{53}^{(7)}p_{32}p_{15}p_{48}p_{01}) - p_{15}p_{26}p_{54}^{(7)}p_{48}p_{01}(p_{62} + p_{63}^{(7)}p_{32}) + p_{41}p_{15}p_{26}p_{38}p_{02}(p_{53}^{(7)}p_{64}^{(7)} - p_{63}^{(7)}p_{54}^{(7)}) + p_{15}p_{54}^{(7)}p_{48}p_{01} + p_{41}p_{13}p_{26}p_{38}p_{64}^{(7)}p_{02}] \quad (45)$$

and D_1 is already specified in equation 24.

IV. Expected No. of Repairs

Let $V_i(t)$ notate no. of repairs performed by repairman in the time interval (0 to t] when the system is at regenerative state S_i at time $t = 0$. The general formula for $V_i(t)$ is given by

$$V_i(t) = \sum_j Q_{ij}^{(n)}(t) \circledS [\alpha_j + V_i(t)] \quad (46)$$

Where $Q_{ij}(t)$ is the probability of system transition from the regenerative state i to regenerative j and $\alpha_j = 1$ if the repairman starts new job at regenerative state j, otherwise $\alpha_j = 0$. Using Equation 46 the following recursive relations are obtained:

$$V_0(t) = Q_{01}(t) \circledS [1 + V_1(t)] + Q_{02}(t) \circledS [1 + V_2(t)] \quad (47)$$
$$V_1(t) = Q_{10}(t) \circledS V_0(t) + Q_{13}(t) \circledS V_3(t) + Q_{15}(t) \circledS V_5(t) \quad (48)$$
$$V_2(t) = Q_{20}(t) \circledS V_0(t) + Q_{24}(t) \circledS V_4(t) + Q_{26}(t) \circledS V_6(t) \quad (49)$$
$$V_3(t) = Q_{32}(t) \circledS V_2(t) + Q_{38}(t) \circledS V_8(t) \quad (50)$$
$$V_4(t) = Q_{41}(t) \circledS V_1(t) + Q_{48}(t) \circledS V_8(t) \quad (51)$$
$$V_5(t) = Q_{51}(t) \circledS V_1(t) + Q_{53}^{(7)}(t) \circledS V_3(t) + Q_{54}^{(7)}(t) \circledS V_4(t) \quad (52)$$
$$V_6(t) = Q_{62}(t) \circledS V_2(t) + Q_{63}^{(7)}(t) \circledS V_3(t) + Q_{64}^{(7)}(t) \circledS V_4(t) \quad (53)$$
$$V_8(t) = Q_{83}(t) \circledS V_3(t) + Q_{84}(t) \circledS V_4(t) \quad (54)$$

Taking Laplace Stieltjes Transformations of the equations (47-54) to get ${V_0}^{**}(s)$.
we have

$$V_0(t) = \lim_{s \to 0} (s V_0^{**}(s)) \quad (55)$$

The expected no. of repairs by repairman are given by

$$V_0 = \lim_{t \to 0} (V_0(t)) = N_4 / D_1$$

where

$$N_4 = (1 - p_{15}p_{51})(p_{38}p_{83}p_{26}p_{62} - p_{24}p_{32}p_{83}p_{48} + p_{84}p_{48}p_{26}p_{62} - p_{64}^{(7)}p_{26}p_{32}p_{83}p_{48} - p_{26}p_{62} - p_{83}p_{38} - p_{84}p_{48} + 1) + (1 - p_{26}p_{62})(-p_{84}p_{38}p_{13}p_{41} + p_{54}^{(7)}p_{15}p_{41}p_{38}p_{83} - p_{84}p_{15}p_{41}p_{38}p_{53}^{(7)} - p_{54}^{(7)}p_{15}p_{41}) + p_{26}p_{32}p_{63}^{(7)}(p_{54}^{(7)}p_{15}p_{41} + p_{15}p_{51} + p_{84}p_{48} - p_{84}p_{48}p_{15}p_{51} - 1) - p_{64}^{(7)}p_{15}p_{41}p_{26}p_{32}p_{53}^{(7)} - p_{64}^{(7)}p_{26}p_{32}p_{13}p_{41} - p_{24}p_{32}p_{53}^{(7)}p_{15}p_{41} - p_{24}p_{32}p_{13}p_{41} \quad (56)$$

and D_1 is already specified in equation 24.

IV. Profit Analysis

The expected total profit per unit time incurred to the system in steady state is given by

$$P_0 = C_0 A_0^F + C_1 A_0^R - C_2 B_0 - C_3 V_0 \tag{57}$$

Where
C_0 = revenue per unit up time at full capacity.
C_1 = revenue per unit up time at reduced capacity.
C_2 = cost per unit time when repairman is busy.
C_3 = cost per repair.

V. Particular Cases

For evaluation of above described various system performance measures and their graphical representation, the following particular cases are considered, where distribution of repair times has been taken as exponential. Let us assume that $g_1(t) = \alpha_1 e^{-\alpha_1 t}, g_2(t) = \alpha_2 e^{-\alpha_2 t}, g_3(t) = \alpha_3 e^{-\alpha_3 t}$ and remaining distributions same as in general case. Therefore we have

$$p_{01} = \frac{\lambda_1}{\lambda_1+\lambda_2}, \quad p_{02} = \frac{\lambda_2}{\lambda_1+\lambda_2}, \quad p_{10} = \frac{\alpha_1}{\lambda_2+\lambda_3+\alpha_1}, \quad p_{13} = \frac{\lambda_2}{\lambda_2+\lambda_3+\alpha_1},$$

$$p_{15} = \frac{\lambda_3}{\lambda_2+\lambda_3+\alpha_1}, \quad p_{20} = \frac{\alpha_2}{\lambda_1+\lambda_3+\alpha_2}, \quad p_{24} = \frac{\lambda_1}{\lambda_1+\lambda_3+\alpha_2}, \quad p_{26} = \frac{\lambda_3}{\lambda_1+\lambda_3+\alpha_2},$$

$$p_{32} = \frac{\alpha_1}{\alpha_1+\lambda_3}, \quad p_{38} = \frac{\lambda_3}{\lambda_3+\alpha_1}, \quad p_{41} = \frac{\alpha_2}{\alpha_2+\lambda_3}, \quad p_{48} = \frac{\lambda_3}{\lambda_3+\alpha_2},$$

$$p_{51} = \frac{\alpha_3}{\alpha_3+\lambda_2}, \quad p_{57} = \frac{\lambda_2}{\lambda_2+\alpha_3}, \quad p_{53}^{(7)} = a[\frac{\lambda_2}{\alpha_3+\lambda_2}], \quad p_{54}^{(7)} = b[\frac{\lambda_2}{\alpha_3+\lambda_2}],$$

$$p_{63}^{(7)} = a[\frac{\lambda_1}{\alpha_3+\lambda_1}], \quad p_{64}^{(7)} = b[\frac{\lambda_1}{\alpha_3+\lambda_1}], \quad p_{62} = \frac{\alpha_3}{\alpha_3+\lambda_1}, \quad p_{67} = \frac{\lambda_1}{\lambda_1+\alpha_3},$$

$$p_{83} = a, \quad p_{84} = b, \quad \mu_0 = \frac{1}{\lambda_1+\lambda_2}, \quad \mu_1 = \frac{1}{\lambda_2+\lambda_3+\alpha_1},$$

$$\mu_2 = \frac{1}{\lambda_1+\lambda_3+\alpha_1}, \quad \mu_3 = \frac{1}{\lambda_3+\alpha_1}, \quad \mu_4 = \frac{1}{\lambda_3+\alpha_2}, \quad \mu_5 = \frac{1}{\lambda_2+\alpha_3},$$

$$\mu_6 = \frac{1}{\lambda_1+\alpha_3}, \quad \mu_8 = \frac{1}{\alpha_3}$$

Estimation of Parameters

The various parameters regarding failure and repair rates involved in our studies are estimated as follows in table 2

Table 2: *Failure & repair rates*

Various rates	corresponding values
Failure rate of STG 1 (λ_1)	0.00043/hr
Failure rate of STG 2 (λ_2)	0.00043/hr
Failure rate of gridline (λ_3)	0.0067/hr
Repair rate of STG 1 (α_1)	0.0065/hr
Repair rate of STG 2 (α_2)	0.0063/hr
Repair rate of gridline (α_3)	0.34/hr

The various costs/revenue amounts involved in our studies are assumed hypothetically. The computed values of Various reliability measures for system performance are given in table 3.

Table 3: *Evaluation of various system effectiveness measures*

Mean time to system failure	39140 hrs.
Availability of the system at full capacity	0.50093/hr
Availability of the system at reduced capacity	0.001044/hr
Busy period of repairman for repair time only	0.12964/hr
Expected no. of repairs	0.000350/hr

VI. Results and Discussion

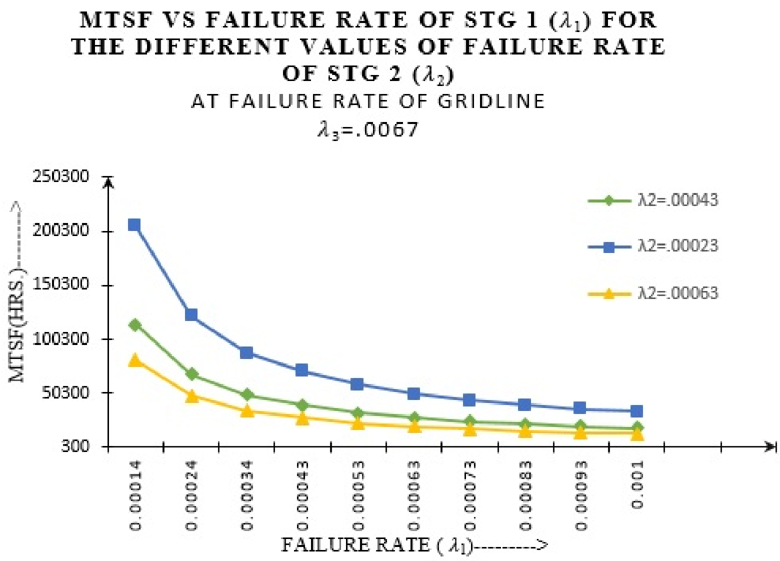


Figure 2: *MTSF vs* (λ_1)*Failure rate of STG*

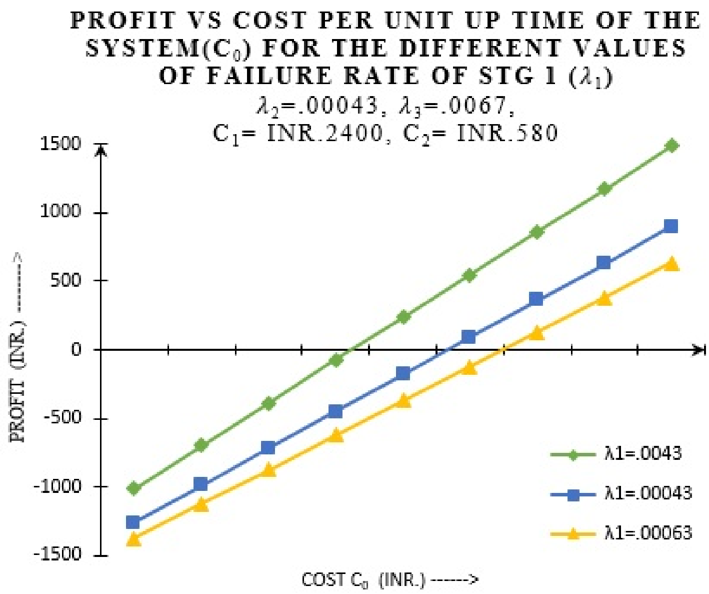


Figure 3: *profit vs cost per unit up time of system*

The figure 2, indicates that MTSF decreases as the falure rate of STG 1 increases and also gives lowering values for the greater values of failure rates of gridline. The graph in figure 3 interpreted that the profit increases with increasing the cost per unit up time of the system and decreases when failure rate of the STG 1 increases.

Table 4: *Cut Point for profit w.r.t. Revenue per unit up-time of the system.*

Failure rate of STG(/hr)	Revenue per unit up time(C_0) (Rs.)	Profit (Rs.)
$\lambda_1 = .00034$	$C_0 < or = or > 161$	negative or zero or positive
$\lambda_1 = .0034$	$C_0 < or = or > 380$	negative or zero or positive
$\lambda_1 = .034$	$C_0 < or = or > 606.$	negative or zero or positive

VII. Conclusion

In this paper self electricity generating System is Studied. The graphical study reveals the negative relationship between failure rates of units of captive power plant and profit gained by the plant. Adding the working of system at reduced capacity results in increasing its availability and profit. The derived results enable us to find acceptable values of revenue per unit up time of the system (Table 4) corresponding to failure rates of units of system. By using this analysis and graphical representations one can procure various system effectiveness measures for similar electricity generation plant.

References

[1] Chandrasekhar, P., Natarajan, R., and Yadavalli, V. (2004), A study on a two unit standby system with erlangian repair time, *Asia-Pacific Journal of Operational Research*, 21(03):271–277.

[2] Goyal, A., Taneja, G., and Singh, D. V. (2010). 3-unit system comprising two types of units with first come first served repair pattern except when both types of units are waiting for repair, *Journal of Mathematics and Statistics*, 6(3):316–320.

[3] Gupta, R. and Goel, R. (1991). Profit analysis of a two-unit cold standby system with abnormal weather condition, *Microelectronics Reliability*, 31(1):1–5.

[4] Parashar, B., Naithani, A., and Bhatia, P. (2012). Analysis of a 3-unit (induced draft fan) system with one warm standby, *International Journal of Engineering Science and Technology*, 4(11):4620–4628.

[5] Rizwan, S., Khurana, V., and Taneja, G. (2010). Reliability analysis of a hot standby industrial system, *International Journal of Modelling and Simulation*,30(3):315–322.

[6] Smith, W. L. (1955). Regenerative stochastic processes, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 232(1188), 6-31.

[7] Singh D. and Taneja G. (2013). Reliability analysis of a power generating system through gas and steam turbines with scheduled inspection, *Aryabhatta Journal of Mathematics & Informatics*, 5(2), 373-380.

[8] Singh, D., & Taneja, G. (2014). Reliability and cost-benefit analysis of a power plant comprising two gas and one steam turbines with scheduled inspection. *International Journal of Soft Computing and Engineering*, 4(3), 30-37.

[9] Rajesh, Taneja G. and Prasad J. (2018) [In Press]. Reliability of a gas turbine system with change in weather and optimisation of electricity price when working in single cycle,*International Journal of Agricultural and Statistical Sciences*.

[10] Rajesh, G. T., & Prasad, J. (2018). Reliability and availability analysis for a three-unit gas turbine power generating system with seasonal effect and FCFS repair pattern. *International Journal of Applied Engineering Research*, 13(12), 10948-10964.

RT&A, No 2 (68)
Volume 17, June 2022

MAP/PH/1 Queue with Vacation, Customer Induced Interruption, Optional Service, Breakdown and Repair Completion

G.Ayyappan, S.Sankeetha

•

Department of Mathematics, Pondicherry Engineering College, India.
Department of Mathematics, Saradha Gangadharan College, India.
ayyappanpec@hotmail.com sangeetha.sivarajp@gmail.com

Abstract

The paper considers a single server that provides consumers with both regular and optional services. The system's inter-arrival time is determined by a Markovian Arrival Process (MAP), the service time is determined by a phase type distribution, and the remaining random variables are distributed exponentially. This system was represented as a QBD process, with the block elements of the generated matrix having finite dimensions, to investigate steady state. Additionally, we addressed the busy period and waiting time distribution for our concept. The system's performance parameters are calculated and graphically shown.

Keywords: Markovian Arrival Process, PH distribution, Vacation, Optional service, Breakdown and Repair.

1. Introduction

Contribution of Nuets (1979) is immeasurable in the field of stochastic process. He pioneered the Markovian Point Process, which led to the development of the Markovian Arrival Process and the Batch Markovian Arrival Process. In his concept of communication and computer application, Lucantoni (1990) established these two arrival processes. One of the most important characteristics of MAP is that it can be used to solve stochastic models using matrix analytic solutions. Chakravarty (2010) in the Encyclopedia of Operations Research and Management Science streamlined this useful tool to make it easier to understand. The discrete and continuous cases of MAP are defined. The parameters utilised in MAP are D_0 and D_1 of dimension m in continuous time, where D_0 is a non-singular stable matrix that rules the transition corresponding to no arrival and D_1 governs the transition relating to arrival. The generator matrix $Q = D_0 + D_1$.The stationary distribution vector π satisfies the system $\pi(D_0 + D_1) = 0$ and $\pi e = 1$.The constant $\lambda = \pi D_1$ is called the fundamental rate of a MAP defined by Latuche at al (1999).

The concept of vacation in the queuing system was introduced by Levy and Yechiali (1975). Vacation is a time where the server is unavailable for service, for a short period of time due to many reasons like filing up the bills or document, verifying with other data etc or even a break. In this real world, there are many occasions where the server is busy or continues to work with low speed during his vacation. Servi (2002) and Finn classify this type of vacation as a working vacation. Doshi (1986), Takagi (1991) and Tian and Zhang (2006) has contributed an excellent survey on the vacation model. Ket et al (2010) and Tian et al (2009) have also recently contemplated on vacation and working vacation. This working vacation concept was introduced and further studied in a retail queue by Do (2010).

A detailed study has been performed by Doshi (1986) for the queueing model with vacations, breakdown and repair in his survey with demonstration. Ayyappan and Thamizhselvi (2018) have reviewed a priority retail model with vacation and also studied the time dependent PGF. Further, second optional service under mixed priority service was studied by K.Jeganathan (2015) in linear retail inventory system. Breakdown and repair process are unavoidable concepts in production unit, service stations etc. When the server gets breakdown then the server gets terminated and goes to the repair process. Depending on the model, the server begins serving the customer whose service was interrupted or begins serving a new customer after the repair process is completed. This concepts was studied in various queueing models by Gaver (1959) , Levy and Yechilai (1976) and many more are interested in this concept.

In reality, the service of the server can terminate for a short period of time. This phenomena is named as interruption which is one of the unavoidable aspects faced by both the server and the customer in the system due to many reasons like emergency call or work, the server/machine may get breakdown, external influence, get some suggestions/ ideas from the fellow workers etc. This interruption was studied in priority queue by Jaiswak(1961). Geramsimov(1973) came up with an idea for investigating an interrupted customer with an algorithm where another queue for interrupted customers were formed and served. This concept was developed in Computer and Communication System by Gelenbe and Derochette(1978). Takine and Sengupta (1997) developed this concept in a MAP process. Rakesh Kumar(2014) investigated discouraged arrivals and customer retention in a single server Markovian queueing system. Rakesh Kumar and Bhavneet Singh Soodam(2019) investigated linked imputes and reneging for a single server queueing model. El-Taha(2003) introduced two server in series where the customer gets interrupted by the set of proposed time threshold while getting service from the server one. Server two offer service only to the interrupted customer or else the customer leaves the system. Krishnamoorthy et al.(2009) developed the model in a single server queue in a level-dependent-quasi-birth and death (LDQBD) process. He further generalized this model in (2011) where a super clock is defined for his predetermined threshold time. Kilmenok and Dudin(2012) and Krishnamurthy et al.(2010) have also investigated further where interrupted customer service has been rejected. Varghese et al.(2010) introduced customer induced service interruption where the customer gets self interrupted while being served by the server. Further extension in this concept has been made by them in the year 2012.

2. Model Description

In this classical queuing model, a single server is considered with the infinite capacity queue where the customer arrives according to Markovian Arrival Process with the parameter matrix D_0 and D_1 are of dimension m. The customer in the service station can be self interrupted and moves onto the buffer 1 which is of maximum capacity K with an exponential distribution δ. After the completion of interruption, the customer moves onto the buffer 2 with an exponential distribution θ which is also of maximum capacity K where the customer gets served by the server. Optional service will be provided by the server whenever the customer needs it. The service time of the server offering service for the customer from the queue, buffer 2 and optional service follows phase type distribution $PH(\alpha_1, t_1), PH(\alpha_2, t_2), PH(\alpha_3, t_3)$ respectively of order n. The vector T_1^0, T_2^0, T_3^0 is given by $T_1^0 = -T_1 e, T_2^0 = -T_2 e, T_3^0 = -T_3 e$ respectively.

The interrupted customer will only enter buffer 1 if space is available; else, the customer will be lost indefinitely. The server follows non-preemptive priority for the customer in the buffer 2 over the customer in the queue. Thus a preference will be given to the customer in buffer 2 whenever the free server offers service. The customer in buffer 2 will be served by first in first service order. When the system is empty, the server avails vacation following exponential distribution with the parameter η. While the server is busy, breakdown of the server may occur. It follows exponential distribution with the parameter γ. Consequently, the repair process starts immediately with the phase type distribution $PH(\beta, R)$ of order r. The vector R^0 is given by

$R^0 = -Re$. After receiving the regular service, the customer can opt for optional service. During the optional service availed by the customer, if the server gets breakdown then the interrupted customer is considered to be lost forever from the system. The server is idle when there are no customers in the queue and buffer 2.

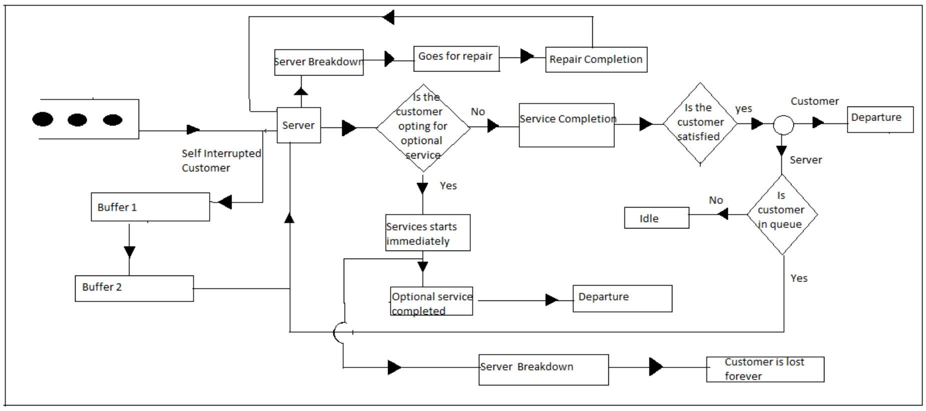


Figure 1

Figure 1: *Schematic Representation of Our Model*

To find a matrix-geometric type solution, this model is investigated as a QBD process. For a thorough examination of Matrix Analytic Methods, see Neuts (1981), Latouche and Ramaswami (1999). The state space under the considered QBD model is defined and the structure of the infinitesimal generator is also investigated using the following notations.

Let

- N(t) be the number of customers in the system at time t
- S(t) be the server status at time t

where

$$S(t) = \begin{cases} 0, & \text{if server is idle} \\ 1, & \text{if server is offering service to the customer in the main queue} \\ 2, & \text{if server is offering service to the customer in the buffer 2} \\ 3, & \text{if server is offering optional service} \\ 4, & \text{if main server is availing vacation} \\ 5, & \text{if the server is under repair} \end{cases}$$

- $N_1(t)$ be the number of customers in the buffer 1 at time t
- $N_2(t)$ be the number of customers in the buffer 2 at time t
- $R(t)$ be the Repair phase at time t
- $C(t)$ be the service phase at time t
- $M(t)$ be the phase of the Markovian Arrival Process at time t
- $M_1 = \frac{K(K+1)}{2}$
- $M_2 = \frac{(K+1)(K+2)}{2}$, where K is the maximum capacity of buffer 1.

$\{N(t), S(t), N_1(t), N_2(t), C(t), R(t), M(t); t \geq 0\}$ is representation of this model by the Continuous Time Markov Process with the state space

$$\Omega = l(0) \cup l(i)$$

where

$$\begin{gathered} l(0) = \{(0,0,0,i_2) : 1 \leq i_2 \leq K\} \\ \cup \{(0,j,i_1,i_2,k,l) : j = 2; 0 \leq i_1 \leq K; 1 \leq i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ \cup \{(0,j,i_1,i_2,k,l) : j = 3,5; 0 \leq i_1, i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ \cup \{(0,4,0,0,l) : 0 \leq l \leq m\} \end{gathered}$$

for $i \geq 1$,

$$\begin{gathered} l(i) = \{(i,j,i_1,i_2,k,l) : j = 1,3,5; 0 \leq i_1, i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ \cup \{(i,j,i_1,i_2,k,l) : j = 2; 0 \leq i_1 \leq K; 1 \leq i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ \cup \{(i,4,0,0,l) : 0 \leq l \leq m\} \end{gathered}$$

The infinitesimal matrix generation of the QBD process is given by

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0\cdots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0\cdots \\ 0 & A_2 & A_1 & A_0 & 0 & 0\cdots \\ 0 & 0 & A_2 & A_1 & A_0 & 0\cdots \\ \cdots & \cdots & \cdots & \ddots & \ddots & \ddots \end{bmatrix}$$

where

$$B_{00} =$$

$$\begin{bmatrix} b^{(00)}_{(11)} & b^{(00)}_{(12)} & 0 & 0 & 0 \\ 0 & T_2 \oplus D_0 - \gamma I_m \otimes I_{M_1} & diag(a_K, a_{K-1}, ..., a_1) & e_{M_1} \otimes qT_2^0 \otimes I_m & diag(e_K, e_{K-1}, ..., e_1) \\ 0 & 0 & b^{(00)}_{(33)} & e_{M_2} \otimes T_3^0 \otimes I_m & \gamma I_{M_2rm} \\ \gamma I_m & 0 & 0 & D_0 - \gamma I_m & 0 \\ b^{(00)}_{(51)} & diag(c_K, c_{K-1}, ..., c_1) & 0 & 0 & (R \oplus D_0) \otimes I_{M_2} \end{bmatrix}$$

$$b^{(00)}_{(11)} = \begin{bmatrix} D_0 & 0 \\ 0 & I_K \otimes (D_0 - \theta I_m) \end{bmatrix}$$

$$b^{(00)}_{(12)} = \begin{bmatrix} 0 \\ diag[\theta I_m \quad 0] \end{bmatrix}$$

$$b^{00}_{(33)} = \begin{bmatrix} T_3 \oplus (D_0 - \gamma I_m) \otimes I_{K+1} & 0 \\ 0 & T_3 \oplus (D_0 - (\gamma + \theta) I_m) \otimes I_{K+1} \end{bmatrix} + \begin{bmatrix} 0 \\ diag(b_K, b_{K-1}, ..., b_1) \end{bmatrix}$$

$$b^{(00)}_{(51)} = \begin{bmatrix} e_n \otimes R^0 \beta I_m & 0 & 0 & 0 & 0 \\ 0 & e_n \otimes R^0 \beta I_m & 0 & 0 & 0 \\ 0 & 0 & e_n \otimes R^0 \beta I_m & 0 & 0 \\ 0 & 0 & 0 & e_n \otimes R^0 \beta I_m & 0 \\ 0 & 0 & 0 & 0 & e_n \otimes R^0 \beta I_m \end{bmatrix}$$

$$B_{01} = \begin{bmatrix} b^{(0,1)}_{(1,1)} & 0 & 0 & 0 & 0 \\ 0 & I_{nM_1} \otimes D_1 & 0 & 0 & 0 \\ 0 & 0 & I_{nM_2} \otimes D_1 & 0 & 0 \\ 0 & 0 & 0 & D_1 & 0 \\ 0 & 0 & 0 & 0 & I_{rM_2} \otimes D_1 \end{bmatrix}$$

$$b^{(01)}_{(11)} = \begin{bmatrix} D_1 & 0 & 0 & 0 & 0 \\ 0 & D_1 & 0 & 0 & 0 \\ 0 & 0 & D_1 & 0 & 0 \\ 0 & 0 & 0 & D_1 & 0 \\ 0 & 0 & 0 & 0 & D_1 \end{bmatrix}$$

$$B_{10} = \begin{bmatrix} b^{(10)}_{(11)} & diag(d_K, d_{K-1}, ..., d_1) & pT_1^0\alpha_1 \otimes I_{M_2} & e_n \otimes qT_1^0\alpha_1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$b^{(10)}_{(11)} = \begin{bmatrix} 0 & e_n \otimes \delta I_{mn} & 0 & 0 & 0 \\ 0 & e_n \otimes pT_1^0\alpha_1 & e_n \otimes \delta I_{mn} & 0 & 0 \\ 0 & 0 & e_n \otimes pT_1^0\alpha_1 & e_n \otimes \delta I_{mn} & 0 \\ 0 & 0 & 0 & e_n \otimes pT_1^0\alpha_1 & 0 \\ 0 & 0 & 0 & 0 & e_n \otimes pT_1^0\alpha_1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} a^2_{(1,1)} & diag(d_K, d_{K-1}, ..., d_1) & pT_1^0\alpha_1 \otimes I_{mM_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$a^2_{(11)} = \begin{bmatrix} pT_1^0\alpha_1 \otimes Im & \delta I_{mn} & 0 & 0 & 0 \\ 0 & pT_1^0\alpha_1 \otimes Im & \delta I_{mn} & 0 & 0 \\ 0 & 0 & pT_1^0\alpha_1 \otimes Im & \delta I_{mn} & 0 \\ 0 & 0 & 0 & pT_1^0\alpha_1 \otimes Im & 0 \\ 0 & 0 & 0 & 0 & pT_1^0\alpha_1 \otimes Im \end{bmatrix}$$

$$A_1 = \begin{bmatrix} a^1_{(11)} & 0 & 0 & 0 & \gamma I_{rmM_2} \\ a^1_{(21)} & a^1_{(22)} & diag(a_K, a_{K-1}, ..., a_1) & 0 & diag(e_K, e_{K-1}, ..., e_1) \\ T_3^0\alpha_3 \otimes I_{M_2} & 0 & a^1_{(3,3)} & 0 & \gamma I_{rmM_2} \\ \eta I_{mM_2} & 0 & 0 & D_0 - \eta I_m & 0 \\ R^0\beta \otimes I_{mM_2} & 0 & 0 & 0 & (R + D_0) \otimes I_{M_2} \end{bmatrix}$$

$$a^1_{(11)} = \begin{bmatrix} T_1 \oplus D_0 - (\gamma+\delta) I_m \otimes I_K & 0 & 0 & 0 \\ 0 & T_1 \oplus D_0 - \gamma I_m & 0 & 0 \\ 0 & 0 & diag(f_K, f_{K-1}, ..., f_1) & 0 \\ 0 & 0 & 0 & T_1 \oplus D_0 - (\gamma+\theta) I_m \otimes I_{K-1} \end{bmatrix} + diag(b_K, b_{K-1}, ..., b_1)$$

$$a^1_{(21)} = diag \begin{bmatrix} I_K \otimes qT_2^0\alpha_2 \otimes I_n & 0 \end{bmatrix}$$

$$a^1_{(22)} = \begin{bmatrix} T_2 \oplus D_0 - \gamma I_m \otimes I_K & 0 \\ 0 & T_2 \oplus D_0 - (\gamma+\theta) I_m \otimes I_{M_1} \end{bmatrix} + \begin{bmatrix} 0 \\ diag(g_{K-1}, ..., g_1) \end{bmatrix} + diag \begin{bmatrix} 0 \\ qT_2^0\alpha_2 \otimes I_{mn(i-1)} \end{bmatrix}$$

$$a^1_{(33)} = \begin{bmatrix} T_3 \oplus (D_0 - \gamma) I_m \otimes I_{K+1} & 0 \\ 0 & diag(h_K, h_{K_1}, ..., h_1) \end{bmatrix} + \begin{bmatrix} 0 \\ diag(g_{K-1}, ..., g_1) \end{bmatrix}$$

$$A_0 = \begin{bmatrix} I_{nM_2} \otimes D_1 & 0 & 0 & 0 & 0 \\ 0 & I_{nM_1} \otimes D_1 & 0 & 0 & 0 \\ 0 & 0 & I_{nM_2} \otimes D_1 & 0 & 0 \\ 0 & 0 & 0 & D_1 & 0 \\ 0 & 0 & 0 & 0 & I_{rM_2} \otimes D_1 \end{bmatrix}$$

$$a_i = pT_2^0\alpha_2 \otimes I_{im}$$

$$b_i = \theta I_{inm}$$

$$c_i = \begin{bmatrix} 0 \\ R^0\beta \otimes I_{im} \end{bmatrix}$$

$$d_i = \begin{bmatrix} 0 & 0 \\ pT_1^0\alpha_1 \otimes I_i & \delta I_{mn(i-1)} \end{bmatrix}$$

$$e_i = \begin{bmatrix} 0 & \gamma I_{irm} \end{bmatrix}$$

$$f_i = \begin{bmatrix} T_1 \oplus D_0 - (\gamma + \delta + \theta) I_m \otimes I_{i-1} & 0 \\ 0 & T_1 \oplus D_0 - (\gamma + \theta) I_m \end{bmatrix}$$

$$g_i = \begin{bmatrix} 0 & \theta I_{inm} \end{bmatrix}$$

$$h_i = T_3 \oplus D_0 - (\gamma + \theta) I_m \otimes I_K, \ K \le i \le 1.$$

3. Analysis of the Stability Condition

The square matrix $A = A_0 + A_1 + A_2$ is defined of order $m[M_2(2n + r) + (nM_1 + 1)]$ as an irreducible infinitesimal generator matrix. The invariant probability vector $\wp$ defined as $\wp = (\wp_0, \wp_1, \wp_2, \wp_3, \wp_4)$ satisfies the condition $\wp A = 0$ and $\wp e = 1$. The vector $\wp$ can be computed by solving the following equations.

$$\wp_0(I_{nM_2} \otimes D_1 + a^1_{(1,1)} + a^2_{(1,1)}) + \wp_1(a^1_{(2,1)}) + \wp_2(T_3^0\alpha_3 \otimes I_{M_2}) + \wp_3(\eta I_m) + \wp_4(R^0\beta \otimes I_{M_2}) = 0$$
$$\wp_0(diag(d_K, d_{K-1}, ..., d_1)) + \wp_1(I_{nM_1} \otimes D_1 + a^1_{(2,2)}) = 0$$
$$\wp_0(pT_1^0\alpha_1 \otimes I_{M_2}) + \wp_1(pT_2^0\alpha_2 \otimes I_{M_1}) + \wp_2(I_{nM_2} \otimes D_1 + a^1_{(3,3)}) = 0$$
$$\wp_3(D_1 + D_0 - \eta I_m) = 0$$
$$\wp_0(diag(\gamma I_{rmM_2}) + \wp_1(diag(e_K, e_{K-1}, ..., e_1)) + \wp_2(\gamma I_{rmM_2}) + \wp_4(I_{rM_2} \otimes D_1 + (R + D_0) \otimes I_{M_2}) = 0$$

For the system to attain stability, the necessary and sufficient condition is $\wp A_0 e \le \wp A_2 e$

$$(\wp_0 + \wp_2)(e_{nM_2} \otimes D_1 e_m) + \wp_1(e_{nM_1} \otimes D_1 e_m) + \wp_3(D_1 e_m) + \wp_4(e_{rM_2} \otimes D_1 e_m) < \wp_0(a^2_{(1,1)}) + \wp_1(diag(d_K, d_{K-1}, ..., d_1)) + \wp_2(pT_1^0\alpha_1 \otimes e_{M_2})$$

4. The Invariant Probability Vector

The unique solution to $XQ = 0$ and $Xe = 1$ is the transition probability vector of the infinitesimal generator Q. This X can be partitioned into $(X_0, X_1, X_2....)$ where each X_i is the row vector corresponding to the server status. The dimension of X_0 is $m[(K+1) + nM_1 + M_2(n+r) + 1]$, and the remaining probability vectors $X_1, X_2, X_3, ...$ are of equal dimension $m[M_2(2n+r) + (nM_1 + 1)]$. The steady state probability vector has a matrix geometric structure satisfying the condition for stability is as follows,

$$X_i = X_1 R^{i-1}, i = 2, 3, 4, ...,$$

where R, the rate matrix, is the minimal non-negative solution to the matrix quadratic equation

$$R^2 A_2 + RA_1 + A_0 = 0$$

and the boundary states X_0 and X_1 is the result of solving the equations

$$X_0 B_{00} + X_1 B_{10} = 0$$
$$X_0 B_{01} + X_1 (A_1 + RA_2) = 0$$

subject to normalizing condition

$$X_0 e + X_1 (I - R)^{-1} e = 1$$

Lautouche and Ramaswamy(1999) have embellished the calculation of rate matrix R by developing Logarithmic Reduction Algorithm, which helps us to obtained R easily.

$$\text{Step 1} : H \leftarrow (-A_1)^{-1} A_0, L \leftarrow (-A_1)^{-1} A_2, G = L \text{and } T = H.$$

$$\begin{gathered}\text{Step 2} : U = HL + LH; \\ M = H^2; \\ H = (I - U)^{-1} M; \\ M = L^2; \\ L = (I - U)^{-1} M; \\ G = G + TL; \\ T = TH;\end{gathered}$$

continue Step 1 until $\|e - Ge\|_\infty < \epsilon$.

$$\text{Step 3} : R = -A_0 (A_1 + A_0 G)^{-1}.$$

5. Analysis of Busy Period

In a classical queueing model, the busy period is defined as the time between a customer arriving at an empty queue and the first epoch after that when the queue becomes empty again. Considering a QBD process, Latouche.G (1978) has coined the term fundamental period defined as the first passage time from level i to level $i-1$, $i \geq 2$. For the boundary states $i = 0, 1$ has to be dealt separately. We can also observe that for all level i, where $i \geq 2$ there are $m[M_2(2n+r) + (nM_1 + 1)]$ states.

Notations:

- $G_{vv'}(k,x)$: The conditional probability that the QBD process enters the level $u-1$ at time $t=0$, by making merely k transition to the left and also by entering the state (u,v') conditioned that it only started in the state (u,v) at time $t=0$.
- The transition matrix $\bar{G}_{vv'}(z,s)=\sum_0^\infty z^k\int_0^\infty e^{-sx}dG_{vv'}(k,x):|z|\leq 1, Re(s)\geq 0$
- $\bar{G}(z,s)$: The matrix $(G_{vv'}(z,s))$, satisfying $\bar{G}(z,s)=z\left[sI-A_1\right]^{-1}A_2+\left[sI-A_1\right]^{-1}A_0\bar{G}^2(z,s)$
- $G=G_{vv'}=\bar{G}(0,1)$ is the first passage time without the boundary states.
- $\bar{G}^{(1,0)}_{(vv')}(k,x)$ is the conditional probability that enters the level 0 from 1 at time $t=0$.
- $\bar{G}^{(0,0)}_{(vv')}(k,x)$ is the first conditional probability returning to level 0.
- $\Re_{1v}$ is the expected first passage time between the levels u and $u-1$, the process in the state (u,v),at time $t=0$.
- $\bar{\Re}_1$ is the column vector $\Re_{1v}$ as its entries.
- $\Re_{2v}$ is the average number of customer who received service in the first passage time between the levels u and $u-1$, begins in the state (u,v), at time $t=0$.
- $\bar{\Re}_2$ is the column vector $\Re_{2v}$ as its entries.
- $\bar{\Re}_1^{(1,0)}$ is the average first passage times from the level 1 to 0.
- $\bar{\Re}_2^{(1,0)}$ is the average number of service completions during the first passage time from the level 1 to 0.
- $\bar{\Re}_1^{(0,0)}$ is the average first return time to level 0.
- $\bar{\Re}_2^{(0,0)}$ is the average number of completed services in the initial return time to level 0.

G matrix can be computed with the help of the result $G=-[A_1+RA_2]^{-1}A_2$ where the rate matrix R is already evaluated using Logarithmic Reduction Algorithmic technique. For the boundary states namely 1 and 0 we have the equations satisfied by $\bar{G}^{(1,0)}(z,s)$ and $\bar{G}^{(0,0)}(z,s)$ respectively.

$$\bar{G}^{(1,0)}(z,s)=z\left[sI-A_1\right]^{-1}B_{10}+\left[sI-A_1\right]^{-1}A_0\bar{G}(z,s)\bar{G}^{(1,0)}(z,s)$$
$$\bar{G}^{(0,0)}(z,s)=z\left[sI-B_{00}\right]^{-1}B_{01}\bar{G}^{(1,0)}(z,s).$$

Since G, $\bar{G}^{(1,0)}(z,s)$, $\bar{G}^{(0,0)}(z,s)$ are stochastic moments can be easily evaluate as follows.

$$\begin{aligned}
\Re_1 &= -\tfrac{\partial}{\partial s}\bar{G}(z,s)|_{s=0,z=1}=-\left[A_0(G+1)+A_1\right]^{-1}e\\
\Re_2 &= \tfrac{\partial}{\partial z}\bar{G}(z,s)|_{s=0,z=1}=-\left[A_0(G+1)+A_1\right]^{-1}A_2e\\
\Re_1^{(1,0)} &= -\tfrac{\partial}{\partial s}\bar{G}^{(1,0)}(z,s)|_{s=0,z=1}=-\left[A_1+A_0G\right]^{-1}\left[A_0\Re_1+e\right]\\
\Re_2^{(1,0)} &= \tfrac{\partial}{\partial z}\bar{G}^{(1,0)}(z,s)|_{s=0,z=1}=-\left[A_1+A_0G\right]^{-1}\left[B_{10}e+A_0\Re_2\right]\\
\Re_1^{(0,0)} &= -\tfrac{\partial}{\partial s}\bar{G}^{(0,0)}(z,s)|_{s=0,z=1}=-B_{00}^{-1}\left[e+B_{01}\Re_1^{(1,0)}\right]\\
\Re_2^{(0,0)} &= \tfrac{\partial}{\partial z}\bar{G}^{(0,0)}(z,s)|_{s=0,z=1}=-B_{00}^{-1}B_{01}\Re_2^{(1,0)}.
\end{aligned}$$

6. Analysis of Waiting Time Distribution

Analysis of distribution of waiting time period of the customer in the queue has been performed in this section by using the first passage time analysis. Let $W(t)$, where $t \geq 0$, denote the distribution function of the waiting time of the tagged incoming customer in the system. The customer in the system has to wait in order to get service from the server if the server is busy or availing vacation or under repair. Otherwise, the customer in the system can get immediate service without any delay when the server is idle. The state space of absorbing time in a Markov chain is given by

$$(*) \cup \bar{0}, \bar{1}, \bar{2}, \bar{3}, ...$$

where $(*)$ denotes the absorbing state, in which the tagged customer gets service from the server without delay and it is defined as

$$(*) = (0,0,0,i_2) : 1 \leq i_2 \leq K$$

The state space of level 0 in a Markov chain is given by

$$\begin{aligned}\bar{0} = &\{(0,j,i_1,i_2,k,l) : j = 2; 0 \leq i_1 \leq K; 1 \leq i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ &\cup \{(0,j,i_1,i_2,k,l) : j = 3,5; 0 \leq i_1, i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ &\cup \{(0,4,0,0,l) : 0 \leq l \leq m\}\end{aligned}$$

The state space of level $i \geq 1$ in a Markov chain is given by

$$\begin{aligned}\bar{i} = &\{(i,j,i_1,i_2,k,l) : j = 1,3,5; 0 \leq i_1, i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ &\cup \{(0,j,i_1,i_2,k,l) : j = 2; 0 \leq i_1 \leq K; 1 \leq i_2 \leq K; i_1 + i_2 \leq K; 0 \leq k \leq n; 0 \leq l \leq m\} \\ &\cup \{(i,4,0,0,l) : 0 \leq l \leq m\}\end{aligned}$$

The transition matrix $\bar{Q}$ is given by

$$\bar{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \cdots \\ H_0 & F_0 & 0 & 0 & 0 \cdots \\ H_1 & F_{10} & F & 0 & 0 \cdots \\ 0 & 0 & F_2 & F & 0 \cdots \\ \cdots & \cdots & \ddots & \ddots & \cdots \end{bmatrix}$$

where the block matrix are as follows.

$$H_0 = \begin{bmatrix} 0 \\ 0 \\ \gamma \\ R^0 \beta \otimes e_n \end{bmatrix}$$

$$F_0 = \begin{bmatrix} f^0_{(1,1)} & 0 & 0 & 0 \\ T_2 \oplus D_0 - \gamma I_N \otimes I_{M_1} & diag(a_K, a_{K-1}, ..., a_1) & e_{M_1} \otimes q T^0_2 & f^0_{(2,5)} \\ 0 & f^0_{(3,3)} & e_{M_2} \otimes T^0_3 & \gamma I_{M_2 r} \\ 0 & 0 & D_0 - \gamma I_n & 0 \\ diag(c_K, c_{K-1}, ..., c_1) & 0 & 0 & (R \oplus D_0) \otimes I_{M_2} \end{bmatrix}$$

where

$$f^0_{(1,1)} = \begin{bmatrix} 0 \\ diag[I_K \otimes \theta \quad 0] \end{bmatrix}$$

$$f^0_{(1,1)} = \begin{bmatrix} D_0 & 0 \\ 0 & I_K \otimes (D_0 - \theta I_n) \end{bmatrix}$$

$$f^0_{(2,5)} = diag \begin{bmatrix} 0 & \gamma I_{rM_1} \end{bmatrix}$$

$$f^0_{(3,3)} = \begin{bmatrix} T_3 \oplus (D_0 - \gamma I_m) \otimes I_{K+1} & 0 \\ 0 & T_3 \oplus (D_0 - (\gamma + \theta)) \otimes I_{K+1} \end{bmatrix} + \begin{bmatrix} 0 \\ diag(b_K, b_{K-1}, ..., b_1) \end{bmatrix}$$

$$H_1 = \begin{bmatrix} h^1_{(1,1)} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$h^1_{(1,1)} = \begin{bmatrix} 0 & \delta I_n & 0 & 0 & 0 \\ 0 & pT^0_1\alpha_1 & \delta I_n & 0 & 0 \\ 0 & 0 & pT^0_1\alpha_1 & \delta I_n & 0 \\ 0 & 0 & 0 & pT^0_1\alpha_1 & 0 \\ 0 & 0 & 0 & 0 & pT^0_1\alpha_1 \end{bmatrix}$$

$$F_{10} = \begin{bmatrix} diag(d_K, d_{K-1}, ..., d_1) & pT^0_1\alpha_1 \otimes I_{M_2} & qT^0_1\alpha_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F = \begin{bmatrix} f_{(1,1)} & 0 & 0 & 0 & \gamma I_{rM_2} \\ f_{(2,1)} & f_{(2,2)} & diag(a_K, a_{K-1}, ..., a_1) & 0 & diag(e_K, e_{K-1}, ..., e_1) \\ T^0_3\alpha_3 \otimes I_{M_2} & 0 & f_{(3,3)} & 0 & \gamma I_{rM_2} \\ \eta & 0 & 0 & D_0 - \eta I_n & 0 \\ R^0\beta \otimes I_{M_2} & 0 & 0 & 0 & (R + D_0) \otimes I_{M_2} \end{bmatrix}$$

$$f_{(1,1)} = \begin{bmatrix} T_1 \oplus D_0 - (\gamma + \delta) I_n \otimes I_K & 0 & 0 & 0 \\ 0 & T_1 \oplus (D_0 - \gamma I_n) & 0 & 0 \\ 0 & 0 & diag(f_K, f_{K-1}, ..., f_1) & 0 \\ 0 & 0 & 0 & T_1 \oplus D_0 - (\gamma + \theta) I_n \otimes I_{K-1} \end{bmatrix} + diag(b_K, b_{K-1}, ..., b_1)$$

$$f_{(2,1)} = diag \begin{bmatrix} qT^0_2\alpha_2 \otimes I_n & 0 \end{bmatrix}$$

$$f_{(2,2)} = \begin{bmatrix} T_2 \oplus (D_0 - \gamma) \otimes I_K & 0 \\ 0 & T_2 \oplus D_0 - (\gamma + \theta) I_n \otimes I_{M_1} \end{bmatrix} + \begin{bmatrix} 0 \\ diag(g_{K-1}, ..., g_1) \end{bmatrix} + diag \begin{bmatrix} 0 \\ qT^0_2\alpha_2 \otimes I_{n(K-1)} \end{bmatrix}$$

$$f_{(3,3)} = \begin{bmatrix} T_3 \oplus (D_0 - \gamma) \otimes I_{K+1} & 0 \\ 0 & diag(h_K, h_{K_1}, ..., h_1) \end{bmatrix} + \begin{bmatrix} 0 \\ diag(b_{K-1}, ..., b_1) \end{bmatrix}$$

$$F_2 = \begin{bmatrix} f^2_{(1,1)} & diag(d_K, d_{K-1}, ..., d_1) & pT^0_1\alpha_1 \otimes I_{M_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$f^2_{(1,1)} = \begin{bmatrix} pT_1^0\alpha_1 \otimes I_n & \delta I_n & 0 & 0 & 0 \\ 0 & pT_1^0\alpha_1 \otimes I_n & \delta I_n & 0 & 0 \\ 0 & 0 & pT_1^0\alpha_1 \otimes I_n & \delta I_n & 0 \\ 0 & 0 & 0 & pT_1^0\alpha_1 \otimes I_n & 0 \\ 0 & 0 & 0 & 0 & pT_1^0\alpha_1 \otimes I_n \end{bmatrix}$$

$$a_i = pT_2^0\alpha_2 \otimes I_i$$

$$b_i = \theta I_{in}$$

$$c_i = \begin{bmatrix} 0 \\ R^0\beta \otimes I_i \end{bmatrix}$$

$$d_i = \begin{bmatrix} 0 & 0 \\ pT_1^0\alpha_1 \otimes I_i & \delta I_{n(i-1)} \end{bmatrix}$$

$$e_i = \begin{bmatrix} 0 & \gamma I_{ir} \end{bmatrix}$$

$$f_i = \begin{bmatrix} T_1 \oplus D_0 - (\gamma + I + \theta)I_n \otimes I_{i-1} & 0 \\ 0 & T_1 \oplus D_0 - (\gamma + \theta)I_n \end{bmatrix}$$

$$g_i = \begin{bmatrix} 0 & \theta I_{in} \end{bmatrix}$$

$$h_i = T_3 \oplus D_0 - (\gamma + \theta)I_n \otimes I_K,\ K \leq i \leq 1.$$

Let $z(0) = (z_0(0), z_1(0), z_2(0), z_3(0), ...)$ be defined as a conditional probability distribution of the system at the arrival time of the tagged customer is given by

$$z_0(0) = x_0 \left[\frac{D_1 e_(K+1)}{\lambda} \right]$$
$$z_i(0) = x_i \left[I_{[M_2(2n+r)+(nM_1+1)]} \frac{D_1 e_(K+1)}{\lambda} \right], \text{ for } i \geq 1$$

where λ is the fundamental arrival rate of the Markov Arrival Process. Now, on defining $z(t) = (z_*(t), z_0(t), z_1(t), z_2(t), ...)$, where

$$z_0(t) : \text{a } 1 \times 1 \text{ vector}$$
$$z_i(t) : \text{a row vector of order } 1 \times (M_2(2n + r) + (nM_1 + 1))$$

The chance that the continuous time Markov chain with the generator matrix $\bar{Q}$ is in the corresponding state of level i at instant t is given by their entries. Since $z * (t)$ denotes the likelihood that the tagged customer is in the absorbing state at time t, we have $W(t) = z_*(t)$, where $t \geq 0$.

The differential equation $z'(t) = z(t)\bar{Q}$ for $t \geq 0$ becomes

$$z'_*(t) = z_0(t)H_0 + z_1(t)H_1;$$
$$z'_0(t) = z_0(t)F_0 + z_1(t)F_{10};$$
$$z'_i(t) = z_i(t)F + z_{i+1}(t)F_2; i \geq 1$$

where $'$ denotes the derivative with respect to t. Let us calculate the Laplace-Stieltjes Transform for $W(t)$ using the Neuts et al. (1990) technique, where the initial probability row vector $w(s)$ denotes the first passage time to level 1 as follows

$$w(s) = \sum_{i=1}^{\infty} z_i(0)[(sI - F)^{-1}F_2]^{i-1} \tag{1}$$

Let $\varphi(i, s)$ be the LST of the time it takes to get absorbed into the state $(*)$, with the constraint that the process starts at level $i = 0, 1$. We have

$$\varphi(0,s) = [sI - F_0]^{-1} H_0 \tag{2}$$

$$\varphi(1,s) = [sI - F]^{-1} F_{10} \varphi(0,s) + [sI - F]^{-1} H_1. \tag{3}$$

Thus, it can easily seen that the Laplace-Stieltjes Transform for the distribution of sojourn time is as

$$\bar{W}(s) = z_0(0)\varphi(0,s) + w(s)\varphi(1,s). \tag{4}$$

The Expected waiting time is

$$E(W) = -\bar(W')(0) = -z_0(0)\varphi(0,0) - w'(0)e_{(M_2(2n+r)+(nM_1+1))} - w(0)\varphi'(1,0) \tag{5}$$

The first term in the preceding equation denotes the average time to enter the absorption state (*) assuming the system is in the level state i=0. On differentiating both the equation (2) and (3) , and setting s=0 , we have,

$$\varphi'(0,0) = (-1)[-F_0]^{-2} H_0 \tag{6}$$

$$\varphi'(1,0) = (-1)[-F_0]^{-2} F_{10} \varphi(0,0) + [-F_1]^{-1} F_{10} \varphi'(0,0) - [-F_1]^{-2} H_1 \tag{7}$$

By using equation (6) and (7) along with the primary conditions $z(t) = (z_0(0), z_1(0), z_2(0), ...)$, it can be easily evaluated the initial terms of (5). From (1) we have

$$w(s) = \sum_{i=1}^{\infty} z_i(0) U^{i-1} \tag{8}$$

where the stochastic matrix $U = [-F]^{-1} F_2$. We have

$$w(0)e_{(M_2(2n+r)+(nM_1+1))} = 1 - z_0(0). \tag{9}$$

Along with the primary conditions $z(t) = (z_0(0), z_1(0), z_2(0), ...)$, using (7) and (8), the last term of equation (5) can be evaluated. Differentiating (1) and substituting $s = 0$, we get,

$$w'(0) = (-1) \sum_{i=1}^{\infty} z_{i+1}(0) \sum_{j=0}^{i-1} U^j [-F]^{-1} U^{i-j} \tag{10}$$

by the condition U is stochastic , we have

$$(-1)w'(0)e_{(M_2(2n+r)+(nM_1+1))} = \sum_{i=1}^{\infty} z_{i+1}(0) \sum_{j=0}^{i-1} U^j [-F]^{-1} e_{(M_2(2n+r)+(nM_1+1))} \tag{11}$$

Defining an irreducible matrix U_2 satisfying two conditions such that $1 - U + U_2$ is non singular and the generalized inverse is of the form $(I - K_1)$. Then the matrix $U_2 = u_0 e_{(M_2(2n+r)+(nM_1+1))}$ where u_0 represents the stationary probability vector of U such that $u_0 U = u_0$ and $u_0 e_{(M_2(2n+r)+(nM_1+1))} = 1$. Moreover U_2 satisfies the property $UU_2 = U_2 U = U_2$. Then we have,

$$\sum_{j=0}^{i-1} U^j (I - U + U_2) = 1 - U^i + iU_2, for \;\; i \geq 1 \tag{12}$$

substituting (12) in (11) and simplifying we get the following

$$\begin{aligned}(-1)w'(0)e_{(M_2(2n+r)+(nM_1+1))} = &\{x_1(I-R)^{-1}\left\{I_{(M_2(2n+r)+(nM_1+1))} \otimes \frac{D_1 e_n}{\lambda}\right\} \\ &- w(0) + x_1 R(I-R)^{-2}\left\{I_{(M_2(2n+r)+(nM_1+1))} \otimes \frac{D_1 e_n}{\lambda}\right\}\} \\ &\times [I - U + U_2]^{-1}[-F]^{-1} e_{(M_2(2n+r)+(nM_1+1))}.\end{aligned} \tag{13}$$

As a result, we have obtained all of the terms in (5), which aids in determining the expected waiting time.

7. Performance Measures

To investigate the behaviour of our model under a steady state condition, a few performance measure of the system are computed.

- Probability that the server is idle
  $P_i = \sum_{k=0}^{K} x_{00k}$
- Probability that the server is busy with the main customer
  $P_{BM} = \sum_{i=1}^{\infty} \sum_{k=0}^{K} x_{i1k}$
- Probability that the server is busy with the customer in buffer 2
  $P_{BB2} = \sum_{i=0}^{\infty} \sum_{k=0}^{K} x_{i2k}$
- Probability that the server is busy with optional service
  $P_{BO} = \sum_{i=0}^{\infty} \sum_{k=0}^{K} x_{i3k}$
- Probability that the server is on vacation
  $P_V = \sum_{i=0}^{\infty} x_{i40}$
- Probability that the server is under repair
  $P_R = \sum_{i=0}^{\infty} \sum_{k=0}^{K} x_{i5k}$
- Expected system size
  $E_{system} = \sum_{p=1}^{\infty} \sum_{i=1}^{j} \sum_{k=0}^{K} p x_{pik} = x_1(1-R)^{-2}e$

8. Numerical Results

The qualitative behaviour of this model will be understood in this section with the help of a few illustrations, both numerically and graphically, by changing various model parameters such as the arrival process and service time distribution. For both the arrival process and the service time distribution, three sets of values from the literature are used as input.

Erlang of order 2 (ERL-A)

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}; D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

Exponential (Exp-A)

$$D_0 = \begin{bmatrix} -1 \end{bmatrix}; D_1 = \begin{bmatrix} 1 \end{bmatrix}$$

Hyperexponential (HYP-EXP-A)

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix}; D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}$$

Considering three phase type distributions for the service process which was suggested by Chakravarthy() **Erlang of order 2 (ERL-S)**

$$\alpha_1 = \alpha_2 = \alpha_3 = \beta = (1,0); T_1 = T_2 = T_3 = R = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

Exponential (Exp-A)

$$\alpha_1 = (1); T_1 = \begin{bmatrix} -44 \end{bmatrix}$$

$$\alpha_2 = (1); T_2 = \begin{bmatrix} -46 \end{bmatrix}$$

$$\alpha_3 = (1); T_3 = \begin{bmatrix} -34 \end{bmatrix}$$

$$\beta = (1);\ R = \begin{bmatrix} -12 \end{bmatrix}$$

Hyperexponential (HYP-EXP-A)

$$\alpha_1 = (0.3, 0.7);\ T_1 = \begin{bmatrix} -9 & 3 \\ 2 & -8 \end{bmatrix}$$

$$\alpha_2 = (0.4, 0.6);\ T_2 = \begin{bmatrix} -12 & 6 \\ 5 & -10 \end{bmatrix}$$

$$\alpha_3 = (0.4, 0.6);\ T_3 = \begin{bmatrix} -6 & 4 \\ 3 & -4 \end{bmatrix}$$

$$\beta = (0.5, 0.5);\ R = \begin{bmatrix} -12 & 3 \\ 3 & -12 \end{bmatrix}$$

Illustration 1

With the aid of 2D graphs, figure 2 to 10 represents the vacation rate versus probability of the server is idle for all possible ordering of arrival and service time by fixing $\eta = 5$, $\gamma = 7$, $\theta = 5$, $\delta = 2$, $P = 0.3$, $q = 0.7$, $K = 4$. An increase in vacation rate implies that the server will pay more attention to serve the customer which has a direct proportion on probability of the server is idle.

Illustration 2

The effect of the customer entering buffer 2 from buffer 1 after completion of interruption with rate θ and breakdown rate verses the expected system size has been investigated by fixing $\eta = 5$, $\gamma = 7$, $\theta = 5$, $\delta = 2$, $P = 0.3$, $q = 0.7$, $K = 4$. In figure 38 to 46, an increase in both the breakdown rate and self interrupted customer moving onto buffer 2 at the rate θ along with the expected system size with distinct group of arrival and service time has been observed briefly.

While there is an increase in self interrupted customer moving onto buffer 2 at the rate θ implies that arrival of customer to buffer 2 increases rapidly and an increase in breakdown rate implies an increase in the waiting time of the customer both in the main queue as well as buffer 2. In both the scenario it is obvious that the expected system size increases due to the minimal availability of the server.

Illustration 3

Table 1 to 3 represents the vacation rate versus expected system size by fixing $\eta = 5$, $\gamma = 7$, $\theta = 5$, $\delta = 2$, $p = 0.3$, $q = 0.7$, $K = 4$. As long as the vacation rate increases it is evident that the expected system size reduces gradually. The effect of increasing the vacation rate leads to more availability of the server in the system which in turn reduces the expected system size in the table. It is obviously that the expected system size decreases rapidly for hyper exponential service compared to a Erlang service which is pretty gradual.

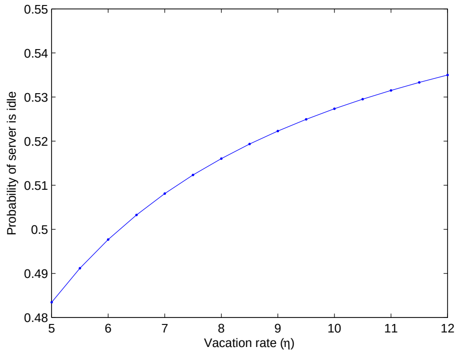


Figure 2: *Vacation rate* (η) *vs Probability of server is Idle - M/M/1*

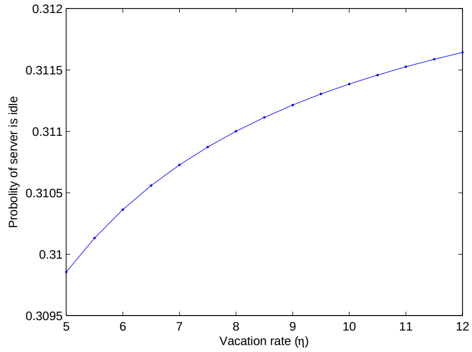


Figure 3: *Vacation rate* (η) *vs Probability of server is Idle - M/Ek/1*

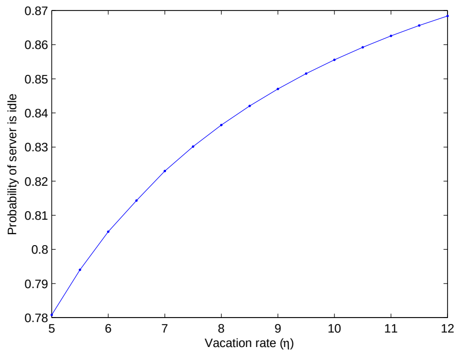


Figure 4: *Vacation rate* (η) *vs Probability of server is Idle - M/Hk/1*

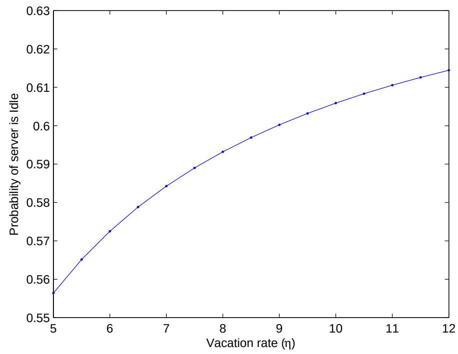


Figure 5: *Vacation rate* (η) *vs Probability of server is Idle - Ek/M/1*

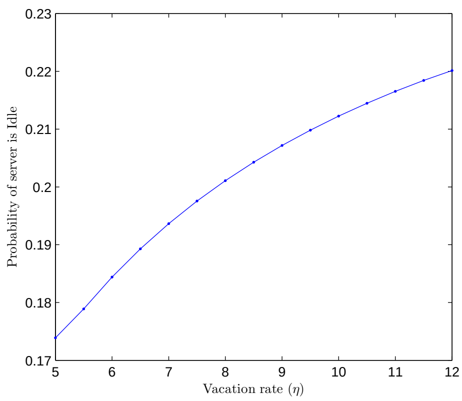


Figure 6: *Vacation rate* (η) *vs Probability of server is Idle - Ek/Ek/1*

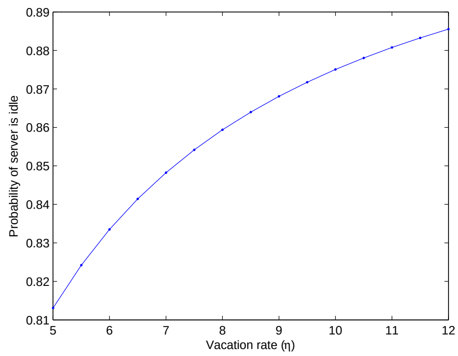


Figure 7: *Vacation rate* (η) *vs Probability of server is Idle - Ek/Hk/1*

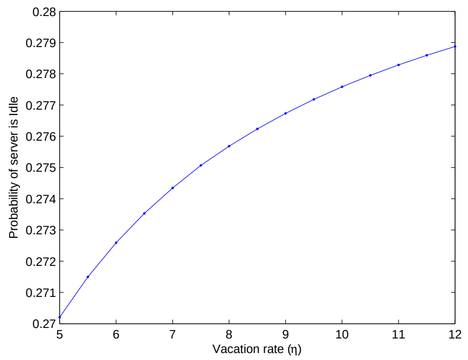


Figure 8: *Vacation rate* (η) *vs Probability of server is Idle - Hk/M/1*

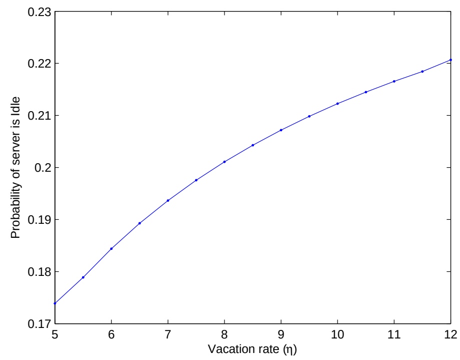


Figure 9: *Vacation rate* (η) *vs Probability of server is Idle - Hk/Ek/1*

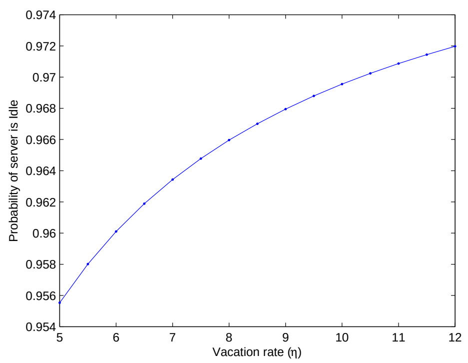


Figure 10: *Vacation rate* (η) *vs Probability of server is Idle - Hk/Hk/1*

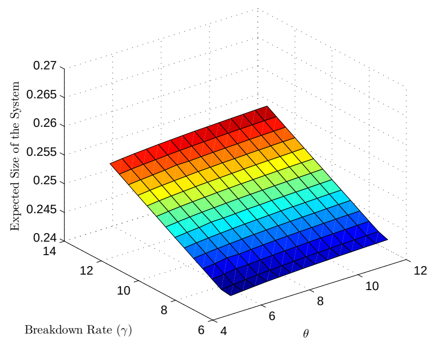


Figure 11: *The customer moving from buffer 1 to buffer 2 after interruption with the rate* θ *and Breakdown rate* (γ) *vs Expected Size of the System - M/M/1*

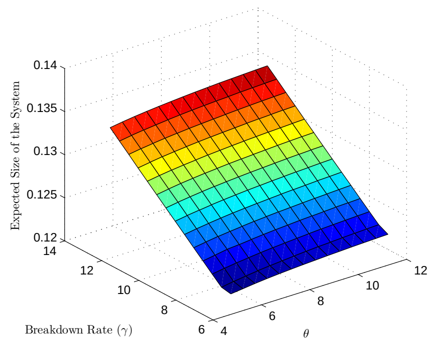


Figure 12: *The customer moving from buffer 1 to buffer 2 after interruption with the rate* θ *and Breakdown rate* (γ) *vs Expected Size of the System - M/Ek/1*

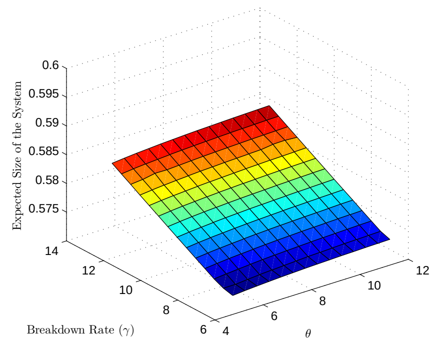


Figure 13: *The customer moving from buffer 1 to buffer 2 after interruption with the rate* θ *and Breakdown rate* (γ) *vs Expected Size of the System - M/Hk/1*

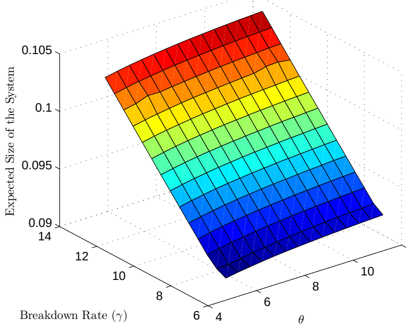


Figure 14: *The customer moving from buffer 1 to buffer 2 after interruption with the rate θ and Breakdown rate (γ) vs Expected Size of the System - Ek/M/1*

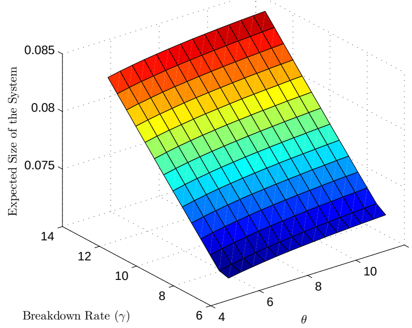


Figure 15: *The customer moving from buffer 1 to buffer 2 after interruption with the rate θ and Breakdown rate (γ) vs Expected Size of the System - Ek/Ek/1*

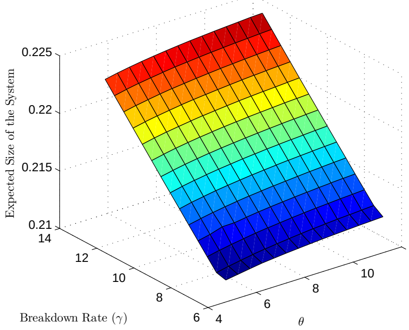


Figure 16: *The customer moving from buffer 1 to buffer 2 after interruption with the rate θ and Breakdown rate (γ) vs Expected Size of the System - Ek/Hk/1*

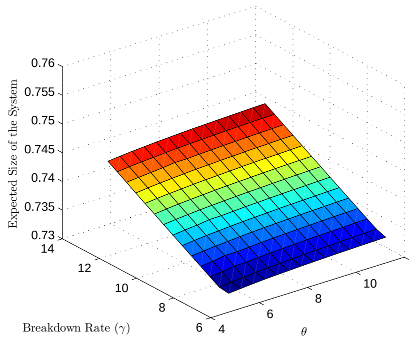


Figure 17: *The customer moving from buffer 1 to buffer 2 after interruption with the rate θ and Breakdown rate (γ) vs Expected Size of the System - Hk/M/1*

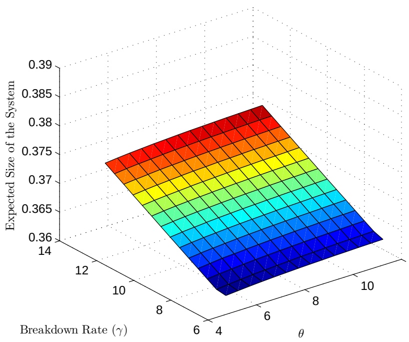


Figure 18: *The customer moving from buffer 1 to buffer 2 after interruption with the rate θ and Breakdown rate (γ) vs Expected Size of the System - Hk/Ek/1*

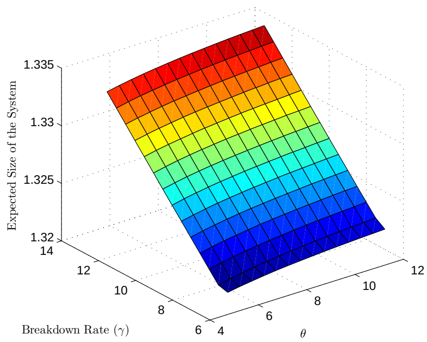


Figure 19: *The customer moving from buffer 1 to buffer 2 after interruption with the rate θ and Breakdown rate (γ) vs Expected Size of the System - Hk/Hk/1*

Table 1 Vacation rate (η) vs expected system size - Exponential-A

	service		
η	Exponential	Erlang	Hyperexponential
5	0.120659907	0.335551494	0.798654237
5.5	0.117283553	0.331603054	0.790156984
6	0.114733332	0.328581189	0.785684522
6.5	0.112762892	0.326217823	0.778725485
7	0.111210648	0.324335072	0.776131839
7.5	0.109967234	0.322811193	0.773832593
8	0.108956588	0.321560623	0.77193052
8.5	0.108124546	0.320521817	0.770586385
9	0.10743173	0.319649605	0.768578623
9.5	0.106848976	0.318910231	0.767878243
10	0.106354332	0.318278064	0.766823475
10.5	0.105931023	0.317733363	0.765928795
11	0.105566063	0.317260725	0.765248458
11.5	0.105249277	0.316847992	0.764572653
12	0.104972601	0.316485461	0.764568136

Table 2 Vacation rate (η) vs Expected system size - Erlang-A

	service		
η	Exponential	Erlang	Hyperexponential
5	0.057197966	0.070334672	0.09364885
5.5	0.054411688	0.064934106	0.091947737
6	0.052360653	0.060794972	0.090637074
6.5	0.05081441	0.057554181	0.109606071
7	0.049624632	0.054970138	0.108780591
7.5	0.048692757	0.052877138	0.108109511
8	0.04795146	0.051158507	0.107556655
8.5	0.04735363	0.049730205	0.107095843
9	0.046865592	0.048530471	0.106707752
9.5	0.046462815	0.047513107	0.106377871
10	0.046127126	0.046643004	0.106095132
10.5	0.045844857	0.045893101	0.105850975
11	0.045605584	0.045242268	0.105638696
11.5	0.045401257	0.044673818	0.105452986
12	0.045225591	0.04417443	0.105289594

Table 3 Vacation rate (η) vs Expected system size - Hyperexponential-A

	service		
η	Exponential	Erlang	Hyperexponential
5	0.675602876	0.695049695	0.983254166
5.5	0.67458552	0.695038233	0.974170849
6	0.67380935	0.695029517	0.972452185
6.5	0.673203901	0.695022735	0.968948456
7	0.67272264	0.695017355	0.966131839
7.5	0.672015304	0.695013015	0.963832593
8	0.672015304	0.695009464	0.96193052
8.5	0.671751078	0.695006521	0.960338541
9	0.671529505	0.695004055	0.958992295
9.5	0.671341885	0.695001969	0.957843366
10	0.671181629	0.695012002	0.956854752
10.5	0.671043669	0.684998654	0.955997773
11	0.670924057	0.684997326	0.955249916
11.5	0.670819681	0.684996166	0.954593302
12	0.670728061	0.684995149	0.954013588

References

[1] Ayyappan, G. Thamizhselvi, P. (2018). Transient Analysis of $M^{[x_1]}, M^{[x_2]}/G_1, G_2/1$ retail queueing system with priority services ,working vacation, vacation interruption,emergency vacation,negative arrival and delayed repair. *Int. J. Appl. Comput. Math, doi.org/10.1007/s40819-018-0509-7* .

[2] Chakravarthy, S. R. (2010). Markovian arrival processes. *In: Wiley Encyclopaedia of Operations Research and Management Science, doi.org/10.1002/9780470400531.eorms0499* .

[3] Do, T. V. (2010). M/M/1 retrial queue with working vacations. *Acta Informatica*, Vol. 47, pp.67–75.

[4] Doshi, B. T. (1986). Queueing systems with vacations-a survey. *Queueing Syst, Theory Appl* , Vol. 1, No. 1, pp.26–66.

[5] El-Taha, M. (2003). Allocation of service time in a two-server system. *Comput Oper Res*, Vol. 30, No. 5, pp.683–693.

[6] Gaver, D. P. (1959). Imbedded Marvov chain analysis of a waiting line process in continuous time. *Ann.Math.Stat.,*, Vol. 30, pp.698–720.

[7] Gelenbe, E. and Derochette, D. (1978). Performance of rollback recovery systems under intermittent failures. *Commun ACM*, Vol. 21, No. 6, pp.493–799.

[8] Gerasimov VI (1973). Optimum time-sharing algorithm for a servicing device in a queuing system with interruptions. *Autom Control Comput Sci*, Vol. 7, No. 6, pp.50–55.

[9] Jaiswal, NK. (1961). Preemptive resume priority queue. *Oper Res*, Vol. 9, pp.732–742.

[10] Jeganathan, K. (2015). Linear Retrial Inventory System with Second Optional Service under Mixed Priority Service. *TWMS Journal of Applied and Engineering Mathematics*, Vol. 5, No. 2, pp.249–268.

[11] Ke, J. C., Wu, C. H. and Zhang, Z. G. (2003). Recent developments in vacation queueing models: A short survey. *International Journal of Operations Research*, Vol. 7, pp.3–8.

[12] Klimenok, V. K. and Dudin, A .N. (2012). A BMAP/PH/N queue with negative customers and partial protection of service. *Commun Stat Simul Comput,doi:10.1080/03610918.2012.625802.*

[13]]A13 Krishnamoorthy, A., Gopakumar, B. and Viswanath, C. N.(2009). A queueing model with interruption resumption/ restart and reneging. *Bull Kerala Math Assoc*, Special issue pp.29–45.

[14] Krishnamoorthy, A., Pramod, P. K. Chakravarthy, SR. (2013). A note on characterizing service interruptions with phase type distribution.*Stoch. Anal. Appl.* , Vol. 31, No. 4, pp.671–683.

[15] Latouche,G. and Ramaswami,V. (1999). Introduction to Matrix Analytic Methods in Stochastic Modeling. *American Statistical Association,Virginia and the Society for Industrial and Applied Mathematics, Pennsylvania. doi.org/10.1137/1.9780898719734*.

[16] Levy, Y. and Yechiali, U. (1975). Utilization of idle time in an M/G/1 queueing system. *Manage. Sci.*, Vol. 22, pp.202–211.

[17]]A17 Levy, Y. and Yechiali, U. (1976). An M/M/s queue with server vacations. *Infor*, Vol. 14 ,No. 2, pp.153–163.

[18] Lucantoni, D., Meier-Hellstern, K. S. and Neuts, M. F. (1990). A single- server queue with server vacations and a class of nonrenewal arrival processes. *Adv. Appl. Probab.*, Vol. 22, pp.676–705.

[19] Neuts, M. F. and Lucantoni, D.M. (1979). A Markovian queue with N servers subject to breakdowns and repairs. *Manag. Sci*, Vol. 25, No. 9, pp.849–861.

[20] Neuts, M. F. (1981). Matrix-Geometric Solutions in Stochastic Models An Algorithmic Approach. *Johns Hopkins University Press, Baltimore and London*.

[21] Rakesh Kumar (2014). A single-server Markovian queuing system with discouraged arrivals and retention of reneged customers. *Yugoslav Journal of Operations Research, DOI: 10.2298/YJOR120911019K*.

[22] Rakesh Kumar and Bhavneet Singh Soodan (2019). Transient analysis of a single-server queuing system with correlated inputs and reneging. *Reliability: Theory and Applications*, Vol. 14, No.1, 52.

[23] Servi, L. D. and Finn, S. D. (2002). M/M/1 queues with working vacations (M/M/1/WV). *Perform. Eval. doi:10.1016/S0166-5316(02)00057-3*.

[24] Takagi, H. (1991). Queueing analysis: A foundation of performance evaluation (Vol. 1). *North- Holland*.

[25] Takine, T. and Sengupta, B. (1997). A single server queue with service interruptions. *Queueing Syst*, Vol. 26, No. 5, pp.285–300.

[26] Tian,N. and Zhang, Z. G. (2006). Vacation queueing models: Theory and applications. *New York: Springer-Verlag*.

[27] Tian, N., Li, J. and Zhang, Z. G. (2009). Matrix analytic method and working vacation queues A survey. *Int. Journal of Information and Management Sciences*, Vol. 20, pp.603–633.

[28] Varghese, J., Chakravarthy, SR. and Krishnamoorthy, A. (2010). On a customer induced interruption in a service system. *In: 8th International workshop on retrial queues*.

[29] Varghese, J., Chakravarthy, SR. and Krishnamoorthy, A. (2012). On a customer induced interruption in a service system. *J Stoch Anal Appl*, Vol. 30, No. 6, pp.949–962.

On a wide plurimodal class of distributions suitable for asymmetric data sets

C. Satheesh Kumar and G. V. Anila

•

Department of Statistics, University of Kerala, Trivandrum-695581, India.
drcsatheeshkumar@gmail.com
Department of Statistics, University of Kerala, Trivandrum-695581, India.
anilasaru@gmail.com

Abstract

Asymmetric normal distributions have received much attention in the literature during the last three decades. But, plurimodal asymmetric normal distributions are not much studied in the literature even though it has much relevance in practical situations. Here we propose a new class of plurimodal, asymmetric normal distribution and investigate its several statistical properties, including certain reliability aspects. A location-scale extension of the proposed model is developed and studied their properties. The maximum likelihood estimation method is employed for estimating the parameters of the proposed extended class of distributions and conducted generalized likelihood ratio test procedure for testing the parameters of the distribution. Three real-life data sets are considered for illustrating the usefulness of the model and a brief simulation study is carried out for examining the performance of maximum likelihood estimators of the proposed model.

Keywords: Asymmetric distributions, Maximum likelihood estimation, Model selection, Plurimodality, Simulation

1. Introduction

The normal distribution is the most important and most widely used distribution in statistics. It is an inevitable tool for the analysis and interpretation of data. But in many practical applications it has been observed that real life data sets are not symmetric. So normal distribution is not an acceptable model for modeling such data sets. In order to overcome this drawback, [2] considered an asymmetric form of normal distribution by introducing a skewness parameter into its probability density function (p.d.f) and named it as "the skew normal distribution". The skew normal distribution defined by [2] as follows:

Let $\phi(.)$ and $\Phi(.)$ be the p.d.f and cumulative distribution function (c.d.f) of a standard normal variate. Then a random variable X is said to follow the skew normal distribution with parameter $\lambda \in R = (-\infty, \infty)$ if its p.d.f $g(x; \lambda)$, for $x \in R$, is given by

$$g(x; \lambda) = 2\phi(x)\Phi(\lambda x). \tag{1}$$

A distribution with p.d.f. (1) we denoted as $SND(\lambda)$ through out the manuscript. The $SND(\lambda)$ has been further studied by several authors such as [3],[4], [5], [6], [7], [8] and [10].

A generalized form of skew normal distribution is developed by [1] through the following p.d.f.

$$g_1(x; \lambda_1, \lambda_2) = 2\phi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \tag{2}$$

in which $x \in R$, $\lambda_1 \in R$, $\lambda_2 \geq 0$. A distribution with pdf (2) we denoted as $SGND(\lambda_1, \lambda_2)$.

The $SGND(\lambda_1, \lambda_2)$ of [1] is log-concave and hence it is not suitable for plurimodal data. To overcome this drawback, [11] considered an extended version of $SGND(\lambda_1, \lambda_2)$ through the name "extended skew generalized normal distribution $(ESGND(\lambda_1, \lambda_2, \alpha))$"which has the following p.d.f.

$$g_2(x; \lambda_1, \lambda_2, \alpha) = \frac{2}{\alpha+2}\phi(x)\left[1 + \alpha\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right)\right], \quad (3)$$

where $x \in R$, $\lambda_1 \in R$, $\lambda_2 \geq 0$ and $\alpha \geq -1$. Through the present work our intention is to propose a wide class of plurimodal asymmetric normal distributions as a modified version of the $ESGND(\lambda_1, \lambda_2, \alpha)$ and named it as the "modified skew generalized normal distribution(MSGND)". In section 2 we present the definition and properties of MSGND. In section 4 we present the characteristic function and moments of MSGND. In section 5 certain reliability measures such as reliability function, mean residual life function etc are derived along with some conditions for unimodal and plurimodal situations are obtained. In section 6 a location scale extension of the MSGND is defined and obtained its properties such as characteristic function, reliability measures etc. In section 7 maximum likelihood estimation of the parameters of the distribution is discussed and in section 8 we constructed a generalized likelihood ratio test (GLRT) procedure. Real life data applications are given for illustrating the usefulness in section 9, a brief simulation study is attempted in section 10. While modelling certain real life data sets ESGND will not give better fit, the MSGND gives better fits. For example see the illustrations given in section 9, where the MSGND is found to be suitable for modelling data sets arising from athletic as well as agricultural data sets.

2. Modified skew generalized normal distribution

Here we define a new class of skew normal distribution namely the "modified skew generalized normal distribution (MSGND)"and derive its distributional important properties.

Definition 2.1. A random variable X is said to follow modified skew generalized normal distribution if its p.d.f is of the following, in which $x \in R$, $\lambda_1 \in R$, $\lambda_2 \geq 0$, $\beta \in R$ and $\alpha \geq -1$.

$$f(x; \lambda_1, \lambda_2, \alpha, \beta) = \frac{\phi(x)}{\alpha+2}\left[2 + \alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(x)\right)\right], \quad (4)$$

where $\lambda(x) = \frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}$, $\phi(.)$ and $\Phi(.)$ are the p.d.f and c.d.f of a standard normal variate. A distribution with p.d.f (4) we denoted as MSGND($\lambda_1, \lambda_2, \alpha, \beta$). Note that

1. When $\alpha = 0$ or $\lambda_1 = 0$, $MSGND(\lambda_1, \lambda_2, \alpha, \beta)$ reduces to the standard normal distribution $N(0, 1)$.
2. When $\beta = 0$, MSGND($\lambda_1, \lambda_2, \alpha, \beta$) reduces to the $ESGND(\lambda_1, \lambda_2, \alpha)$.
3. When $\alpha = -1$ and $\beta = 0$, MSGND($\lambda_1, \lambda_2, \alpha, \beta$) reduces to the SGND$(\lambda_1, \lambda_2)$.

For some particular choices of $\alpha, \lambda_1, \lambda_2$ and β, the p.d.f. $f(x; \lambda_1, \lambda_2, \alpha, \beta)$ given in (4) of $MSGND(\lambda_1, \lambda_2, \alpha, \beta)$ is plotted as given in Figure 1.

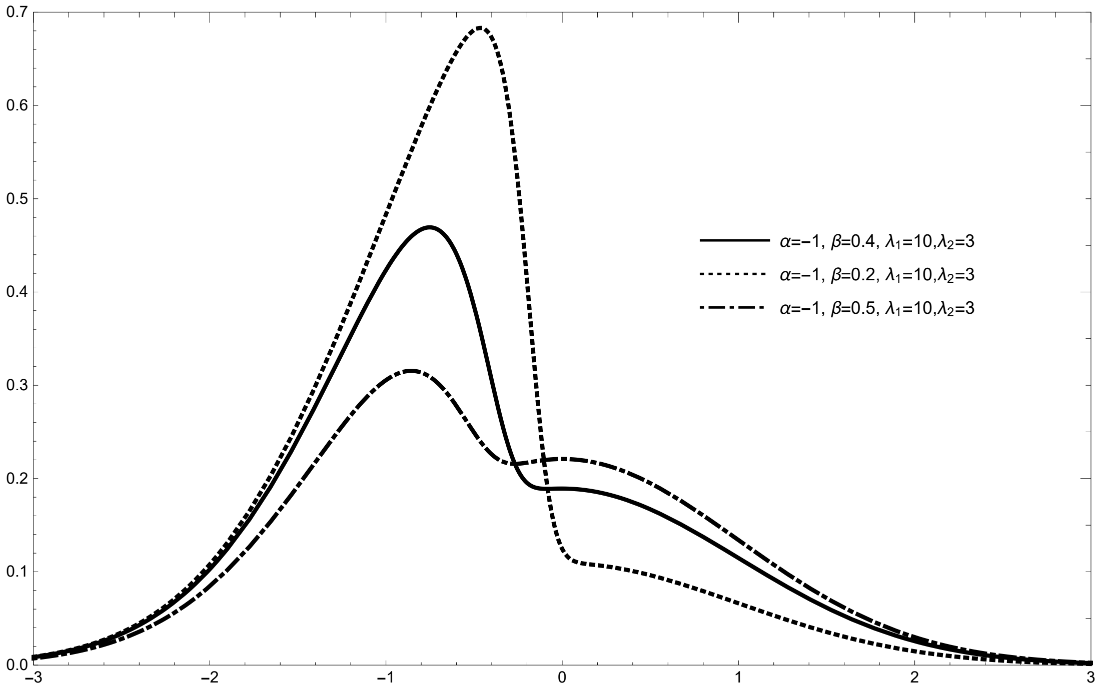


Figure 1: *Probability plots of MSGND($\lambda_1, \lambda_2, \alpha, \beta$) for fixed values of $\lambda_1, \lambda_2, \alpha$ and various values of β*

.

3. Results

Result 3.1. If X has MSGND($\lambda_1, \lambda_2, \alpha, \beta$), then $Y_1 = -X$ has MSGND($-\lambda_1, \lambda_2, \alpha, \beta$).

Proof. The p.d.f $f_1(y)$ of $Y_1 = -X$ is the following, for $y \in R, \lambda_1 \in R$, $\lambda_2 \geq 0$, $\beta \in R$ and $\alpha \geq -1$.

$$\begin{aligned} f_1(y) &= f(-y; -\lambda_1, \lambda_2, \alpha, \beta)|\frac{dx}{dy}| \\ &= \frac{\phi(-y)}{\alpha+2}\phi(-y)\left[2 + \alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(-y)\right)\right] \\ &= f(y; -\lambda_1, \lambda_2, \alpha, \beta) \end{aligned}$$

■

Result 3.2. If X has MSGND($\lambda_1, \lambda_2, \alpha, \beta$) then $Y_2 = |X|$ has the p.d.f (5), in which $\Delta(y) = \Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(y)\right) + \Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(-y)\right)$.

Proof. The p.d.f. $f_2(y)\, of\, Y_2 = |X|$ is the following, for $y > 0$.

$$f_2(y) = f(y; \lambda_1, \lambda_2, \alpha, \beta)|\frac{dx}{dy}| + f(-y; \lambda_1, \lambda_2, \alpha, \beta)|\frac{dx}{dy}|$$

in the light of Result 3.1 we have,

$$\begin{aligned}
f_2(y) &= f(y;\lambda_1,\lambda_2,\alpha,\beta) + f(y;-\lambda_1,\lambda_2,\alpha,\beta) \\
&= \frac{\phi(y)}{\alpha+2}\left[2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(y)\right)\right] + \\
&\quad \frac{\phi(y)}{\alpha+2}\left[2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(-y)\right)\right] \\
&= \frac{\phi(y)}{\alpha+2}\left[4+\alpha[\Phi(\beta)]^{-1}\left\{\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(y)\right)\right.\right. \\
&\quad \left.\left.+\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(-y)\right)\right\}\right] \\
&= \frac{\phi(y)}{\alpha+2}\left[4+\alpha[\Phi(\beta)]^{-1}\Delta(y)\right]. \qquad (5)
\end{aligned}$$

■

Result 3.3. If X has MSGND$(\lambda_1,\lambda_2,\alpha,\beta)$ then $Y_3 = X^2$ has pdf (6), in which $\Delta(y)$ is as defined in Result 3.2.

Proof. For $y > 0$, the p.d.f of $f_3(y)$ of $Y_3 = X^2$ is

$$\begin{aligned}
f_3(y) &= f(\sqrt{y};\lambda_1,\lambda_2,\alpha,\beta)|\frac{dx}{dy}| + f(-\sqrt{y};\lambda_1,\lambda_2,\alpha,\beta)|\frac{dx}{dy}| \\
&= \frac{\phi(\sqrt{y})}{\alpha+2}\left[2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda^2}+\lambda\left(\sqrt{(y)}\right)\right)\right]\frac{1}{2\sqrt{y}} + \\
&\quad \frac{\phi(-\sqrt{y}}{\alpha+2}\left[2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda^2}+\lambda(-\sqrt{(y)})\right)\right]\frac{1}{2\sqrt{y}} \\
&= \frac{\phi(\sqrt{y})}{(\alpha+2)2\sqrt{y}}\left[4+\alpha[\Phi(\beta)]^{-1}\Delta(\sqrt{y})\right]. \qquad (6)
\end{aligned}$$

■

Result 3.4. The c.d.f of MSGND$(\lambda_1,\lambda_2,\alpha,\beta)$ with p.d.f (4) is the following, for $x \in R$.

$$F(x) = \frac{\Phi(x)}{\alpha+2}\left[2+\alpha\frac{[\Phi(\beta)]^{-1}}{2}\right] - \frac{\alpha[\Phi(\beta)]^{-1}}{\alpha+2}\xi_\beta(x;\lambda(v)), \qquad (7)$$

where

$$\xi_\beta(x,\lambda(v)) = \int_x^\infty \int_0^{\beta\sqrt{1+\lambda^2}+\lambda(v)} \phi(v)\phi(u)\,dvdu, \qquad (8)$$

with $\lambda(v) = \frac{\lambda_1 v}{\sqrt{1+\lambda_2 v^2}}$. For particular values of $\lambda_1, \lambda_2, \beta$ and x, we can evaluate (8) by using the mathematical softwares such as MATHCAD, MATHEMATICA, etc.

Proof.

$$\begin{aligned}
F(x) &= \int_{-\infty}^x f(v;\lambda_1,\lambda_2,\alpha,\beta)dv \\
&= \frac{2}{\alpha+2}\Phi(x) + \frac{\alpha[\Phi(\beta)]^{-1}}{\alpha+2}\int_{-\infty}^x \phi(v)\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(v)\right)dv \\
&= \frac{2\Phi(x)}{\alpha+2} + \frac{\alpha[\Phi(\beta)]^{-1}}{\alpha+2}\left[\frac{1}{2}\Phi(x) - \xi_\beta(x,\lambda(v))\right] \\
&= \frac{\Phi(x)}{\alpha+2}\left[2+\frac{\alpha[\Phi(\beta)]^{-1}}{2}\right] - \frac{\alpha[\Phi(\beta)]^{-1}}{\alpha+2}\xi_\beta(x,\lambda(v)).
\end{aligned}$$

■

4. Characteristic function and Moments

In this section we obtain the characteristic function and moments of MSGND.

Result 4.1. The characteristic function $\psi_X(t)$ of MSGND$(\lambda_1, \lambda_2, \alpha, \beta)$ with p.d.f (4) is the following, for any $t \in R$ and $i = \sqrt{-1}$.

$$\psi_X(t) = \frac{e^{\frac{-t^2}{2}}}{\alpha+2}\left\{2 + \alpha[\Phi(\beta)]^{-1} E\left[\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(u+it)\right)\right]\right\} \tag{9}$$

where $\lambda(u+it) = \frac{\lambda_1(u+it)}{\sqrt{1+\lambda_2(u+it)^2}}$.

Proof. Let X follows MSGND$(\lambda_1, \lambda_2, \alpha, \beta)$ with p.d.f (4). Then by the definition of characteristic function, we have the following for any $t \in R$ and $i = \sqrt{-1}$.

$$\begin{aligned}
\psi_X(t) &= E(e^{itX}) \\
&= \frac{2}{\alpha+2}\int_{-\infty}^{\infty} e^{itx}\phi(x)dx + \frac{\alpha[\Phi(\beta)]^{-1}}{\alpha+2}\int_{-\infty}^{\infty} e^{itx}\phi(x)\Phi\left(\beta\sqrt{1+\lambda^2} + \lambda(x)\right)dx \\
&= \frac{e^{\frac{-t^2}{2}}}{\alpha+2}\left\{2 + \alpha[\Phi(\beta)]^{-1}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-(x-it)^2}{2}}\Phi(\beta\sqrt{1+\lambda^2} + \lambda(x))dx\right\} \qquad (10)
\end{aligned}$$

On substituting $x - it = u$ in (10), we obtain

$$\psi_X(t) = \frac{e^{\frac{-t^2}{2}}}{\alpha+2}\left\{2 + \alpha[\Phi(\beta)]^{-1} E\left[\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(u+it)\right)\right]\right\}, \tag{11}$$

which implies (9). ■

The expression for even moments and odd moments of $MSGND(\lambda_1, \lambda_2, \alpha, \beta)$ are obtained through the following results.

Result 4.2. If X follows $MSGND(\lambda_1, \lambda_2, \alpha, \beta)$, then for k=1,2,...,

$$E(X^{2k}) = \frac{2^{k+\frac{1}{2}}}{(\alpha+2)\sqrt{2\pi}}\Gamma(k+\frac{1}{2}) + \frac{\alpha[\Phi(\beta)]^{-1}}{2(\alpha+2)}A_k(\beta, \lambda_1, \lambda_2), \tag{12}$$

in which

$$A_k(\lambda_1, \lambda_2, \beta) = \int_0^{\infty} u^{k-\frac{1}{2}}\phi(\sqrt{u})\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(\sqrt{u})\right)du,$$

where $\lambda(\sqrt{u}) = \frac{\lambda_1\sqrt{u}}{\sqrt{1+\lambda_2 u}}$, $\lambda_1 \in R$, $\lambda_2 \geq 0$, $\beta \in R$ which can be easily evaluated by using the softwares MATHCAD and MATHEMATICA.

Proof. By the definition of raw moments, for any $k \geq 0$, integer,

$$E(X^{2k}) = \int_{-\infty}^{\infty} x^{2k} f(x; \lambda_1, \lambda_2, \alpha, \beta)dx. \tag{13}$$

On substituting $x^2 = u$ in (13) we obtain the following in the light of (4) we have,

$$\begin{aligned}
E(X^{2k}) &= \frac{1}{\alpha+2}\int_0^{\infty} u^k\phi(\sqrt{u})\frac{1}{\sqrt{u}}du + \frac{\alpha[\Phi(\beta)]^{-1}}{2(\alpha+2)} \\
&\quad \int_0^{\infty} u^k\phi(\sqrt{u})\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(\sqrt{u})\right)\frac{1}{\sqrt{u}}du \\
&= \frac{1}{(\alpha+2)}\left[\int_0^{\infty} u^{k-\frac{1}{2}}\phi(\sqrt{u})du + \frac{\alpha[\Phi(\beta)]^{-1}}{2}\right. \\
&\quad \left. u^{k-\frac{1}{2}}\phi(\sqrt{u})\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(\sqrt{u})\right)\right]du,
\end{aligned}$$

which leads to (12). ■

Result 4.3. If X follows $MSGND(\lambda_1, \lambda_2, \alpha, \beta)$, then for k=0,1,2,...,

$$E(X^{2k+1}) = \frac{2^{k+1}}{(\alpha+2)\sqrt{2\pi}}\Gamma(k+1) + \frac{\alpha[\Phi(\beta)]^{-1}}{2(\alpha+2)}A_{k+\frac{1}{2}}(\lambda_1, \lambda_2, \beta), \tag{14}$$

in which

$$A_{k+\frac{1}{2}}(\lambda_1, \lambda_2, \beta) = \int_0^\infty u^k \phi(\sqrt{u})\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda(?\sqrt{u})\right) du,$$

for $\lambda_1 \in R$, $\lambda_2 \geq 0$, $\beta \in R$ which can be easily evaluated using the softwares MATHCAD and MATHEMATICA.

Proof. By definition of raw moments,

$$E(X^{2k+1}) = \int_{-\infty}^\infty x^{2k+1} f(x; \lambda_1, \lambda_2, \alpha, \beta) dx. \tag{15}$$

On substituting $x^2 = u$ in (15) in the light of (4), we get

$$\begin{aligned} E(X^{2k+1}) &= \frac{1}{\alpha+2}\int_0^\infty u^{k+\frac{1}{2}}\phi(\sqrt{u})\frac{1}{\sqrt{u}}du + \frac{\alpha[\Phi(\beta)]^{-1}}{2(\alpha+2)} \\ &\quad \int_0^\infty u^{k+\frac{1}{2}}\phi(\sqrt{u})\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(\sqrt{u})\right)\frac{1}{\sqrt{u}}du \\ &= \frac{1}{(\alpha+2)}\left[\int_0^\infty u^k\phi(\sqrt{u})du + \frac{\alpha[\Phi(\beta)]^{-1}}{2}\right. \\ &\quad \left.\int_0^\infty u^k\phi(\sqrt{u})\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(\sqrt{u})\right)\right] du, \end{aligned}$$

which leads to (14). ■

5. Reliability measures and mode

Here we obtain some properties of MSGND($\lambda_1, \lambda_2, \alpha, \beta$) with p.d.f. (4) useful in reliability studies. Let X follows MSGND($\lambda_1, \lambda_2, \alpha, \beta$) with p.d.f (4). Now, from the definition of reliability function $R(t)$, failure rate $r(t)$ and mean residual life function $\mu(t)$ of X, we obtain the following results.

Result 5.1. The reliability function $R(t)$ of X is the following, in which $\xi_\beta(t, \lambda(x))$ is as defined in Result 3.4.

$$R(t) = \frac{[1-\Phi(t)]}{\alpha+2}\left\{2 + \frac{\alpha[\Phi(\beta)]^{-1}}{2}\right\} + \frac{\alpha[\Phi(\beta)]^{-1}}{\alpha+2}\xi_\beta(t, \lambda(x))$$

Result 5.2. The failure rate $r(t)$ of X is given by

$$r(t) = \frac{\phi(t)[2+\alpha[\Phi(\beta)]^{-1}\Phi(\beta\sqrt{1+\lambda^2}+\lambda(x))]}{(1-\Phi(t))[2+\frac{\alpha[\Phi(\beta)]^{-1}}{2}] + \alpha[\Phi(\beta)]^{-1}\xi_\beta(t, \lambda(x))}.$$

Result 5.3. The mean residual life function of MSGND($\lambda_1, \lambda_2, \alpha, \beta$) is

$$\begin{aligned} M(t) &= \frac{2\phi(t)}{(\alpha+2)R(t)} + \frac{\alpha[\Phi(\beta)]^{-1}}{(\alpha+2)R(t)}\left[\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(x)\right)\phi(t)\right. \\ &\quad \left.+\Lambda_\beta(t; \lambda_1, \lambda_2)\right] - t \end{aligned} \tag{16}$$

where

$$\Lambda_\beta(t; \lambda_1, \lambda_2) = \int_t^\infty \phi(x)\left[\frac{d}{dx}\left(\int_0^{\beta\sqrt{1+\lambda_1^2}+\lambda(x)}\phi(u)du\right)\right]dx.$$

Proof. By definition, the mean residual life function (MRLF) of X is given by

$$\begin{aligned} M(t) &= E(X-t|X>t) \\ &= E(X|X>t)-t, \end{aligned} \tag{17}$$

where

$$\begin{aligned} E(X|X>t) &= \frac{2}{R(t)(\alpha+2)}\int_t^{\infty} x\phi(x)dx \\ &+\frac{\alpha[\Phi(\beta)]^{-1}}{R(t)}\int_t^{\infty} x\phi(x)\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(x)\right)dx. \end{aligned} \tag{18}$$

Since $\phi(.)$ is the p.d.f of standard normal variate $\phi'(x) = -x\phi(x)$. Therefore (18) becomes,

$$\begin{aligned} E(X|X>t) &= \frac{2}{(\alpha+2)R(t)}\int_t^{\infty} -\phi'(x)dx \\ &+\frac{\alpha[\Phi(\beta)]^{-1}}{(\alpha+2)R(t)}\int_t^{\infty} -\phi'(x)\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda(x)\right)dx. \end{aligned} \tag{19}$$

On integrating (19), we obtain the following

$$\begin{aligned} E(X|X>t) &= \frac{2}{(\alpha+2)R(t)}\phi(t)+\frac{\alpha[\Phi(\beta)]^{-1}}{(\alpha+2)R(t)}\left(-\Phi(\lambda(x)+\beta\sqrt{1+\lambda^2})\phi(x)\right)_t^{\infty} \\ &-\frac{\alpha[\Phi(\beta)]^{-1}}{R(t)(\alpha+2)}\int_t^{\infty} -\phi(x)\left[\frac{d}{dx}\left(\int_{-\infty}^{\beta\sqrt{1+\lambda_1^2}+\lambda(x)}\phi(u)du\right)\right]dx. \end{aligned} \tag{20}$$

On solving (20) and substituting in (17), we get (16).

The functions $R(t), r(t)$ and $\mu(t)$ are equivalent in the sense that if one of them is given, the other two can be uniquely determined. ■

Next, through the following result we derive certain conditions under which the MSGND($\lambda_1, \lambda_2, \alpha, \beta$) is log-concave.

Result 5.4. The p.d.f of MSGND($\lambda_1, \lambda_2, \alpha, \beta$) is log-concave under the following two cases.
Case 1: For $x > 0$,

(i) when $\lambda_1 < 0$ provided for all $\alpha > 0$ and $\beta > 0$ and

(ii) when $\lambda_1 > 0$ provided $|\frac{3\lambda_1\lambda_2^2x^3}{(1+\lambda_2x^2)^{\frac{5}{2}}}| < |\frac{3\lambda_1\lambda_2x}{(1+\lambda_2x^2)^{\frac{3}{2}}}|$

Case 2: For $x < 0$, the p.d.f of MSGND($\lambda_1, \lambda_2, \alpha, \beta$) is log concave

(i) when $\lambda_1 > 0$ provided for all $\alpha > 0$ and $\beta > 0$ and

(i) when $\lambda_1 < 0$ provided $|\frac{3\lambda_1\lambda_2^2x^3}{(1+\lambda_2x^2)^{\frac{5}{2}}}| < |\frac{3\lambda_1\lambda_2x}{(1+\lambda_2x^2)^{\frac{3}{2}}}|$.

Proof. To prove $log[f(x;\lambda_1,\lambda_2,\alpha,\beta)]$ is a concave function of x, it is enough to show that its second derivative is negative for all x. Thus

$$\frac{d}{dx}log[f(x;\lambda_1,\lambda_2,\alpha,\beta)] = -x+\frac{\alpha[F(\beta)]^{-1}f(\eta)\eta'}{2+\alpha[F(\beta)]^{-1}F(\eta)}$$

and

$$\frac{d^2}{dx^2}log[f(x;\lambda_1,\lambda_2,\alpha,\beta)] = -1-\Delta_1-\Delta_2+\Delta_3$$

in which,

$$\Delta_1 = \frac{\alpha[F(\beta)]^{-1}\eta'^2 f(\eta)\eta}{2+\alpha[F(\beta)]^{-1}F(\eta)} \quad (21)$$

$$\Delta_2 = \frac{\alpha^2[F(\beta)]^{-2}(f(\eta))^2\eta'^2}{[2+\alpha[F(\beta)]^{-1}F(\eta)]^2} \quad (22)$$

and

$$\Delta_3 = \frac{\alpha[F(\beta)]^{-1}f(\eta)\eta''}{2+\alpha[F(\beta)]^{-1}F(\eta)}, \quad (23)$$

where

$$\begin{aligned}\eta &= \lambda(x)+\beta\sqrt{1+\lambda_1^2}\\ \eta' &= \frac{\lambda_1}{\sqrt{1+\lambda_2 x^2}} - \frac{\lambda_1\lambda_2 x^2}{(1+\lambda_2 x^2)^{\frac{3}{2}}}\\ \eta'' &= \frac{3\lambda_1\lambda_2^2 x^3}{(1+\lambda_2 x^2)^{\frac{5}{2}}} - \frac{3\lambda_1\lambda_2 x}{(1+\lambda_2 x^2)^{\frac{3}{2}}}.\end{aligned}$$

Note that $\Delta_1 > 0$, for $\alpha > 0$ and $\eta > 0$. Here $\eta > 0$ for all values of $\lambda_1 > 0$ and $\beta > 0$. Consequently $\Delta_2 > 0$ for all values of $\alpha > 0, \beta > 0$ and $\lambda_1 > 0$. Also, $\Delta_3 < 0$ for either when $\alpha < 0$ and $\eta'' > 0$ or when $\alpha > 0$ and $\eta'' < 0$. Hence (4) is log-concave in these situations. ■

As a consequence of the above result, we have the following result.

Result 5.5. MSGND($\lambda_1, \lambda_2, \alpha, \beta$) density is strongly unimodal under the following two cases. Case 1: For $x > 0$,

(i) if $\lambda_1 < 0$ provided for all $\alpha > 0$ and $\beta > 0$ and

(ii) if $\lambda_1 > 0$ provided $|\frac{3\lambda_1\lambda_2^2 x^3}{(1+\lambda_2 x^2)^{\frac{5}{2}}}| < |\frac{3\lambda_1\lambda_2 x}{(1+\lambda_2 x^2)^{\frac{3}{2}}}|$

Case 2: For $x < 0$,

(i) if $\lambda_1 > 0$ provided for all $\alpha > 0$ and $\beta > 0$ and

(i) if $\lambda_1 < 0$ provided $|\frac{3\lambda_1\lambda_2^2 x^3}{(1+\lambda_2 x^2)^{\frac{5}{2}}}| < |\frac{3\lambda_1\lambda_2 x}{(1+\lambda_2 x^2)^{\frac{3}{2}}}|$.

Result 5.6. MSGND($\lambda_1, \lambda_2, \alpha, \beta$) density is plurimodal under the following two cases. Case 1: For $x > 0$,

(i) if $\lambda_1 < 0$ provided for all $\alpha < 0$ and $\beta > 0$ and

(ii) if $\lambda_1 > 0$ provided $|\frac{3\lambda_1\lambda_2^2 x^3}{(1+\lambda_2 x^2)^{\frac{5}{2}}}| > |\frac{3\lambda_1\lambda_2 x}{(1+\lambda_2 x^2)^{\frac{3}{2}}}|$

Case 2: For $x < 0$,

(i) if $\lambda_1 > 0$ provided for all $\alpha < 0$ and $\beta > 0$ and

(i) if $\lambda_1 < 0$ provided $|\frac{3\lambda_1\lambda_2^2 x^3}{(1+\lambda_2 x^2)^{\frac{5}{2}}}| > |\frac{3\lambda_1\lambda_2 x}{(1+\lambda_2 x^2)^{\frac{3}{2}}}|$.

6. Extended form of MSGND

In this section we discuss an extended form of MSGND$(\lambda_1, \lambda_2, \alpha, \beta)$ by introducing the location parameter μ and scale parameter σ.

Definition 6.1. Let $X \sim MSGND(\lambda_1, \lambda_2, \alpha, \beta)$ with p.d.f given in (4). Then $Y = \mu + \sigma X$ is said to have an extended MSGND with the following p.d.f.

$$\begin{aligned} f^*(y, \mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta) \quad = \quad & \frac{1}{\sigma(\alpha+2)}\phi(\frac{y-\mu}{\sigma})\left[2 + \alpha[\Phi(\beta)]^{-1}\right. \\ & \left.\Phi(\beta\sqrt{1+\lambda_1^2} + \lambda^*(y))\right], \end{aligned} \tag{24}$$

in which $\lambda^*(y) = \frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}$, $y \in R$, $\mu \in R$, $\lambda_1 \in R$, $\beta \in R$, $\sigma > 0$, $\lambda_2 \geq 0$ and $\alpha \geq -1$. A distribution with p.d.f (24) is denoted as EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$. Clearly when

(i) When β=0, the EMSGND $(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$ reduces to ESGND $(\mu,\sigma;\lambda_1, \lambda_2, \alpha)$ of [11].

(ii) β=0 and $\lambda_1 = 0$, the EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$ reduces to the p.d.f of normal distribution.

(iii) When $\beta = 0$ and $\lambda_2 = 0$, the EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$ reduces to EGMNSN ($\mu, \sigma; \alpha, \lambda$) of [9].

Now, we obtain the following results of EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$, in a similar way as we defined in section 2 and 4.

Result 6.1. The cumulative distribution function (c.d.f) $F^*(y)$ of EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$ with p.d.f (24) is the following, for $y \in R$.

$$F^*(y) = \left[2 + \frac{\alpha\Phi[(\beta)]^{-1}}{2}\right]\frac{\Phi(\frac{y-\mu}{\sigma})}{\sigma(\alpha+2)} - \frac{\alpha[\Phi(\beta)]^{-1}}{\sigma(\alpha+2)}\xi_\beta\left(y, \lambda^*(y)\right)$$

where $\xi_\beta\left(y, \lambda^*(y)\right)$, is as defined in Result 3.4.

Result 6.2. The characteristic function of EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$ is given by

$$\psi_Y^*(t) = \frac{e^{it\mu - \frac{t^2\sigma^2}{2}}}{\alpha+2}\left\{2 + \alpha[\Phi(\beta)]^{-1}E\left[\Phi\left(\beta\sqrt{1+\lambda_1^2} + \lambda^*(z+\sigma^2 it)\right)\right]\right\},$$

where $\lambda^*(z + \sigma^2 it) = \frac{\lambda_1(z+\sigma^2 it)}{\sqrt{\sigma^2+\lambda_2(z+\sigma^2 it)^2}}$.

Result 6.3. The reliability function $R^*(t)$ of Y is the following, in which $\xi_\beta(t, \lambda^*(y))$ is as defined in Result 3.4, with $\lambda^*(y) = \frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}$

$$\begin{aligned} R^*(t) \quad = \quad & \frac{1}{\sigma(\alpha+2)}\left[1 - F(\frac{t-\mu}{\sigma})\right]\left\{2 + \frac{\alpha}{2}[F(\beta)]^{-1}\right\} + \frac{\alpha[F(\beta)]^{-1}}{\sigma(\alpha+2)} \\ & \xi_\beta\left(t, \lambda^*(y)\right). \end{aligned}$$

Result 6.4. The failure rate $r^*(t)$ of Y is given by

$$r^*(t) = \frac{f(\frac{t-\mu}{\sigma})\left[2 + \alpha[F(\beta)]^{-1}F\left(\beta\sqrt{1+\lambda_1^2} + \lambda^*(t)\right)\right]}{\frac{1}{\sigma(\alpha+2)}\left[1 - F(\frac{t-\mu}{\sigma})\right]\left\{2 + \frac{\alpha}{2}[F(\beta)]^{-1}\right\} + \frac{\alpha[F(\beta)]^{-1}}{\sigma(\alpha+2)}\xi_\beta\left(t, \lambda^*(t)\right)},$$

where $\lambda^*(t) = \frac{\lambda_1(t-\mu)}{\sqrt{\sigma^2+\lambda_2(t-\mu)^2}}$.

7. Maximum likelihood estimation

The log-likelihood function, $\ln L$ of the random sample of size n from a population following EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$ is the following.

$$\begin{aligned} ln\,L &= n\ln\left(\frac{1}{\sqrt{2\pi}}\right) - n\ln\sigma - n\ln(\alpha+2) - \frac{1}{2}\sum_{i=1}^{n}\frac{(y_i-\mu)^2}{\sigma^2} \\ &\quad + \sum_{i=1}^{n}\ln\left(2+\alpha\left[\Phi(\beta)\right]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)\right). \end{aligned} \tag{25}$$

On differentiating (25) with respect to parameters $\mu, \sigma, \lambda_1, \lambda_2, \alpha$ and β and then equating to zero, we obtain the following normal equations.

$$\begin{aligned} \sum_{i=1}^{n}\frac{(y_i-\mu)}{\sigma^2} &- \sum_{i=1}^{n}\frac{\alpha[\Phi(\beta)]^{-1}\phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)\left(\frac{\lambda_1}{\sqrt{\sigma^2+\lambda_2(y_i-\mu)^2}}\right)}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} \\ &+ \sum_{i=1}^{n}\frac{\alpha[\Phi(\beta)]^{-1}\phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)\left(\frac{\lambda_1\lambda_2(y_i-\mu)^2}{[\sigma^2+\lambda_2(y_i-\mu)^2]^{\frac{3}{2}}}\right)}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} = 0, \end{aligned} \tag{26}$$

$$\frac{n}{\sigma} = \sum_{i=1}^{n}\frac{(y_i-\mu)^2}{\sigma^3} + \sum_{i=1}^{n}\frac{\alpha\lambda_1\Phi[(\beta)]^{-1}\phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)\left(\frac{(y_i-\mu)\sigma}{[\sigma^2+\lambda_2(y_i-\mu)^2]^{\frac{3}{2}}}\right)}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)}, \tag{27}$$

$$\sum_{i=1}^{n}\frac{\alpha[\Phi(\beta)]^{-1}\phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)\left[\frac{\lambda^*(y)}{\lambda_1}+\frac{\beta\lambda_1}{\sqrt{1+\lambda_1^2}}\right]}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} = 0, \tag{28}$$

$$\sum_{i=1}^{n}\frac{\alpha[\Phi(\beta)]^{-1}\phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)\left[\frac{\lambda_1(y_i-\mu)^3}{[\sigma^2+\lambda_2(y_i-\mu)^2]^{\frac{3}{2}}}\right]}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} = 0, \tag{29}$$

$$\sum_{i=1}^{n}\frac{[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} = 0 \tag{30}$$

and

$$\begin{aligned} &\sum_{i=1}^{n}\frac{\alpha[\Phi(\beta)]^{-1}\phi(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y))\sqrt{1+\lambda_1^2}}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} - \\ &\sum_{i=1}^{n}\frac{\alpha[\Phi(\beta)]^{-2}\phi(\beta)\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)}{2+\alpha[\Phi(\beta)]^{-1}\Phi\left(\beta\sqrt{1+\lambda_1^2}+\lambda^*(y)\right)} = 0. \end{aligned} \tag{31}$$

On solving the equations (26) to (31), we get the maximum likelihood estimate of the parameters of EMSGND$(\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta)$.

8. Generalized likelihood ratio test

In this section we discuss a test procedure for testing the parameter β of $EMSGND$. For testing the null hypothesis $H_0 : \beta = 0$ against the alternative hypothesis $H_1 : \beta \neq 0$ by using the generalized likelihood ratio test, the test statistic is

$$-2ln\lambda(x) = 2[lnL(\hat{\Theta}; x) - lnL(\hat{\Theta}^*; x)],$$

where $\hat{\Theta}$ is the maximum likelihood estimator of $\Theta = (\mu, \sigma, \lambda_1, \lambda_2, \alpha, \beta)$ with no restriction, and $\hat{\Theta}^*$ is the maximum likelihood estimator of Θ when $\beta = 0$. The test statistic given is asymptotically distributed as χ^2 with 1 degrees of freedom. For further details see [12].

9. Applications

In this section we consider three real life data applications of the EMSGND. The first data is taken from [9]. The data gives the Otis IQ scores for 52 non-white males hired by a large insurance company in 1971. The observed data is given below:
Data set 1:
91, 102, 100, 117, 122, 115, 97, 109, 108, 104, 108, 118, 103, 123, 123, 103, 106, 102, 118, 100, 103, 107, 108, 107, 97, 95, 119, 102, 108, 103, 102, 112, 99, 116, 114, 102, 111, 104, 122, 103, 111, 101, 91, 99, 121, 97, 109, 106, 102, 104, 107, 95.
The second data is taken from [8]. This data is related to the milk production of 28 cows in which the variable under study is the daily milk production in kilogram and the variable recorded for three times milking cows. Data set 2:
34.6, 27.7, 29.2, 25.3, 27.6, 37.9, 32.6, 32, 30.7, 29.6, 38.3, 32.9, 30.8, 32.2, 32.9, 28.1, 33.9, 28.6, 28.1, 35.9, 34.8, 40.3, 30.9, 34.4, 19.8, 25.8, 37.3, 32.4.
The third data is taken from [8]. The data includes 100 females and 102 males with 13 variables such as height, weight, body mass index (BMI) etc. We choose for the variable under study is the BMI values for the second 50 females. The data is given below: Data set 3:
24.47, 23.99, 26.24, 20.04, 25.72, 25.64, 19.87, 23.35, 22.42, 20.42, 22.13, 25.17, 23.72, 21.28, 20.87, 19.00, 22.04, 20.12, 21.35, 28.57, 26.95, 28.13, 26.85, 25.27, 31.93, 16.75, 19.54, 20.42, 22.76, 20.12, 22.35, 19.16, 20.77, 19.37, 22.37, 17.54, 19.06, 20.30, 20.15, 25.36, 22.12, 21.25, 20.53, 17.06, 18.29, 18.37, 18.93, 17.79, 17.05, 20.31.
We obtained the maximum likelihood estimate (MLE) of the parameters by using the data sets with the help of the MATHCAD software. The numerical results obtained are presented in Table 1, which includes the estimated values of the parameters and the corresponding Kolmogorov-Smirnov Statistics (KSS) values of models: ESGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha$) and EMSGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta$). Also, its Akaike's Information Criterion(AIC), Bayesian Information Criterion(BIC) and corrected Akaike's Information Criterion(AICc) values are obtained.

Table 1: *Estimated values of the parameters for the model: ESGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha$) and EMSGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta$) with respective values of KSS, P-value, log-likelihood, AIC, BIC and AICc in case of Data Set 1, 2 and 3.*

Data set	**Estimates of the parameters**	**ESGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha$)**	**EMSGND ($\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta$)**
1	$\hat{\mu}$	102.18167	106.654
	$\hat{\sigma}$	4.94535	8
	$\hat{\alpha}$	1.89352	2
	$\hat{\beta}$	-	8
	$\hat{\lambda}_1$	6.23351	0.809
	$\hat{\lambda}_2$	0.38068	0.349
	KSS	0.5	0.129961
	P-Value	3.91952×10^{-12}	0.315666
	Log-likelihood	-193.257	-183.43
	AIC	396.515	378.859
	BIC	406.271	390.567
	AICc	397.819	380.726
2	$\hat{\mu}$	31.43211	31.5934
	$\hat{\sigma}$	2.0174	4.464
	$\hat{\alpha}$	0.96229	0.5
	$\hat{\beta}$	-	20.006
	$\hat{\lambda}_1$	0.64746	6.002
	$\hat{\lambda}_2$	15.91043	8
	KSS	0.451625	0.0783872
	P-Value	9.83689×10^{-6}	0.98998
	Log-likelihood	-100.348	-81.1221
	AIC	210.697	174.244
	BIC	217.358	182.237
	AICc	213.424	178.244
3	$\hat{\mu}$	20.1321	21.865
	$\hat{\sigma}$	1.07639	3.33
	$\hat{\alpha}$	1.17325	0.6
	$\hat{\beta}$	-	8
	$\hat{\lambda}_1$	7.85813	7
	$\hat{\lambda}_2$	10.13911	5
	KSS	0.5	0.121453
	P-Value	1.07824×10^{-11}	0.418815
	Log-likelihood	-324.56	-130.59
	AIC	659.119	273.18
	BIC	668.68	284.652
	AICc	660.483	275.134

It is clear from Table 1, that the EMSGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta$) is a more appropriate model to all the three data sets compared to the existing model ESGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha$). We have plotted the histogram of the respective data sets along with the corresponding fitted values of the $ESGND$ and $EMSGND$ in Figures 2, 3 and 4 respectively. It shows that $EMSGND$ yields a better fit than $ESGND$ in all the cases. Thus, the model discussed in this paper provides more flexibility in modeling in case of all the three datasets due to the presence of the extra parameter.

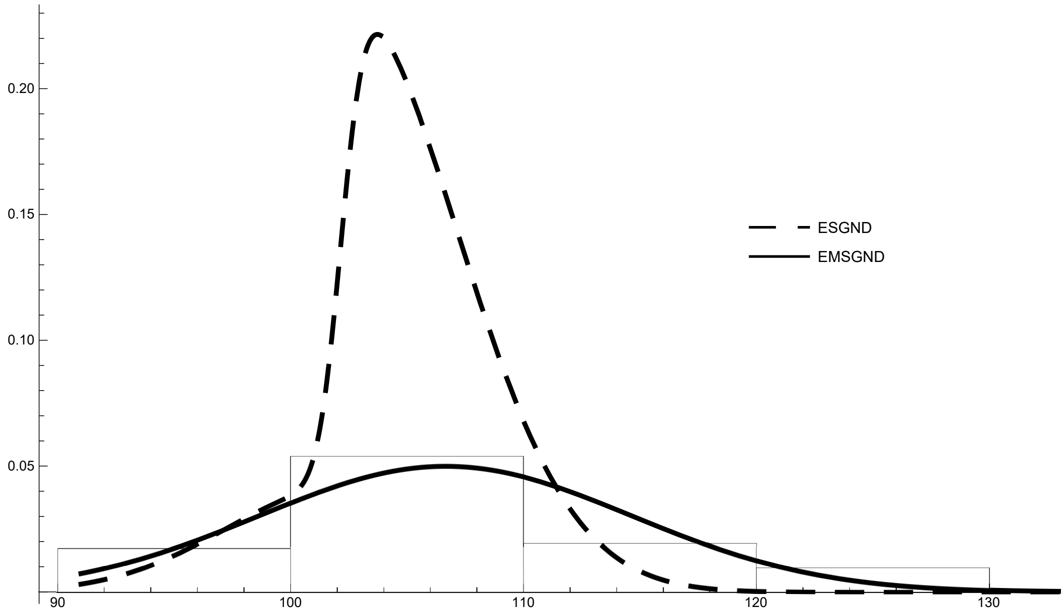


Figure 2: *Histogram of Data set 1 and fitted distributions*

.

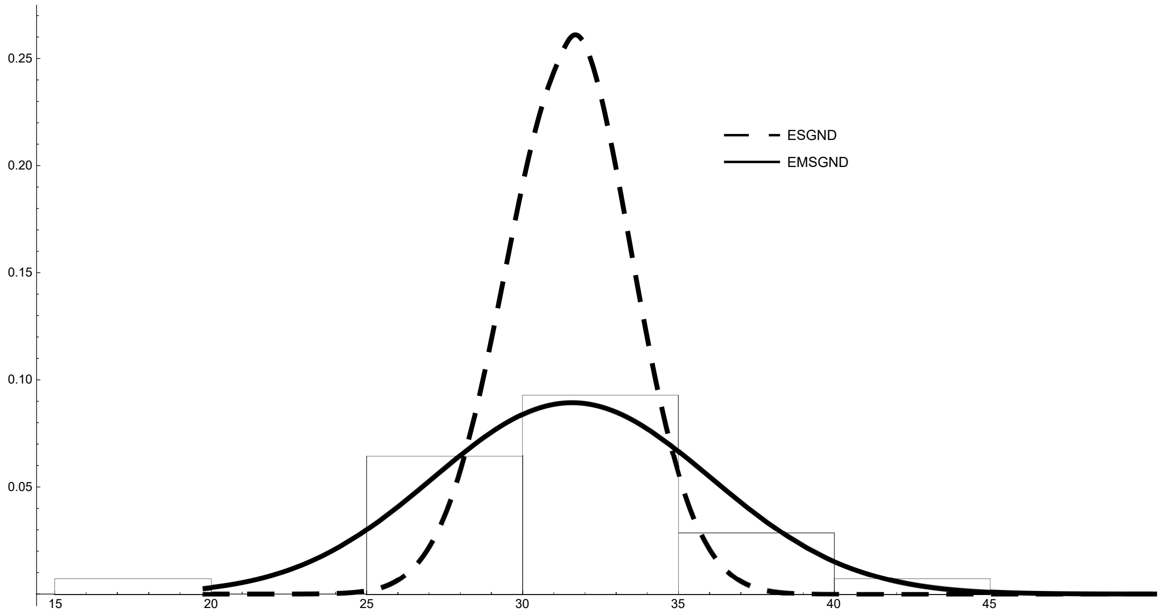


Figure 3: *Histogram of Data set 2 and fitted distributions*

.

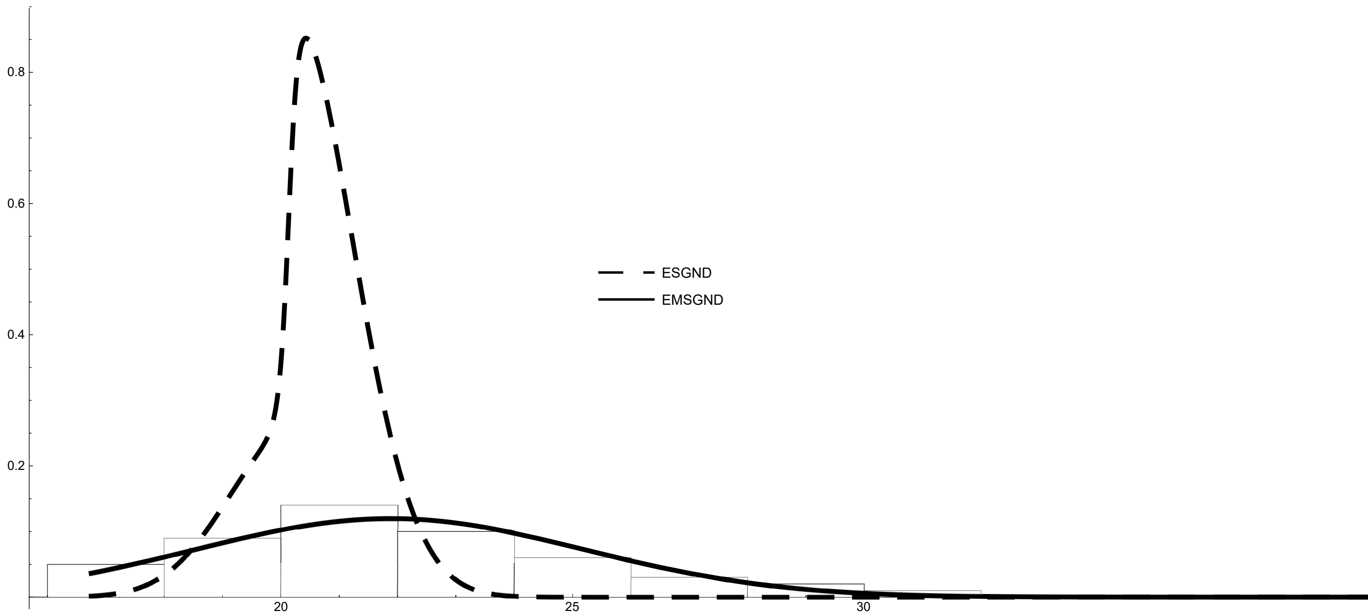


Figure 4: *Histogram of Data set 3 and fitted distributions*

.

Also, we conduct a generalized likelihood ratio test for illustrating the usefulness of the model, which is described as follows.
Let us consider the problem of testing a hypothesis $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$ in the case of Data set 1. The MLEs and values of the likelihood for $ESGND$ and $EMSGND$ are

$$\hat{\mu} = 102.18167,\ \hat{\sigma} = 4.94535,\ \hat{\lambda_1} = 6.23351,\ \hat{\lambda_2} = 0.38068,\ \hat{\alpha} = 1.89352,$$

$L(\hat{\Theta}^*; x) = 1.17315 \times 10^{-84}$ and

$$\hat{\mu} = 106.654,\ \hat{\sigma} = 8,\ \hat{\lambda_1} = 0.809,\ \hat{\lambda_2} = 0.349,\ \hat{\alpha} = 2,\ \hat{\beta} = 8,$$

$L(\hat{\Theta}; x) = 2.17542 \times 10^{-80}$, respectively. The value of likelihood ratio (LR) test statistic is 19.6557. Since the critical value for the test with significance level 0.05 at one degrees of freedom is 3.84, the null hypothesis is rejected.

Similarly we consider the problem of testing $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$ using the Data set 2. The MLEs and values of the likelihood for $ESGND$ and $EMSGND$ are

$$\hat{\mu} = 31.43211,\ \hat{\sigma} = 2.0174,\ \hat{\lambda_1} = 0.64746,\ \hat{\lambda_2} = 15.91043,\ \hat{\alpha} = 0.96229,$$

$L(\hat{\Theta}^*; x) = 2.62584 \times 10^{-44}$ and

$$\hat{\mu} = 31.5934,\ \hat{\sigma} = 4.464,\ \hat{\lambda_1} = 6.002,\ \hat{\lambda_2} = 8,\ \hat{\alpha} = 0.5,\ \hat{\beta} = 20.006,$$

$L(\hat{\Theta}; x) = 5.87655 \times 10^{-36}$, respectively. The value of likelihood ratio (LR) test statistic is 38.4525. Since the critical value for the test with significance level 0.05 at one degrees of freedom is 3.84, the null hypothesis is rejected.

Similarly we consider the problem of testing $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$ using the Data set 3. The MLEs and values of the likelihood for $ESGND$ and $EMSGND$ are

$$\hat{\mu} = 20.1321,\ \hat{\sigma} = 1.07639,\ \hat{\lambda_1} = 7.85813,\ \hat{\lambda_2} = 10.13911,\ \hat{\alpha} = 1.17325,$$

$L(\hat{\Theta}^*; x) = 1.11045 \times 10^{-141}$ and

$$\hat{\mu} = 21.865,\ \hat{\sigma} = 3.33,\ \hat{\lambda_1} = 7,\ \hat{\lambda_2} = 5,\ \hat{\alpha} = 0.6,\ \hat{\beta} = 8,$$

$L(\hat{\Theta}; x) = 1.92965 \times 10^{-}57$, respectively. The value of likelihood ratio (LR) test statistic is 387.939. Since the critical value for the test with significance level 0.05 at one degrees of freedom is 3.84, the null hypothesis is rejected.

10. Simulation Study

In order to assess the performance of the maximum likelihood estimators of the parameters of the EMSGND($\mu, \sigma; \lambda_1, \lambda_2, \alpha, \beta$), we have conducted a brief simulation study as follows. We have simulated data sets of sizes 30, 50 and 100 from the EMSGND for the parameter values $\mu = 2, \sigma = 0.5, \lambda_1 = 0.8, \lambda_2 = 0.3, \alpha = 1$ and $\beta = 8$. We obtain likelihood estimates of these parameters and computed bias and mean square errors (MSE). The results obtained are presented in Table 2.

Table 2: *Estimate of the parameters and corresponding bias and mean square error(MSE).*

Sample size	Parameters	Estimate	Bias	MSE
30	$\hat{\mu}$	1.988331	0.1883313	0.03546867
	$\hat{\sigma}$	0.4833721	-0.0166279	0.0002764871
	$\hat{\lambda_1}$	0.83	0.54	0.2916
	$\hat{\lambda_2}$	0.29	-1.71	2.9241
	$\hat{\beta}$	7.98	1.98	3.9204
	$\hat{\alpha}$	1.37	1.07	1.1449
50	$\hat{\mu}$	1.99147	-0.008530334	7.27666×10^{-05}
	$\hat{\sigma}$	0.4910136	-0.00898637	8.075485×10^{-05}
	$\hat{\lambda_1}$	0.78	-0.02	4×10^{-04}
	$\hat{\lambda_2}$	0.285	-0.015	0.000225
	$\hat{\beta}$	7.87	-0.13	0.0169
	$\hat{\alpha}$	1.36	0.36	0.1296
100	$\hat{\mu}$	1.994749	-0.005251242	2.757554×10^{-05}
	$\hat{\sigma}$	0.5006707	0.0006706597	4.497845×10^{-07}
	$\hat{\lambda_1}$	0.795	-0.005	2.5×10^{-05}
	$\hat{\lambda_2}$	0.29	-0.01	1×10^{-04}
	$\hat{\beta}$	7.9	-0.1	0.01
	$\hat{\alpha}$	1	$-3.248735 \times 10^{-12}$	1.055428×10^{-23}

From Table 2, it can be observed that both the bias and MSE are in decreasing order as sample size increases.

11. Summary and Conclusion

Through this paper we proposed a wide class of distributions which are suitable for asymmetric as well as plurimodal situations. Certain structural properties of the distribution are derived and discussed its reliability properties as well as unimodal and plurimodal properties. A location scale extension of this class of distribution is also considered and obtained its analogous properties. Further we discussed the maximum likelihood estimation of the parameters of the model and thereby illustrated the procedures through certain real life applications using three real data sets and shown that the model is suitable for all the data sets compared to the existing model. Also, we constructed a test procedure for establishing the significance of the additional parameter β. In order to assess the performance of the maximum likelihood estimation procedure, we carried out a brief simulation study. The proposed model is shown to be more appropriate for asymmetric as well as plurimodal data sets. Certain characteristic properties as well as inferential aspects of the model are yet to study, which we hope to publish in another article. Even though there is flexibility in the proposed model compared to the existing model from the practical point of view, there is scope for developing a further generalized version of the proposed model so as to model more complicated data sets. Such possibilities are under investigation and hope to publish through another article shortly.

Acknowledgments

Authors would like to express their sincere thanks to the Editor in Chief, the Associate Editor and anonymous referees for their valuable comments that have helped to improved the quality and presentation of this paper.

References

[1] Arellano-Valle, R.B Gómez, H. W., and Quintana, F. A. (2004). A new class of skew-normal distributions. *Communications in Statistics-Theory and Methods*, 33:1465–1480.

[2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian-Journal of Statistics*, 12:171–178.

[3] Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. *Statistica*, 46:199–208.

[4] Azzalini, A and Dalla Valle, A (1996). The multivariate skew-normal distribution. *Biometrika*, 83:715–726.

[5] Branco, M. D. and Dey, D. K.(2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, 79:99–113.

[6] Henze, N.(1986). A probabilistic representation of the'skew-normal'distribution. *Scandinavian Journal of Statistics*, 13:271–275.

[7] Kumar, C. S. and Anila, G. V(2018). Asymmetric Curved Normal Distribution. *Journal of Statistical Research*, 52:173–186.

[8] Kumar, C. S. and Anila, G. V(2018). On Some Aspects of a Generalized Asymmetric Normal Distribution. *Statistica*, 77:161–179.

[9] Kumar, C. S. and Anusree, M. R.(2011). On a generalized mixture of standard normal and skew normal distributions. *Statistics & Probability Letters*, 81:1813–1821.

[10] Kumar, C. S. and Anusree, M. R. (2014). On a modified class of generalized skew normal distribution. *South African Statistical Association*, 48:111–124.

[11] Kumar, C. S. and Anusree, M. R. (2015). On an extended version of skew generalized normal distribution and some of its properties. *Communications in Statistics-Theory and Methods*, 44:573–586.

[12] Rohatgi, V. K and Saleh, A. M. E. (2015). An Introduction to Probability and Statistics. *John Wiley & Sons, New York*.

PERFORMANCE MODELING AND DSS FOR ASSEMBLY LINE SYSTEM OF LEAF SPRING MANUFACTURING PLANT

Shanti Parkash

•

Ph.D. Scholar, Department of Production and Industrial Engineering, National Institute of Technology, Kurukshetra-136119, INDIA, Shanti_62000012@nitkkr.ac.in

P.C.Tewari

•

Professor, Department of Mechanical Engineering, National Institute of Technology, Kurukshetra-136119, INDIA, pctewari1@gmail.com

Abstract

This work deals with the Performance Modelling and purposed the Decision Support System (DSS) for maintenance priorities of an assembly line system using a probabilistic approach. This system consists of four subsystems i.e. Shot Peening, Painting Machine, Assembly Platform and Riveting Machine. Performance modelling among various subsystems has been done by Markovian approach. Steady state probabilities are determined by drawing transition diagram and solving the differential equations. Decision matrices are formed with the help of different combinations of failure and repair rates of all the subsystems. The key finding of this work is that painting machine is the most critical subsystem.

Keywords: Markovian Approach, Availability, DSS, Reliability, Maintainability, Performability, RAMS.

I. Introduction

Automobile sector becomes a driver for the growth of a country like India. Leaf springs are the important part of vehicle suspension system which support the overall weight of the vehicle and help to maintain a safe and comfortable ride. Leaf spring manufacturing plants have usually very complex systems in their higher production units. The maintenance of these complex systems becomes costly and time consuming in today's industrial scenario. These challenges have now taken by engineers as an opportunity. An appropriate decision can reduce the operating as well as maintenance costs and also improves the performability of the system. Performability of the plant reduces when the system becomes unavailable for longer period of time. Reliability, Availability, Maintainability and Safety (RAMS) approach plays a significant role to take better and quick decisions in a proper time frame. This is a four dimensional approach which can helps both engineers and managers to enhance the performability of the system by utilizing the best combination of failure and repair rates. RAMS reduces the cost of the plant which helps to achieve the breakeven point rapidly. DSS has been developed using various statistical based techniques such as Reliability Hazard Analysis, Failure Mode and Effects Analysis (FMEA), Reliability Block Diagram, Root Cause Analysis, Fault Tree Analysis, Finite Element Analysis, Markov Analysis, Petri Nets etc.

In this work, Markovian approach has been used for performance modeling and analysis of the system. Markov birth-death analysis is used to predict a random variable, based upon the current state not on the previous activities. It defines the future action on the basis of the current state of a variable. In engineering, this approach has been used to predict the performability of system on the basis of their current state. The probability of any variable has been determined by a decision tree, called transition diagram.

II. Literature Review

Over the past decade, many researchers using this markovian approach for performance modeling of different complex systems. Zhang and Cao [1] determined the reliability of a heat exchanger in deep-sea submersibles using Markov analysis method. Jiang et al. [2] applied Markov chain method on the measured sensors to get their reliability degradation over time in drilling machines. Salari and Makis [3] proposed a modelling for a multi-unit production system using Markov renewal theory. Galagedarage and Khan [4] introduced a methodology which detect and diagnosis the fault using hidden markovian method. Malik and Tewari [5] had done the performance modelling for the Water Flow System (WFS) of a thermal power plant. Alizadeh and Srinivas [6] developed a reliability redundant model for safety systems using markov method. Hassan et al. [7] purposed a stochastic model for liquefied natural gas plant using Markov analysis. Shichang et al. [8] developed a Markov model for multistage manufacturing plant for performance analysis. Kumar [9] had done availability analysis of air circulation system of a thermal plant by markov modelling. Liu et al. [10] discussed double 2-out-of-2 system to obtained time dependent safety and reliability of the system. Kumar et al. [11] evaluated the availability of a thermal power plant using markov birth-death technique. Vora and Tewari [12] described stochastic modelling and analysis of condensate system of a thermal plant using markovian approach. Ge and Asgarpoor [13] developed algorithm for reliability evaluation of equipment with fuzzy markov model.

III. System Description

Assembly line system of a leaf spring manufacturing plant has four major subsystems: Shot Peening (A), Painting Machine (B), Assembly Platform(C) and Riveting Machine (D). Out of these subsystems, only painting machine subsystem has 4 lines in parallel arrangement. Failure of any line of this subsystem reduces the capacity of the system. Other subsystems have no provision of redundancy.

The nomenclatures used for the subsystems (as shown in fig. 1) are described as:

A, B, C and D	: Represent all subsystems are operating in full capacity.
B′, B′′, B′′′	: Represent subsystem B is operating in reduces capacity.
a, b, c and d	: Represent the failure state of all subsystems.
$\lambda_{i,}$, i=1,2,3,4	: Mean constant failure rates for different subsystems A,B, C and D respectively.
$\mu_{i,}$ i =1, 2, 3, 4	: Mean constant repair rates for different subsystems A, B, C and D respectively.
$P_j(t)$, j= 0,1,2……19	: Probability at time ‘t’ the system is in j^{th} state

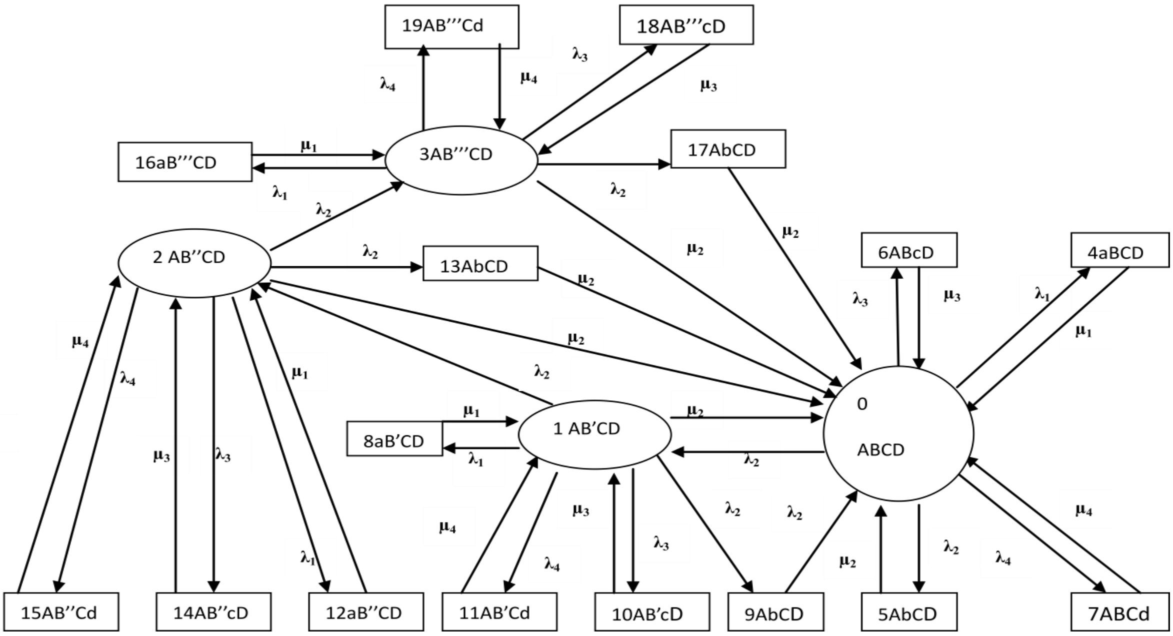


Fig 1:*Transition Diagram of Assembly Line System*

This transition diagram shows the total 20 states ('0' to '19') out of which state '0' represents the full capacity operation, 3 states (i.e., '1' to '3') represents the reduced capacity operation, while 15 states (i.e., '4' to '19') represents the failure state in the transition diagram.

IV. Performance Modelling of the Assembly Line System

To determine the performability of an assembly line system of a Leaf spring manufacturing plant, the mathematical formulation has been carried out using mnemonic rule for all the subsystems. Following mathematical equations represent the two states of the system, transient and steady states.

I. Transient State

The following first order differential equations associated with the transition diagram of the system at time (t+Δt):

$$P_0(t+\Delta t) - P_0(t) = [-\lambda_1\Delta t - 2\lambda_2\Delta t - \lambda_3\Delta t - \lambda_4\Delta t]\,P_0(t) + \mu_1\Delta t P_4(t) + \mu_2\Delta t\,\{P_1(t) + P_2(t) + P_3(t) + P_5(t) + P_9(t) + P_{13}(t) + P_{17}(t)\} + \mu_3\Delta t P_6(t) + \mu_4\Delta t P_7(t)]$$

Taking $\Delta t \to 0$, we get:

$$P'_0(t) = -X_0P_0(t) + \mu_1\Delta t P_4(t) + \mu_2\Delta t\,\{P_1(t) + P_2(t) + P_3(t) + P_5(t) + P_9(t) + P_{13}(t) + P_{17}(t)\} + \mu_3\Delta t\, P_6(t) + \mu_4\Delta t P_7(t)$$

$$P'_0(t) + X_0P_0(t) = \mu_1\Delta t P_4(t) + \mu_2\Delta t\,\{P_1(t) + P_2(t) + P_3(t) + P_5(t) + P_9(t) + P_{13}(t) + P_{17}(t)\} + \mu_3\Delta t\, P_6(t) + \mu_4\Delta t\, P_7(t) \quad (1)$$

Similarly,

$$P'_1(t) + X_1P_1(t) = \lambda_2P_0(t) + \mu_1P_8(t) + \mu_3P_{10}(t) + \mu_4P_{11}(t) \quad (2)$$

$P'_2(t) + X_1P_2(t) = \lambda_2P_1(t) + \mu_1P_{12}(t) + \mu_3P_{14}(t) + \mu_4P_{15}(t)$ (3)

$P'_3(t) + X_2P_3(t) = \lambda_2P_2(t) + \mu_1P_{16}(t) + \mu_3P_{18}(t) + \mu_4P_{19}(t)$ (4)

Where

$X_0 = \lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4$

$X_1 = \lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 + \mu_2$

$X_2 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \mu_2$

$P'_i(t) + \mu_j P_i(t) = \lambda_j P_0(t)$, where, i = 4, 5, 6, 7; j = 1, 2, 3, 4 (5)

$P'_i(t) + \mu_j P_i(t) = \lambda_j P_1(t)$, where, i = 8, 9, 10, 11; j = 1, 2, 3, 4 (6)

$P'_i(t) + \mu_j P_i(t) = \lambda_j P_2(t)$, where, i = 12, 13, 14, 15; j = 1, 2, 3, 4 (7)

$P'_i(t) + \mu_j P_i(t) = \lambda_j P_3(t)$, where, i = 16, 17, 18, 19; j = 1, 2, 3, 4 (8)

II. Steady State

Steady state probabilities of the system are obtained by imposing the following restriction: as $t\rightarrow\infty$, $d/dt\rightarrow 0$.Final results for long run availability are obtained from steady state.

In this state, equation (1) to (8) reduced to the following system of equations:

$X_0P_0 = \mu_1P_4 + \mu_2(P_1 + P_2 + P_3 + P_5 + P_9 + P_{13} + P_{17}) + \mu_3P_6 + \mu_4P_7$ (9)

Similarly,

$X_1P_1 = \lambda_2P_0 + \mu_1P_8 + \mu_3P_{10} + \mu_4P_{11}$ (10)

$X_2P_2 = \lambda_2P_1 + \mu_1P_{12} + \mu_3P_{14} + \mu_4P_{15}$ (11)

$X_2P_3 = \lambda_2P_2 + \mu_1P_{16} + \mu_3P_{18} + \mu_4P_{19}$ (12)

$\mu_jP_i = \lambda_jP_0$, where, i = 4, 5, 6, 7; j = 1, 2, 3, 4 (13)

$\mu_jP_i = \lambda_jP_1$, where, i = 8, 9, 10, 11; j = 1, 2, 3, 4 (14)

$\mu_jP_i = \lambda_jP_2$, where, i = 12, 13, 14, 15; j = 1, 2, 3, 4 (15)

$\mu_jP_i = \lambda_jP_3$, where, i = 16, 17, 18, 19; j = 1, 2, 3, 4 (16)

By solving these equations, we get

Where $N_1 = \lambda_1/\mu_1$, $N_2 = \lambda_2/\mu_2$, $N_3 = \lambda_3/\mu_3$, $N_4 = \lambda_4/\mu_4$

$P_{19} = N_4P_3 = N_4(\lambda_2/2\lambda_2+\mu_2)^2(\lambda_2/\lambda_2+\mu_2)P_0$,

$P_{18} = N_3P_3 = N_3(\lambda_2/2\lambda_2+\mu_2)^2(\lambda_2/\lambda_2+\mu_2)P_0$,

$P_{17} = N_2P_3 = N_2(\lambda_2/2\lambda_2+\mu_2)^2(\lambda_2/\lambda_2+\mu_2)P_0$,

$P_{16} = N_1P_3 = N_1(\lambda_2/2\lambda_2+\mu_2)^2(\lambda_2/\lambda_2+\mu_2)P_0$,

$P_{15} = N_4P_2 = N_4(\lambda_2/2\lambda_2+\mu_2)^2P_0$,

$P_{14} = N_3P_2 = N_3 (\lambda_2/2\lambda_2+\mu_2)^2 P_0$,

$P_{13} = N_2P_2 = N_2 (\lambda_2/2\lambda_2+\mu_2)^2 P_0$,

$P_{12} = N_1P_2 = N_1 (\lambda_2/2\lambda_2+\mu_2)^2 P_0$,

$P_{11} = N_4P_1 = N_4 (\lambda_2/2\lambda_2+\mu_2) P_0$,

$P_{10} = N_3P_1 = N_3 (\lambda_2/2\lambda_2+\mu_2) P_0$,

$P_9 = N_2P_1 = N_2 (\lambda_2/2\lambda_2+\mu_2) P_0$,

$P_8 = N_1P_1= N_1 (\lambda_2/2\lambda_2+\mu_2) P_0$,

$P_7 = N_4P_0$,

$P_6 = N_3P_0$,

$P_5 = N_2P_0$,

$P_4 = N_1P_0$,

$P_3 = (\lambda_2/\lambda_2+\mu_2) P_2 = (\lambda_2/\lambda_2+\mu_2)(\lambda_2/2\lambda_2+\mu_2)^2 P_0$,

$P_2 = (\lambda_2/2\lambda_2+\mu_2) P_1 = (\lambda_2/2\lambda_2+\mu_2)^2 P_0$

$P_1 = (\lambda_2/2\lambda_2+\mu_2) P_0$

Now under the normalizing condition, summation of all the state probabilities is equal to one,

$\Sigma P_i = 1$, i.e. $P_0+P_1+P_2+\ldots\ldots\ldots\ldots\ldots+P_{19}=1$

We get from the above equations,

$P_0 = [1 + (\lambda_2/2\lambda_2+\mu_2) + (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda_2/\lambda_2+\mu_2) (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda 1/ \mu_1) + (\lambda_2/ \mu_2)+ (\lambda_3/ \mu_3)+ (\lambda_4/ \mu_4)+ (\lambda 1/\mu_1) (\lambda_2/2\lambda_2+\mu_2) + (\lambda_2/\mu_2) (\lambda_2/2\lambda_2+\mu_2) + (\lambda_3/\mu_3) (\lambda_2/2\lambda_2+\mu_2) + (\lambda_4/\mu_4) (\lambda_2/2\lambda_2+\mu_2) + (\lambda 1/\mu_1) (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda_2/\mu_2) (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda_3/\mu_3) (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda_4/\mu_4) (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda 1/\mu_1) (\lambda_2/2\lambda_2+\mu_2)^2 (\lambda_2/\lambda_2+\mu_2) + (\lambda_2/\mu_2) (\lambda_2/2\lambda_2+\mu_2)^2 (\lambda_2/\lambda_2+\mu_2) + (\lambda_3/\mu_3) (\lambda_2/2\lambda_2+\mu_2)^2 (\lambda_2/\lambda_2+\mu_2) + (\lambda_4/ \mu_4) (\lambda_2/2\lambda_2+\mu_2)^2 (\lambda_2/\lambda_2+\mu_2)]^{-1}$

The long run performability of the system in terms of availability A (∞) can now be determined using the following equation:

$A (\infty) = P_0 + P_1 + P_2 + P_3 = P_0 + (\lambda_2/2\lambda_2+\mu_2) P_0 + (\lambda_2/2\lambda_2+\mu_2)^2 P_0 + (\lambda_2/\lambda_2+\mu_2) (\lambda_2/2\lambda_2+\mu_2)^2 P_0$

$$= [1+ (\lambda_2/2\lambda_2+\mu_2) + (\lambda_2/2\lambda_2+\mu_2)^2 + (\lambda_2/\lambda_2+\mu_2)(\lambda_2/2\lambda_2+\mu_2)^2] P_0 \quad (17)$$

Failure and repair data for study were obtained from the maintenance logbook of the plant. Failure and repair data follow the exponential distribution.

Table 1: *Failure and repair rates of assembly line system*

Name of the SUBSYSTEM	Exponential Distribution	
	Mean Failure Rate (λ)	Mean Repair Rate (μ)
SHOT PEENING	0.0003 (λ_1)	0.0050 (μ_1)
PAINTING MACHINE	0.0030 (λ_2)	0.0310 (μ_2)
ASSEMBLY	0.0002 (λ_3)	0.0076 (μ_3)
RIVETING	0.0030 (λ_4)	0.2000 (μ_4)

V. Results and Discussion

Table 2, 3, 4 and 5 represent the performability matrices for various subsystems of the assembly line system according to the best possible combinations of failure and repair rates of various subsystems. Tables and figs. 2 to 5 show the effect of failure and repair rates of Shot Peening, Painting Machine, Assembly Platform and Riveting Machine on the steady state performance of the system respectively. Table 2 and fig. 2 reveal the effect of various failure and repair rates of shoot peening on the performability of the system in the terms of availability while other parameters remain constant. As the failure rate increases from 0.0001 to 0.0005, the performability of the system decreases sharply from 0.8077 to 0.6105 (approx.20%) in terms of availability. Similarly, when the repair rate increases from 0.0010 to 0.0090, the performability of the system increases from 0.8077 to 0.8702 (approx. 8%) in terms of availability.

Table 2: *Effect of the failure and repair rates of shot peening subsystem on system performability (%)*

Failure Rates	Repair Rates of Shot Peening					Constant Parameters
	0.0010	0.0030	0.0050	0.0070	0.0090	
0.0001	0.8077	0.8537	0.8635	0.8678	0.8702	λ_2=0.0030
0.0002	0.7473	0.8300	0.8488	0.8571	0.8618	μ_2=0.0310
0.0003	0.6954	0.8077	0.8347	0.8468	0.8537	λ_3=0.0002
0.0004	0.6502	0.7865	0.8210	0.8367	0.8456	μ_3=0.0076
0.0005	0.6105	0.7664	0.8077	0.8268	0.8378	λ_4=0.0030 μ_4=0.2000

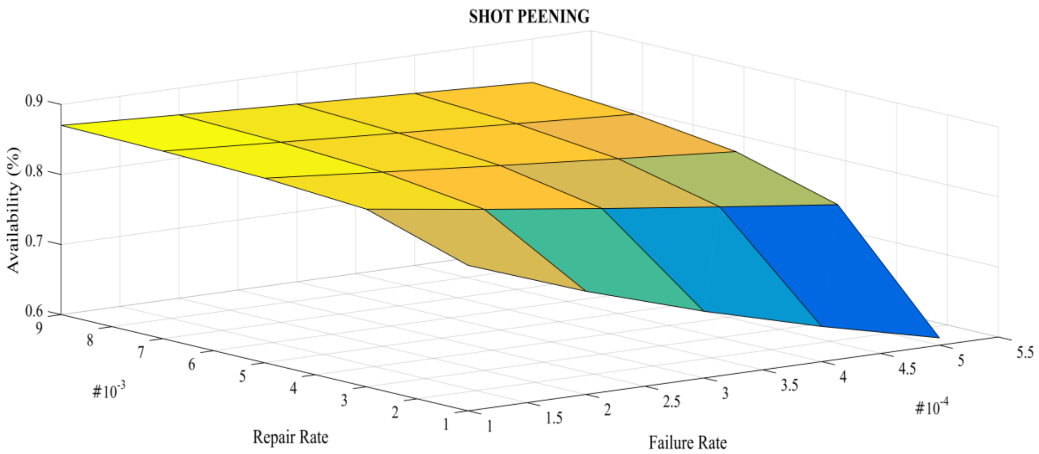


Fig. 2: *Effect of varying failure and repair rates of shot peening subsystem on system performability*

Table 3 and fig. 3 describe the effect of various failure and repair rates of painting machine on the performability of the system in terms of availability while other parameters remain constant. It is observed that when the failure rate increases from 0.0010 to 0.0090, the performability of the system decreases sharply from 0.8388 and 0.5755 (approx. 27%) and similarly when the repair rate increase from 0.0110 to 0.0910, the performability of the system increases from 0.8388 to 0.8990 (approx. 6%).

Table 3: *Effect of the failure and repair rates of painting machine subsystem on system performability (%)*

Failure Rates	Repair Rates of Painting Machine					Constant Parameters
	0.0110	0.0310	0.0510	0.0710	0.0910	
0.0010	0.8388	0.8822	0.8921	0.8965	0.8990	λ_1=0.0003
0.0030	0.7278	0.8347	0.8620	0.8745	0.8816	μ_1=0.0050
0.0050	0.6427	0.7920	0.8338	0.8534	0.8649	λ_3=0.0002
0.0070	0.5755	0.7535	0.8074	0.8334	0.8487	μ_3=0.0076
0.0090	0.5210	0.7186	0.7826	0.8143	0.8332	λ_4=0.0030 μ_4=0.2000

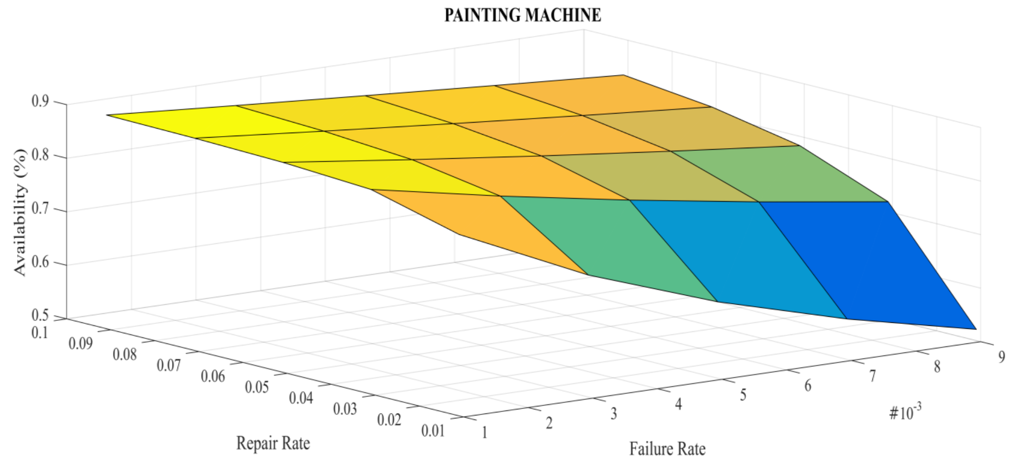


Fig.3: *Effect of varying failure and repair rates of painting machine subsystem on system performability*

Similarly, for the assembly platform the performability of the system in terms of availability varies between 67.37% to 84.59% (approx. 17%) for different combination of failure and repair rates of respective subsystem when other parameters remain constant as shown in the table 4 and fig.4.

Table 4: *Effect of the failure and repair rates of assembly platform subsystem on system performability (%)*

Failure Rates	Repair Rates of Assembly Platform					Constant Parameters
	0.0016	0.0036	0.0056	0.0076	0.0096	
0.0001	0.8102	0.8336	0.8406	0.8439	0.8459	λ_1=0.0003
0.0002	0.7711	0.8148	0.8282	0.8347	0.8385	μ_1=0.0050
0.0003	0.7357	0.7967	0.8161	0.8256	0.8312	λ_2=0.0030
0.0004	0.7033	0.7795	0.8044	0.8167	0.8241	μ_2=0.0310
0.0005	0.6737	0.7630	0.7930	0.8080	0.8171	λ_4=0.0030 μ_4=0.2000

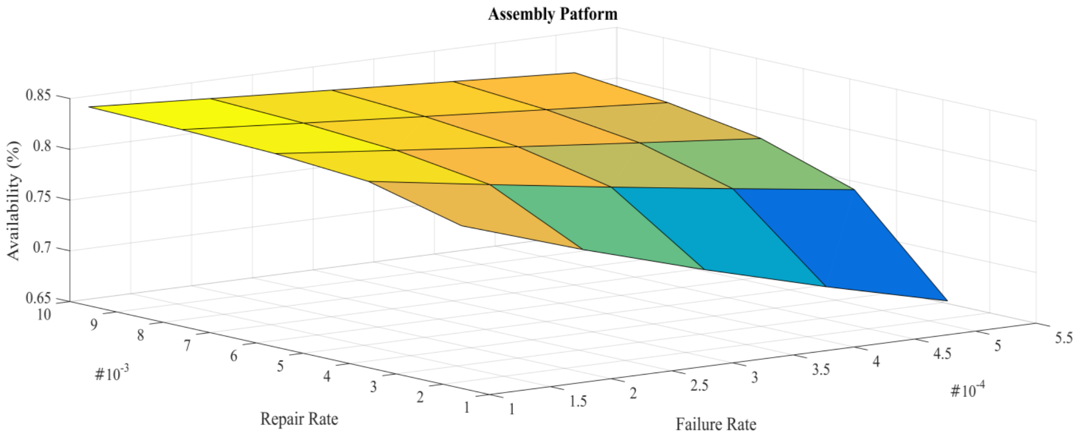


Fig. 4: *Effect of varying failure and repair rates of assembly platform subsystem on system performability*

Table 5 and fig. 5 reveal that variation in failure and repair rates of the riveting machine subsystem increases the system performability from 81.10% to 84.38% (approx. 3%) in terms of availability when other parameters remain constant.

Table 5: *Effect of the failure and repair rates of riveting machine subsystem on system performability (%)*

Failure Rates	Repair Rates of Riveting Machine					Constant Parameters
	0.1000	0.2000	0.3000	0.4000	0.5000	
0.0010	0.8382	0.8417	0.8429	0.8435	0.8438	λ_1=0.0003
0.0020	0.8312	0.8382	0.8405	0.8417	0.8424	μ_1=0.0050
0.0030	0.8243	0.8347	0.8382	0.8399	0.8410	λ_2=0.0030
0.0040	0.8176	0.8312	0.8358	0.8382	0.8396	μ_2=0.0310
0.0050	0.8110	0.8278	0.8335	0.8364	0.8382	λ_3=0.0002 μ_3=0.0076

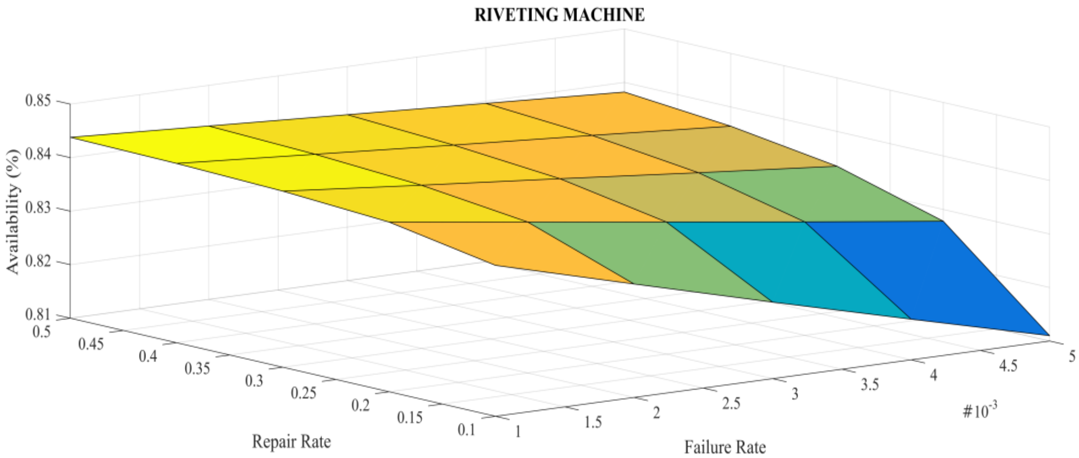


Fig. 5: *Effect of varying failure and repair rates of assembly platform subsystem on system performability*

These performability matrices (Table 2 to Table 5) are very helpful to propose the maintenance priorities for assembly line system. Painting machine has the highest impact on the performance of the system having a variation of 37.82% whereas the lowest impact is done by riveting machine where the variation of 3.28% occurs.

VI. Conclusions and Future Scope

The present work is a case study of assembly line system of leaf spring manufacturing plant. Performance analysis in term of availability is carried out using the Markova method. The study reveals that painting machine subsystem being the most critical component of assembly line system whereas riveting machine system being the lowest contributor in the performance of the system. On the basis of this detailed analysis, a DSS (Decision Support System) has been proposed for maintenance priorities for various subsystem of assembly line system due to which system performance will be enhanced. It is presented in the Table 6.

Table 6: *DSS for assembly line system*

Subsystem	Variation in Failure Rates λ (Repair Rates μ)	Effect on System Performability (%)	Recommended Maintenance Priority
Shot Peening	0.0001-0.0005(0.001-0.009)	0.8702-0.6105(25.97)	II
Painting Machine	0.0010-0.0090(0.011-0.091)	0.8990-0.5210(37.82)	I
Assembly Platform	0.0001-0.0005(0.0016-0.0096)	0.8459-0.6737(17.22)	III
Riveting Machine	0.001-0.005(0.10-0.30)	0.8438-0.8110(3.28)	IV

This work enhances the system performance using the Markovian approach. Markovian approach has some limitations; literature review reveals that use of Petri Nets overcomes these limitations. In fact, selection of appropriate technique had done an impact on maintenance costs. Further the results can be validated with some other robust techniques such as Genetic Algorithm (GA), Teacher Learning Based Optimization (TLBO), Ant Colony Algorithm (ACA), Particle Swarm Optimization (PSO) etc. for such industrial systems.

References

[1] Zhang, K. and Cao, H., "Reliability analysis of heat exchanging system of deep-sea manned submersibles using Markov model", Quality Engineering, 2021.

[2] Jiang, Q., Gao, D., Zhong, Li., Guo, S. and Xiao, A., "Quantitative sensitivity and reliability analysis of sensor networks for well kick detection based on dynamic Bayesian networks and Markov chain", Journal of Loss Prevention in the Process Industries, Volume 66, 2020.

[3] Salari, N. and Makis, V., "Application of Markov renewal theory and semi-Markov decision processes in maintenance modeling and optimization of multi-unit systems", Naval Research Logistics, 2020.

[4] Galagedarage, M. and Khan, F., "Dynamic process fault detection and diagnosis based on a combined approach of hidden Markov and Bayesian network model", Chemical Engineering Science, Volume 201, 2019, pp. 82-96.

[5] Malik, S. and Tewari, P.C., "Performance modeling and maintenance priorities decision for the Waterflow system of a coal-based thermal power plant", IJQRM, Volume 35, 2017, pp. 996-1010.

[6] Alizadeh, S. and Srinivas, S., "Reliability modeling of redundant safety systems without automatic diagnostics incorporating common cause failures and process demand", ISA Transactions, Volume 71, Part 2, 2017, pp. 599-614.

[7] Hassan, J., Thodi, P. and Khan, F., "Availability analysis of a LNG processing plant using the Markov process", Journal of Quality in Maintenance Engineering, Volume 22 , 2016, pp. 302-320.

[8] Shichang, Du., Rui, Xu., Delin, H. and Xufeng, Y., "Markov modeling and analysis of multi-stage manufacturing systems with remote quality information feedback", Computers & Industrial Engineering, Volume 88, 2015, pp. 13-25.

[9] Kumar, R., "Availability analysis of thermal power plant boiler air circulation system using Markov approach", Growing Science Ltd., 2013, doi: 10.5267/j.dsl.2013.08.001.

[10] Liu, Z., Liu, Y., Cai, B., Li, J. and Tian, X., "Markov Modeling of Double 2-out-of-2 System with Imperfect Detection and Common Cause Failures", International Conference on Computer, Networks and Communication Engineering, 2013.

[11] Kumar, R., Sharma, A. and Tewari, P. C., "Markov approach to evaluate the availability simulation model for power generation system in a thermal power plant", International Journal of Industrial Engineering Computations, Volume 3, 2012, pp.743–750.

[12] Vora, Y. and Tewari, P. C., "Simulation model for stochastic analysis and performance evaluation of steam generator system of a thermal power plant", International Journal of Engineering Science and Technology, Volume 3, 2011, pp. 5141-5149.

[13] Ge, H. and Asgarpoor, S., "Reliability Evaluation of Equipment and Substations with Fuzzy Markov Processes", IEEE Transactions on Power Systems, volume 25, 2010, pp. 1319-1328.

CERTAIN CURVATURE CONDITIONS ON LORENTZIAN PARA-KENMOTSU MANIFOLDS

S. Sunitha Devi[1], K. L. Sai Prasad[2,*], T. Satyanarayana[3]

•

Department of Mathematics
[1] Vignan's Institute of Information Technology, Visakhapatnam, 530 049, INDIA
[2,*] Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam, 530 048, INDIA
[3] Pragati Engineering College, Surampalem, Near Peddapuram, Andhra Pradesh, INDIA
sunithamallakula@yahoo.com[1] klsprasad@yahoo.com[2,*] tsn9talluri@gmail.com[3]

Abstract

We classify Lorentzian para-Kenmotsu manifolds which satisfy the curvature conditions $W_2.C = 0$, $Z.C = L_C Q(g, C)$, $W_2.Z - Z.W_2 = 0$ and $W_2.Z + Z.W_2 = 0$, where W_2 is the Weyl- projective tensor, Z is the concircular tensor, and C is the Weyl conformal curvature tensor. We study and have shown that the manifold M is η-Einstein provided that the Weyl-projective curvature tensor W_2 meets the condition $W_2.Z - Z.W_2 = 0$, and it is an Einstein manifold if $W_2.Z + Z.W_2 = 0$. Finally, in this article, we derive the conditions in relation to conformally flatness of the manifold, whenever the LP-Kenmotsu manifold satisfies the condition $Z.C = L_C Q(g, C)$.

Keywords: Para-contact metric manifold, *LP*-Kenmotsu manifold, concircular curvature tensor, conformal curvature tensor, Weyl-projective tensor.
2010 Mathematics Subject Classification: 53C07, 53C25

I. Introduction

In 1989, K. Matsumoto [7] 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 Matsumoto and Mihai [8], Mihai and Rosca [6], Mihai, Shaikh and De [5], Venkatesha and Bagewadi [15], Venkatesha, Pradeep Kumar and Bagewadi [16, 17] and obtained several results of these manifolds.

In 1995, Sinha and Sai Prasad [2] 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 defined a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly *LP*-Kenmotsu) manifolds [1] and they studied ϕ-semisymmetric *LP*-Kenmotsu manifolds with a quarter-symmetric non-metric connection admitting Ricci solitons [13].

On the other hand, In 1970 [4], Pokhariyal and Mishra introduced new tensor fields, called the Weyl-projective curvature tensor W_2 of type $(1, 3)$ and the tensor field E on a Riemannian manifold.The Weyl-projective curvature tensor W_2 with respect to Riemannian connection on a Riemannian manifold M is given by:

$$W_2(X, Y)W = R(X, Y)W + \frac{1}{n-1}[g(X, W)QY - g(Y, W)QX], \tag{1}$$

where $QX = (n-1)X$, which plays an important role in the theory of the projective transformations of connections.

Further, Pokhariyal [3] studied the properties of these tensor fields on a Sasakian manifold. Matsumoto, Ianus and Mihai extended these concepts to almost para-contact structures and studied para-Sasakian manifolds admitting these tensor fields [9] in 1986 and these results were generalised by De and Sarkar, in 2009 [14]. Sai Prasad and Satyanarayana studied the W_2-tensor field in an SP-Kenmotsu manifold [10]. In our earlier work, we consider LP-Kenmotsu manifolds admitting the Weyl-projective curvature tensor W_2 and shown that these manifolds admiting 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 [11, 12].

Inspired by these studies, in the present work, we explore a class of Lorentzian para-Kenmotsu manifolds that admits certain curvature conditions. The current study is arranged as follows: Section 2 has certain prerequisites. In section 3, it is illustrated that the manifold M is η-Einstein provided that the Weyl-projective curvature tensor W_2 meets the condition $W_2.Z - Z.W_2 = 0$, and it is an Einstein manifold if $W_2.Z + Z.W_2 = 0$. Finally, we derive the conditions in relation to conformally flatness of the manifold, whenever the LP-Kenmotsu manifold satisfying the condition $Z.C = L_C Q(g, C)$, where the concircular curvature tensor $Z(X,Y)$ is given by:

$$Z(X,Y)W = R(X,Y)W - \frac{r}{n(n-1)}[g(Y,W)X - g(X,W)Y]. \tag{2}$$

II. Preliminaries

An n-dimensional differentiable manifold M admitting a (1, 1) tensor field ϕ, contravariant vector field ξ, a 1-form η and the Lorentzian metric $g(X,Y)$ satisfying

$$\phi^2 X = X + \eta(X)\xi, \;\; g(\phi X, \phi Y) = g(X,Y) + \eta(X)\eta(Y) \tag{3}$$

and

$$\eta(\xi) = -1, \;\; \phi\xi = 0, \;\; \eta(\phi X) = 0, \;\; g(X,\xi) = \eta(X), \; rank\phi = n-1. \tag{4}$$

is called Lorentzian almost paracontact manifold [7].

In a Lorentzian almost paracontact manifold, we have

$$\Phi(X,Y) = \Phi(Y,X), \tag{5}$$

where $\Phi(X,Y) = g(X, \phi Y)$.

A Lorentzian almost paracontact manifold M is called Lorentzian para-Kenmotsu manifold if [1]

$$(\nabla_X \phi)Y = -g(\phi X, Y)\xi - \eta(Y)\phi X, \tag{6}$$

for any vector fields X and Y on M and ∇ is the operator of covariant differentiation with respect to the Lorentzian metric g.

In the Lorentzian para-Kenmotsu manifold, the following relations hold good:

$$\nabla_X \xi = -\phi^2 X = -X - \eta(X)\xi \tag{7}$$

and

$$(\nabla_X \eta)Y = -g(X,Y)\xi - \eta(X)\eta(Y). \tag{8}$$

Further, on a Lorentzian para-Kenmotsu manifold M, the following relations hold [1]:

$$g(R(X,Y)Z,\xi) = \eta(R(X,Y)Z) = g(Y,Z)\eta(X) - g(X,Z)\eta(Y), \tag{9}$$

$$R(\xi,X)Y = g(X,Y)\xi - \eta(Y)X, \tag{10}$$

$$R(X,Y)\xi = \eta(Y)X - \eta(X)Y; \text{ when } X \text{ is orthogonal to } \xi, \tag{11}$$

$$R(\xi,X)\xi = X + \eta(X)\xi, \tag{12}$$

$$S(X,\xi) = (n-1)\eta(X), \tag{13}$$

$$S(\phi X,\phi Y) = S(X,Y) + (n-1)\eta(X)\eta(Y). \tag{14}$$

A Lorentzian para-Kenmotsu manifold M is said to be an η-Einstein manifold if its Ricci tensor $S(X,Y)$ is of the form

$$S(X,Y) = ag(X,Y) + b\eta(X)\eta(Y), \tag{15}$$

where a and b are scalar functions on M.

Next we define endomorphisms $R(X,Y)$ and $X \wedge_A Y$ by

$$R(X,Y)W = \nabla_X\nabla_Y W - \nabla_Y\nabla_X W - \nabla_[X,Y]W, \tag{16}$$

$$(X \wedge_A Y)W = A(Y,W)X - A(X,W)Y, \tag{17}$$

A is the symmetric $(0,2)$- tensor.

For a $(0,k)$-tensor field T, $K \geq 1$, on (M_n, g) we define $W_2.T$, $Z.T$ and $Q(g,T)$ by

$$\begin{aligned}(W_2(X,Y).T(X_1,X_2,...,X_k) = &- T(W_2(X,Y)X_1,X_2,...,X_k)\\ &- T(X_1,W_2(X,Y)X_2,...,X_k)\\ &- ... - T(X_1,X_2,...,W_2(X,Y)X_k),\end{aligned} \tag{18}$$

$$\begin{aligned}(Z(X,Y).T(X_1,X_2,...,X_k) = &- T(Z(X,Y)X_1,X_2,...,X_k)\\ &- T(X_1,Z(X,Y)X_2,...,X_k)\\ &- ... - T(X_1,X_2,...,Z(X,Y)X_k),\end{aligned} \tag{19}$$

$$\begin{aligned}Q(g,T)(X_1,X_2,...,X_k;X,Y) = &- T((X\wedge Y)X_1,X_2,...,X_k)\\ &- T(X_1,(X\wedge Y)X_2,...,X_k)\\ &- ... - T(X_1,X_2,...,(X\wedge Y)X_k),\end{aligned} \tag{20}$$

respectively.

By definition the Weyl Conformal curvature tensor C is given by

$$\begin{aligned}C(X,Y)Z =& R(X,Y)Z - \frac{1}{n-2}[S(Y,Z)X - S(X,Z)Y + g(Y,Z)QX - g(X,Z)QY]\\ &\frac{r}{(n-1)(n-2)}[g(Y,Z)X - g(X,Z)Y],\end{aligned} \tag{21}$$

where Q denotes the Ricci operator, i.e., $S(X,Y) = g(QX,Y)$ and r is scalar curvature. The Weyl conformal curvature tensor C is defined by $C(X,Y,Z,W) = g(C(X,Y)Z,W)$. If $C = 0, n \geq 4$, then M is conformally flat.

III. MAIN RESULTS

In the present section we consider the LP-Kenmotsu manifold satisfying the curvature conditions $W_2.C = 0$, $Z.C = L_C Q(g, C)$, $W_2.Z - Z.W_2 = 0$, and $W_2.Z + Z.W_2 = 0$.

First we give the following proposition.

Proposition 1. Let M be an n-dimensional $(n > 3)$ LP-Kenmotsu manifold. If the condition $W_2.C = 0$ holds on M, then

$$S^2(X,U) = (n-1)(r-2)\eta(X)\eta(U) + (r+n-2)S(U,X) - (n-1)g(X,U)$$

is satisfied on M, where $S^2(X,U) = S(QX,U)$.

Proof: Assume that M is an n-dimensional, $n > 3$, LP-Kenmotsu manifold satisfying the condition $W_2.C = 0$. From (18) we have

$$\begin{aligned}(W_2(V,X).C)(Y,U)W = &- W_2(V,X)C(Y,U)W\\ &- C(W_2(V,X)Y,U)W - C(Y,W_2(V,X)U)W\\ &- C(Y,U)W_2(V,X)W = 0,\end{aligned} \tag{22}$$

where $X, Y, U, V, W \in \chi(M)$. Taking $V = \xi$ in (22), we have

$$\begin{aligned}(W_2(\xi,X).C)(Y,U)W = &- W_2(\xi,X)C(Y,U)W\\ &- C(W_2(\xi,X)Y,U)W - C(Y,W_2(\xi,X)U)W\\ &- C(Y,U)W_2(\xi,X)W = 0,\end{aligned} \tag{23}$$

Furthermore, substituting (1), (9), (13), (21) into (23) and multiplying with ξ, we get.

$$\begin{aligned}&- g(X,C(Y,U)W) - g(X,Y)\eta(C(\xi,U)W + \eta(Y)\eta(C(X,U)W)\\ &- g(X,U)\eta(C(Y,\xi)W) + \eta(U)\eta(C(Y,X)W) - g(X,W)\eta(C(Y,U)\xi)\\ &+ \eta(W)\eta(C(Y,U)X) + \frac{1}{n-1}[\eta(C(Y,U)W) - \eta(Y)\eta(C(QX,U)W)\\ &+ g(X,Y)\eta(C(Q\xi,U)W) + g(X,U)\eta(C(Y,Q\xi)W) - \eta(U)\eta(C(Y,QX)W)\\ &- \eta(W)\eta(C(Y,U)QX) + g(X,W)\eta(C(Y,U)Q\xi)] = 0.\end{aligned} \tag{24}$$

Thus replacing W with ξin (24), we have

$$- g(X,C(Y,U)\xi) - \eta(C(Y,U)X) + \frac{1}{n-1}[\eta(C(Y,U)QX)] = 0. \tag{25}$$

Again taking $Y = \xi$ in (25)and after some calculations, since $n > 3$, we get

$$S^2(U,X) = (n-1)(r-2)\eta(X)\eta(U) + (r+n-2)S(U,X) - (n-1)g(X,U).$$

Theorem 2. Let M be an n-dimensional $(n > 3)$ LP-Kenmotsu manifold. If the condition $Z.C = L_C Q(g, C)$ holds on M, then either M is conformally flat or $L_C = \frac{r}{n(n-1)} - 1$.

Proof. Let M be an LP-Kenmotsu manifold. So we have

$$(Z(V,X).C)(Y,U)W = L_C Q(g,C)(Y,U,W;V,X).$$

Then from (19) and (20) we can write,

$$\begin{aligned}Z(V,X)C(Y,U)W &- C(Z(V,X)Y,U)W - C(Y,Z(V,X)U)W\\ &- C(Y,U)Z(V,X)W\\ &= L_C[(V\wedge X)C(Y,U)W - C((V\wedge X)Y,U)W\\ &- C(Y,(V\wedge X)U)W - C(Y,U)(V\wedge X)W].\end{aligned} \tag{26}$$

Therefore, replacing v with ξ in (26), we have

$$\begin{aligned} Z(\xi,X)C(Y,U)W - C(Z(\xi,X)Y,U)W - C(Y,Z(\xi,X)U)W \\ - C(Y,U)Z(\xi,X)W \\ = L_C[(\xi\wedge X)C(Y,U)W - C((\xi\wedge X)Y,U)W \\ - C(Y,(\xi\wedge X)U)W - C(Y,U)(\xi\wedge X)W]. \end{aligned} \tag{27}$$

Using (20), (9) and taking the inner product of (27) with ξ, we get

$$\begin{aligned} &\big[1-\frac{r}{n(n-1)}-L_C\big][-g(X,C(Y,U)W)-\eta(X)\eta(C(Y,U)W) \\ &-g(X,Y)\eta(C(\xi,U)W)+\eta(Y)\eta(C(X,U)W) \\ &-g(X,U)\eta(C(Y,\xi)W)+\eta(U)\eta(C(Y,X)W)+\eta(W)\eta(C(Y,U)X)\big]=0. \end{aligned} \tag{28}$$

Putting $X=Y$ in (28), we have

$$\begin{aligned} &\big[1-\frac{r}{n(n-1)}-L_C\big][-g(Y,C(Y,U)W)+\eta(W)\eta(C(Y,U)Y) \\ &-g(Y,Y)\eta(C(\xi,U)W)-g(Y,U)\eta(C(Y,\xi)W)\big]=0. \end{aligned} \tag{29}$$

A contraction of (29) with respect to Y gives us

$$\big[1-\frac{r}{n(n-1)}-L_C\big]\eta(C(\xi,U)W)=0. \tag{30}$$

If $L_C \neq 1-\frac{r}{n(n-1)}$, then eq.(30) is reduced to

$$\eta(C(\xi,U)W)=0, \tag{31}$$

which gives

$$S(U,W)=\big(\frac{r}{(n-1)}-1\big)g(U,W)+\big(\frac{r}{(n-1)}-n\big)\eta(U)\eta(W). \tag{32}$$

Therefore, M is a η-Einstein manifold. So, using (31) and (32), we have eq. (28) in the form

$$C(Y,U,W,X)=0,$$

which means that M is conformally flat.

If $L_C \neq 0$ and $\eta(C(\xi,U)W)\neq 0$, then $1-\frac{r}{n(n-1)}-L_C=0$, which gives $L_C=1-\frac{r}{n(n-1)}$. This completes the proof of the theorem.

Corollary 3. Every n-dimensional $(n>3)$ nonconformally flat *LP*-Kenmotsu manifold satisfies $Z.C=\big(1-\frac{r}{n(n-1)}\big)Q(g,C)$.

Theorem 4. Let M be an n-dimensional $(n>3)$ *LP*-Kenmotsu manifold. M satisfies the condition

$$W_2.Z - Z.W_2 = 0$$

if and only if M is a η-Einstein manifold.

Proof. Let M satisfy the condition $W_2.Z - Z.W_2 = 0$. Then we can write

$$\begin{aligned}
W_2.Z - Z.W_2 =& R(V,X)R(Y,U)W + \frac{1}{n-1}\left[g(V,R(Y,U)W)QX - g(X,R(Y,U)W)QV\right] \\
&- \frac{r}{n(n-1)}g(U,W)\left[R(V,X)Y + \frac{1}{n-1}\left(g(V,Y)QX - g(X,Y)QV\right)\right] \\
&+ \frac{r}{n(n-1)}g(Y,W)\left[R(V,X)U + \frac{1}{n-1}\left(g(V,U)QX - g(X,U)QV\right)\right] \\
&- R(V,X)R(Y,U)W + \frac{r}{n(n-1)}\left[g(X,R(Y,U)W)V - g(V,R(Y,U)W)X\right] \\
&- \frac{1}{n-1}g(Y,W)\left[R(V,X)QU - \frac{r}{n(n-1)}\left(g(X,QU)V - g(V,QU)X\right)\right] \\
&+ \frac{1}{n-1}g(U,W)\left[R(V,X)QY - \frac{r}{n(n-1)}\left(g(X,QY)V - g(V,QY)X\right)\right] = 0.
\end{aligned} \tag{33}$$

Therefore, replacing V with ξ in (33), we have

$$\begin{aligned}
W_2.Z - Z.W_2 =& \frac{1}{n-1}\left[g(\xi,R(Y,U)W)QX - g(X,R(Y,U)W)Q\xi\right] \\
&- \frac{r}{n(n-1)}g(U,W)\left[R(\xi,X)Y + \frac{1}{n-1}\left(g(\xi,Y)QX - g(X,Y)Q\xi\right)\right] \\
&+ \frac{r}{n(n-1)}g(Y,W)\left[R(\xi,X)U + \frac{1}{n-1}\left(g(\xi,U)QX - g(X,U)Q\xi\right)\right] \\
&- R(\xi,X)R(Y,U)W + \frac{r}{n(n-1)}\left[g(X,R(Y,U)W)\xi - g(\xi,R(Y,U)W)X\right] \\
&- \frac{1}{n-1}g(Y,W)\left[R(\xi,X)QU - \frac{r}{n(n-1)}\left(g(X,QU)\xi - g(\xi,QU)X\right)\right] \\
&+ \frac{1}{n-1}g(U,W)\left[R(\xi,X)QY - \frac{r}{n(n-1)}\left(g(X,QY)\xi - g(\xi,QY)X\right)\right] = 0.
\end{aligned} \tag{34}$$

Using (10), (13), we get

$$\begin{aligned}
W_2.Z - Z.W_2 =& \frac{1}{n-1}\left[g(\xi,R(Y,U)W)QX - g(X,R(Y,U)W)Q\xi\right] \\
&- \frac{r}{n(n-1)}g(U,W)\left[g(X,Y)\xi - \eta(Y)X\right] - \frac{r}{n(n-1)}g(U,W)\eta(Y)X \\
&+ \frac{r}{n(n-1)}g(U,W)g(X,Y)\xi + \frac{r}{n(n-1)}g(Y,W)\left[g(X,U)\xi - \eta(U)X\right] \\
&- \frac{r}{n(n-1)}g(Y,W)\eta(U)X - \frac{r}{n(n-1)}g(Y,W)g(X,U)\xi \\
&+ \frac{r}{n(n-1)}\left[g(X,R(Y,U)W)\xi - g(\xi,R(Y,U)W)X\right. \\
&- \frac{1}{(n-1)}g(Y,W)\left[g(X,QU)\xi - \eta(QU)X\right] + \frac{r}{n(n-1)^2}g(Y,W)g(X,QU)\xi \\
&- \frac{r}{n(n-1)^2}g(Y,W)\eta(QU)X + \frac{1}{(n-1)}g(U,W)\left[g(X,QY)\xi - \eta(QY)X\right] \\
&- \frac{r}{n(n-1)^2}g(U,W)g(X,QY)\xi - \frac{r}{n(n-1)^2}g(U,W)\eta(QY)X = 0.
\end{aligned} \tag{35}$$

Again, taking $U = \xi$ in (35), we get

$$
\begin{aligned}
&\frac{1}{n-1}\left[g(\xi, g(Y,W)\xi - \eta(W)Y)(n-1)X - g(X, g(Y,W)\xi - \eta(W)Y)(n-1)\xi\right] \\
&- \frac{r}{n(n-1)}\eta(W)\left[g(X,Y)\xi - \eta(Y)X\right] - \frac{r}{n(n-1)}\eta(Y)\eta(W)X \\
&+ \frac{r}{n(n-1)}g(X,Y)\eta(W)\xi + \frac{r}{n(n-1)}g(Y,W)\left[\eta(X)\xi + X\right] \\
&- \frac{r}{n(n-1)}g(Y,W)\eta(U)X - \frac{r}{n(n-1)}g(Y,W)g(X,U)\xi \\
&+ \frac{r}{n(n-1)}g(Y,W)X - \frac{r}{n(n-1)}g(Y,W)\eta(X)\xi \\
&+ \frac{r}{n(n-1)}\left[g(X, g(Y,W)\xi - \eta(W)Y)\xi - g(\xi, g(Y,W)\xi - \eta(W)Y)x\right. \\
&- \frac{1}{(n-1)}g(Y,W)\left[(n-1)\eta(X)\xi - (n-1)X\right] + \frac{r}{n(n-1)}g(Y,W)\eta(X)\xi \\
&- \frac{r}{n(n-1)^2}g(Y,W)X + \frac{1}{(n-1)}\eta(W)\left[(n-1)g(X,Y)\xi - (n-1)\eta(Y)X\right] \\
&- \frac{r}{n(n-1)^2}\eta(W)S(X,Y)\xi - \frac{r}{n(n-1)}\eta(W)\eta(Y)X = 0.
\end{aligned}
\tag{36}
$$

Taking the inner product of (36) with ξ, we find

$$
\begin{aligned}
-2\eta(W)\eta(Y)\eta(X) - 2\eta(W)g(X,Y) + \frac{r}{n(n-1)}\eta(W)g(X,Y) + \frac{2r}{n(n-1)}\eta(W)\eta(Y)\eta(X) \\
+ \frac{2r}{n(n-1)}\eta(X)g(Y,W) + \frac{r}{n(n-1)^2}\eta(W)S(X,Y) = 0.
\end{aligned}
\tag{37}
$$

Again, taking $W = \xi$ and using (4) in (37), we get

$$
S(X,Y) = [\frac{2(n-1)}{r}\eta(X)\eta(Y) + \frac{(n-r)(n-1)}{r}g(X,Y) \tag{38}
$$

So, M is a η-Einstein manifold.

Conversely, if M is a η-Einstein manifold, then it is easy to show that $W_2.Z - Z.W_2 = 0$. Our theorem is thus proved.

Theorem 5. Let M be an n-dimensional $(n > 3)$ *LP*-Kenmotsu manifold. M satisfies the condition

$$
W_2.Z + Z.W_2 = 0
$$

if and only if M is an Einstein manifold.

Proof. Let M satisfy the condition $W_2.Z + Z.W_2 = 0$. Then from (33) and (34) we can write

$$
\begin{aligned}
&2R(V,X)R(Y,U)W + \frac{1}{n-1}\left[g(V, R(Y,U)W)QX - g(X, R(Y,U)W)QV\right] \\
&- \frac{r}{n(n-1)}g(U,W)\left[R(V,X)Y + \frac{1}{n-1}\left(g(V,Y)QX - g(X,Y)QV\right)\right] \\
&+ \frac{r}{n(n-1)}g(Y,W)\left[R(V,X)U + \frac{1}{n-1}\left(g(V,U)QX - g(X,U)QV\right)\right] \\
&- \frac{r}{n(n-1)}\left[g(X, R(Y,U)W)V - g(V, R(Y,U)W)X\right] \\
&+ \frac{1}{n-1}g(Y,W)\left[R(V,X)QU - \frac{r}{n(n-1)}\left(g(X,QU)V - g(V,QU)X\right)\right] \\
&- \frac{1}{n-1}g(U,W)\left[R(V,X)QY - \frac{r}{n(n-1)}\left(g(X,QY)V - g(V,QY)X\right)\right] = 0.
\end{aligned}
\tag{39}
$$

Therefore, replacing V with ξ in (39), we have

$$\begin{aligned} &2R(\xi,X)R(Y,U)W+\frac{1}{n-1}\left[g(\xi,R(Y,U)W)QX-g(X,R(Y,U)W)Q\xi\right]\\ &-\frac{r}{n(n-1)}g(U,W)\left[R(\xi,X)Y+\frac{1}{n-1}\left(g(\xi,Y)QX-g(X,Y)Q\xi\right)\right]\\ &+\frac{r}{n(n-1)}g(Y,W)\left[R(\xi,X)U+\frac{1}{n-1}\left(g(\xi,U)QX-g(X,U)Q\xi\right)\right]\\ &-\frac{r}{n(n-1)}\left[g(X,R(Y,U)W)\xi-g(\xi,R(Y,U)W)X\right]\\ &+\frac{1}{n-1}g(Y,W)\left[R(\xi,X)QU-\frac{r}{n(n-1)}\left(g(X,QU)\xi-g(\xi,QU)X\right)\right]\\ &-\frac{1}{n-1}g(U,W)\left[R(\xi,X)QY-\frac{r}{n(n-1)}\left(g(X,QY)\xi-g(\xi,QY)X\right)\right]=0. \end{aligned} \tag{40}$$

Again, taking $Y=\xi$ in (40), we get

$$\begin{aligned} &2R(\xi,X)R(\xi,U)W+\frac{1}{n-1}\left[g(\xi,R(\xi,U)W)QX-g(X,R(\xi,U)W)Q\xi\right]\\ &-\frac{r}{n(n-1)}g(U,W)\left[R(\xi,X)\xi-\frac{1}{n-1}QX-\frac{1}{n-1}\eta(X)Q\xi\right]\\ &+\frac{r}{n(n-1)}\eta(W)\left[R(\xi,X)U+\frac{1}{n-1}\eta(U)QX-\frac{1}{n-1}g(X,U)Q\xi\right]\\ &-\frac{r}{n(n-1)}\left[g(X,R(\xi,U)W)\xi-g(\xi,R(\xi,U)W)X\right]\\ &+\frac{1}{n-1}\eta(W)\left[R(\xi,X)QU-\frac{r}{n(n-1)}S(X,U)\xi+\frac{r}{n(n-1)}(n-1)\eta(U)\eta(X)\right]\\ &-\frac{1}{n-1}g(U,W)\left[(n-1)R(\xi,X)\xi-\frac{r}{n(n-1)}(n-1)\eta(X)\xi-\frac{r}{n(n-1)}(n-1)X\right]=0. \end{aligned} \tag{41}$$

Taking the inner product of (41) with ξ and using (7), (10), we get

$$eta(W)g(X,U)-\frac{r}{n(n-1)}\eta(W)g(X,U)-\frac{1}{n-1}\eta(W)S(X,U)+\frac{r}{n(n-1)^2}S(X,U)=0. \tag{42}$$

Again, taking $W=\xi$ and using (4) in (42), we get

$$-g(X,U)-\frac{r}{n(n-1)}g(X,U)+\frac{1}{n-1}S(X,U)+\frac{r}{n(n-1)^2}S(X,U)=0. \tag{43}$$

Thus, from (43), we have

$$S(X,U)=(n-1)g(X,U) \tag{44}$$

So, M is an Einstein manifold.

Conversely, if M is an Einstein manifold, then it is easy to show that $W_2.Z+Z.W_2=0$. Our theorem is thus proved.

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.

Declarations of conflicting interests: The authors declares that there is no conflict of interest.

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] B. B. Sinha and K. L. Sai Prasad, A class of almost para contact metric Manifold, Bulletin of the Calcutta Mathematical Society, 87(1995), 307-312.

[3] G. P. Pokharial, Study of a new curvature tensor in a Sasakian manifold, Tensor (N.S.), 36(1982), 222-225.

[4] G. P. Pokhariyal and R. S. Mishra, The curvature tensors and their relativistic significance, Yokohoma Math. J., 18(1970), 105-108.

[5] I Mihai, A. A. Shaikh and U. C. De, On Lorentzian para-Sasakian manifolds, Rendiconti del Seminario Matematico di Messina, (1999),Serie II.

[6] I Mihai and R. Rosca, On Lorentzian P-Sasakian manifolds, Classical Analysis, World Scientific Publ., Singapore, (1992), 155-169.

[7] K. Matsumoto, On Lorentzian Paracontact manifolds, Bulletin of the Yamagata University Natural Science, 12(2) (1989), 151-156.

[8] K. Matsumoto and I Mihai, On a certain transformation in a Lorentzian para-Sasakian manifold, Tensor, N.S., 47(1988), 189-197.

[9] K. Matsumoto, S. Ianus and Ion Mihai, On *P*-Sasakian manifolds which admit certain tensor fileds, Publ. Math. Debrece, 33(1986), 61-65.

[10] K. L. Sai Prasad and T. Satyanarayana, Some curvature properties on a Special paracontact Kenmotsu manifold with respect to Semi-symmetric connection, *Turkish Journal of Analysis and Number Theory*, 3(2015), no.4, 94-96.

[11] K. L. Sai Prasad, S. Sunitha Devi and G. V. S. R. Deekshitulu, 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), 990-1001.

[12] K. L. Sai Prasad, S. Sunitha Devi and G. V. S. R. Deekshitulu, On a class of Lorentzian paracontact metric manifolds, Italian Journal of Pure and Applied Mathematics, Accepted for publication in 2021, to appear in 2022 issue.

[13] Rajendra Prasad, Abdul Haseeb and Umesh Kumar Gautam, On $\check{\phi}$-semisymmetric *LP*-Kenmotsu manifolds with a QSNM-connection admitting Ricci solitons, Kragujevac Journal of Mathematics, 45(5),(2021), 815-827.

[14] U. C. De and Avijit Sarkar, On a type of P-Sasakian manifolds, Math. Reports, 11(2009), no.2, 139-144.

[15] Venkatesha and C. S. Bagewadi, On concircular ϕ-recurrent LP-Sasakian manifolds, Differ. Geom. Dyn. Syst., 10 (2008), 312-319.

[16] Venkatesha and C. S. Bagewadi and K. T. Pradeep Kumar, Some results on Lorentzian para-Sasakian manifolds, ISRN Geometry, Vol. 2011, Article ID 161523, 9 pages, (2011). doi: 10.5402/2011/161523.

[17] Venkatesha, K. T. Pradeep Kumar and C. S. Bagewadi, On Lorentzian Para-Sasakian manifolds satisfying W_2 curvature tensor, IOSR J. of Mathematics, 9(6), (2014), 124-127.

Bayesian Survival Modeling of Marshal Olkin Generalized-G family with random effects using R and STAN

Shazia Farhin, Firdoos Yousuf and Athar Ali Khan

•

Department of Statistics and Operations Research
Aligarh Muslim University, Aligarh-202002, India
farhinshazia@gmail.com
firdoos1990@gmail.com
atharkhan1962@gmail.com

Abstract

*The purpose of this paper is to fit the Marshall-Olkin generalized-G(MOG-G) family to censored survival data with random effect in the Bayesian environment. Three special distrbution based on MOG-G family are obtained, namely Marshall-Olkin generalized-exponential, Marshall-Olkin generalized-Weibull, and Marshall-Olkin generalized-Lomax. The probabilistic programming language **STAN** is used for the fitting of these three distrbution to the survival data. **STAN** offers full Bayesian inference and implements via Hamiltonian Monte Carlo algorithm and No-U-Turn Sampler(NUTS) algorithm of MCMC. We compared the models with the help of leave one out cross-validation information criteria and Watanabe Akaike information criteria. Stan codes for the analysis are provided.*

Keywords: Bayesian modeling, Marshall-Olkin generalized-G family, censored survival data, random effect, Leave one out information criteria, STAN

1. Introduction

In the survival analysis, researchers are using the extended version of standard distribution to analyze the lifetime data and problems related to the modeling of the aging or failure process. In this paper, we have used the Marshall-Olkin generalized-G (MOG-G) family to fit censored survival data, including the random effect. [1] proposed the MOG-G family and studied its mathematical properties along with application in the fitting of lifetime data.The Marshall Olkin distribution has been extended by using the genesis of other distributions to create a wider family of distribution. see for example, Marshall-Olkin-G family [2], Kumaraswamy marshal-Olkin family [3], beta Marshall-Olkin family [4], Beta Generalized Marshall-Olkin-G family [5], Exponentiated Marshall-Olkin family [6], The generalized Marshall-Olkin-Kumaraswamy-G family [7], The Beta generalized Marshall-Olkin Kumaraswamy-G [8], The exponentiated generalized Marshall–Olkin family [9], The Weibull Marshall–Olkin family [10].

We have considered the three models based on the MOG-G family and are fitted to the survival data. The first model is Marshall-Olkin Generalized-Exponential(MOG-E), the second is Marshall-Olkin Generalized-Weibull (MOG-W) model, and the third one is Marshall-Olkin Generalized-Lomax (MOG-L) model. The data with random effect significantly affects the distribution of the patients' survival time and accounts heterogeneity among the patients. Fitting a large number of random effects in a non-Bayesian setting requires a large amount of data. Often, the data is too small to estimate random-effects parameters reliably. However, Bayesian modeling can be used if there is not enough data for inferential statistics. So, the above three models have been fitted to the censored survival data under the Bayesian setup in `R` [11] using the probabilistic programming language `STAN` [12], which offers full Bayesian inference. `STAN` uses Hamiltonian Monte Carlo (HMC) sampling [13],[14] and its extension. No-U-Turn Sampler(NUTS) [15] algorithm of MCMC for the simulation and computation of posterior estimate. HMC is a more efficient and sophisticated MCMC algorithm, and it is the combination of MCMC and deterministic simulation methods. To find the region of posterior distribution with high mass, HMC uses the gradient of the log posterior density. After that, it jumps around the posterior distribution [16]. Whether the priors are conjugate or not, the above algorithms converge at a fast rate to high dimensional target distributions as compared to other algorithms of MCMC [15].

The purpose of this paper is to fit the three models, namely MOG-E, MOG-W, and MOG-L, to the censored survival data containing random effects under the Bayesian environment using the `R` and `STAN` and select the best model for the real survival data.

2. Marshal-Olkin Generalized-G family

Suppose that $G(t,\psi)$ and $g(t,\psi)$ be baseline cdf and pdf of a continuous random variable T with parameter vector ψ. The cdf, pdf, survival function, and hazard function of the MOG-G family are respectively given by

$$F(t,a,\alpha,\psi) = \frac{1-[1-G(t,\psi)]^a}{1-(1-\alpha)[1-G(t,\psi)]^a}, \qquad t \in R \tag{1}$$

$$f(t,a,\alpha,\psi) = \frac{\alpha a g(t,\psi)[1-G(t,\psi)]^{a-1}}{[1-(1-\alpha)[1-G(t,\psi)]^a]^2}, \qquad t \in R \tag{2}$$

$$S(t,a,\alpha,\psi) = \frac{\alpha[1-G(t,\psi)]^a}{1-(1-\alpha)[1-G(t,\psi)]^a}, \qquad t \in R \tag{3}$$

$$h(t,a,\alpha,\psi) = \frac{a g(t,\psi)[G(t,\psi)]^{-1}}{1-(1-\alpha)[1-G(t,\psi)]^a}, \qquad t \in R \tag{4}$$

Hence forth a random variable T with pdf (2) is denoted by $T\sim$MOG-G(α,a,ψ), where α and a are two positive shape parameter.

2.1. Marshall-Olkin Generalized Exponential model

Consider T as a continous random variable follow an exponential distribution with scale parameter $\lambda > 0$, whose pdf and cdf is given by $g(t) = \frac{1}{\lambda}e^{-\frac{t}{\lambda}}$ and $G(t) = 1 - e^{-\frac{t}{\lambda}}$, $t > 0$. Then the pdf and cdf of MOG-E model are respectively given by

$$f(t) = \frac{\alpha a \frac{1}{\lambda} exp(-a\frac{t}{\lambda})}{\left[1-(1-\alpha)exp(-a\frac{t}{\lambda})\right]^2} \tag{5}$$

$$F(t) = \frac{1-exp(-a\frac{t}{\lambda})}{1-(1-\alpha)exp(-a\frac{t}{\lambda})} \tag{6}$$

The survival function corresponding to Equation (6) is given as

$$S(t) = \frac{\alpha\ exp(-a\frac{t}{\lambda})}{1-(1-\alpha)exp(-a\frac{t}{\lambda})} \tag{7}$$

Hazard function of the MOG-E model is written as

$$h(t) = \frac{a\frac{1}{\lambda}}{1-(1-\alpha)exp(-a\frac{t}{\lambda})} \tag{8}$$

In survival analysis, random generation of time variable from a survival model is done by putting $u = S(t)$, where U is a random variable follow $Uniform(0,1)$. So, the generation of time variable from MOG-E model is obtained by

$$t = \frac{\lambda}{a} log\left(\frac{\alpha}{u} + (1-\alpha)\right) \tag{9}$$

Following the [17], the joint likelihood function for right censored data is given as

$$L = \prod_{i=0}^{n} Pr(t_i, \delta_i) = \prod_{i=0}^{n} \{h(t_i)\}^{\delta_i} S(t_i) \tag{10}$$

here δ_i is an indicator variable

$$\delta_i = \begin{cases} 0, censored \\ 1, observed \end{cases}$$

The likelihood function for the MOG-E survival model is given by

$$L = \prod_{i=0}^{n} \left\{ \frac{a\frac{1}{\lambda}}{1-(1-\alpha)exp(-a\frac{t}{\lambda})} \right\}^{\delta_i} \times \frac{\alpha\ exp(-a\frac{t}{\lambda})}{1-(1-\alpha)exp(-a\frac{t}{\lambda})} \tag{11}$$

2.2. Marshall-Olkin Generalized Weibull model

Let $g(t)$ and $G(t)$ be the pdf and cdf of Weibull distribution with shape parameter $\gamma > 0$ and scale parameter $\lambda > 0$. Where, $g(t) = \frac{\gamma}{\lambda^\gamma} t^{\gamma-1} e^{-(\frac{t}{\lambda})^\gamma}$ and $G(t) = 1 - e^{-(\frac{t}{\lambda})^\gamma}$, $t > 0$. Then the pdf of MOG-W model is given by

$$f(t) = \frac{\alpha a \gamma \frac{1}{\lambda} t^{\gamma-1} exp(-a(\frac{t}{\lambda})^\gamma)}{\left[1-(1-\alpha)exp(-a(\frac{t}{\lambda})^\gamma)\right]^2} \tag{12}$$

Therefore, random variable T is denoted by $T \sim$MOG-W(α,a,γ,λ). The cdf of MOG-W model is written as

$$F(t) = \frac{1 - exp(-a(\frac{t}{\lambda})^\gamma)}{1-(1-\alpha)exp(-a(\frac{t}{\lambda})^\gamma)} \tag{13}$$

Survival function and hazard function of the MOG-W model are given respectively

$$S(t) = \frac{\alpha exp(-a(\frac{t}{\lambda})^\gamma)}{1-(1-\alpha)exp(-a(\frac{t}{\lambda})^\gamma)} \tag{14}$$

$$h(t) = \frac{a\gamma\frac{1}{\lambda}t^{\gamma-1}}{1-(1-\alpha)exp(-a(\frac{t}{\lambda})^\gamma)} \tag{15}$$

Random generation from the MOG-W model is done by the expression given below

$$t = \lambda \left[\frac{1}{a} log\left(\frac{\alpha}{u} + (1-\alpha)\right)\right]^{\frac{1}{\gamma}} \tag{16}$$

Using the Equation (10), the joint likelihood function for the MOG-W model based on right censored is written as

$$L = \prod_{i=0}^{n} \left\{ \frac{a\gamma\frac{1}{\lambda}t^{\gamma-1}}{1-(1-\alpha)exp(-a(\frac{t}{\lambda})^\gamma)} \right\}^{\delta_i} \times \frac{\alpha exp(-a(\frac{t}{\lambda})^\gamma)}{1-(1-\alpha)exp(-a(\frac{t}{\lambda})^\gamma)} \tag{17}$$

2.3. Marshall-Olkin Generalized Lomax model

Taking Lomax distribution with parameters $\gamma > 0$ and $\lambda > 0$ having pdf $g(t) = \frac{\gamma}{\lambda}\left(1+\frac{t}{\lambda}\right)^{-(\gamma+1)}$ and cdf $G(t) = 1-\left(1+\frac{t}{\lambda}\right)^{-\gamma}$, $t > 0$. Then the pdf and cdf of a random variable $T\sim$MOG-L(α,a,γ,λ) model are given respectively

$$f(t) = \frac{\alpha a \frac{\gamma}{\lambda}\left(1+\frac{t}{\lambda}\right)^{-1}\left(1+\frac{t}{\lambda}\right)^{-a\gamma}}{\left[1-(1-\alpha)\left(1+\frac{t}{\lambda}\right)^{-a\gamma}\right]^2} \tag{18}$$

$$F(t) = \frac{1-\left(1+\frac{t}{\lambda}\right)^{-a\gamma}}{1-(1-\alpha)\left(1+\frac{t}{\lambda}\right)^{-a\gamma}} \tag{19}$$

Survival function of the MOG-L is given by

$$S(t) = \frac{\alpha\left(1+\frac{t}{\lambda}\right)^{-a\gamma}}{1-(1-\alpha)\left(1+\frac{t}{\lambda}\right)^{-a\gamma}} \tag{20}$$

Hazard function of the MOG-L is written as

$$h(t) = \frac{\alpha a \frac{\gamma}{\lambda}\left(1+\frac{t}{\lambda}\right)^{-1}}{1-(1-\alpha)\left(1+\frac{t}{\lambda}\right)^{-a\gamma}} \tag{21}$$

Generation of survival time from the MOG-L model is given by

$$t = \lambda\left[\left(\frac{\alpha}{u}+(1-\alpha)\right)^{\frac{1}{a\gamma}}-1\right] \tag{22}$$

The joint likelihood function for the MOG-L model is written as

$$L = \prod_{i=0}^{n}\left\{\frac{\alpha a \frac{\gamma}{\lambda}\left(1+\frac{t}{\lambda}\right)^{-1}}{1-(1-\alpha)\left(1+\frac{t}{\lambda}\right)^{-a\gamma}}\right\}^{\delta_i} \times \frac{\alpha\left(1+\frac{t}{\lambda}\right)^{-a\gamma}}{1-(1-\alpha)\left(1+\frac{t}{\lambda}\right)^{-a\gamma}} \tag{23}$$

3. Kidney catheter data

This dataset, originally discussed in [18]. The study concerns with the recurrence times to infection, at the point where the catheter is inserted, for kidney patients using portable dialysis equipment. The data consist of times until the first and second recurrence of kidney infection in 38 patients. Each patient has exactly two observations. Each survival time is the time until infection since the insertion of the catheter. A Catheter may be removed for reasons other than infection, in which case the observation is censored. There are about 24% censored observations in the dataset. This data set has unmeasured or 'random' effect that is an identification code of patients, which accounts heterogeneity among the patients. This data set available in the package `survival` [19] of `R` [11].

Discription of kidney catheter data variables are given below:
time: time to infection in days
status: event status, 1=infection occurs or 0=censored
age: age in years
sex: 1=male, 2=female
disease: disease type(0=GN, 1=AN, 2=PKD, 3=Other)
id: identification code of the patients

3.1. Construction of data frame in R

Fitting of Bayesian models to the kidney catheter data with `stan` function requires data in a listed form, which we have created as below;

```
require(survival)
data(cancer, package="survival")
head(kidney)
y=kidney$time
x1=kidney$age
x2=kidney$sex
kidney$disease1=as.numeric(kidney$disease)
x3=kidney$GN=as.numeric(kidney$disease1==2)
x4=kidney$AN=as.numeric(kidney$disease1==3)
x5=kidney$PKD=as.numeric(kidney$disease1==4)
x=cbind(1,x1,x2,x3,x4,x5)
N=nrow(x)
M=ncol(x)
J=38
event=kidney$status
Id=as.integer(kidney$id)##identity of subject
datk=list(y=y,x=x,N=N,M=M,event=event,J=J,Id=Id)
```

4. Bayesian Analysis of MOG-G family

4.1. Prior Specification

For the construction of the Bayesian regression model, we need to specify a prior distribution to the parameters of the model. We have chosen half-Cauchy prior for shape and scale parameters and regularizing prior for regression coefficient.

4.1.1 Half-Cauchy prior distribution

The probability density function of half-Cauchy distribution with scale γ is given by

$$f(x) = \frac{2\gamma}{\pi(x^2+\gamma^2)}, \qquad x > 0, \gamma > 0$$

The mean and variance of half-cauchy distribution does not exist, but its mode is equal to zero. The half-cauchy distribution with scale γ=25 is nearly flat prior but not completely, the prior distribution that are not completely flat provides enough information for the numerical approximation algorithm to continue to explore the target density, the posterior distribution [20],[21]. [22] support the use half cauchy prior for scale parameter because of its excellent frequentist risk properties, and its sensible behaviour in the presence of sparsity compared to the usual conjugate aternative. [20] have also discussed the points in support of half cauchy prior.

4.1.2 Gaussian prior distribution

The probability density function of Gaussian distribution with mean μ and variance σ^2 is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x-\mu)^2}{2\sigma^2}), \qquad -\infty < x < \infty, \sigma > 0, \mu > 0$$

In this paper, we have chosen Gaussian prior with mean 0, and standard deviation 5 for β coefficient as a regularizing prior because this prior prevent a model from getting too excited by the data that slows the rate of over excitement of model and reduce the overfitting of data to the model [23].

4.2. Model Specification

Following the [24] to build a regression model, we have introduced covariates including random intercept through the log link function i.e.

$$log(\lambda_i) = \beta_1 + w_{[subj_i]} + \beta_2 x_{i1} + \beta_3 x_{i2} + \beta_4 x_{i3} + \beta_5 x_{i4} + \beta_6 x_{i5}$$

$$\lambda_i = exp(\beta_1 + w_{[subj_i]} + \beta_2 x_{i1} + \beta_3 x_{i2} + \beta_4 x_{i3} + \beta_5 x_{i4} + \beta_6 x_{i5})$$

or,

$$\lambda_i = exp(w_{[subj_i]} + x_i\beta)$$

where, $w_{[subj_i]}$ is the variability accounted by subject or patients called as the random intercept, $w \sim N(0, \sigma_w)$, and $\beta \sim N(0, \sigma = 5)$

4.2.1 Posterior density of MOG-E

By using bayes theorem, the joint posterior distribution is given as

$$P(a, \alpha, \beta|X, t) \propto L(t|a, \alpha, \beta, X) \times P(a) \times P(\alpha) \times P(\beta) \tag{24}$$

$$P(a,\alpha,\beta|X,t) \propto \prod_{i=0}^{n} \left\{ \frac{a \frac{1}{exp(w_{[subj_i]}+x_i\beta)}}{1-(1-\alpha)exp(-a\frac{t}{exp(w_{[subj_i]}+x_i\beta)})} \right\}^{\delta_i} \times \frac{\alpha\ exp(-a\frac{t}{exp(w_{[subj_i]}+x_i\beta)})}{1-(1-\alpha)exp(-a\frac{t}{exp(w_{[subj_i]}+x_i\beta)})} \tag{25}$$
$$\times \frac{2\times 25}{\pi(a^2+25^2)} \times \frac{2\times 25}{\pi(\alpha^2+25^2)} \times \frac{1}{\sigma_w\sqrt{2\pi}} exp\left(-\frac{w_i^2}{2\sigma_w^2}\right) \times \prod_{j=0}^{J} \frac{1}{5\sqrt{2\pi}} exp\left(-\frac{1}{2\times 25}\beta_j^2\right)$$

4.2.2 Posterior density of MOG-W

By using bayes theorem, the joint posterior distribution is given as

$$P(a, \alpha, \gamma, \beta|X, t) \propto L(t|a, \alpha, \gamma, \beta, X) \times P(a) \times P(\alpha) \times P(\gamma) \times P(\beta) \tag{26}$$

$$P(a,\alpha,\gamma,\beta|X,t) \propto \prod_{i=0}^{n} \left\{ \frac{a\gamma \frac{1}{exp(w_{[subj_i]}+x_i\beta)} t^{\gamma-1}}{1-(1-\alpha)exp(-a(\frac{t}{exp(w_{[subj_i]}+x_i\beta)})^{\gamma})} \right\}^{\delta_i} \tag{27}$$
$$\times \frac{\alpha exp(-a(\frac{t}{exp(w_{[subj_i]}+x_i\beta)})^{\gamma})}{1-(1-\alpha)exp(-a(\frac{t}{exp(w_{[subj_i]}+x_i\beta)})^{\gamma})} \times \frac{2\times 25}{\pi(a^2+25^2)} \times \frac{2\times 25}{\pi(\alpha^2+25^2)}$$
$$\times \frac{2\times 25}{\pi(\gamma^2+25^2)} \times \frac{1}{\sigma_w\sqrt{2\pi}} exp\left(-\frac{w_i^2}{2\sigma_w^2}\right) \times \prod_{j=0}^{J} \frac{1}{5\sqrt{2\pi}} exp\left(-\frac{1}{2\times 25}\beta_j^2\right)$$

4.2.3 Posterior density of MOG-L

By using bayes theorem, the joint posterior distribution is given as

$$P(a, \alpha, \gamma, \beta|X, t) \propto L(t|a, \alpha, \gamma, \beta, X) \times P(a) \times P(\alpha) \times P(\gamma) \times P(\beta) \tag{28}$$

$$P(a,\alpha,\gamma,\beta|X,t) \propto \prod_{i=0}^{n} \left\{ \frac{\alpha a \frac{\gamma}{exp(w_{[subj_i]}+x_i\beta)} \left(1+\frac{t}{exp(w_{[subj_i]}+x_i\beta)}\right)^{-1}}{1-(1-\alpha)\left(1+\frac{t}{exp(w_{[subj_i]}+x_i\beta)}\right)^{-a\gamma}} \right\}^{\delta_i} \tag{29}$$

$$\times \frac{\alpha\left(1+\frac{t}{exp(w_{[subj_i]}+x_i\beta)}\right)^{-a\gamma}}{1-(1-\alpha)\left(1+\frac{t}{exp(w_{[subj_i]}+x_i\beta)}\right)^{-a\gamma}} \times \frac{2\times 25}{\pi(a^2+25^2)} \times \frac{2\times 25}{\pi(\alpha^2+25^2)}$$

$$\times \frac{2\times 25}{\pi(\gamma^2+25^2)} \times \frac{1}{\sigma_w\sqrt{2\pi}} exp\left(-\frac{w_i^2}{2\sigma_w^2}\right) \times \prod_{j=0}^{J} \frac{1}{5\sqrt{2\pi}} exp\left(-\frac{1}{2\times 25}\beta_j^2\right)$$

4.3. Implementation using Stan

Bayesian modeling of MOG-G family in STAN language includes the creation of blocks: functions block, data block, transformed data block, parameters block, transformed parameters block, model block, and generated quantities block. To run STAN code in R requires package rstan that is an interface of R and STAN.

4.3.1 Stan code for MOG-E model

```
modelMOGE="functions{
vector log_moegs(vector t, real a, real alpha, vector lambda){
vector[num_elements(t)]log_moegs;
for(i in 1:num_elements(t)){
log_moegs[i]=log(alpha)-a*t[i]/lambda[i]-log(1-(1-alpha)*exp(-a*t[i]/lambda[i]));
}
return log_moegs;
}
vector log_moegh(vector t, real a, real alpha, vector lambda){
vector[num_elements(t)]log_moegh;
for(i in 1:num_elements(t)){
log_moegh[i]=log(a)-log(lambda[i])-log(1-(1-alpha)*exp(-a*t[i]/lambda[i]));
}
return log_moegh;
}
real surv_MOEG_lpdf(vector t, vector d, real a, real alpha, vector lambda){
vector[num_elements(t)] llikmoeg;
real prob;
llikmoeg=d .* log_moegh(t,a,alpha,lambda)+log_moegs(t,a,alpha,lambda);
prob=sum(llikmoeg);
return prob;
}}
data{
int N;
vector<lower=0>[N] y;
vector<lower=0,upper=1>[N] event;
int M;
matrix[N,M] x;
int<lower=1>J;
int<lower=1,upper=J>Id[N];
}
parameters{
```

```
real<lower=0>a;
vector[M] beta;
real<lower=0> alpha;
vector[J] w;
real<lower=0>sigma_w;
}
transformed parameters{
vector[N] linpred;
vector<lower=0>[N] lambda;
linpred=x*beta;
for(i in 1:N){
lambda[i]=exp(w[Id[i]]+linpred[i]);
}}
model{
target+=cauchy_lpdf(alpha|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=cauchy_lpdf(a|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=normal_lpdf(beta|0,5);
target+=normal_lpdf(w|0,sigma_w);
target+=cauchy_lpdf(sigma_w|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=surv_MOEG_lpdf(y|event,a,alpha,lambda);
}
generated quantities{
vector[N] log_lik;
vector[N] yrepmoeg;
real dev;
dev=0;
for(n in 1:N) log_lik[n]=event[n]*(log(a)-log(lambda[n])-log(1-(1-alpha)*exp(-a*y[n]/
lambda[n])))+log(alpha)-a*y[n]/lambda[n]-log(1-(1-alpha)*exp(-a*y[n]/lambda[n]));
{real u;
u=uniform_rng(0,1);
for(n in 1:N) yrepmoeg[n]=(lambda[n]/a)*log(alpha/u+(1-alpha));
}
dev=dev+(-2)*surv_MOEG_lpdf(y|event,a,alpha,lambda);
}"
```

4.3.2 Stan code for MOG-W model

```
modelMOGW="functions{
vector log_mogws(vector t, real a, real alpha,real gamma, vector lambda){
vector[num_elements(t)]log_mogws;
for(i in 1:num_elements(t)){
log_mogws[i]=log(alpha)-a*(t[i]/lambda[i])^(gamma)-log(1-(1-alpha)
*exp(-a*(t[i]/lambda[i])^(gamma)));
}
return log_mogws;
}
vector log_mogwh(vector t, real a, real alpha, real gamma, vector lambda){
vector[num_elements(t)]log_mogwh;
for(i in 1:num_elements(t)){
log_mogwh[i]=log(a)+log(gamma)-gamma*log(lambda[i])+(gamma-1)*log(t[i])
-log(1-(1-alpha)*exp(-a*(t[i]/lambda[i])^(gamma)));
}
return log_mogwh;
}
```

```
real surv_MOGW_lpdf(vector t, vector d, real a, real alpha,real gamma,vector lambda){
vector[num_elements(t)] llikmogw;
real prob;
llikmogw=d .* log_mogwh(t,a,alpha,gamma,lambda)+log_mogws(t,a,alpha,gamma,lambda);
prob=sum(llikmogw);
return prob;
}}
data{
int N;
vector<lower=0>[N] y;
vector<lower=0,upper=1>[N] event;
int M;
matrix[N,M] x;
int<lower=1>J;
int<lower=1,upper=J>Id[N];
}
parameters{
real<lower=0>a;
vector[M] beta;
real<lower=0> alpha;
real<lower=0> gamma;
vector[J] w;
real<lower=0>sigma_w;
}
transformed parameters{
vector[N] linpred;
vector<lower=0>[N] lambda;
linpred=x*beta;
for(i in 1:N){
lambda[i]=exp(w[Id[i]]+linpred[i]);
}}
model{
target+=cauchy_lpdf(alpha|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=cauchy_lpdf(a|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=cauchy_lpdf(gamma|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=normal_lpdf(beta|0,5);
target+=normal_lpdf(w|0,sigma_w);
target+=cauchy_lpdf(sigma_w|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=surv_MOGW_lpdf(y|event,a,alpha,gamma,lambda);
}
generated quantities{
vector[N] log_lik;
vector[N] yrepmogw;
real dev;
dev=0;
for(n in 1:N) log_lik[n]=event[n]*(log(a)+log(gamma)-gamma*log(lambda[n])+(gamma-1)*
log(y[n])-log(1-(1-alpha)*exp(-a*(y[n]/lambda[n])^(gamma))))+log(alpha)
-a*(y[n]/lambda[n])^(gamma)-log(1-(1-alpha)*exp(-a*(y[n]/lambda[n])^(gamma)));
{real u;
u=uniform_rng(0,1);
for(n in 1:N) yrepmogw[n]=lambda[n]*((1/a)*log(alpha/u+(1-alpha)))^(1/gamma);
}
dev=dev+(-2)*surv_MOGW_lpdf(y|event,a,alpha,gamma,lambda);
```

```
}"
```

4.3.3 Stan code for MOG-L model

```
modelMOGL="functions{
vector log_mogls(vector t, real a, real alpha,real gamma, vector lambda){
vector[num_elements(t)]log_mogls;
for(i in 1:num_elements(t)){
log_mogls[i]=log(alpha)-a*gamma*log(1+t[i]/lambda[i])-log(1-(1-alpha)
*(1+t[i]/lambda[i])^(-a*gamma));
}
return log_mogls;
}
vector log_moglh(vector t, real a, real alpha, real gamma, vector lambda){
vector[num_elements(t)]log_moglh;
for(i in 1:num_elements(t)){
log_moglh[i]=log(a)+log(gamma)-log(lambda[i])-log(1+t[i]/lambda[i])-
log(1-(1-alpha)*(1+t[i]/lambda[i])^(-a*gamma));
}
return log_moglh;
}
real surv_MOGL_lpdf(vector t, vector d, real a, real alpha,real gamma,vector lambda){
vector[num_elements(t)] llikmogl;
real prob;
llikmogl=d .* log_moglh(t,a,alpha,gamma,lambda)+log_mogls(t,a,alpha,gamma,lambda);
prob=sum(llikmogl);
return prob;
}}
data{
int N;
vector<lower=0>[N] y;
vector<lower=0,upper=1>[N] event;
int M;
matrix[N,M] x;
int<lower=1>J;
int<lower=1,upper=J>Id[N];
}
parameters{
real<lower=0>a;
vector[M] beta;
real<lower=0> alpha;
real<lower=0> gamma;
vector[J] w;
real<lower=0>sigma_w;
}
transformed parameters{
vector[N] linpred;
vector<lower=0>[N] lambda;
linpred=x*beta;
for(i in 1:N){
lambda[i]=exp(w[Id[i]]+linpred[i]);
}}
model{
target+=cauchy_lpdf(alpha|0,25)- 1 * cauchy_lccdf(0|0,25);
```

```
target+=cauchy_lpdf(a|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=cauchy_lpdf(gamma|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=normal_lpdf(beta|0,5);
target+=normal_lpdf(w|0,sigma_w);
target+=cauchy_lpdf(sigma_w|0,25)- 1 * cauchy_lccdf(0|0,25);
target+=surv_MOGL_lpdf(y|event,a,alpha,gamma,lambda);
}
generated quantities{
vector[N] log_lik;
vector[N] yrepmogl;
real dev;
dev=0;
for(n in 1:N) log_lik[n]=event[n]*(log(a)+log(gamma)-log(lambda[n])-log(1+y[n]/lambda[n])
-log(1-(1-alpha)*(1+y[n]/lambda[n])^(-a*gamma)))+log(alpha)-a*gamma*log(1+y[n]/lambda[n])
-log(1-(1-alpha)*(1+y[n]/lambda[n])^(-a*gamma));
{real u;
u=uniform_rng(0,1);
for(n in 1:N) yrepmogl[n]=lambda[n]*((alpha/u+(1-alpha))^(1/(a*gamma))-1);
}
dev=dev+(-2)*surv_MOGL_lpdf(y|event,a,alpha,gamma,lambda);
}"
```

4.4. Fitting with Stan

To fit the survival models based on MOG-G family, the function `stan` is used, and list `datk` of data pass into the function `stan`. `STAN` used C^{++} compiler to samples the posterior distrbution of the model parameters, including random intercepts w_j for each patient J. To get summary of result, the function `print` is used.

4.4.1 Fitting of MOG-E model

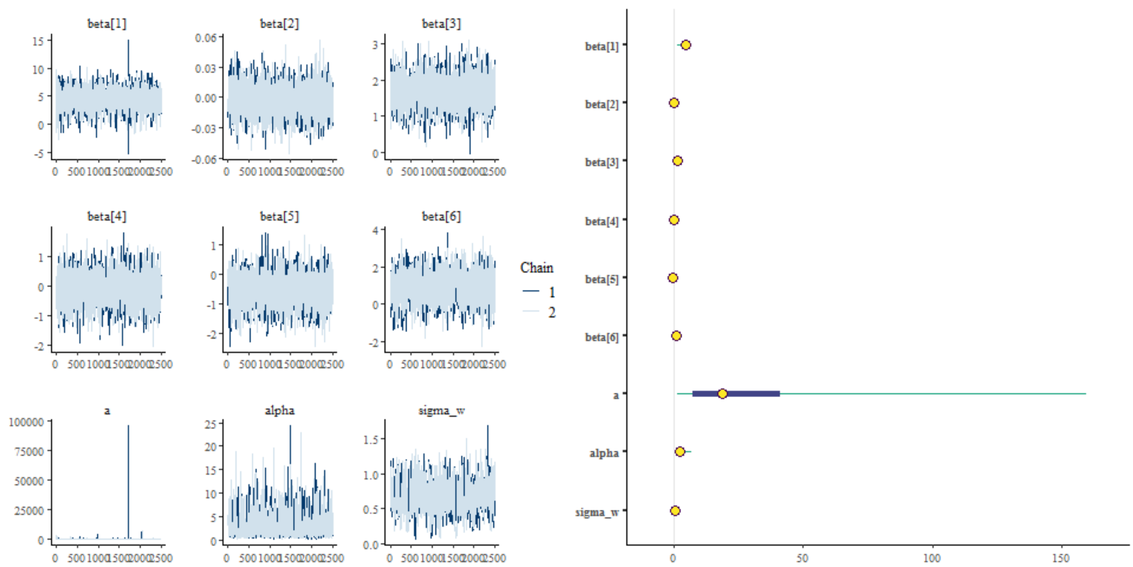
```
MOGE=stan(model_code = modelMOGE,data=datk,iter=5000,chains = 2)
print(MOGE)
```

Summarizing Output: After fitting of MOG-E survival model to the kidney data set, we get the results in tabular form are given in Table 1. It contains posterior estimates, standard deviation, credible interval, n_eff(crude estimate of effective sample size), and Rhat called as potential scale reduction factor [16], which estimate the convergence of Markov chain to the target distribution. Besides $\hat{R}$, Traceplot also shows the convergence of the Markov chain. According to [16] the acceptable limit of n_eff is >100 and $\hat{R}$ values lower than 1.1. Rhat for all parameters of the MOG-E model is close to 1, which means Markov chains converge to the target distribution, the Monte Carlo error is acceptable, and the effective sample size is reasonable. Here, we can see that the posterior estimate of parameters β_1 (Intercept) is 4.440, and β_3 (Sex) is 1.678 are statistically significant as 95% credible interval (CI) does not contains 0 respectively. The positive value of β_3 inferred that the male patients have more chance to get infected at the place where a catheter is inserted than the female patients. The posterior estimate of parameters β_2 (Age) is -0.002, β_4 (AN) is -0.116, β_5 (GN) is -0.544 and β_6 (PKD) is 0.906 are not statistically significant as corresponding CI includes 0.

Table 1: *Posterior summary of MOG-E model parameters*

parametrs	mean	se_mean	sd	2.5%	50%	97.5%	n_eff	Rhat
beta[1]	4.440	0.033	1.728	0.833	4.468	7.612	2699	1.001
beta[2]	-0.002	0.000	0.013	-0.029	-0.002	0.025	2584	1.000
beta[3]	1.678	0.007	0.387	0.917	1.678	2.427	3309	1.000
beta[4]	-0.116	0.009	0.483	-1.085	-0.112	0.836	2678	1.000
beta[5]	-0.544	0.009	0.484	-1.533	-0.533	0.397	2777	1.000
beta[6]	0.906	0.014	0.717	-0.493	0.910	2.317	2456	1.000
a	129.344	78.878	2544.490	0.671	18.762	275.779	1041	1.002
alpha	2.825	0.059	2.186	0.534	2.225	8.820	1397	1.001
sigma_w	0.651	0.012	0.221	0.195	0.660	1.091	367	1.003

Graphical Analysis



(a) **(b)**

Figure 1: *(a) Traceplot of MOG-E model parameters, two chains were run depicted in different color and mixing of two chains is good means Markov chains converge to the target distribution, and (b) Caterpillar plot of the MOG-E model parameters.*

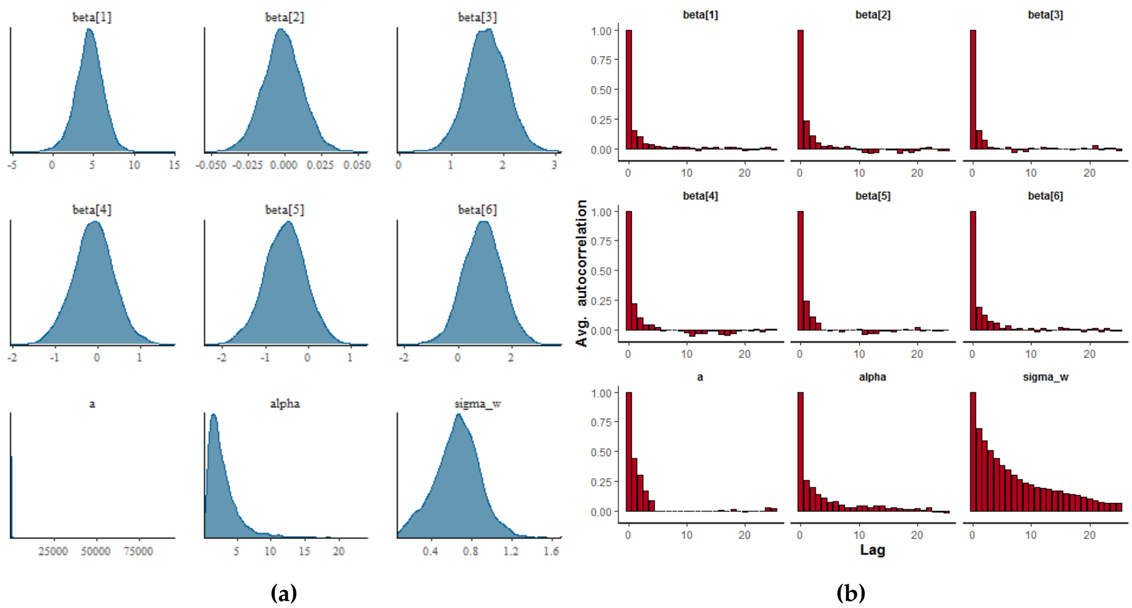


Figure 2: *(a) Posterior density plot MOG-E model parameters, (b) Autocorrelation plot of MOG-E model parameters, after 20 lag autocorrelation declining towards zero.*

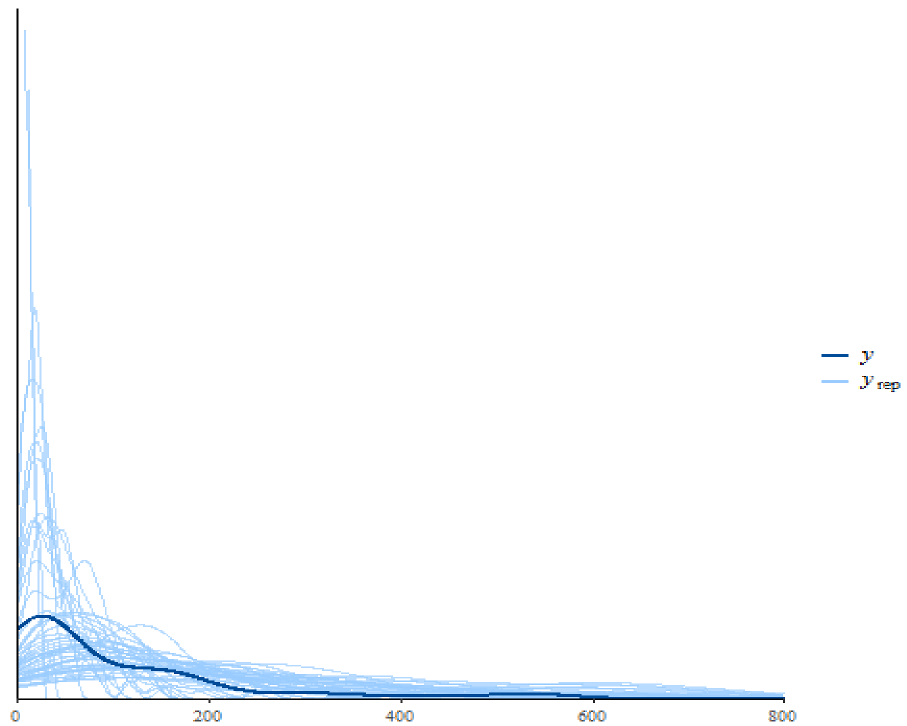


Figure 3: *The posterior predictive density (PPD) plot of the MOG-E model is done by plotting the data y and then overlaying the density of the predicted values y_{rep}, which are generated from the posterior predictive distribution of the given model. PPD plot of the MOG-E model shows that the posterior predictive density fits the data well.*

4.4.2 Fitting of MOG-W model

```
MOGW=stan(model_code = modelMOGW,data=datk,iter=5000,chains = 2)
```

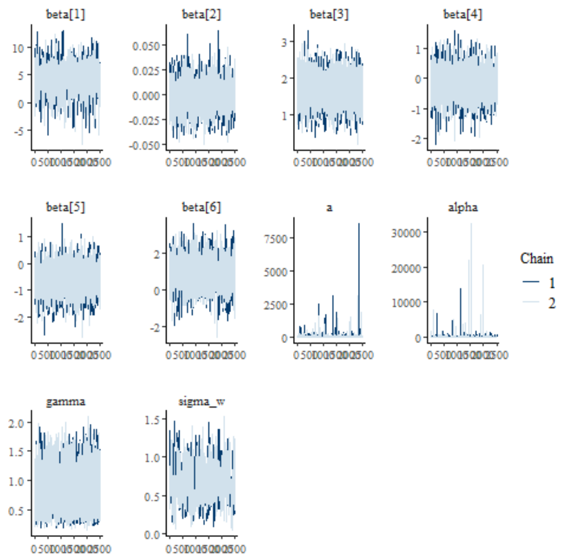
```
print(MOGW)
```

Summarizing Output: It is an evident from Table 2 that the Rhat of the MOG-W model parameters are close to 1, which indicates Markov chain converges to the target distribution and effective sample size is enough to get conversion.

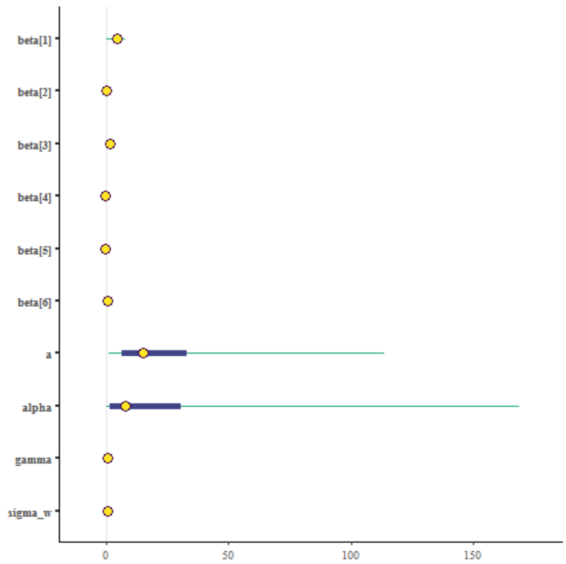
Table 2: *Posterior summary of MOG-W model parameters*

parametrs	mean	se_mean	sd	2.5%	50%	97.5%	n_eff	Rhat
beta[1]	4.289	0.053	2.252	-0.936	4.468	8.381	1807	1.000
beta[2]	-0.001	0.000	0.014	-0.028	-0.001	0.027	1892	1.001
beta[3]	1.688	0.007	0.391	0.930	1.686	2.451	2978	1.000
beta[4]	-0.164	0.010	0.479	-1.130	-0.155	0.744	2102	1.002
beta[5]	-0.586	0.011	0.489	-1.582	-0.577	0.350	1935	1.001
beta[6]	0.841	0.020	0.742	-0.625	0.858	2.323	1383	1.002
a	38.413	2.699	169.093	0.786	15.061	185.706	3926	1.000
alpha	72.737	12.350	726.957	0.109	7.801	335.523	3465	1.001
gamma	0.800	0.010	0.357	0.289	0.744	1.562	1332	1.000
sigma_w	0.622	0.013	0.227	0.175	0.630	1.063	288	1.002

Graphical Analysis



(a)



(b)

Figure 4: *(a) Traceplot of MOG-W model parameters, two chains were run depicted in different color and (b) caterpillar plot of the MOG-W model parameters.*

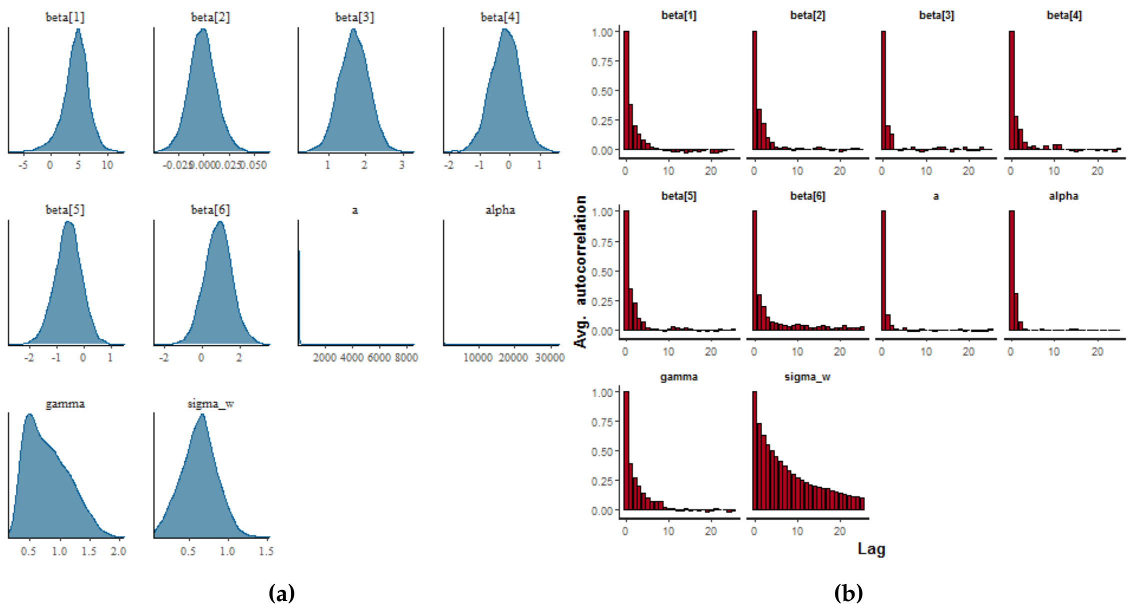


Figure 5: *(a) Posterior density plot MOG-W model parameters, (b) Autocorrelation plot of MOG-W model parameters, after 20 lag autocorrelation declining towards zero.*

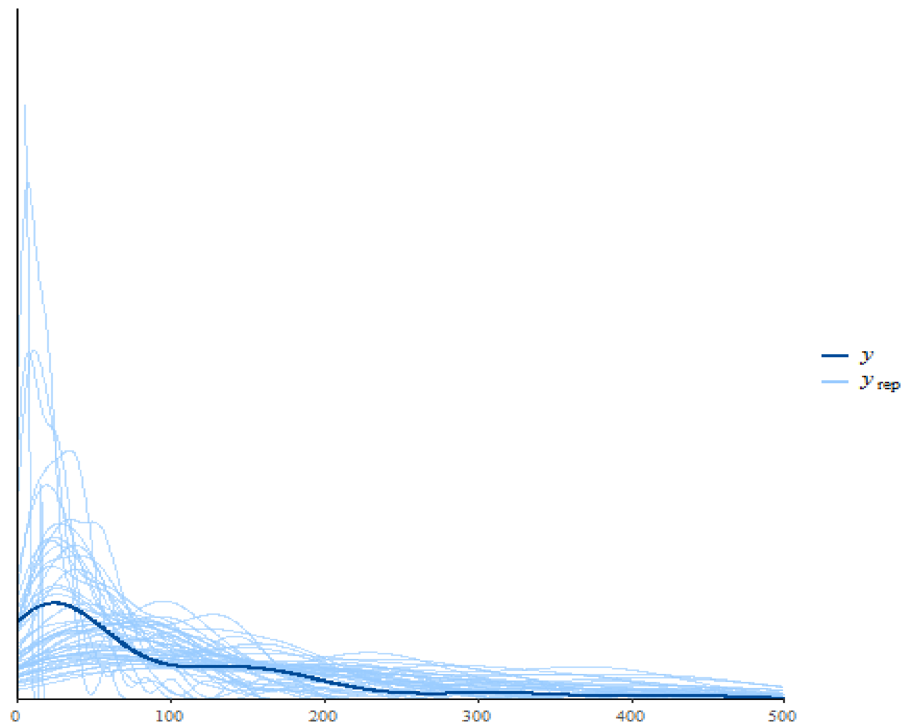


Figure 6: *PPD plot of the MOG-W model shows that the posterior predictive density fits the data well and good compatibility of the model to the data.*

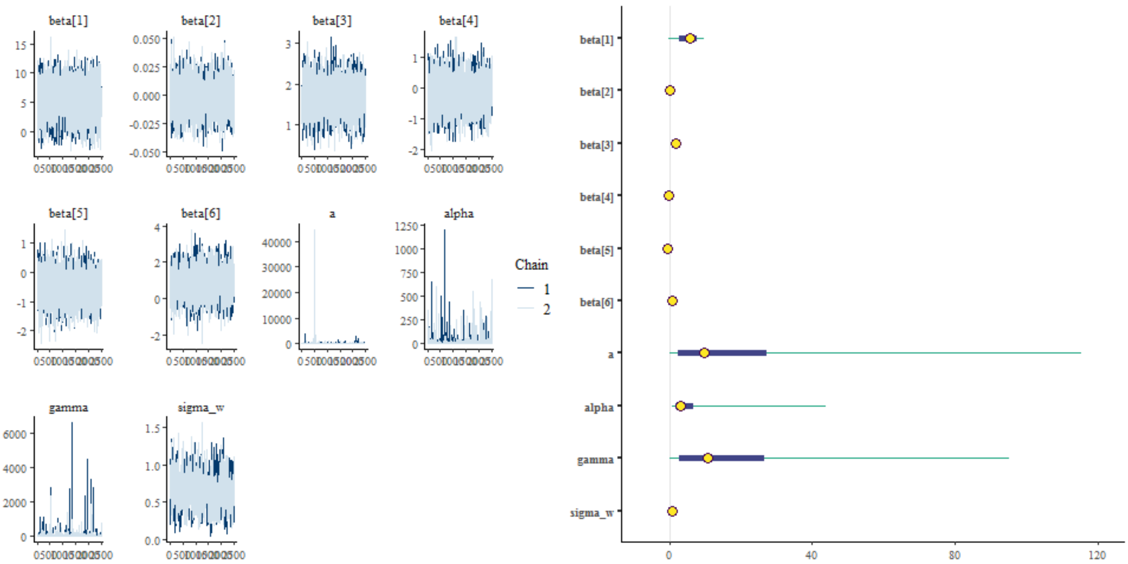
4.4.3 Fitting of MOG-L model

```
MOGW=stan(model_code = modelMOGW,data=datk,iter=5000,chains = 2)
print(MOGW)
```

Table 3: *Posterior summary of MOG-L model parameters*

parametrs	mean	se_mean	sd	2.5%	50%	97.5%	n_eff	Rhat
beta[1]	5.180	0.115	3.164	-0.916	5.636	10.534	753	1.012
beta[2]	-0.002	0.000	0.013	-0.027	-0.002	0.025	2500	1.000
beta[3]	1.693	0.007	0.377	0.944	1.700	2.433	3111	1.000
beta[4]	-0.153	0.010	0.494	-1.146	-0.161	0.839	2577	1.000
beta[5]	-0.562	0.009	0.470	-1.478	-0.567	0.394	2478	1.000
beta[6]	0.869	0.014	0.713	-0.530	0.867	2.274	2501	1.000
a	41.617	9.152	641.724	0.119	9.765	182.772	4916	1.000
alpha	11.755	1.010	41.049	0.639	3.193	82.591	1653	1.004
gamma	32.510	2.419	163.682	0.129	10.641	152.084	4580	1.000
sigma_w	0.611	0.014	0.226	0.201	0.612	1.058	253	1.013

Graphical Analysis



(a) **(b)**

Figure 7: *(a) Traceplot of the MOG-L model parameters, two chains were run depicted in different color and (b) caterpillar plot of the MOG-L model parameters.*

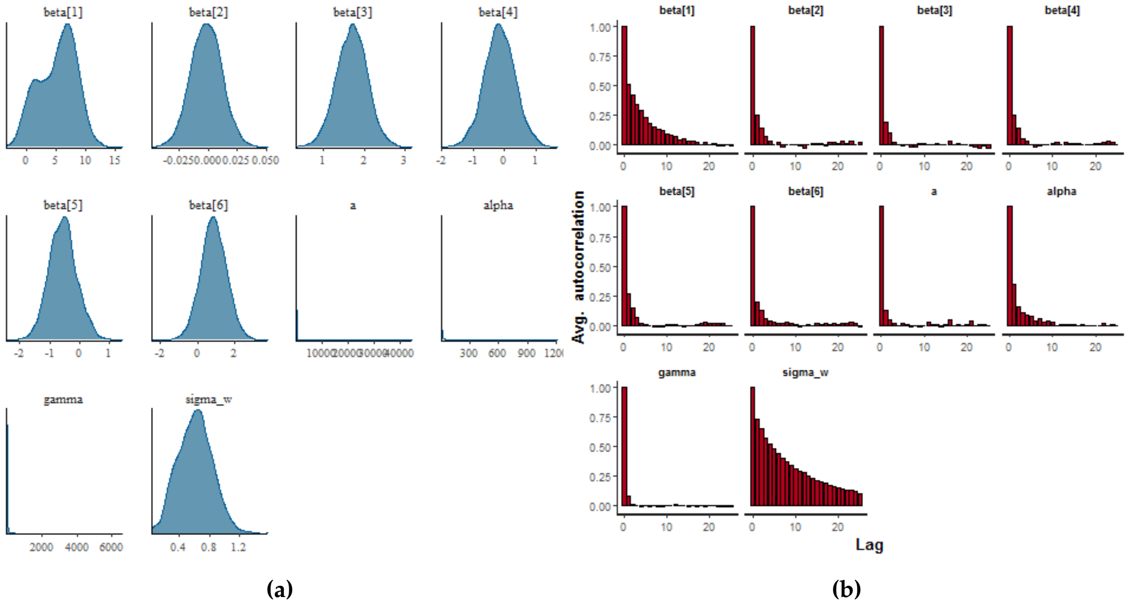


Figure 8: *(a) Posterior density plot the MOG-L model parameters, (b) Autocorrelation plot of the MOG-L model parameters, after 20 lag autocorrelation declining towards zero.*

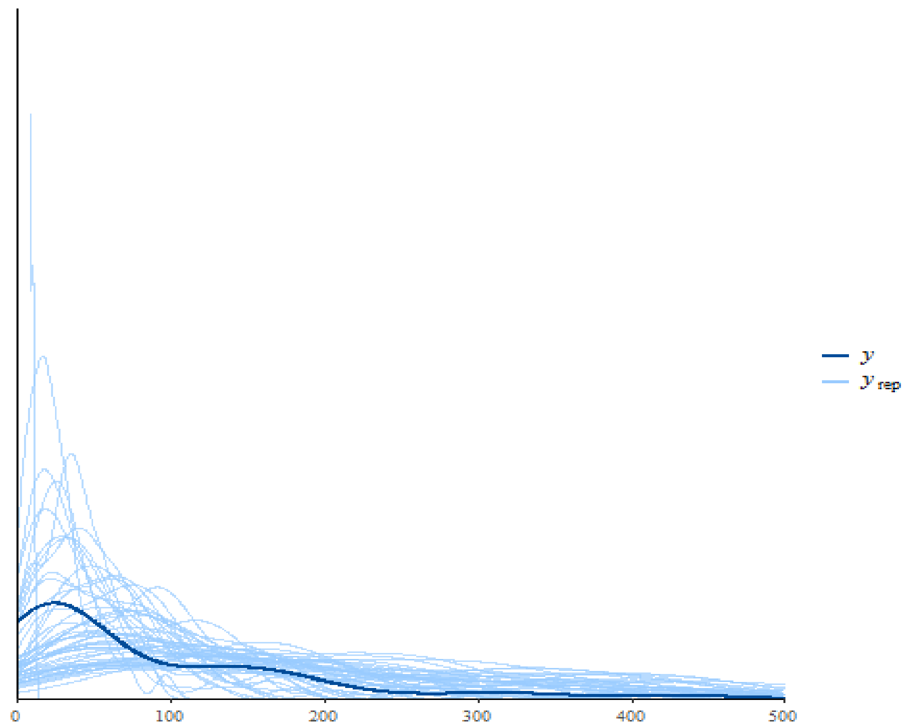


Figure 9: *PPD plot of the MOG-L model shows that the posterior predictive density fits the data well, and the model is compatible with the given data.*

4.5. Bayesian model Comparison

In order to compare the fitted models, we consider the model selection criteria like Watanabe Akaike information criteria [25] and leave one out cross-validation information criteria (LOOIC). However, the lower value of these selection methods indicates a better model fit. The WAIC and

LOOIC are two fully Bayesian selection methods than the others information criteria. They are methods for estimating pointwise out of sample prediction accuracy from a fitted Bayesian model using the log-likelihood evaluated at the posterior simulations of the parameters [16]. Although WAIC is asymptotically equal to LOOIC, it is preferable to use LOOIC because of its robustness in finite cases with weak priors or influential observation.

Table 4: *WAIC and LOOIC values for all models.*

Model	No of parameters	WAIC	LOOIC
MOG-E	3	667.9	673.3
MOG-W	4	671.1	675.8
MOG-L	4	670.5	674.0

From Table 4, it is evident that the value of WAIC and LOOIC of the MOG-E model is less than the MOG-W and MOG-L, which means the MOG-E model is a better survival model as compared to other models for Kidney catheter data.

5. Discussion and Conclusion

In this paper, the MOG-G family are fitted to the real survival data includes random effect in Bayesian setup, which is implemented by `STAN` language using package `rstan` of `R`. For all models, the Markov chains converges to the target distribution, and covariate Sex is significant. The Posterior predictive check has been computed using the posterior predictive density plot for the MOG-E, MOG-W, and MOG-L models. We have seen in the PPD plot, the data y and replicated data set y^{rep} are showing the same behavior and looks similar which means the replicated data sets are coming from the same model as the given data set, and all are the adequate model for predicting the future value. Upon comparison with the results obtained through LOOIC and WAIC, it can be concluded that the MOG-E model fits the data better than MOG-W and MOG-L.

References

[1] Yousof, H. M., Afify, A. Z., Nadarajah, S., Hamedani, G., and Aryal, G. R. (2018). The marshall-olkin generalized-g family of distributions with applications. *Statistica*, 78(3):273-295.

[2] 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(3):641-652.

[3] Alizadeh, M., Tahir, M., Cordeiro, G. M., Mansoor, M., Zubair, M., and Hamedani, G. (2015). The kumaraswamy marshal-olkin family of distributions. *Journal of the Egyptian Mathematical Society*, 23(3):546-557.

[4] Alizadeh, M., Cordeiro, G. M., Brito, E. d., and B Demeétrio, C. G. (2015). The beta marshall-olkin family of distributions. *Journal of Statistical Distributions and Applications*, 2(1):1-18.

[5] Handique, L. and Chakraborty, S. (2016). The beta generalized marshall-olkin-g family of distributions. arXiv preprint arXiv:1608.05985.

[6] B Dias, C. R., Cordeiro, G. M., Alizadeh, M., Diniz Marinho, P. R., and Campos Coelho, H. F. (2016). Exponentiated marshall-olkin family of distributions. *Journal of Statistical Distributions and Applications*, 3(1):1-21.

[7] Chakraborty, S. and Handique, L. (2017). The generalized marshall-olkin-kumaraswamy-g family of distributions. *Journal of data Science*, 15(3):391-422.

[8] Handique, L. and Chakraborty, S. (2017). The beta generalized marshall-olkin kumaraswamy-g family of distributions with applications. *Int. J. Agricult. Stat. Sci*, 13(2):721-733.

[9] Handique, L., Chakraborty, S., and de Andrade, T. A. (2019). The exponentiated generalized marshallolkin family of distribution: its properties and applications. *Annals of Data Science*, 6(3):391-411.

[10] Korkmaz, M. C., Cordeiro, G. M., Yousof, H. M., Pescim, R. R., Afify, A. Z., and Nadarajah, S. (2019). The weibull marshallolkin family: Regression model and application to censored data. *Communications in Statistics-Theory and Methods*, 48(16):4171-4194.

[11] R Core Team (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[12] Team, S. D. et al. (2017). Stan modeling language users guide and reference manual. *Technical report*.

[13] Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D. (1987). Hybrid monte carlo. *Physics letters B*, 195(2):216-222.

[14] Neal, R. M. et al. (2011). Mcmc using hamiltonian dynamics. Handbook of markov chain monte carlo, 2(11):2.

[15] Hoffman, M. D. and Gelman, A. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593-1623.

[16] Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis*. Chapman and Hall/CRC.

[17] Collett, D. (2015). *Modelling survival data in medical research*. CRC press.

[18] McGilchrist, C. and Aisbett, C. (1991). Regression with frailty in survival analysis. *Biometrics*, 461-466.

[19] Therneau, T. M. and Lumley, T. (2014). Package 'survival'. Survival analysis Published on CRAN, 2:3.

[20] Khan, N. and Khan, A. A. (2018). Bayesian analysis of topp-leone generalized exponential distribution. *Austrian Journal of Statistics*, 47(4):1-15.

[21] Akhtar, M. T. and Khan, A. A. (2014). Bayesian analysis of generalized log-burr family with R. *SpringerPlus*, 3(1):185.

[22] Gelman, A. et al. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian analysis*, 1(3):515-534.

[23] McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

[24] Lawless, J. F. (2011). *Statistical models and methods for lifetime data*. John Wiley and Sons.

[25] Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11(12):3571:3594.

RT&A, No 2 (68)
Volume 17, June 2022

On the Use of Entropy as a Measure of Dependence of Two Events. Part 2

Valentin Vankov Iliev

•

Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia, Bulgaria
viliev@math.bas.bg

Abstract

The joint experiment $\mathfrak{J}_{(A,B)}$ of two binary trials $A \cup A^c$ and $B \cup B^c$ in a probability space can be produced not only by the ordered pair (A, B) but by a set consisting, in general, of 24 ordered pairs of events (named Yule's pairs). The probabilities $\xi_1, \xi_2, \xi_3, \xi_4$ of the four results of $\mathfrak{J}_{(A,B)}$ are linear functions in three variables $\alpha = \Pr(A)$, $\beta = \Pr(B)$, $\theta = \Pr(A \cap B)$, and constitute a probability distribution. The symmetric group S_4 of degree four has an exact representation in the affine group $\mathrm{Aff}(3, \mathbb{R})$, which is constructed by using the types of the form $[\alpha, \beta, \theta]$ of those 24 Yule's pairs. The corresponding action of S_4 permutes the components of the probability distribution $(\xi_1, \xi_2, \xi_3, \xi_4)$, and, in particular, its entropy function is S_4-invariant. The function of degree of dependence of two events, defined in the first part of this paper via modifying the entropy function, turns out to be a relative invariant of the dihedral group of order 8.

Keywords: probability space; experiment in a sample space; probability distribution; entropy; degree of dependence; relative invariant.

1. Introduction

The initial idea of this work was to describe all symmetries of the sequence of Yule's pairs from (1) which produce one and the same experiment [3, 4.1,(1)]. If we consider the equivalence classes of the form $[(\alpha, \beta, \theta)]$ that contain the members of (1), then the naturally constructed in terms of coordinate functions α, β, θ affine automorphisms of the linear space $\mathbb{R}^3$ form a group which is isomorphic to the symmetric group S_4, see Section 2, Theorem 1. The components $\xi_1, \xi_2, \xi_3, \xi_4$ of the probability distribution [3, 4.1,(2)] are linear functions in α, β, θ. The group S_4 naturally acts via above isomorphism and permutes ξ_i's. As a consequence we obtain Theorem 2 which asserts that the entropy function $E_{\alpha,\beta}(\theta) = E(\alpha, \beta, \theta)$ of the probability distribution $(\xi_1, \xi_2, \xi_3, \xi_4)$ (see [3, 5.1]) is an absolute S_4-invariant.

In Section 3, Theorem 3, we show that the degree of dependence function $e_{\alpha,\beta}(\theta)$, defined in [3, 5.2] via "normalization" of the entropy function $E_{\alpha,\beta}(\theta)$, is a relative invariant of the dihedral group D_8, see [2, Ch.1,1.]. The proof uses the embedding of D_8 as one of the three Sylow 2-subgroups of S_4.

We use definitions and notation from [3, 2].

2. Methods

In this paper we are using fundamentals of:

- Affine geometry and Real algebraic geometry
- Invariant Theory.

3. The Group of Symmetry of an Experiment

3.1. Yule's Pairs and Experiments

Let $A, B \in \mathcal{A}$. We define $A \Diamond B = (A \triangle B)^c$, where $A \triangle B = (A^c \cap B) \cup (A \cap B^c)$ is the symmetric difference of A and B.

Any ordered pair $(A, B) \in \mathcal{A}^2$ produces the experiment $\mathfrak{J} = \mathfrak{J}_{(A,B)}$ from [3, 4.1,(1)], which is naturally identified with the partition $\{A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c\}$ of Ω (cf. [4, I,§5]). The proof of the next Lemma is straightforward.

Lemma 1. Yule's pairs from the sequence with members

$$\begin{gathered}
(A,B) \text{ of type } (\alpha,\beta,\theta),\\
(A,B^c) \text{ of type } (\alpha,1-\beta,\alpha-\theta),\\
(A^c,B) \text{ of type } (1-\alpha,\beta,\beta-\theta),\\
(A^c,B^c) \text{ of type } (1-\alpha,1-\beta,1-\alpha-\beta+\theta),\\
(B,A) \text{ of type } (\beta,\alpha,\theta),\\
(B,A^c) \text{ of type } (\beta,1-\alpha,\beta-\theta),\\
(B^c,A) \text{ of type } (1-\beta,\alpha,\alpha-\theta),\\
(B^c,A^c) \text{ of type } (1-\beta,1-\alpha,1-\alpha-\beta+\theta),\\
(A,A\Diamond B) \text{ of type } (\alpha,1-\alpha-\beta+2\theta,\theta),\\
(A\Diamond B,A) \text{ of type } (1-\alpha-\beta+2\theta,\alpha,\theta),\\
(B,A\Diamond B) \text{ of type } (\beta,1-\alpha-\beta+2\theta,\theta),\\
(A\Diamond B,B) \text{ of type } (1-\alpha-\beta+2\theta,\beta,\theta),\\
(A^c,A\Diamond B) \text{ of type } (1-\alpha,1-\alpha-\beta+2\theta,1-\alpha-\beta+\theta),\\
(A\Diamond B,A^c) \text{ of type } (1-\alpha-\beta+2\theta,1-\alpha,1-\alpha-\beta+\theta),\\
(B^c,A\Diamond B) \text{ of type } (1-\beta,1-\alpha-\beta+2\theta,1-\alpha-\beta+\theta),\\
(A\Diamond B,B^c) \text{ of type } (1-\alpha-\beta+2\theta,1-\beta,1-\alpha-\beta+\theta),\\
(A,A\triangle B) \text{ of type } (\alpha,\alpha+\beta-2\theta,\alpha-\theta),\\
(A\triangle B,A) \text{ of type } (\alpha+\beta-2\theta,\alpha,\alpha-\theta),\\
(B,A\triangle B) \text{ of type } (\beta,\alpha+\beta-2\theta,\beta-\theta),\\
(A\triangle B,B) \text{ of type } (\alpha+\beta-2\theta,\beta,\beta-\theta),\\
(A^c,A\triangle B) \text{ of type } (1-\alpha,\alpha+\beta-2\theta,\beta-\theta),\\
(A\triangle B,A^c) \text{ of type } (\alpha+\beta-2\theta,1-\alpha,\beta-\theta),\\
(B^c,A\triangle B) \text{ of type } (1-\beta,\alpha+\beta-2\theta,\alpha-\theta),\\
(A\triangle B,B^c) \text{ of type } (\alpha+\beta-2\theta,1-\beta,\alpha-\theta),
\end{gathered} \tag{1}$$

are exactly the pairs that produce the experiment $\mathfrak{J}_{(A,B)}$.

Remark 1. (i) According to [1, 2.1, 2.7.1, 2.8.4], the set of points (α, β, θ) in $\mathbb{R}^3$ where the types from Lemma 1 are pair-wise different is semi-algebraic, open, and three-dimensional. Its trace U_3 on the interior $\mathring{T}_3$ of the classification tetrahedron T_3 from [3, 4.1] is not empty because otherwise $\mathring{T}_3$ would be subset of a finite union of planes. Theorem 2.2.1 from [1, 2.1] guaranties that the open two dimensional projection U_2 of U_3 onto $\alpha\beta$-plane is semi-algebraic. Note that "openness" is with respect to the standard topology in $\mathbb{R}^3$.

(ii) Under some "plentifulness" condition on Boolean algebra $\mathcal{A}$ (for example, if it is non-atomic), there exist plenty of Yule's pairs (A, B) of type $(\alpha, \beta) \in U_2$. In this case (we call it "general") the sequence from Lemma 1 consists of 24 Yule's pairs.

3.2. The Group of Symmetry

Let $\mathcal{E}$ be the set of all experiments in the probability space $(\Omega, \mathcal{A}, \Pr)$, that is, the set of all finite partitions of Ω with members from $\mathcal{A}$. The rule $(A, B) \mapsto \mathfrak{J}_{(A,B)}$ defines a map $\mathfrak{J} \colon \mathcal{A}^2 \to \mathcal{E}$ and Lemma 1 implies that the inverse image $\mathfrak{J}^{-1}(\mathfrak{J}_{(A,B)})$ coincides with the associated set of the sequence (1). Let us denote by $\mathcal{I}_{(A,B)}$ the set of equivalence classes in $\mathcal{A}^2$ of the form $[(\alpha, \beta, \theta)]$, which contain the members of $\mathfrak{J}^{-1}(\mathfrak{J}_{(A,B)})$. If $\alpha = \Pr(A)$, $\beta = \Pr(B)$, $\theta = \Pr(A \cap B)$, then (A, B) is a Yule's pair of type (α, β, θ), (A, B^c) is a Yule's pair of type $(\alpha, 1 - \beta, \alpha - \theta)$, (A^c, B) is a Yule's pair of type $(1 - \alpha, \beta, \beta - \theta)$, etc. Considering α, β, θ as coordinate functions in $\mathbb{R}^3$, the members of $\mathcal{I}_{(A,B)}$ produce the set $\mathfrak{S}_4$ consisting of 24 affine automorphisms of $\mathbb{R}^3$ from the following list:

$$\begin{aligned}
\varphi_{(1)}(\alpha, \beta, \theta) &= (\alpha, \beta, \theta),\\
\varphi_{(12)(34)}(\alpha, \beta, \theta) &= (\alpha, 1 - \beta, \alpha - \theta),\\
\varphi_{(13)(24)}(\alpha, \beta, \theta) &= (1 - \alpha, \beta, \beta - \theta),\\
\varphi_{(14)(23)}(\alpha, \beta, \theta) &= (1 - \alpha, 1 - \beta, 1 - \alpha - \beta + \theta),\\
\varphi_{(23)}(\alpha, \beta, \theta) &= (\beta, \alpha, \theta),\\
\varphi_{(1342)}(\alpha, \beta, \theta) &= (\beta, 1 - \alpha, \beta - \theta),\\
\varphi_{(1243)}(\alpha, \beta, \theta) &= (1 - \beta, \alpha, \alpha - \theta),\\
\varphi_{(14)}(\alpha, \beta, \theta) &= (1 - \beta, 1 - \alpha, 1 - \alpha - \beta + \theta),\\
\varphi_{(34)}(\alpha, \beta, \theta) &= (\alpha, 1 - \alpha - \beta + 2\theta, \theta),\\
\varphi_{(243)}(\alpha, \beta, \theta) &= (1 - \alpha - \beta + 2\theta, \alpha, \theta),\\
\varphi_{(234)}(\alpha, \beta, \theta) &= (\beta, 1 - \alpha - \beta + 2\theta, \theta),\\
\varphi_{(24)}(\alpha, \beta, \theta) &= (1 - \alpha - \beta + 2\theta, \beta, \theta),\\
\varphi_{(142)}(\alpha, \beta, \theta) &= (1 - \alpha, 1 - \alpha - \beta + 2\theta, 1 - \alpha - \beta + \theta),\\
\varphi_{(1423)}(\alpha, \beta, \theta) &= (1 - \alpha - \beta + 2\theta, 1 - \alpha, 1 - \alpha - \beta + \theta),\\
\varphi_{(143)}(\alpha, \beta, \theta) &= (1 - \beta, 1 - \alpha - \beta + 2\theta, 1 - \alpha - \beta + \theta),\\
\varphi_{(1432)}(\alpha, \beta, \theta) &= (1 - \alpha - \beta + 2\theta, 1 - \beta, 1 - \alpha - \beta + \theta),\\
\varphi_{(12)}(\alpha, \beta, \theta) &= (\alpha, \alpha + \beta - 2\theta, \alpha - \theta),\\
\varphi_{(123)}(\alpha, \beta, \theta) &= (\alpha + \beta - 2\theta, \alpha, \alpha - \theta),\\
\varphi_{(132)}(\alpha, \beta, \theta) &= (\beta, \alpha + \beta - 2\theta, \beta - \theta),\\
\varphi_{(13)}(\alpha, \beta, \theta) &= (\alpha + \beta - 2\theta, \beta, \beta - \theta),\\
\varphi_{(1324)}(\alpha, \beta, \theta) &= (1 - \alpha, \alpha + \beta - 2\theta, \beta - \theta),\\
\varphi_{(134)}(\alpha, \beta, \theta) &= (\alpha + \beta - 2\theta, 1 - \alpha, \beta - \theta),\\
\varphi_{(124)}(\alpha, \beta, \theta) &= (1 - \beta, \alpha + \beta - 2\theta, \alpha - \theta),\\
\varphi_{(1234)}(\alpha, \beta, \theta) &= (\alpha + \beta - 2\theta, 1 - \beta, \alpha - \theta).
\end{aligned}$$

The above affine automorphisms of $\mathbb{R}^3$ are indexed by the permutations σ from the symmetric group S_4 because of the theorem below.

The operator of symmetry

$$\sigma \colon H \to H, (\xi_1, \xi_2, \xi_3, \xi_4) \mapsto (\xi_{\sigma^{-1}(1)}, \xi_{\sigma^{-1}(2)}, \xi_{\sigma^{-1}(3)}, \xi_{\sigma^{-1}(4)}),$$

permutes the components of the probability distribution [3, 4.1,(2)] produced by the experiment $\mathfrak{J}_{(A,B)}$ and we have

Theorem 1. (i) One has $\iota \circ \varphi_{\sigma^{-1}} = \sigma \circ \iota$.

(ii) The map

$$S_4 \to \mathrm{Aff}(3, \mathbb{R}), \sigma \mapsto \varphi_{\sigma^{-1}}, \tag{2}$$

is a group anti-monomorphism with image $\mathfrak{S}_4$.

(iii) The group $\mathfrak{S}_4$ is the affine symmetry group of the classification tetrahedron T_3.

Proof. (i) It is enough to check the equality $\varphi_{\sigma^{-1}} = \iota^{-1} \circ \sigma \circ \iota$ for all $\sigma \in S_4$. For example, let $\sigma = (1243)$, so $\sigma^{-1} = (1342)$. We have

$$(\sigma \circ \iota)(\alpha, \beta, \theta) = (\xi_{\sigma^{-1}(1)}, \xi_{\sigma^{-1}(2)}, \xi_{\sigma^{-1}(3)}, \xi_{\sigma^{-1}(4)}) = (\xi_3, \xi_1, \xi_4, \xi_2),$$

$$(\iota^{-1} \circ \sigma \circ \iota)(\alpha, \beta, \theta) = \iota^{-1}(\xi_3, \xi_1, \xi_4, \xi_2) = (\beta, 1 - \alpha, \beta - \theta) =$$

$$\varphi_{(1342)}(\alpha, \beta, \theta) = \varphi_{\sigma^{-1}}(\alpha, \beta, \theta).$$

(ii) The map (2) is injective; moreover, it is a group anti-homomorphism because $\varphi_{(1)} = \iota^{-1} \circ (1) \circ \iota = (1)$ and $\varphi_{\tau^{-1}\sigma^{-1}} = \varphi_{(\sigma\tau)^{-1}} = \iota^{-1} \circ (\sigma\tau) \circ \iota = \iota^{-1} \circ \sigma \circ \tau \circ \iota = \iota^{-1} \circ \sigma \circ \iota \circ \iota^{-1} \circ \tau \circ \iota = \varphi_{\sigma^{-1}} \circ \varphi_{\tau^{-1}}$.

(iii) In accord with part (i), for any $\sigma \in S_4$ we have $\iota(\varphi_\sigma(T_3)) = \sigma^{-1}(\iota(T_3)) = \sigma^{-1}(\Delta_3) = \Delta_3$, hence $\varphi_\sigma(T_3) = \iota^{-1}(\Delta_3) = T_3$. On the other hand, S_4 is the symmetry group of the regular tetrahedron (see, for example, [5, 8.4]). Since both tetrahedrons are isomorphic as affine spans, the proof is done.

■

For any $\sigma \in S_4$ we write down the affine automorphism φ_σ in terms of coordinates in $\mathbb{R}^3$: $\varphi_\sigma(\alpha, \beta, \theta) = (\alpha^{(\sigma)}, \beta^{(\sigma)}, \theta^{(\sigma)})$ and obtain that φ_σ maps the components of the partition $T_3 = \cup_{(\alpha,\beta)\in[0,1]^2} \{\alpha\} \times \{\beta\} \times I(\alpha, \beta)$ onto the corresponding components of the partition $T_3 = \cup_{(\alpha,\beta)\in[0,1]^2} \{\alpha^{(\sigma)}\} \times \{\beta^{(\sigma)}\} \times I(\alpha^{(\sigma)}, \beta^{(\sigma)})$. Moreover, φ_σ maps the components of the partition $\mathring{T}_3 = \cup_{(\alpha,\beta)\in(0,1)^2} \{\alpha\} \times \{\beta\} \times \mathring{I}(\alpha, \beta)$ onto the corresponding components of the partition $\mathring{T}_3 = \cup_{(\alpha,\beta)\in(0,1)^2} \{\alpha^{(\sigma)}\} \times \{\beta^{(\sigma)}\} \times \mathring{I}(\alpha^{(\sigma)}, \beta^{(\sigma)})$.

Let us set $\hat{T}_3 = \cup_{(\alpha,\beta)\in(0,1)^2} \{\alpha\} \times \{\beta\} \times I(\alpha, \beta)$. In particular, we obtain the following

Lemma 2. Let $(\alpha, \beta) \in (0,1)^2$, $\sigma \in S_4$. (i) The automorphism φ_σ maps the set

$$\{(\alpha, \beta, \ell(\alpha, \beta)), (\alpha, \beta, r(\alpha, \beta))\}$$

of endpoints of the segments $\{\alpha\} \times \{\beta\} \times I(\alpha, \beta)$ onto the set

$$\{(\alpha^{(\sigma)}, \beta^{(\sigma)}, \ell(\alpha^{(\sigma)}, \beta^{(\sigma)})), (\alpha^{(\sigma)}, \beta^{(\sigma)}, r(\alpha^{(\sigma)}, \beta^{(\sigma)}))\}$$

of endpoints of their images $\{\alpha^{(\sigma)}\} \times \{\beta^{(\sigma)}\} \times I(\alpha^{(\sigma)}, \beta^{(\sigma)})$.

(ii) One has $\varphi_\sigma(\hat{T}_3) = \hat{T}_3$.

In accord with Theorem 1, (ii), the group S_4 acts on the real functions $F \colon \mathbb{R}^3 \to \mathbb{R}$ via the rule $\sigma \cdot F = F \circ \varphi_{\sigma^{-1}}$. Let

$$G \colon \mathring{\Delta}_3 \to \mathbb{R}, G(\xi_1, \xi_2, \xi_3, \xi_4) = -\xi_1 \ln \xi_1 - \xi_2 \ln \xi_2 - \xi_3 \ln \xi_3 - \xi_4 \ln \xi_4,$$

$$E \colon \mathring{T}_3 \to \mathbb{R}, E = G \circ \iota.$$

The function G is continuously differentiable on the interior $\mathring{\Delta}_3$ and can be extended under the name $\hat{G}$ as continuous on $\hat{\Delta}_3 = \iota(\hat{T}_3)$. The function E is continuously differentiable on the interior $\mathring{T}_3$ and can be extended under the name $\hat{E}$ as continuous on $\hat{T}_3$ (cf. [3, 5.1,Theorem 2, (iii)]). Moreover, $\hat{G} = \hat{G} \circ \sigma$ (that is, $\hat{G}$ is an absolute S_4-invariant) and $\hat{E} = \hat{G} \circ \iota$. Lemma 2, (ii), allows us to extend the action of the symmetric group S_4 on $\hat{T}_3$ via the rule $\sigma \cdot \hat{E} = \hat{E} \circ \varphi_{\sigma^{-1}}$.

Throughout the end of the paper, with an abuse of the language, we designate $\hat{G}$ via G and $\hat{E}$ via E.

Theorem 2. The function $E\colon \hat{T}_3 \to \mathbb{R}$ is an (absolute) invariant of the symmetric group S_4.

Proof. Theorem 1, (i), yields $E = G \circ \iota = G \circ \sigma \circ \iota = G \circ \iota \circ \varphi_{\sigma^{-1}} = E \circ \varphi_{\sigma^{-1}} = \sigma \cdot E$ for all $\sigma \in S_4$.

■

4. Degree of Dependance: Further Properties

4.1. The Groups of Symmetry

Let us suppose $(\alpha, \beta) \in (0,1)^2$ and set

$$e(\alpha,\beta,\theta) = \begin{cases} -\frac{E(\alpha,\beta,\alpha\beta)-E(\alpha,\beta,\theta)}{E(\alpha,\beta,\alpha\beta)-E(\alpha,\beta,\ell(\alpha,\beta))} & \text{if } \ell(\alpha,\beta) \le \theta \le \alpha\beta \\ \frac{E(\alpha,\beta,\alpha\beta)-E(\alpha,\beta,\theta)}{E(\alpha,\beta,\alpha\beta)-E(\alpha,\beta,r(\alpha,\beta))} & \text{if } \alpha\beta \le \theta \le r(\alpha,\beta), \end{cases} \tag{3}$$

where $I(\alpha,\beta) = [\ell(\alpha,\beta), r(\alpha,\beta)]$. Note that in [3, 5.2] the function $e_{\alpha,\beta}(\theta) = e(\alpha,\beta,\theta)$ is said to be the degree of dependence of events A and B with $\alpha = \Pr(A)$, $\beta = \Pr(B)$, and $\theta = \Pr(A \cap B)$.

Let us consider the dihedral subgroup $D_8 = \langle (1342), (14) \rangle$ of S_4 and let $\chi\colon D_8 \to \mathbb{R}^*$ be its Abelian character with kernel $K = \langle (14), (23) \rangle$ and image $\{1, -1\}$.

Theorem 3. The function e from (3) is a relative invariant of weight χ of the dihedral group D_8.

Proof. Given $\sigma \in S_4$ we have

$$(\sigma^{-1} \cdot e)(\alpha,\beta,\theta) = e(\varphi_\sigma(\alpha,\beta,\theta)) = e(\alpha^{(\sigma)},\beta^{(\sigma)},\theta^{(\sigma)}) =$$

$$\begin{cases} -\frac{E(\alpha^{(\sigma)},\beta^{(\sigma)},\alpha^{(\sigma)}\beta^{(\sigma)})-E(\alpha^{(\sigma)},\beta^{(\sigma)},\theta^{(\sigma)})}{E(\alpha^{(\sigma)},\beta^{(\sigma)},\alpha^{(\sigma)}\beta^{(\sigma)})-E(\alpha^{(\sigma)},\beta^{(\sigma)},\ell(\alpha^{(\sigma)},\beta^{(\sigma)}))} & \text{if } \ell(\alpha^{(\sigma)},\beta^{(\sigma)}) \le \theta^{(\sigma)} \le \alpha^{(\sigma)}\beta^{(\sigma)} \\ \frac{E(\alpha^{(\sigma)},\beta^{(\sigma)},\alpha^{(\sigma)}\beta^{(\sigma)})-E(\alpha^{(\sigma)},\beta^{(\sigma)},\theta^{(\sigma)})}{E(\alpha^{(\sigma)},\beta^{(\sigma)},\alpha^{(\sigma)}\beta^{(\sigma)})-E(\alpha^{(\sigma)},\beta^{(\sigma)},\ell(\alpha^{(\sigma)},\beta^{(\sigma)}))} & \text{if } \alpha^{(\sigma)}\beta^{(\sigma)} \le \theta^{(\sigma)} \le r(\alpha^{(\sigma)},\beta^{(\sigma)}), \end{cases}$$

where $I(\alpha^{(\sigma)},\beta^{(\sigma)}) = [\ell(\alpha^{(\sigma)},\beta^{(\sigma)}), r(\alpha^{(\sigma)},\beta^{(\sigma)})]$. For any $\sigma \in D_8$ we have $\varphi_\sigma(\alpha,\beta,\alpha\beta) = (\alpha^{(\sigma)},\beta^{(\sigma)},\alpha^{(\sigma)}\beta^{(\sigma)})$. On the other hand, given $\sigma \in K$, the inequalities $\ell(\alpha,\beta) \le \theta \le \alpha\beta$ are equivalent to the inequalities $\ell(\alpha^{(\sigma)},\beta^{(\sigma)}) \le \theta^{(\sigma)} \le \alpha^{(\sigma)}\beta^{(\sigma)}$ and the inequalities $\alpha\beta \le \theta \le r(\alpha,\beta)$ are equivalent to the inequalities $\alpha^{(\sigma)}\beta^{(\sigma)} \le \theta^{(\sigma)} \le r(\alpha^{(\sigma)},\beta^{(\sigma)})$. Given $\sigma \in D_8 \backslash K$, the inequalities $\ell(\alpha,\beta) \le \theta \le \alpha\beta$ are equivalent to the inequalities $\alpha^{(\sigma)}\beta^{(\sigma)} \le \theta^{(\sigma)} \le r(\alpha^{(\sigma)},\beta^{(\sigma)})$ and the inequalities $\alpha\beta \le \theta \le r(\alpha,\beta)$ are equivalent to the inequalities $\ell(\alpha^{(\sigma)},\beta^{(\sigma)}) \le \theta^{(\sigma)} \le \alpha^{(\sigma)}\beta^{(\sigma)}$. The corresponding equalities hold simultaneously because of Lemma 2, (i). Now, Theorem 2 yields that $\sigma \cdot e = \chi(\sigma)e$ for all permutations $\sigma \in D_8$.

■

We obtain immediately the following

Corollary 1. For any $(\alpha,\beta) \in (0,1)^2$ and for any $\theta \in I(\alpha,\beta)$ one has

$$e_{\alpha,\beta}(\theta) = e_{\beta,\alpha}(\theta) = e_{1-\alpha,1-\beta}(1-\alpha-\beta+\theta) = e_{1-\beta,1-\alpha}(1-\alpha-\beta+\theta),$$

$$-e_{\alpha,\beta}(\theta) = e_{\alpha,1-\beta}(\alpha-\theta) = e_{1-\alpha,\beta}(\beta-\theta) = e_{\beta,1-\alpha}(\beta-\theta) = e_{1-\beta,\alpha}(\alpha-\theta).$$

Acknowledgements

I would like to thank the referees for their very useful remarks.
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Conflicting Interests

The Author declares that there is no conflict of interest.

References

[1] Bochnak J., Coste M., Roy M-F. (1998). Real Algebraic Geometry, Springer.
[2] Dieudonne J. A., Carrell J. B. (1970). Invariant Theory, Old and New, Academic Press, Inc.
[3] Iliev V. V. (2021). On the Use of Entropy as a Measure of Dependence of Two Events. *Reliability: Theory & Applications*, 16(2): 237 - 248.
[4] Kolmogorov A. N. (1956). Foundations of the Theory of Probability, Chelsea Publishing Company, New Yourk.
[5] Smith J. T. (2000). Methods of Geometry, John Wiley & Sons.

RT&A, No 2 (68)
Volume 17, June 2022

Analysis of $MAP/PH/1$ Queueing model with Multiple Vacations, Optional Service, Close-down, Setup, Breakdown, Phase Type Repair and Impatient Customers

G. Ayyappan, G. Archana @ Gurulakshmi*, B. Somasundaram

•

Department of Mathematics,
Puducherry Technological University,
Puducherry, India.
Department of Mathematics,
Vel Tech Rangarajan Dr. Sagunthala
R & D Institution of Science and Technology,
Tamilnadu, India.
ayyappanpec@hotmail.com, archanagurulakshmi@gmail.com, somu.b92@gmail.com

Abstract

The purpose of this paper is to analyse a single server queueing model with multiple vacations, optional service, close-down, setup, balking, breakdown and repair under the assumption that the customers arrive according to a Markovian Arrival Process (MAP). The service and repair times follow the phase-type distributions. At the completion of service, in case there are no customers in the system, the server closes down the system and goes for vacation. After completion of the vacation, the server has to start the setup process if a minimum of one customer is present in the system or else the server goes for another vacation. The server provides optional service to the customers those who are in need of additional services. By employing the matrix analytic method, the stationary probability vector has been evaluated. The stability condition, busy period analysis, distribution function for waiting time and some of the system performance measures concerning this model are derived. The outcome arising out of numerical values and graphical representations are also presented for this model.

Keywords: Multiple vacations, Optional service, Close-down, Setup, Breakdown, Phase type repair, Balking.

AMS Subject Classification (2010): 60K25, 68M30, 90B22 .

1. Introduction

During almost all the day-to-day activities of our life, we come across various queues in many places like shopping malls, traffic signals, billing section, railway counters, communication networks, telecommunication systems, etc in which the queues are either visible or invisible. Mostly people do not prefer to stand in a queue for a long time. Considering the attendant consequences of spending enormous time in queues, it becomes imperative to employ appropriate queueing models to offer remedial measures for these congestion situations.

In the analysis of a queue, Markovian Arrival Process (MAP) proves to be a very useful appliance in the point process which includes the Markov Modulated Poisson Process and phase type renewal process. Neuts [24] has introduced Versatile Markovian Point Processes (VMPP) through which the arrival process is formulated. Later, MAP and BMAP were introduced by Lucantoni et al. [17] . A method of analysis was provided by Chakravarthy [4] who considered the Markovian arrival process with various types of arrivals in the representation of parameter matrices (D_0, D_1) with dimension n where D_0 governs for no arrivals and D_1 governs for arrivals. Let D be the generator matrix defined by $D = D_0 + D_1$.

A queueing system with a single server and K waiting rooms with finite capacity was examined by Niu et al. [28]. They considered the arrival of a batch of customers with the server taking multiple vacations during the idle time. Jau-Chuan Ke et al. [10] have investigated a multi server retrial queue with single and multiple vacations. In this system, they have analysed the cost function to determine the optimum value of the server at minimum cost.

Artalejo et al. [1] have carried out a busy period analysis by considering the distribution function of the waiting time with multi-servers and finite retrial group. They have dealt with a system possessing finite buffer for retrial group and provided certain numerical illustrations in their model.

Chakravarthy [7] explored the working of a single server queueing model with multiple vacations and optional secondary services. Choudhury and Paul [8] have recognized a single server queueing system with Bernoulli schedule and multiple vacations policy. They have evaluated the distribution function of the waiting time with busy period analysis and gone into the performance measures of effectiveness.

Kulkarni and Choi [14] have investigated a single server retrial queue with breakdown and repair. Wang and Zhang [27] have analysed a queueing system with balking, reneging and motivating. They also have performed a busy period analysis.

Chakravarthy and Agarwal [5] have analysed a machine repair problem with an unreliable server and phase type repairs and services. Maragathasundari et al. [18] have examined a single server queueing model with compulsory short vacation and considered the reneging during long vacation as optional.

Wang and Zhang [26] studied a single server Discrete-time Retrial G-queue with breakdown and repairs due to negative arrivals. They considered the negative arrival of a customer which distracts the positive arrival of a customer. A single server queue with Bernoulli vacation, setup time, reneging, balking, Bernoulli feedback, breakdown and repair have been examined by Ayyappan and Gowthami [2]. With the help of matrix-geometric method, they have evaluated the probability vector and the rate matrix. They have derived the probability of the server to be idle, busy and the repair time.

MAP arrivals, impatient customers and a perishable inventory system with N-policy have been discussed by Suganya et al. [25]. They have analysed the cost function and presented numerical illustrations of their queueing model. Kumar and Sharma [13] studied a finite capacity single server queueing model with retention of reneging and balking. In addition, they derived the stability condition and the probability of server being idle, busy and on vacation.

A multi-server queueing system with Bernoulli feedback, impatient customers, single and multiple vacation policies have been examined by Kadi et al. [20]. Further, cost analysis and performance measure are also presented for their model. Ayyappan and Thilagavathy [3] have examined a single server queue with vacation, immediate feedback, breakdown, delayed repair, starting failures, stand-by server, and impatient customers.

An overview of the remaining part of this article is as follows: In section 2, we briefly explain the implementation of our model. The description of the model is provided in section 3. In section 4, we present a generation of the matrix and the notations of our model. In section 5, the stability condition, the stationary probability vector and the rate matrix R are evaluated. We have performed the analysis of busy period in section 6. In section 7, the features of some performance measures are examined. Particular case of our model have been dealt with in section 8. In section 9, we have evaluated the distribution of the waiting time in our model. In section 10, the numerical illustrations and graphical representation are provided. The conclusion part of this model has been presented in section 11.

2. Implementation of our model

As a measure of illustration of our model, let us consider a nationalized bank which has more than one serving counters. We choose any one of these counters for our model. In the counter, the server deals with many transaction processes listed below:

1. Deposit or withdrawal.
2. Foreign currency transaction.
3. Loan process, etc.

A customer may demand money transaction process in any of the ways as mentioned above. When the customer arrives, if he finds the availability of the server, then he receives the service immediately; otherwise, the customer has to wait in the line until he reaches the service point. After the customer gets service, he may exit the counter or he has an option to go for another service (optional service). After providing the service, the server goes for vacation (like receiving the telephone calls, arranging the money, checking the transactions, etc.). After the completion of the service, the server will put down the system and go for vacation (close-down). At the completion of vacation, if there is no customer in the system for receiving the service, then the server takes another vacation (multiple vacation). Once the vacation is completed, the server will do some settings in the system and refresh the system to give the service (setup). During the period of vacation, the incoming customer may balk in the particular counter due to the impatience (balking). In the time of busy period, the server can attain breakdown (like power problem, lack of network, hanging the system etc.). After carrying out the repair process, the server will be ready to provide the service to the waiting customers, those are in interruption of service in front of the queue. Our model has been formulated to hold in all these situations.

3. Model Description

A single server queueing system with infinite capacity has been dealt with in this model. The arrival of customers is according to Markovian arrival process with (D_0, D_1) as its parameter matrices of order m. The matrix D_0 is governed for the transition which deals for no arrival and the matrix D_1 deals for the arrival of customers.
The time duration of both normal and optional services follow PH-distributions with the notation (α_1, T_1) and (α_2, T_2) of order t_1 and t_2 where $T_1^0 + T_1 e = 0$ and $T_2^0 + T_2 e = 0$. The repair times of the server during both the normal and optional services are based on the PH-distributions with notation (β_1, S_1) and (β_2, S_2) of order s_1 and s_2, respectively where $S_1^0 + S_1 e = 0$ and $S_2^0 + S_2 e = 0$.
Upon the completion of the process of service, the customer may either leave the system with probability q, or he may need optional service by the server with probability p, with $p + q = 1$.
After the service is completed, the system will be close-down by the server only if the system becomes empty in which case the server moves on to vacation.
After the completion of vacation, if there is a customer waiting in the system for the service then the server starts the setup process. Otherwise, the server goes on to another vacation.
During the busy period (both normal and optional), the server may encounter failure of its service and it would start the repair process immediately. At that time, the customer who are getting the service from the server have to join the head of the queue.
After the completion of the repair process, the server begins the service to the customers those who are waiting in the queue.
During the vacation period, the arriving customers may balk the system with probability b or they may join the system with probability $(1 - b)$.
The close-down times, vacation times, setup times, breakdown times all follow exponential distribution with the parameters $\varphi, \eta, \psi, \tau$ respectively.

4. Generation of the matrix under QBD process

Let us list down some notations concerning this model and describe the construction of the generator matrix of the Quasi-Birth and Death process in this section:

Notations:

- $\otimes$ - Kronecker product of two matrices with various orders.
- $\oplus$ - Kronecker sum of two matrices with various orders.
- I_m - An Identity matrix of m-dimensional.
- $e = e_{(t_1+t_2+s_1+s_2+3)m}$.
- $e_2(g)$ - $2m \times 1$ vector with $m+1$ to $2m$ elements are 1 and the remaining elements are zero.
- $e(a)$ - $(t_1+t_2+s_1+s_2+3)m \times 1$ vector with first $t_1 m$ elements are 1 and the remaining elements are zero.
- $e(b)$ - $(t_1+t_2+s_1+s_2+3)m \times 1$ vector with $t_1 m+1$ to $t_1 m+t_2 m$ elements are 1 and the remaining elements are zero.
- $e(c)$ - $(t_1+t_2+s_1+s_2+3)m \times 1$ vector with $t_1 m+t_2 m+1$ to $t_1 m+t_2 m+s_1 m$ elements are 1 and the remaining elements are zero.
- $e(d)$ - $(t_1+t_2+s_1+s_2+3)m \times 1$ vector with $t_1 m+t_2 m+s_1 m+1$ to $t_1 m+t_2 m+s_1 m+s_2 m$ elements are 1 and the remaining elements are zero.
- $e(g)$ - $(t_1+t_2+s_1+s_2+3)m \times 1$ vector with $t_1 m+t_2 m+s_1 m+s_2 m+m+1$ to $t_1 m+t_2 m+s_1 m+s_2 m+2m$ elements are 1 and the remaining elements are zero.
- Let λ be the fundamental arrival rate defined by $\lambda = \pi D_1 e_m$ where π is the stationary probability vector.
- The normal and optional service rates of the server are indicated as $\mu_1 = [\alpha_1(-T_1^{-1})e_{t_1}]^{-1}$ and $\mu_2 = [\alpha_2(-T_2^{-1})e_{t_2}]^{-1}$.
- The repair rates (breakdown occurred during normal and optional services) for the server are indicated as
  $\sigma_1 = [\beta_1(-S_1^{-1})e_{s_1}]^{-1}$ and $\sigma_2 = [\beta_2(-S_2^{-1})e_{s_2}]^{-1}$ respectively.
- $N(t)$ is the number of customers in the system at time t.
- $C(t)$ be the status of the server at time t, where

$$C(t) = \begin{cases} 0, & \text{if the server is busy with normal service} \\ 1, & \text{if the server is busy with optional service} \\ 2, & \text{if the server is under PH repair process (breakdown occurred during normal service)} \\ 3, & \text{if the server is under PH repair process (breakdown occurred during optional service)} \\ 4, & \text{if the server is on close-down process} \\ 5, & \text{if the server is on vacation} \\ 6, & \text{if the server is on setup process} \end{cases}$$

- $S(t)$ is the service phase.
- $K(t)$ is the repair phase.
- $A(t)$ is the Markovian arrival process phase.

Let $\{N(t), C(t), S(t), K(t), A(t), \quad t \geq 0\}$ be the Continuous Time Markov Chain ($CTMC$) with state level independent QBD structure for which the state space is provided by

$$\Omega = l(0) \cup l(i)$$

where

$l(0) = \{(0, j, a) : j = 4, 5 : \ 1 \leq a \leq m\}$

and for $\quad i \geq 1$,

$$
\begin{aligned}
l(i) = &\{(i,0,k_1,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le k_1 \le t_1,\ 1 \le a \le m\} \cup \{(i,1,k_2,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le k_2 \le t_2,\ 1 \le a \le m\} \\
&\cup \{(i,2,r_1,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le r_1 \le s_1,\ 1 \le a \le m\} \cup \{(i,3,r_2,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le r_2 \le s_2,\ 1 \le a \le m\} \\
&\cup \{(i,4,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le a \le m\} \cup \{(i,5,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le a \le m\} \cup \{(i,6,a):\ i\epsilon\mathbb{Z}^+,\ 1 \le a \le m\}
\end{aligned}
$$

4.1. The Infinitesimal Matrix Generation

The QBD process has infinitesimal generator matrix Q is given by

$$
Q = \begin{bmatrix}
B_{00} & B_{01} & 0 & 0 & 0 & 0 & \dots \\
B_{10} & A_1 & A_0 & 0 & 0 & 0 & \dots \\
0 & A_2 & A_1 & A_0 & 0 & 0 & \dots \\
0 & 0 & A_2 & A_1 & A_0 & 0 & \dots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \dots
\end{bmatrix}
$$

The entries of the Q matrix are defined by

$$
B_{00} = \begin{bmatrix} D_0 - \varphi I_m & \varphi I_m \\ 0 & D_0 + bD_1 \end{bmatrix}, \quad
B_{01} = \begin{bmatrix} 0 & 0 & 0 & 0 & D_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-b)D_1 & 0 \end{bmatrix},
$$

$$
B_{10} = \begin{bmatrix}
qT_1^0 \otimes I_m & 0 \\
T_2^0 \otimes I_m & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}, \quad
A_0 = \begin{bmatrix}
I_{t_1} \otimes D_1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_{t_2} \otimes D_1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_{s_1} \otimes D_1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_{s_2} \otimes D_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & D_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & (1-b)D_1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & D_1
\end{bmatrix},
$$

$$
A_1 = \begin{bmatrix}
T_1 \oplus D_0 - \tau_1 I_{mt_1} & \alpha_2 \otimes pT_1^0 \otimes I_m & e_2 \otimes \tau_1 \beta_1 \otimes I_m & 0 & 0 & 0 & 0 \\
0 & T_2 \oplus D_0 - \tau_2 I_{mt_2} & 0 & e_2 \otimes \tau_2 \beta_2 \otimes I_m & 0 & 0 & 0 \\
\alpha_1 \otimes S_1^0 \otimes I_m & 0 & S_1 \oplus D_0 & 0 & 0 & 0 & 0 \\
0 & \alpha_2 \otimes S_2^0 \otimes I_m & 0 & S_2 \oplus D_0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & D_0 - \varphi I_m & \varphi Im & 0 \\
0 & 0 & 0 & 0 & 0 & bD_1 + D_0 - \eta I_m & \eta I_m \\
\psi(\alpha_1 \otimes I_m) & 0 & 0 & 0 & 0 & 0 & D_0 - \psi I_m
\end{bmatrix},
$$

$$
A_2 = \begin{bmatrix}
qT_1^0 \otimes \alpha_1 \otimes I_m & 0 & 0 & 0 & 0 & 0 & 0 \\
T_2^0 \otimes \alpha_1 \otimes I_m & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
$$

5. System Analysis

To ensure that the system is stable, we have to evaluate our model under certain conditions are described in the sequel:

5.1. Analysis of stability condition

Let A be the matrix defined by $A = A_0 + A_1 + A_2$ and ς be the stationary probability vector of A which satisfies the condition

$$\varsigma A = 0, \quad \varsigma e = 1.$$

where the vector ς represents the state given by $\varsigma = (\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4, \varsigma_5, \varsigma_6)$.

The vector ς, partitioned as $\varsigma = (\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4, \varsigma_5, \varsigma_6)$ is evaluated with the help of the equation:

$$\begin{aligned}
&\varsigma_0[(I_{t_1} \otimes D_1) + (T_1 + D_0 - \tau_1 I_{mt_1}) + (qT_1^0 \otimes \alpha_1 \otimes I_m)] + \varsigma_1[T_2^0 \otimes \alpha_2 \otimes I_m] \\
&\qquad\qquad + \varsigma_2[\alpha_1 \otimes S_1^0 \otimes I_m] + \varsigma_6[\psi \alpha_1 \otimes I_m] = 0, \\
&\varsigma_0[\alpha_2 \otimes pT_1^0 \otimes I_m] + \varsigma_1[(I_{t_2} \otimes D_1) + (T_2 + D_0 - \tau_2 I_{mt_2})] + \varsigma_3[\alpha_2 \otimes S_2^0 \otimes I_m] = 0, \\
&\varsigma_0[e_2 \otimes \tau_1 \beta_1 \otimes I_m] + \varsigma_2[(I_{s_1} \otimes D_1) + (S_1 \otimes D_0)] = 0, \\
&\varsigma_1[e_2 \otimes \tau_2 \beta_2 \otimes I_m] + \varsigma_3[(I_{s_2} \otimes D_1) + (S_2 \otimes D_0)] = 0, \\
&\varsigma_4[D - \varphi I_m] = 0, \\
&\varsigma_4[\varphi I_m] + \varsigma_5[D - \eta I_m] = 0, \\
&\varsigma_5[\eta I_m] + \varsigma_6[D - \psi I_m] = 0
\end{aligned}$$

subject to

$$\varsigma_0 e_{mt_1} + \varsigma_1 e_{mt_2} + \varsigma_2 e_{ms_1} + \varsigma_3 e_{ms_2} + \varsigma_4 e_m + \varsigma_5 e_m + \varsigma_6 e_m = 1.$$

The necessary and sufficient condition for the stability of the system is that the QBD process satisfies the condition $\varsigma A_0 e < \varsigma A_2 e$.
i.e.,

$$\begin{aligned}
&\varsigma_0[I_{t_1} \otimes D_1] + \varsigma_1[I_{t_2} \otimes D_1] + \varsigma_2[I_{s_1} \otimes D_1] + \varsigma_3[I_{s_2} \otimes D_1] + (\varsigma_4 + \varsigma_6)[D_1] + \varsigma_5[(1-b)D_1] \\
&\qquad\qquad < \varsigma_0[qT_1^0 \otimes \alpha_1 \otimes I_m] + \varsigma_1[T_2^0 \otimes \alpha_1 \otimes I_m].
\end{aligned}$$

5.2. Analysis of stationary probability vector

Let X be the stationary probability vector. It is subdivided as $X=(X_0, X_1, X_2, ...)$ which is the steady state probability vector of Q.

The dimensions of X_0 and X_i , $i \geq 1$ are $2m$ and $(t_1 + t_2 + s_1 + s_2 + 3)m$ respectively. The vector X of Q satisfies the condition

$$XQ = 0 \text{ and } Xe = 1.$$

However, once the stability condition is satisfied, we find the vector X as invariant probability with the help of the equation,

$$X_i = X_1 R^{i-1}, \quad i \geq 2.$$

where the matrix R is the minimal non-negative solution to

$$R^2 A_2 + R A_1 + A_0 = 0.$$

By means of the equations

$$X_0 B_{00} + X_1 B_{10} = 0,$$

$$X_0 B_{01} + X_1[A_1 + RA_2] = 0$$

we can find the vectors namely X_0 and X_1 subject to normalizing condition

$$X_0 e_{2m} + X_1[I - R]^{-1} e_{(t_1+t_2+s_1+s_2+3)m} = 1.$$

The "Logarithmic Reduction Algorithm" can be used to quickly calculate the rate matrix R as specified by Latouche and Ramaswami [16].

6. Busy Period Analysis

A busy period can be measured as the interval between the customers entering into an empty system and when the system size reduces to empty for the first time. As a result, this is the first passage time between the level i and level $i-1$, $i \geq 2$ under the consideration of QBD process. The first return time to level 0 with minimum one visit to a state in any other level is known as busy cycle. It is necessary to deal with $i = 0, 1$ independently for the boundary states. For each and every level i, where $i \geq 1$, we can observe that there are $m(t_1 + t_2 + s_1 + s_2 + 3)$ states.

Notations:

1. $G_{j,j'}(k, x)$ - The probability that the QBD moves by making k left transitions to the level $(i-1)$ and entering the state (i, j'), with the condition of beginning from the state (i, j) at time $t = 0$.
2. Let the joint transform matrix be

$$\tilde{G}_{j,j'}(z, s) = \sum_{k=1}^{\infty} z^k \int_0^{\infty} e^{-sx} dG_{j,j'}(k, x); \quad |z| \leq 1, \quad Re(s) \geq 0$$

3. The matrix $\tilde{G}(z, s) = (\tilde{G}_{j,j'}(z, s))$. (Neuts [21])
4. The matrix $G = (G_{j,j'}) = \tilde{G}(1, 0)$ concerns the first passage times except for the boundary states.
5. $G^{(1,0)}_{j,j'}(k, x)$ - The probability that the system moves from level 0 to level 1 at time $t = 0$.
6. $G^{(0,0)}_{j,j'}(k, x)$ - The probability that the system returns to level 0 at time $t = 0$.
7. S_{1j} - The average first passage time among the levels i and $i-1$, the process in the state (i, j) at time $t = 0$.
8. $\tilde{S}_1$ - The column vector with S_{1j} as its entries.
9. S_{2j} - The average number of customers who receive the service at the first passage time among the levels i and $i-1$, beginning in the state (i, j) at time $t = 0$.
10. $\tilde{S}_2$ - The column vector with S_{2j} as its entries.
11. $\tilde{S}^{(1,0)}_1$ - The average first return time from level 1 to 0.
12. $\tilde{S}^{(1,0)}_2$ - The average number of services completed during the first return time from level 1 to 0.
13. $\tilde{S}^{(0,0)}_1$ - The average first return time to level 0.
14. $\tilde{S}^{(0,0)}_2$ - The average number of services completed during the first return time to level 0.

We can compute the matrix $\tilde{G}(z,s)$ with the equation

$$\tilde{G}(z,s) = z(sI - A_1)^{-1}A_2 + (sI - A_1)^{-1}A_0\tilde{G}^2(z,s).$$

After evaluating the rate matrix R, the G matrix can be computed with the help of logarithmic reduction algorithm (Lautouche and Ramaswami, [16]) provided by

$$G = -(A_1 + RA_2)^{-1}A_2.$$

The equations

$$\tilde{G}^{(1,0)}(z,s) = z(sI - A_1)^{-1}B_{10} + (sI - A_1)^{-1}A_0\tilde{G}(z,s)\tilde{G}^{(1,0)}(z,s)$$
$$\tilde{G}^{(0,0)}(z,s) = (sI - B_{00})^{-1}B_{01}\tilde{G}^{(1,0)}(z,s).$$

which are satisfied by $\tilde{G}^{(1,0)}(z,s)$ and $\tilde{G}^{(0,0)}(z,s)$ lead to the boundary states namely 1 and 0, respectively.

Since $G, \tilde{G}^{(1,0)}(1,0)$ and $\tilde{G}^{(0,0)}(1,0)$ are all stochastic matrices, we can calculate the moments as follows:

$$\tilde{S}_1 = -\frac{\partial \tilde{G}(z,s)}{\partial s}\Big|_{s=0,z=1} = -[A_0(G+I) + A_1]^{-1}e,$$
$$\tilde{S}_2 = \frac{\partial \tilde{G}(z,s)}{\partial z}\Big|_{s=0,z=1} = -[A_0(G+I) + A_1]^{-1}]A_2e,$$
$$\tilde{S}_1^{(1,0)} = -\frac{\partial \tilde{G}^{(1,0)}(z,s)}{\partial s}\Big|_{s=0,z=1} = -[A_1 + A_0G]^{-1}[e + A_0\tilde{S}_1],$$
$$\tilde{S}_2^{(1,0)} = \frac{\partial \tilde{G}^{(1,0)}(z,s)}{\partial z}\Big|_{s=0,z=1} = -[A_1 + A_0G]^{-1}[B_{10}e + A_0\tilde{S}_2],$$
$$\tilde{S}_1^{(0,0)} = -\frac{\partial \tilde{G}^{(0,0)}(z,s)}{\partial s}\Big|_{s=0,z=1} = -B_{00}^{-1}[e + B_{01}\tilde{S}_1^{(1,0)}],$$
$$\tilde{S}_2^{(0,0)} = \frac{\partial \tilde{G}^{(0,0)}(z,s)}{\partial z}\Big|_{s=0,z=1} = -B_{00}^{-1}[B_{01}\tilde{S}_2^{(1,0)}].$$

7. Performance Measure

- Probability that the server is busy with the normal service

$$P_{BNS} = X_1(I - R)^{-1}e(a)$$

- Probability that the server is busy with the optional service

$$P_{BOS} = X_1(I - R)^{-1}e(b)$$

- Probability that the server is in breakdown during the normal service

$$P_{BDNS} = X_1(I - R)^{-1}e(c)$$

- Probability that the server is in breakdown during the optional service

$$P_{BDOS} = X_1(I - R)^{-1}e(d)$$

- Probability of the server being in vacation

$$P_{VAC} = X_1(I - R)^{-1}e(g) + X_0e_2(g)$$

- Expected size of the system

$$E_S = \sum_{z=1}^{\infty} zX_z e = X_1(I-R)^{-2}e$$

- Expected system size that the server is busy with the normal service

$$E_{BNS} = X_1(I-R)^{-2}e(a)$$

- Expected system size that the server is busy with the optional service

$$E_{BOS} = X_1(I-R)^{-2}e(b)$$

- Expected system size that the server will be in breakdown during the normal service

$$E_{BDNS} = X_1(I-R)^{-2}e(c)$$

- Expected system size that the server will be in breakdown during the optional service

$$E_{BDOS} = X_1(I-R)^{-2}e(d)$$

- Expected system size that the server being in vacation

$$E_{VAC} = X_1(I-R)^{-2}e(g).$$

8. Particular case

We consider an exponential distribution for the arrival, service and repair times. Let us denote:

D_0=[$-\lambda$], D_1=[λ], α_1=[1], T_1=[μ_1], α_2=[1], T_2=[μ_2], β_1=[1], S_1=[σ_1], β_2=[1], S_2=[σ_2]

With our assumption, the infinitesimal generator matrix becomes

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & \dots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & 0 & 0 & \dots \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \dots \end{bmatrix}.$$

The entries of the Q matrix are defined by

$$B_{00} = \begin{bmatrix} -\lambda-\varphi & \varphi \\ 0 & -\lambda+b\lambda \end{bmatrix}, B_{01} = \begin{bmatrix} 0 & 0 & 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-b)\lambda & 0 \end{bmatrix}, B_{10} = \begin{bmatrix} q\mu_1 & 0 \\ \mu_2 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} -\lambda-\mu_1-\tau_1 & p\mu_1 & \tau_1 & 0 & 0 & 0 & 0 \\ 0 & -\lambda-\mu_2-\tau_2 & 0 & \tau_2 & 0 & 0 & 0 \\ \sigma_1 & 0 & -\sigma_1-\lambda & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & -\sigma_2-\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\lambda-\varphi & \varphi & 0 \\ 0 & 0 & 0 & 0 & 0 & b\lambda-\lambda-\eta & \eta \\ \psi & 0 & 0 & 0 & 0 & 0 & -\lambda-\psi \end{bmatrix},$$

$$A_2 = \begin{bmatrix} q\mu_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, A_0 = \begin{bmatrix} \lambda & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-b)\lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda \end{bmatrix},$$

Consequently, the matrix A becomes

$$A = \begin{bmatrix} -\mu_1 - \tau_1 + q\mu_1 & p\mu_1 & \tau_1 & 0 & 0 & 0 & 0 \\ \mu_2 & -\mu_2 - \tau_2 & 0 & \tau_2 & 0 & 0 & 0 \\ \sigma_1 & 0 & -\sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & -\sigma_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\varphi & \varphi & 0 \\ 0 & 0 & 0 & 0 & 0 & -\eta & \eta \\ \psi & 0 & 0 & 0 & 0 & 0 & -\psi \end{bmatrix}.$$

The stationary probability vector ξ of A which satisfies $\xi A = 0$ and $\xi e = 1$ is given by $\xi = (\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6)$, where

$$\xi_0 = \frac{\mu_2\sigma_1\sigma 2}{\mu_2\sigma_1\sigma_2+\mu_2\sigma_2\tau+\mu_1 p\sigma_1\sigma_2+\mu_1 p\sigma_1\tau},\ \xi_1 = \frac{\mu_1 p\sigma_1\sigma 2}{\mu_2\sigma_1\sigma_2+\mu_2\sigma_2\tau+\mu_1 p\sigma_1\sigma_2+\mu_1 p\sigma_1\tau},\ \xi_2 = \frac{\mu_2\sigma_2\tau}{\mu_2\sigma_1\sigma_2+\mu_2\sigma_2\tau+\mu_1 p\sigma_1\sigma_2+\mu_1 p\sigma_1\tau},$$

$$\xi_3 = \frac{\mu_1 p\sigma_1\tau}{\mu_2\sigma_1\sigma_2+\mu_2\sigma_2\tau+\mu_1 p\sigma_1\sigma_2+\mu_1 p\sigma_1\tau},\ \xi_4 = 0,\ \xi_5 = 0,\ \xi_6 = 0.$$

The necessary and sufficient condition required by the system to remain stable is $\xi A_0 e < \xi A_2 e$. Hence

$$\lambda < \frac{\mu_1\mu_2\sigma_1\sigma_2}{\mu_2\sigma_1\sigma_2 + \mu_2\sigma_2\tau + \mu_1 p\sigma_1\sigma_2 + \mu_1 p\sigma_1\tau}.$$

9. Waiting Time Distribution

Using the first passage time, we perform an analysis of waiting time distribution of such of those customers who arrive in the queueing line. Let $W(t)$, $t \geq 0$ denote the distribution function of the waiting time of the incoming tagged customer in the queue. When the server is in busy, repair process or in vacation, the customer has to wait in the queueing line to get service from the server.

The absorbing state $(*)$ corresponds to the upcoming tagged customer to receive the service without waiting in the queue. Let us introduce the absorption time in a continuous time Markov chain with the state space as follows:

$$\tilde{\Omega} = (*) \cup \{0, 1, 2, 3, ...\}$$

where

$$l(0) = \{(0, j) : j = 4, 5\}$$

and for $i \geq 1$,

$$\begin{aligned} l(i) = &\{(i, 0, k_1) : i \epsilon \mathbb{Z}^+,\ 1 \leq k_1 \leq t_1\} \cup \{(i, 1, k_2) : i \epsilon \mathbb{Z}^+,\ 1 \leq k_2 \leq t_2\} \\ &\cup \{(i, 2, r_1) : i \epsilon \mathbb{Z}^+,\ 1 \leq r_1 \leq s_1\} \cup \{(i, 3, r_2) : i \epsilon \mathbb{Z}^+,\ 1 \leq r_2 \leq s_2\} \\ &\cup \{(i, 4) : i \epsilon \mathbb{Z}^+,\ 1 \leq a \leq m\} \cup \{(i, 5) : i \epsilon \mathbb{Z}^+\} \cup \{(i, 6) : i \epsilon \mathbb{Z}^+\} \end{aligned}$$

Let the transition matrix $\tilde{Q}$ of this Markov process be

$$\tilde{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ M_0 & P_{00} & 0 & 0 & 0 & 0 & \dots \\ M_1 & P_1 & P & 0 & 0 & 0 & \dots \\ 0 & 0 & P_2 & P & 0 & 0 & \dots \\ 0 & 0 & 0 & P_2 & P & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \dots \end{bmatrix}$$

where the entries of the $\tilde{Q}$ matrix are defined by

$$M_0 = \begin{bmatrix} 0 \\ \eta \end{bmatrix}, P_{00} = \begin{bmatrix} -\varphi & \varphi \\ 0 & -\eta \end{bmatrix}, M_1 = \begin{bmatrix} qT_1^0 \\ T_2^0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, P_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, P_2 = \begin{bmatrix} qT_1^0 \otimes \alpha_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ T_2^0 \otimes \alpha_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$P = \begin{bmatrix} T_1 - \tau_1 I_{ns} & pT_1^0 \otimes \alpha_2 & e_2 \otimes \tau_1\beta_1 & 0 & 0 & 0 & 0 \\ 0 & T_2 - \tau_2 I_{os} & 0 & e_2 \otimes \tau_2\beta_2 & 0 & 0 & 0 \\ \alpha_1 \otimes S_1^0 & 0 & S_1 & 0 & 0 & 0 & 0 \\ 0 & \alpha_2 \otimes S_2^0 & 0 & S_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\varphi & \varphi & 0 \\ 0 & 0 & 0 & 0 & 0 & -\eta & \eta \\ \psi\alpha_1 & 0 & 0 & 0 & 0 & 0 & -\psi \end{bmatrix}.$$

Let us define $z(0)$=$(z_0(0), z_1(0), z_2(0), ...)$ which is the conditional probability distribution of the system state defined on the arrival of the tagged customers. The probability vectors of $z_0(0)$ and $z_i(0)$ are respectively given by

$$z_0(0) = x_0\big[I_2 \otimes \frac{D_1 e_m}{\lambda}\big]$$
$$z_i(0) = x_i[I_{t_1+t_2+s_1+s_2+3} \otimes \frac{D_1 e_m}{\lambda}], \quad \text{for } i \geq 1$$

where the fundamental arrival rate of the Markovian arrival process is denoted by λ. Let

$$z(t) = (z_*(t), z_0(t), z_1(t), ...),$$

where

$$\begin{aligned} z_0(t) \quad &: \quad \text{vector of order } (1 \times 2), \\ z_i(t),\ i \geq 1 \quad &: \quad \text{vector of order } 1 \times (t_1 + t_2 + s_1 + s_2 + 3). \end{aligned}$$

The elements of $z_i(t),\ i \geq 1$ are the probabilities of the $CTMC$ wherein the respective states of level i with the generator matrix $\tilde{Q}$ are at epoch t. The probability of the process being in the absorbing state at time t is given by $z_*(t)$.

We have $W(t)$ = $z_*(t)$, for $t \geq 0$.

The differential equation $z'(t)$ = $z(t)\tilde{Q}$, where $t \geq 0$ becomes

$$\begin{aligned} z_*'(t) &= z_0(t)M_0 + z_1(t)M_1, \\ z_0'(t) &= z_0(t)P_{00} + z_1(t)P_1, \\ z_i'(t) &= z_i(t)P + z_{i+1}(t)P_2, \quad i \geq 1. \end{aligned}$$

where $'$ denotes the derivative with respect to t.

Let us compute the Laplace Stieltjes Transform (LST) of $W(t)$ with the aid of the technique mentioned by Neuts [23].

The process is commenced at the state i with initial probability vector $z_i(0)$, $i \geq 1$.

Let $w(s)$ be the row vector which specifies the Laplace-Stieltjes transform of the initial transit time to level 1.

As per the scheme specified by Neuts [22], we have

$$w(s) = \sum_{i=1}^{\infty}[(sI - P)^{-1}P_2]^{i-1}. \tag{1}$$

Let the Laplace-Stieltjes transform of absorbing time to the state $(*)$ correspond to the process starting at state level $i = 0, 1$ be indicated by $\phi(i, s)$. Applying a result in Neuts [22], we have

$$\phi(0, s) = [sI - P_{00}]^{-1}M_0, \tag{2}$$

$$\phi(1, s) = [sI - P]^{-1}P_1\phi(0, s) + [sI - P]^{-1}M_1. \tag{3}$$

Thus, the LST of the waiting time distribution $\tilde{W}(s)$ is evaluated as

$$\tilde{W}(s) = z_0(0)\phi(0, s) + w(s)\phi(1, s). \tag{4}$$

9.1. Expected Waiting Time

The expected waiting time is denoted by

$$E(W) = -z_0(0)\phi'(0, 0) - w'(0)e_{t_1+t_2+s_1+s_2+3} - w(0)\phi'(1, 0)e_n. \tag{5}$$

When the system has the state level $i = 0$, the average time to enter into the absorbing state $(*)$ is denoted by the foremost terms of the preceding equation.

Likewise, if the system has the state level $i \geq 1$, the average amount of time to enter into the absorbing state $(*)$ is denoted by the last two terms of the above equation.

On differentiating (2) and (3) with respect to t and substituting $s = 0$, we get

$$\phi'(0, 0) = (-1)[-P_{00}]^{-2}M_0, \tag{6}$$

$$\phi'(1, 0) = (-1)[-P]^{-2}P_1\phi(0, 0) + [-P]^{-1}P_1\phi'(0, 0) - [-P]^{-2}M_1. \tag{7}$$

By using the expression (6) together with the probability vector

$$z(0) = (z_0(0), z_1(0), z_2(0), ...),$$

we can determine the first term of (5). From (1), we have

$$w(0) = \sum_{i=1}^{\infty} z_i(0)V^{i-1}, \tag{8}$$

where $V = [-P]^{-1}P_2$. Since V is a stochastic matrix, we get

$$w(0)e_{t_1+t_2+s_1+s_2+3} = 1 - z_0(0)e_2. \tag{9}$$

With the help of (7) and (9) together with the probability vector $z(0) = (z_0(0), z_1(0), z_2(0), ...)$, we can compute the final term of (5).

On differentiating (1) with respect to t and substituting $s = 0$, we get

$$w'(0) = (-1)\sum_{i=1}^{\infty} z_{i+1}(0)\sum_{j=0}^{i-1} V^j[-P]^{-1}V^{i-j}. \tag{10}$$

The stochastic nature of V implies that

$$(-1)w'(0)e_{t_1+t_2+s_1+s_2+3} = (-1)\sum_{i=1}^{\infty} z_{i+1}(0)\sum_{j=0}^{i-1} V^j[-P]^{-1}e_{t_1+t_2+s_1+s_2+3}. \tag{11}$$

We can evaluate the value of $(-1)w'(0)e_{t_1+t_2+s_1+s_2+3}$, by using the method specified in Neuts [22].

Now, let us consider the stochastic matrix V_2 satisfying two conditions namely $I - V + V_2$ is non-singular and the generalized inverse of the form $I - V$. Then, the matrix V_2 may be chosen as $V_2 = v_0 e_{t_1+t_2+s_1+s_2+3}$, where v_0 is the stationary probability vector of V such that

$$v_0 V = v_0 \text{ and } v_0 e_{t_1+t_2+s_1+s_2+3} = 1.$$

In view of the property

$$VV_2 = V_2V = V_2.$$

we get

$$\sum_{j=0}^{i-1} V^j(I - V + V_2) = I - V^i + iV_2, \quad \text{for } i \geq 1. \tag{12}$$

Substituting (12) in (11) and carrying out some simplifications, we get

$$\begin{aligned}(-1)w'(0)e_{t_1+t_2+s_1+s_2+3} = &\left[x_1[I-R]^{-1}[I_{t_1+t_2+s_1+s_2+3} \otimes \frac{D_1 e_m}{\lambda}]\right.\\ &- w(0) + x_1 R[I-R]^{-2}[I_{t_1+t_2+s_1+s_2+3} \otimes \frac{D_1 e_m}{\lambda} V_2]\\ &\times [I - V + V_2]^{-1}[-P]^{-1}e_{t_1+t_2+s_1+s_2+3}.\end{aligned} \tag{13}$$

Thus, we have determined all the terms of (5). By using (5), the expected waiting time can be easily evaluated.

10. Numerical Results

In this section, we examine the outcome of our system with the utilisation of numerical and graphical methods. The five different MAP representations are given below with distinct variance and correlation structure and their mean values are 1. These values are suggested by Chakravarthy [4] . In the first three process of arrival, like $ERL - A$, $EXP - A$ and $HYP.EXP - A$ correspond to renewal process and thus the correlation is zero.

Arrival in Erlang of order 2 (ERL-A):

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

Arrival in Exponential (EXP-A):

$$D_0 = [-1], D_1 = [1]$$

Arrival in Hyper exponential (HYP-EXP-A):

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}$$

MAP-NC-A: Arrival in MAP - Negative Correlation:

$$D_0 = \begin{bmatrix} -1.00222 & 1.00222 & 0 \\ 0 & -1.00222 & 0 \\ 0 & 0 & -225.75 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.01002 & 0 & 0.99220 \\ 223.4925 & 0 & 2.2575 \end{bmatrix}$$

MAP-PC-A: Arrival in MAP - Positive Correlation:

$$D_0 = \begin{bmatrix} 1.00222 & 1.00222 & 0 \\ 0 & -1.00222 & 0 \\ 0 & 0 & -225.75 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.99220 & 0 & 0.01002 \\ 2.2575 & 0 & 223.4925 \end{bmatrix}.$$

Let us consider the following PH-distributions for the service and repair process which are suggested by Chakravarthy [4].

ERL-S (Service in Erlang of order 2):

$$\alpha_1 = \alpha_2 = (1, 0), \quad T_1 = T_2 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

ERL-R (Repair in Erlang of order 2):

$$\beta_1 = \beta_2 = (1, 0), \quad S_1 = S_2 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

EXP-S (Service in Exponential):

$$\alpha_1 = \alpha_2 = (1), \quad T_1 = T_2 = [-1]$$

EXP-R (Repair in Exponential):

$$\beta_1 = \beta_2 = (1), \quad S_1 = S_2 = [-1]$$

HYP.EXP-S (Service in Hyper exponential):

$$\alpha_1 = \alpha_2 = (0.8, 0.2), \quad T_1 = T_2 = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

HYP.EXP-R (Repair in Hyper exponential):

$$\beta_1 = \beta_2 = (0.8, 0.2), \quad S_1 = S_2 = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}.$$

10.1. Illustrative Example 1

From the Tables 1-5, we explore the effect of the fundamental arrival rate (λ) on the Expected system size (ES) and Expected waiting time (EW). Fix $\mu_1 = 15$, $\mu_2 = 12$, $\sigma_1 = 8$, $\sigma_2 = 6$, $\eta = 8$, $\tau = 1$, $\varphi = 10$, $\psi = 10$, $p = 0.2$, $q = 0.8$, $b = 0.1$.

Table 1: *Fundamental arrival rate (λ) vs ES and EW - EXP-A*

	$EXP-S$		$ERL-S$		$HYP-S$	
λ	ES	EW	ES	EW	ES	EW
1.0	0.15853	0.30160	0.15741	0.30182	0.16332	0.29898
1.2	0.16477	0.36572	0.16339	0.36566	0.17067	0.36396
1.4	0.17091	0.43122	0.16925	0.43075	0.17796	0.43085
1.6	0.17696	0.49815	0.17502	0.49714	0.18522	0.49972
1.8	0.18296	0.56659	0.18071	0.56490	0.19247	0.57067
2.0	0.18892	0.63661	0.18635	0.63409	0.19973	0.64383
2.2	0.19486	0.70832	0.19196	0.70480	0.20703	0.71930
2.4	0.20080	0.78180	0.19755	0.77712	0.21438	0.79723
2.6	0.20677	0.85719	0.20316	0.85116	0.22181	0.87777
2.8	0.21278	0.93460	0.20879	0.92704	0.22935	0.96107

Table 2: *Fundamental arrival rate (λ) vs ES and EW - ERL-A*

	$EXP-S$		$ERL-S$		$HYP-S$	
λ	ES	EW	ES	EW	ES	EW
1.0	0.14153	0.29431	0.14086	0.29510	0.14543	0.29040
1.2	0.14657	0.35594	0.14567	0.35661	0.15163	0.35268
1.4	0.15177	0.41860	0.15063	0.41905	0.15806	0.41660
1.6	0.15706	0.48236	0.15566	0.48243	0.16466	0.48225
1.8	0.16241	0.54724	0.16072	0.54681	0.17138	0.54970
2.0	0.16778	0.61332	0.16579	0.61222	0.17818	0.61904
2.2	0.17317	0.68065	0.17085	0.67871	0.18507	0.69038
2.4	0.17856	0.74932	0.17590	0.74636	0.19203	0.76384
2.6	0.18396	0.81942	0.18094	0.81523	0.19908	0.83955
2.8	0.18939	0.89105	0.18598	0.88543	0.20621	0.91765

Table 3: *Fundamental arrival rate (λ) vs ES and EW - HYP-EXP-A*

	$EXP-S$		$ERL-S$		$HYP-S$	
λ	ES	EW	ES	EW	ES	EW
1.0	0.17911	0.31245	0.17708	0.31172	0.18736	0.31350
1.2	0.18868	0.38164	0.18618	0.38018	0.19877	0.38521
1.4	0.19808	0.45341	0.19507	0.45097	0.21005	0.46028
1.6	0.20740	0.52794	0.20386	0.52429	0.22127	0.53894
1.8	0.21671	0.60549	0.21262	0.60036	0.23251	0.62140
2.0	0.22610	0.68632	0.22143	0.67945	0.24383	0.70793
2.2	0.23564	0.77074	0.23037	0.76186	0.25529	0.79880
2.4	0.24540	0.85911	0.23952	0.84794	0.26695	0.89431
2.6	0.25546	0.95182	0.24894	0.93807	0.27886	0.99479
2.8	0.26588	1.04931	0.25872	1.03271	0.29107	1.10060

Table 4: *Fundamental arrival rate (λ) vs ES and EW - MAP-NC-A*

	$EXP-S$		$ERL-S$		$HYP-S$	
λ	ES	EW	ES	EW	ES	EW
1.0	0.21789	0.32963	0.21810	0.33120	0.21685	0.32110
1.2	0.22094	0.39857	0.22097	0.40024	0.22086	0.38946
1.4	0.22416	0.46873	0.22399	0.47041	0.22513	0.45954
1.6	0.22757	0.54021	0.22718	0.54180	0.22966	0.53149
1.8	0.23116	0.61312	0.23053	0.61448	0.23444	0.60542
2.0	0.23493	0.68755	0.23404	0.68855	0.23947	0.68149
2.2	0.23888	0.76360	0.23771	0.76409	0.24474	0.75981
2.4	0.24301	0.84138	0.24153	0.84121	0.25025	0.84055
2.6	0.24732	0.92102	0.24552	0.92001	0.25600	0.92385
2.8	0.25181	1.00264	0.24968	1.00061	0.26200	1.00989

Table 5: *Fundamental arrival rate (λ) vs ES and EW - MAP-PC-A*

	$EXP-S$		$ERL-S$		$HYP-S$	
λ	ES	EW	ES	EW	ES	EW
1.0	4.93292	5.03451	5.02349	5.12623	4.43675	4.53209
1.2	5.06281	6.18903	5.15683	6.30311	4.54890	5.56542
1.4	5.18831	7.38672	5.28604	7.52488	4.65537	6.63320
1.6	5.31349	8.63171	5.41520	8.79585	4.76012	7.73857
1.8	5.44069	9.92837	5.54670	10.12062	4.86538	8.88481
2.0	5.57147	11.28133	5.68210	11.50408	4.97259	10.07536
2.2	5.70694	12.69553	5.82258	12.95143	5.08272	11.31383
2.4	5.84801	14.17626	5.96905	14.46829	5.19652	12.60403
2.6	5.99546	15.72922	6.12233	16.06069	5.31459	13.95000
2.8	6.15001	17.36055	6.28322	17.73519	5.43745	15.35604

From the above tables 1, 2, 3, 4 and 5, the following observations are made:

- As fundamental arrival rate (λ) increases, the expected system size (ES) increases for various probable sequence of arrival and service times.
- As fundamental arrival rate (λ) increases, the expected waiting time of the system (EW) increases for various probable sequence of arrival and service times.
- With the estimation of the values of various arrival times, the average system size increases much faster for hyper exponential service time and slowly for Erlang service time.

10.2. Illustrative Example 2

The two dimensional graphs are illustrated in the following Figures $1-12$. We explore the effect of the vacation rate (η) on the Expected system size (ES) and Expected waiting time (EW). Fix $\lambda = 1$, $\mu_1 = 15$, $\mu_2 = 12$, $\sigma_1 = 8$, $\sigma_2 = 6$, $\tau = 1$, $\varphi = 10$, $\psi = 10$, $p = 0.2$, $q = 0.8$, $b = 0.1$.

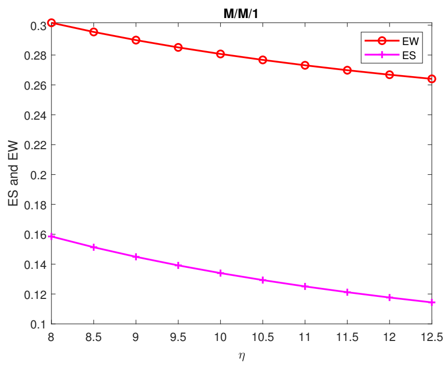


Figure 1: *Vacation rate (η) vs. ES and EW*

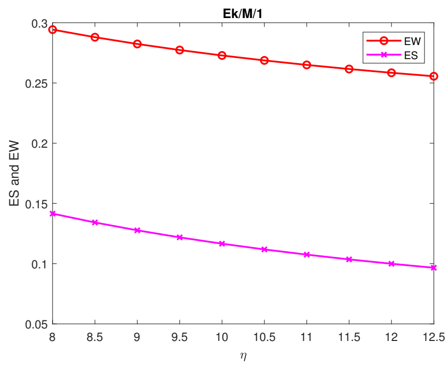


Figure 2: *Vacation rate (η) vs. ES and EW*

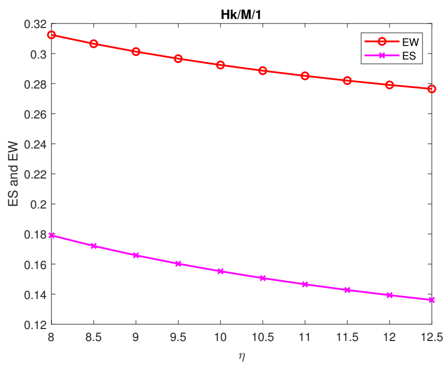


Figure 3: *Vacation rate (η) vs. ES and EW*

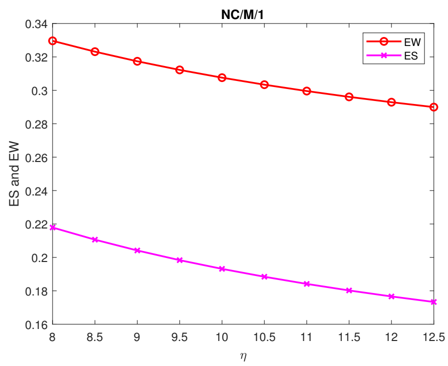


Figure 4: *Vacation rate (η) vs. ES and EW*

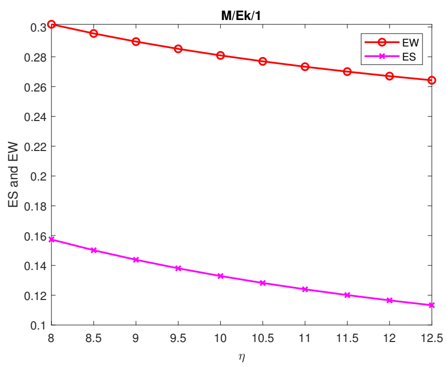


Figure 5: *Vacation rate (η) vs. ES and EW*

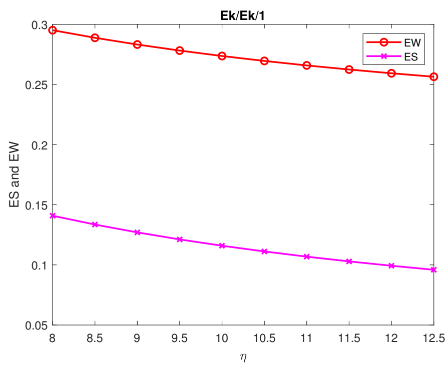


Figure 6: *Vacation rate (η) vs. ES and EW*

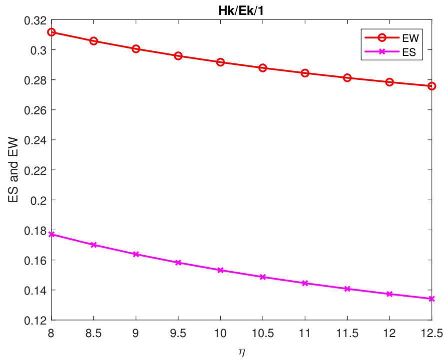


Figure 7: *Vacation rate (η) vs. ES and EW*

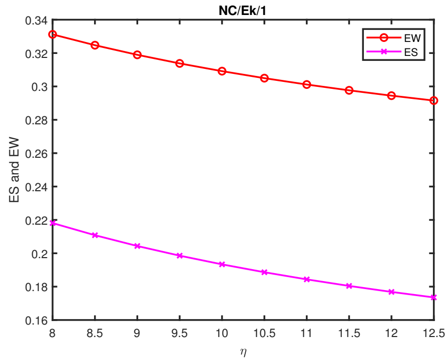


Figure 8: *Vacation rate (η) vs. ES and EW*

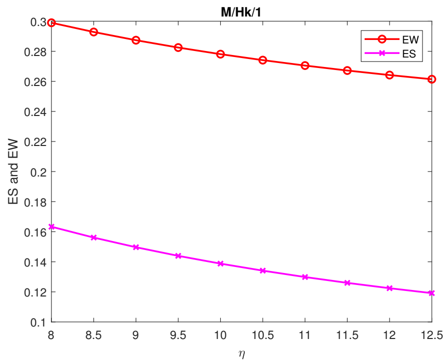


Figure 9: *Vacation rate (η) vs. ES and EW*

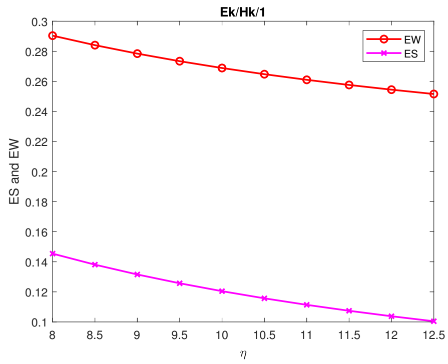


Figure 10: *Vacation rate (η) vs. ES and EW*

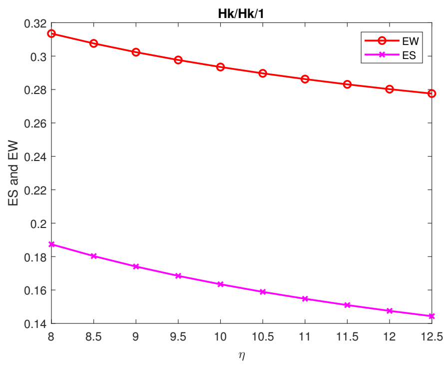


Figure 11: *Vacation rate (η) vs. ES and EW*

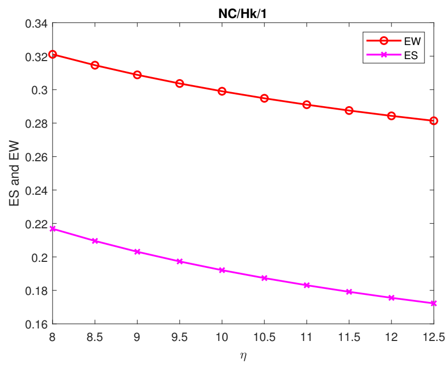


Figure 12: *Vacation rate (η) vs. ES and EW*

We observe from the above Figures $1-12$, that when lifting the vacation rate η, the Expected system size (ES) and Expected waiting time (EW) increase rapidly in the case of arrival by hyper-exponential and slowly in Erlang arrival. Likewise it is high in Erlang service and slow in hyper-exponential. By examining the graphs, we see that the Expected system size (ES) and Expected waiting time (EW) decrease faster for Erlang service rather than those of exponential service and hyper-exponential services.

10.3. Illustrative Example 3

From the three dimensional graphs $13-24$, we explore the effect of the normal service rate (μ_1) and the breakdown rate (τ) on the probability that the server is busy with the normal service

(P_{BNS}). Fix $\lambda = 1$, $\mu_2 = 12$, $\sigma_1 = 8$, $\sigma_2 = 6$, $\eta = 8$, $\varphi = 10$, $\psi = 10$, $p = 0.2$, $q = 0.8$, $b = 0.1$.

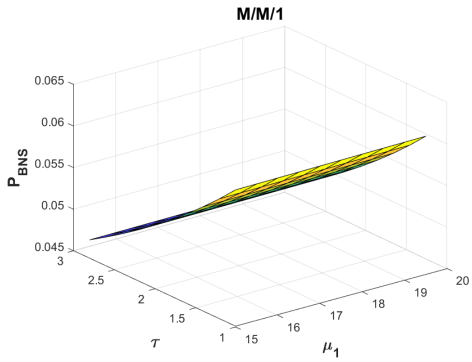


Figure 13: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

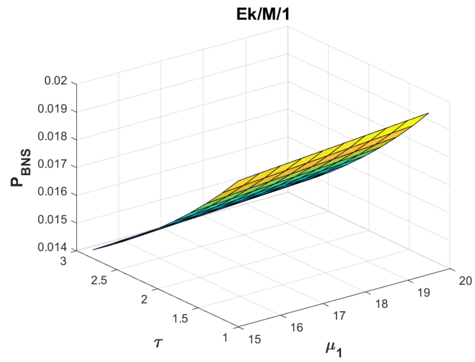


Figure 14: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

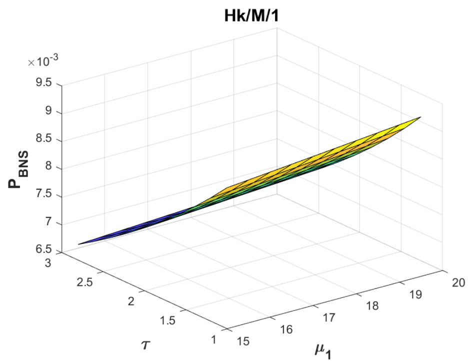


Figure 15: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

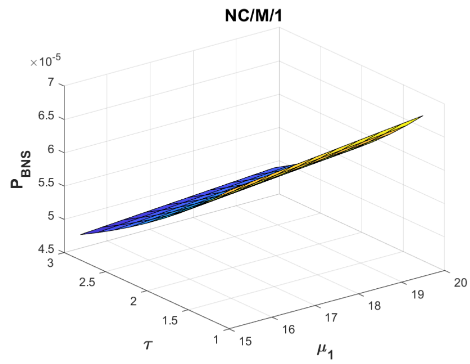


Figure 16: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

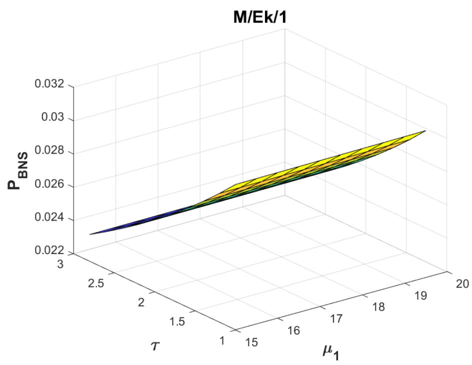


Figure 17: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

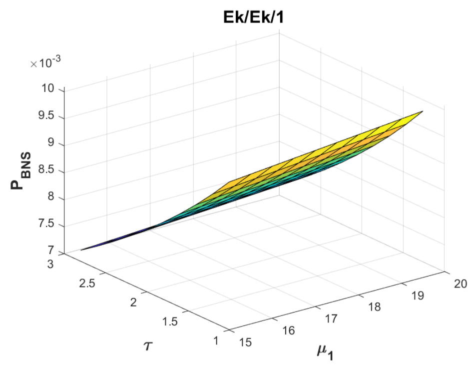


Figure 18: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

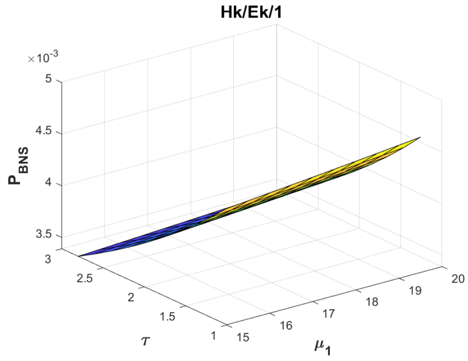


Figure 19: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

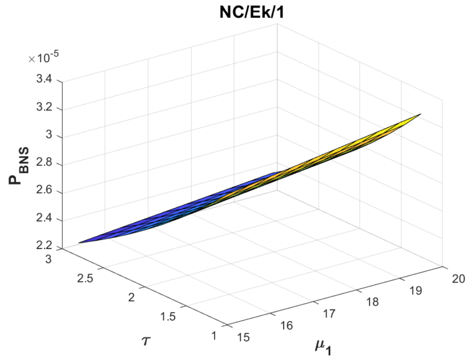


Figure 20: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

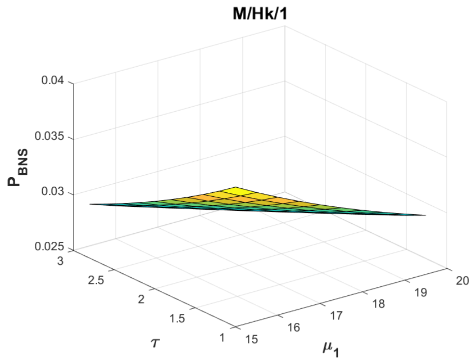


Figure 21: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

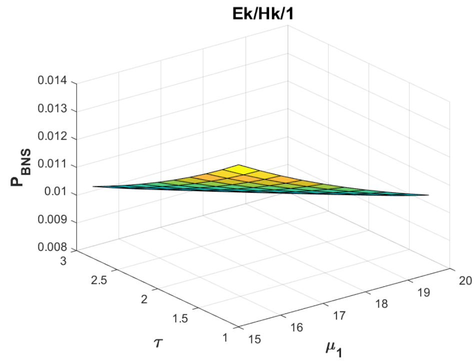


Figure 22: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

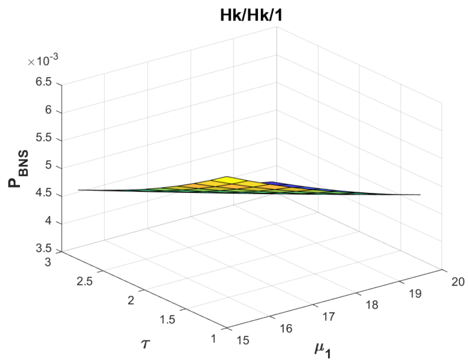


Figure 23: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

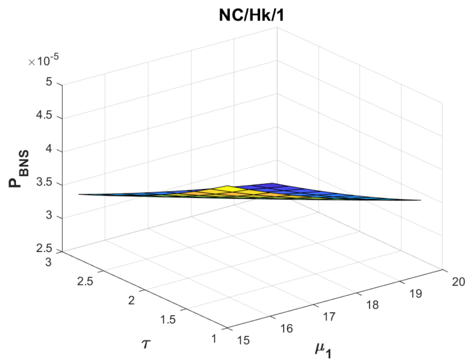


Figure 24: *Normal service rate (μ_1) and Breakdown rate(τ) vs. P_{BNS}*

We observe from the Figures $13 - 24$ that when lifting both the normal service rate (μ_1) and the breakdown rate (τ), the probability that the server is busy with the normal service (P_{BNS}) decreases for various arrival and service patterns. The negative arrival (NC) of MAP increases rapidly rather than that for hyper-exponential arrival. Likewise, the increment rate reduces in Erlang service and increases in hyper-exponential services.

11. Conclusion

In this paper, we have dealt with a single server queueing model with multiple vacations, optional service, close-down, setup, balking, breakdown and phase type repair. The arrival time follows MAP and service and repair times follow phase type distributions. We have presented the busy period analysis and waiting time distribution of this model. We have evaluated the probability of the situation that the server is busy, under repair and on vacation. Particular case has also been discussed. Numerical illustrations and graphical representations are presented in this paper. The future direction of our work can be the investigation of the queueing model by using BMAP for the arrival process with N-policy.

References

[1] Artalejo, J.R., Chakravarthy, S.R. and Lopez-Herrero, M.J. (2007). The Busy period and the Waiting Time Analysis of a $MAP/M/C$ queue with Finite Retrial Group, *Stochastic Analysis and Applications*, 25(2): 445-469.

[2] Ayyappan, G. and Gowthami, R. (2021). A $MAP/PH/1$ Queue with Setup time, Bernoulli vacation, Reneging, Balking, Bernoulli feedback, Breakdown and Repair, *Reliability: Theory and Applications*, 16(2): 191-221.

[3] Ayyappan, G. and Thilagavathy, K. (2020). Anaysis of $MAP/PH/1$ Queueing model with Setup, Closedown, Multiple vacations, Stand by server, Breakdown, Repair and Reneging, *Reliability: Theory and Applications*, 15(2): 104-143.

[4] Chakravarthy, S.R. (2011). Markovian Arrival Process, Wiley Encyclopaedia of Operation Research and Management Science.

[5] Chakravarthy, S.R. and Agarwal, A. (2003). Analysis of a Machine Repair Problem with an Unreliable server and Phase type repairs and services, *Wiley Periodicals, Inc. Naval Research Logistics*, 50(5): 462-480.

[6] Chakravarthy, S.R., Shruti, Kulshrestha, R. (2020). A Queueing model with server Breakdowns, Repairs, Vacations and Backup server, *Operations Research Perspectives*, 7 : 1-13.

[7] Chakravarthy, S.R. (2013). Analysis of $MAP/PH1, PH2/1$ Queue with Vacations and Optional secondary services, *Applied Mathematical Modelling*, 37(20-21) : 8886-8902.

[8] Choudhury, G. and Paul, M. (2006). A two phases queueing system with Bernoulli vacation schedule under Multiple vacation policy, *Statistical Methodology*, 3(2) : 174-185.

[9] Gray, W.J., Wang, P.P. and Scott, M. (2000). A vacation queueing model with service Breakdowns, *Applied Mathematical Modelling*, 24(5-6): 391-400.

[10] Ke, J.-C., Chang, F.-M. and Liu, T.-H. (2017). $M/M/C$ balking retrial queue with vacation, *Quality Technology Quantitative Management*, 16(1) : 54-66.

[11] Krishna Kumar, B., Anbarasu, S. and Anantha Lakshmi, S.R. (2005). Performance analysis for queueing systems with close down periods and server under maintenance, *International Journal of Systems Science*, 46(1) : 88-110.

[12] Kumar, B. (2018). Unreliable bulk queueing model with Optional services, Bernoulli vacation schedule and balking, *International Journal of Mathematics in Operational Research*, 12(3) : 293-316.

[13] Kumar, R. and Sharma, S. (2012). An $M/M/1/N$ Queueing model with Retention of Reneged customers and Balking, *American Journal of Operational Research*, 2(1) : 1-5.

[14] Kulkarni, V.G. and Choi, B.D. (1990). Retrial queues with server subject to breakdowns and repairs, *Queueing Systems*, 7(2) : 191-208.

[15] Latouche, G. and Ramaswami, V. (1999). Introduction of Matrix-Analytic Methods in stochastic modeling, *Society for Industrial and Applied Mathematics*, Philadelphia.

[16] Latouche, G. and Ramaswami, V. (1993). A Logarithmic Reduction Algorithm for Quasi-Birth-Death Processes. *Journal of Applied Probability*, 30 : 650-674.

[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(3) : 676-705.

[18] Maragathasundari, S., Asha, N. and Swedheetha, C. (2018). $M/G/1$ Queue with compulsory Short Vacation and Reneging during Optional Long vacation, *International Journal of Pure and Applied Mathematics*, 118(7) : 317-323.

[19] Medhi,J.(1994). Stochastic processes, *J.Wiley*, New York.

[20] Kadi,M., Bouchentouf, A.A. and Yahiaoui, L. (2020). On a Multiserver Queueing System with Customers Impatience Until the End of Service Under Single and Multiple Vacation Policies, *Applications and Applied Mathematics*, 15(2) : 740-763.

[21] Neuts, M.F. (1981). Matrix-geometric solutions in stochastic models: An algorithmic approach, *The Johns Hopkins University Press*, Baltimore, London.

[22] Neuts, M.F. and Lucantoni, D.M. (1979). A Markovian queue with N servers subject to Breakdowns and Repairs, *Management Science*, 25(9) : 849-861.

[23] Neuts, M.F. and Ramalhoto,M.F. (1984). A Service model in which the server is required to search for customers, *Journal of Applied Probability*, 21(1) : 157-166.

[24] Neuts, M.F. (1979). A Versatile Markovian point process, *Journal of Applied Probability*, 16(4) : 764-779.

[25] Suganya, R., Nkenyereye, L., Anbazhagan, N., Amutha, S., Kameshwari, M., Acharya, S. and Joshi, G.P. (2021). Perishable Inventory system with N-policy, MAP Arrivals and Impatient customers, *Mathematics*, 9(13) : 1-15.

[26] 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.

[27] Wang, Q. and Zhang, B. (2018). Analysis of a Busy period queueing system with Balking, Reneging and Motivating, *Applied Mathematical Modelling*, 64 : 480-488.

[28] Niu, Z., Shu, T. and Takahashi, Y. (2003). A Vacation queue with Setup and Closedown times and batch Markovian Arrival Processes, *Performance Evaluation*, 54(3) : 225-248.

Reliability and Sensitivity Analysis of Two Non-Identical Unit Standby System Subject to Pre-operation Random Inspection of Standby Unit

Amit Manocha

•

Department of Applied Sciences, TITS Bhiwani, Haryana, India
amitmanocha80@yahoo.com

Anil Kumar Taneja*

•

Department of Mathematics, Galgotias University, Greater Noida, UP, India
draniltanejagu@gmail.com

Gulshan Lal Taneja

•

Department of Mathematics, Maharshi Dayanand University, Rohtak, Haryana, India
drgtaneja@gmail.com

*Corresponding Author

Abstract

This paper examines the stochastic behavior of standby redundant system having two non-identical units. The system comprised of main unit and non-identical cold standby unit. When the main unit collapses, standby unit is exposed to operable conditions. Due to long-time and non-use of standby unit, though with small chances, it is observed that standby unit gets corrupt and becomes inoperable even in standby mode. Further, it demands repair/maintenance to make it worth-operating. Henceforth, it is considered to perform random inspection of standby unit to ensure that whether it is in operable condition or not. Inspection as well as repair both the tasks are performed by single repair facility. semi-Markov and regenerative processes are applied to derive expressions for the system performance indices. Profit function and bounds (upper/lower) for various costs involved are evaluated. Numerical study has been performed to illustrate the behavior of model developed. Sensitivity and relative sensitivity analysis has also been done for MTSF and steady-state availability.

Keywords: Reliability, Pre-operation random inspection, Cost-benefit analysis, Bounds, Sensitivity analysis

1. Introduction

Technological advances in recent decades have paved the way for numerous complicated and sophisticated systems. The ever increasing tech savvy inclination of consumers urges industries to introduce automation in their industrial process. Therefore, the need of hour is reduction in failures, availability and improvement in operational capacity of such systems. Redundancy is technique by which a system can be made highly reliable. Standby redundant systems have been used at a large scale in automation industry especially in computer and network, telecommunication and power systems. The two unit standby system and the various issues arising during the usage of such systems like switch over and activation time of standby unit, imperfect switching, random change of standby unit etc. have been addressed very extensively by several researchers. A standby system with switching device for activating standby unit and repaired failed unit for operations was investigated by Singh and Singh [1]. Mokaddis et al. [2] analyzed reliability models for standby system. Different working modes of the operative unit were taken into account. The perfect or imperfect switching of standby unit by assuming arbitrary distributions for failure and repair times were also studied.

Considering the activation time of standby unit, economic study of two-unit standby system was performed by Gupta et al. [3]. El-Said and El-Sherbeny [4] investigated profitability function for standby system, wherein the operative and standby unit interchanged randomly. Parashar and Taneja [5] analyzed stochastically hot standby PLC system. The study was carried out by collecting real data from various industries. Imperfect switching of standby unit as well as repairman patience time was studied by Rashad et al. [6]. A standby system with different failure types was discussed by Mahmoud and Mosherf [7].The preventive maintenance of online unit was also done when its operative time reaches to time t, subject to the availability of standby unit. Mathew et al. [8] analyzed two-unit working in parallel configuration casting plant system. Different kinds of failures were taken into the consideration. Jain and Rani [9] used Markov process to obtain availability characteristics for the standby system having switching failure and reboot delay. Manocha and Taneja [10] discussed two stages of repair for standby system. Jia et al. [11] compared perfect and imperfect switching policies for standby system. Barak et al. [12] investigated standby system, in which inspection of failed standby unit was conducted to confirm its reparability status. Wang et al. [13] investigated a warm standby system. The failures due to hardware and human errors were considered in their study and priority in use was given to main unit. Profit analysis was not done by the authors. El-Sherbeny et al. [14] discussed the idea of change between active unit and standby unit after random amount of time. Eventually, it can be concluded that certain technical issues that affect operational capacity of the system needs to be addressed as a prerequisite for standby units. Keeping this in view, the present article examines a two non-identical unit cold standby system, wherein standby unit may be inspected randomly to see as to whether it is worth useable or not. Sensitivity analysis with regard to MTSF and availability has also been done.The present paper is organised as follows.

System description and assumptions made to carry out the analysis are given in Section 2.Notations, different states and method used in the study are cited in Section 3, 4 and 5 respectively.In Section 6 stochastic model for the system (as defined in Section 2) is developed. Explicit expressions for different performance denoting characteristics of the system, profit and sensitivity function are derived in Section 7, 8 and 9 respectively. Numerical discussions are made in Section 10. Concluding remarks are stated in Section 11.

2. System Description and Assumptions

Proposed system consists of operative main unit and non-identical cold standby unit. Whenever, main unit get fail, the standby unit starts working and main unit goes for repair. There is a possibility that due to long-time non-use of standby unit in non-operative mode, it may be degraded and may become inoperable. The standby unit is inspected randomly to check either it can be made operable or it is inoperable due to degradation. Immediately the inoperable standby unit goes under repair/maintenance of the repairman. The repair process follows the first-come-first served (FCFS) rule. A single repair facility is considered for the system which takes cares of repair as well as inspection related activities. We use regenerative and semi-Markov process to obtain the various performance indicating characteristics of the system like Reliability, MTSF, point wise and transient availability, expected number of visits and time taken by repairman for repairing/inspecting the units. Finally these measures are used to formulate the profit and sensitivity function.The life time distribution of both the units is taken as exponential, whereas other time distribution are considered general. After each repair, unit is supposed to works like new one. The random variables used in developing stochastic model are independent.

3. Nomenclature

The notations for various rates/probabilities/pdf/states are:

λ/α: failure rate of main/standby unit
p/q: probability of operable/inoperable standby unit
p_1/q_1: probability of operable/inoperable standby unit after random inspection
W_{iF}/W_{iI}: P[repairman is engaged in regenerative state i for repair/inspection at instant t without switching to any other state]
®/ ©: symbol of Stieljes/ Laplace convolution.
E_0: Initial state of system
$g(t)(G(t))/g_1(t)(G_1(t))$: pdf (cdf) of repair time of main / standby unit
$h(t)(H(t))/i(t)(I(t))$: pdf (cdf) of time to/ time of inspection of standby unit
$q_{ij}(t)(q_{ij}^{(k)}(t))/Q_{ij}(t)(Q_{ij}^{(k)}(t))$: pdf/cdf of transition time from regenerative state i to j (or via non-regenerative state k).

Refer [5] for rest of the nomenclature used in the study

4. State of the System

The various states of the system at certain time point are described as:

State 0:(M_o, S)	State 1:(M_o, S_i)	State 2:(M_r, S_{wr})
State 3:(M_r, S_o)	State 4:(M_{wr}, S_I)	State 5:(M_o, S_r)
State 6:(M_{wr}, S_r)	State7:(M_R, S_{wr})	State 8:(M_{wr}, S_R)

where,

M_o: main unit is operative
S: standby unit
S_i: standby unit is under inspection
M_r: main unit is under repair
S_{wr}: standby unit is inline to get repaired
S_o: standby unit is operative
M_{wr}: main unit is waiting for repair
S_I: Inspection of standby unit is continued from last state
S_r: repair of standby unit is in progress
M_R: repair of main unit is in progress from last state
S_R: repair of standby unit is in progress from last state

5. Material and Methods

The time point at which system conditions are no longer relevant to system situation before to that time point are referred to as regenerative point, and the corresponding state is known to it as regenerative state otherwise non-regenerative state. In the model being discussed, when the repair/inspection is considered from previous state, the state is non-regenerative. The repair/inspection time distribution has been taken arbitrary; whereas the state where operation is continued from the previous state is the regenerative state as the failure time has been considered to follow exponential distribution which has the memory less property. Therefore the process is not purely Markov and hence semi-Markov (Branson and Shah[15]) process and regenerative process (Srinivasan and Gopalan [16]) have been used.

6. Stochastic Model

The transition between various states as described in Section 4 are shown in Fig.1. The state space is ξ=(0,1,2,3,4,5,6,7,8), where Ω=(2,6,7,8) and ϕ= (4) are failure and down state space respectively. By definition of regenerative process and assumptions made the sets ω=(0,1,2,3,5,6) and $\overline{\omega}$=(4,7,8) represents set of regenerative and non-regenerative states respectively. The transition densities

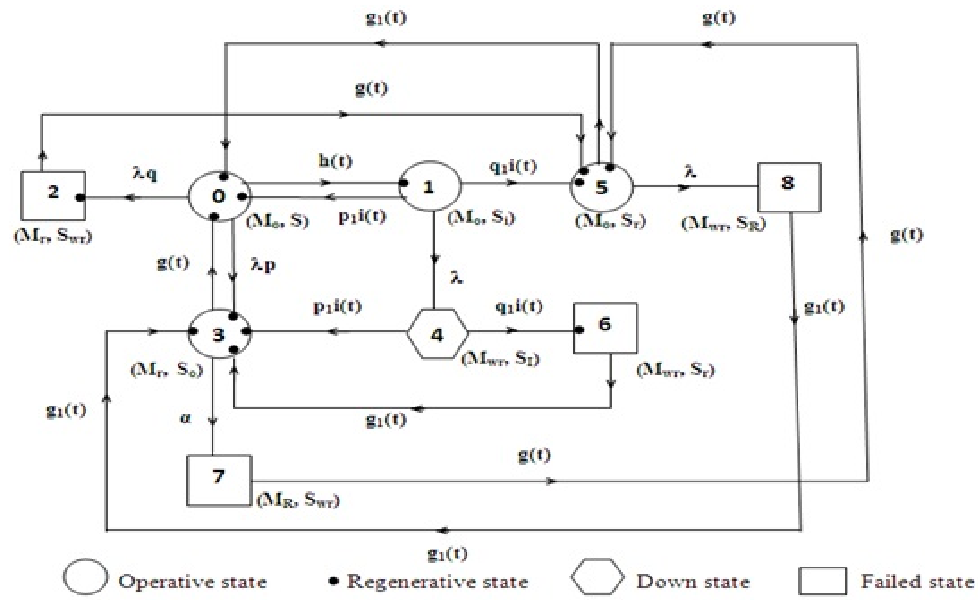


Figure 1: *State transition diagram*

from state i to j (or via k) are:

$$
\begin{aligned}
&q_{01}(t) = e^{-\lambda t}h(t), && q_{02}(t) = \lambda q e^{-\lambda t}\overline{H}(t), && q_{03}(t) = \lambda p e^{-\lambda t}\overline{H}(t)\\
&q_{10}(t) = p_1 e^{-\lambda t}i(t), && q_{13}^{(4)}(t) = p_1(\lambda e^{-\lambda t}©1)i(t), && q_{15}(t) = q_1 e^{-\lambda t}i(t)\\
&q_{16}^{(4)}(t) = q_1(\lambda e^{-\lambda t}©1)i(t), && q_{25}(t) = g(t), && q_{30}(t) = e^{-\alpha t}g(t)\\
&q_{37}(t) = \alpha e^{-\alpha t}\overline{G}(t), && q_{35}^{(7)}(t) = (\alpha e^{-\alpha t}©1)g(t), && q_{50}(t) = e^{-\lambda t}g_1(t)\\
&q_{53}^{(8)}(t) = (\lambda e^{-\lambda t}©1)g_1(t), && q_{58}(t) = \lambda e^{-\lambda t}\overline{G}_1(t), && q_{63}(t) = g_1(t)
\end{aligned}
\tag{1}
$$

Mean sojourn time (μ_i) in state $i \in \omega$ is

$$
\begin{aligned}
&\mu_0 = \int_0^\infty e^{-\lambda t}\overline{H}(t)dt, && \mu_1 = \int_0^\infty e^{-\lambda t}\overline{I}(t)dt, && \mu_2 = \int_0^\infty \overline{G}(t)dt\\
&\mu_3 = \int_0^\infty e^{-\alpha t}\overline{G}(t)dt, && \mu_5 = \int_0^\infty e^{-\lambda t}\overline{G}_1(t)dt, && \mu_6 = \int_0^\infty \overline{G}_1(t)dt
\end{aligned}
\tag{2}
$$

Further, defining

$$m_{ij} = E(q_{ij}(t)) = \int_0^\infty t q_{ij}(t)dt \tag{3}$$

and

$$m_{ij}^{(k)} = E(q_{ij}^{(k)}(t)) = \int_0^\infty t q_{ij}^{(k)}(t)dt \tag{4}$$

we have,

$$\begin{aligned} m_{01}+m_{02}+m_{03}&=\mu_0, & m_{10}+m_{15}+m_{13}^{(4)}+m_{16}^{(4)}&=K_1(say)\\ m_{25}&=\mu_2, & m_{30}+m_{37}&=\mu_3\\ m_{30}+m_{35}^{(7)}&=\mu_2, & m_{50}+m_{58}&=\mu_5\\ m_{50}+m_{53}^{(8)}&=\mu_6, & m_{63}&=\mu_6 \end{aligned} \tag{5}$$

7. System Performability Measures

7.1. System Reliability

If $mf_i(t)$ represents the cdf of time, taken by the system to transit from state $i, i \in \omega$ to a failed state, then from transition diagram we have

$$mf_0(t) = Q_{01}(t)®mf_1(t) + Q_{03}(t)®mf_3(t) + Q_{02}(t) \tag{6}$$

$$mf_1(t) = Q_{10}(t)®\, mf_0(t) + Q_{13}^{(4)}(t)®mf_3(t) + Q_{15}(t)®mf_5(t) + Q_{16}^{(4)}(t) \tag{7}$$

$$mf_3(t) = Q_{30}(t)®mf_0(t) + Q_{37}(t) \tag{8}$$

$$mf_5(t) = Q_{50}(t)®mf_0(t) + Q_{58}(t) \tag{9}$$

Making use of Laplace-Stieljes transformation for eqns. (6)-(9),the expressions obtained for $mf_0^{**}(s)$, reliability{R(t)} of the system and MTSF (mean time to system failure) are

$$mf_0^{**}(s) = L(s)/D(s) \tag{10}$$

$$R(t) = \mathcal{L}^{-1}\{1 - mf_0^{**}(s)/s\} \tag{11}$$

$$MTSF = \int_0^\infty R(t)dt = L/D \tag{12}$$

where

$$L(s) = Q_{01}^{**}(s)\{Q_{13}^{(4)**}(s)Q_{37}^{**}(s) + Q_{15}^{**}(s)Q_{58}^{**}(s) + Q_{16}^{(4)**}(s)\} + Q_{03}^{**}(s)Q_{37}^{**}(s) + Q_{02}^{**}(s) \tag{13}$$

$$D(s) = 1 - Q_{01}^{**}(s)\{Q_{10}^{**}(s) + Q_{13}^{(4)**}(s)Q_{30}^{**}(s) + Q_{15}^{**}(s)Q_{50}^{**}(s)\} - Q_{03}^{**}(s)Q_{30}^{**}(s) \tag{14}$$

$$L = \mu_0 + p_{01}K_1 + (p_{01}p_{13}^{(4)} + p_{03})\mu_3 + p_{01}p_{15}\mu_5 \tag{15}$$

$$D = 1 - p_{01}p_{10} - p_{01}p_{13}^{(4)}p_{30} - p_{01}p_{15}p_{50} - p_{03}p_{30} \tag{16}$$

7.2. System Availability

Let $W_i(t)$= P[system is in operative state $i, i \in \omega$, instead of transferring either to any state $j, j \in \omega$ or to itself via state $k, k \in \overline{\omega}$], Then

$$W_0(t) = e^{-\lambda t}\overline{H}(t) \tag{17}$$

$$W_1(t) = e^{-\lambda t}\overline{I}(t) \tag{18}$$

$$W_3(t) = e^{-\alpha t}\overline{G}(t) \tag{19}$$

$$W_5(t) = e^{-\lambda t}\overline{G}_1(t) \tag{20}$$

Defining $AV_i(t)$=P[system is operative at instant t | $E_0 = i,\ i \in \omega$]. Referring to contentions of regenerative process and from transition state diagram, the availabilities $AV_i(t)$ satisfies the relations

$$AV_0(t) = W_0(t) + q_{01}©AV_1(t) + q_{02}©AV_2(t) + q_{03}©AV_3(t) \tag{21}$$

$$AV_1(t) = W_1(t) + q_{10}©AV_0(t) + q_{13}^{(4)}©AV_3(t) + q_{16}^{(4)}©AV_6(t) + q_{15}©AV_5(t) \tag{22}$$

$$AV_2(t) = q_{25}©AV_5(t) \tag{23}$$

$$AV_3(t) = W_3(t) + q_{30}©AV_0(t) + q_{35}^{(7)}©AV_5(t) \tag{24}$$

$$AV_5(t) = W_5(t) + q_{50}©AV_0(t) + q_{53}^{(8)}©AV_3(t) \tag{25}$$

$$AV_6(t) = q_{63}©AV_3(t) \tag{26}$$

Using Laplace transformation and method of determinants for eqns. (21)-(26), we obtain

$$AV_0^*(s) = L_1(s)/D_1(s) \tag{27}$$

The system's transient and steady-state availability are

$$AV_0(t) = \mathcal{L}^{-1}\{L_1(s)/D_1(s)\} \tag{28}$$

$$AV_\infty = \lim_{t\to\infty} AV_0(t) = \lim_{s\to 0} sAV_0^*(s) = L_1/D_1 \tag{29}$$

where,

$$\begin{aligned} L_1(s) = \{1 - q_{35}^{(7)*}(s)q_{53}^{(8)*}(s)\}\{W_0^*(s) + q_{01}^*(s)W_1^*(s)\} + \{q_{01}^*(s)q_{16}^{(4)*}(s)q_{63}^*(s) + q_{03}^*(s) \\ + q_{01}^*(s)q_{13}^{(4)*}(s)\}\{W_3^*(s) + q_{35}^{(7)*}(s)W_5^*(s)\} + \{q_{01}^*(s)q_{15}^*(s) \\ + q_{02}^*(s)q_{25}^*(s)\}\{q_{53}^{(8)*}(s)W_3^*(s) + W_5^*(s)\} \end{aligned} \tag{30}$$

$$\begin{aligned} D_1(s) = \{1 - q_{35}^{(7)*}(s)q_{53}^{(8)*}(s)\}\{1 - q_{01}^*(s)q_{10}^*(s)\} - q_{02}^*(s)q_{25}^*(s)q_{50}^*(s) - q_{01}^*(s)q_{15}^*(s)q_{50}^*(s) \\ - q_{30}^*(s)\{q_{03}^*(s) + q_{03}^*(s)q_{35}^{(7)*}(s)q_{50}^*(s) + q_{01}^*(s)q_{13}^{(4)*}(s) + q_{01}^*(s)q_{15}^*(s)q_{53}^{(8)*}(s) \\ + q_{02}^*(s)q_{25}^*(s)q_{53}^{(8)*}(s)\} - q_{01}^*(s)q_{16}^{(4)*}(s)q_{63}^*(s)\{q_{30}^*(s) + q_{35}^{(7)*}(s)q_{50}^*(s)\} \end{aligned} \tag{31}$$

$$\begin{aligned} L_1 = (1 - p_{35}^{(7)}p_{53}^{(8)})(\mu_0 + p_{01}\mu_1) + \{p_{01}(p_{13}^{(4)} + p_{16}^{(4)} + p_{15}p_{53}^{(8)}) + p_{03} + p_{02}p_{53}^{(8)}\}\mu_3 + \{p_{01}(p_{13}^{(4)}p_{35}^{(7)} \\ + p_{16}^{(4)}p_{35}^{(7)} + p_{03}p_{35}^{(7)} + p_{15}) + p_{02}p_{25}\}\mu_5 \end{aligned} \tag{32}$$

$$\begin{aligned} D_1 = (1 - p_{35}^{(7)}p_{53}^{(8)})(\mu_0 + p_{01}K_1 + p_{02}\mu_2) + (1 - p_{01}p_{10} - p_{01}p_{15}p_{50} - p_{02}p_{50})\mu_2 + \{(1 - p_{01}p_{10})p_{35}^{(7)} \\ + p_{02}p_{30} + p_{01}p_{15}p_{50} + p_{16}^{(4)}(1 - p_{35}^{(7)}p_{53}^{(8)})\}\mu_6 \end{aligned} \tag{33}$$

Employing the same procedure as discussed in Sub-section 7.2, other performability measures of the system are as follows:

7.3. Busy Period Analysis

7.3.1 Expected Time for Repairing the Failed Unit

Let $B_i(t)$=P[repairman is engaged in repair at instant t | $E_0 = i$, $i \in \omega$].The expected time taken by repairman in repairing the failed unit is

$$B_\infty = \lim_{t\to\infty} B_0(t) = \lim_{s\to 0} sB_0^*(s) = \lim_{s\to 0} s\{L_2(s)/D_1(s)\} = L_2/D_1 \tag{34}$$

where,

$$\begin{aligned} L_2(s) = \{1 - q_{35}^{(7)*}(s)q_{53}^{(8)*}(s)\}\{q_{01}^*(s)q_{16}^{(4)*}(s)W_{6F}^*(s) + q_{02}^*(s)W_{2F}^*(s)\} + q_{01}^*(s)q_{15}^*(s)q_{53}^{(8)*}(s) \\ + \{q_{03}^*(s) + q_{01}^*(s)q_{16}^{(4)*}(s)q_{63}^*(s) + q_{01}^*(s)q_{13}^{(4)*}(s)\}\{W_{3F}^*(s) + q_{35}^{(7)*}(s)W_{5F}^*(s)\} \\ + q_{02}^*(s)q_{25}^*(s)\{q_{53}^{(8)*}(s)W_{3F}^*(s) + W_{5F}^*(s)\} \end{aligned} \tag{35}$$

$$\begin{aligned} L_2 = p_{01}\{(p_{13}^{(4)} + p_{16}^{(4)})(\mu_2 + p_{35}^{(7)}\mu_6) + p_{15}p_{53}^{(8)} + (1 - p_{35}^{(7)}p_{53}^{(8)})\mu_6\} + p_{02}\{(1 - p_{35}^{(7)}p_{53}^{(8)})\mu_2 \\ + p_{53}^{(8)}\mu_2 + \mu_6\} + p_{03}(\mu_2 + p_{35}^{(7)}\mu_6) \end{aligned} \tag{36}$$

7.3.2 Expected Time for Inspection of the Standby Unit

Letting $I_i(t)$=P[repairman remains involved in inspection at time t | $E_0 = i, i \in \omega$].The expected time for which standby unit is under inspection, in steady-state is

$$I_\infty = \lim_{t\to\infty} I_0(t) = \lim_{s\to 0} sI_0^*(s) = \lim_{s\to 0} s\{L_3(s)/D_1(s)\} = L_3/D_1 \tag{37}$$

where,

$$L_3(s) = q_{01}^*(s)\{1 - q_{35}^{(7)*}(s)q_{53}^{(8)*}(s)\}W_{1I}^*(s) \tag{38}$$

$$L_3 = p_{01}\{1 - p_{35}^{(7)}p_{53}^{(8)}\}K_1 \tag{39}$$

7.4. Expected Number of Visits by the Repairman

If M(t) denotes the expected number of visits by repairman in the time interval (0,t] then $NV_i(t)$= E{M(t) | $E_0 = i, i \in \omega$}. In steady-state, the number of visits are

$$NV_\infty = \lim_{t\to\infty} NV_0(t) = \lim_{s\to 0} sNV_0^{**}(s) = \lim_{s\to 0} s\{L_4(s)/D_1(s)\} = L_4/D_1 \tag{40}$$

where,

$$L_4(s) = \{Q_{01}^{**}(s) + Q_{02}^{**}(s) + Q_{03}^{**}(s)\}\{1 - Q_{35}^{(7)**}(s)Q_{53}^{(8)**}(s)\} \tag{41}$$

$$L_4 = (1 - p_{35}^{(7)}p_{53}^{(8)}) \tag{42}$$

$D_1(s)$ and D_1 are specified in eqns. (31) and (33) respectively. Now, derived indexes are used to perform cost-benefit analysis in the succeeding section.

8. Cost-Benefit Analysis

As we know, the profit for any manufacturing system is the difference of expected revenue and expected recurring cost.Utilizing eqns.(29), (34), (37) and (40), the profit function for the defined system, in steady-state, is

$$P_\infty = (R_0 AV_\infty) - (C_B B_\infty + C_I I_\infty + C_V V_\infty) \tag{43}$$

where,R_0=Revenue generated per unit time
C_B/C_I= Recurring cost per unit time for repairing/inspecting the units
C_V= Recurring cost at per visit of repairman
For the system to be profitable, the eq.(43) is used to obtain the bounds for revenue/cost(s), which are shown in Table 1.

Table 1: *Bounds for revenue and various cost(s)*

Revenue/Cost	Bound	Value
R_0	Lower	$(C_B B_\infty + C_I I_\infty + C_V V_\infty)/AV_\infty$
C_B	Upper	$(R_0 AV_\infty - C_I I_\infty - C_V V_\infty)/B_\infty$
C_I	Upper	$(R_0 AV_\infty - C_B B_\infty - C_V V_\infty)/I_\infty$
C_V	Upper	$(R_0 AV_\infty - C_B B_\infty - C_I I_\infty)/V_\infty$

9. Sensitivity and Relative Sensitivity Analysis

Sensitivity analysis is performed to find out how the variation in incoming variable affects the specific outgoing variable under certain specific conditions. Since, there is significance difference between the values of incoming variables, so to compare their effects on outgoing variables, relative sensitivity function is used. Relative sensitivity function is defined as percentage change that results from the percentage change in one of the variable. The sensitivity and relative sensitivity functions for MTSF and availability (AV_∞) are formulated as:

$$\pi_k = \frac{\partial MTSF}{\partial k} \tag{44}$$

$$\delta_k = \pi_k(\frac{k}{MTSF}) \tag{45}$$

$$\rho_k = \frac{\partial AV_\infty}{\partial k} \tag{46}$$

$$\tau_k = \rho_k(\frac{k}{AV_\infty}) \tag{47}$$

where $k = \lambda, \alpha, \beta, \beta_1, \gamma, \theta$.

10. Results and Discussion

In this section numerical analysis is done to illustrate the developed stochastic model.Input/Output variables are specified in the subsections 10.1 and 10.2 respectively for further discussions.

10.1. Input Variables

The repair time of main/standby unit, time to inspection and time for inspection of standby unit are supposed to be exponential with parameters β,β_1,θ and γ respectively. Then
$G(t) = 1 - exp(-\beta t)$,$G_1(t) = 1 - exp(-\beta_1 t)$, $H(t) = 1 - exp(-\theta t)$ and $I(t) = 1 - exp(-\gamma t)$.
Time (t) and various rates/cost(s) are our input variables and their values are taken as:
$\lambda = 0.001, \alpha = 0.008, p = 0.98, q = 0.02, p_1 = 0.95, q_1 = 0.05, \beta_1 = 0.85, \gamma = 10, \beta = 0.65, \theta = 0.004$
$R_0 = 40, C_B = 5000, C_I = 2000, C_V = 2000$.

10.2. Output Variables

Measures including reliability, MTSF, availability, profit and sensitivity functions are output variables as obtained in sections 7, 8 and 9 respectively. Variations in output variables caused by changes in input variables have been investigated and are discussed in the following subsections.

10.3. Trend of Reliability{R(t)} w.r.t. time(t) for varying λ

Taking the other parameter constant, as mentioned in subsection 10.1, the mathematical expressions for reliability {R(t)} of the system for varied λ are as follows:
For $\lambda = 0.001$

$$\begin{aligned} R(t) = 0.991602 + 2.14908 \times 10^{-10} e^{-10.001t} - 3.11 \times 10^{-7} e^{-0.85122t} - 1.81 \times 10^{-5} e^{-0.658973t} \\ + 0.00841596 e^{-0.00382477t} \end{aligned} \tag{48}$$

For $\lambda = 0.002$

$$\begin{aligned} R(t) = 0.983342 + 4.29906 \times 10^{-10} e^{-10.002t} - 6.23 \times 10^{-7} e^{-0.852221t} - 3.6 \times 10^{-5} e^{-0.659946t} \\ + 0.016695 e^{-0.00385044t} \end{aligned} \tag{49}$$

For $\lambda = 0.003$

$$R(t) = 0.975214 + 6.4499 \times 10^{-10} e^{-10.003t} - 9.37 \times 10^{-7} e^{-0.853222t} - 5.4 \times 10^{-5} e^{-0.660919t} + 0.0248404 e^{-0.00387604t} \quad (50)$$

Fig.2 shows the trends of system reliability {R(t)} for varied (t, λ). Clearly, it goes down with the rise in the values of variables t and λ respectively.

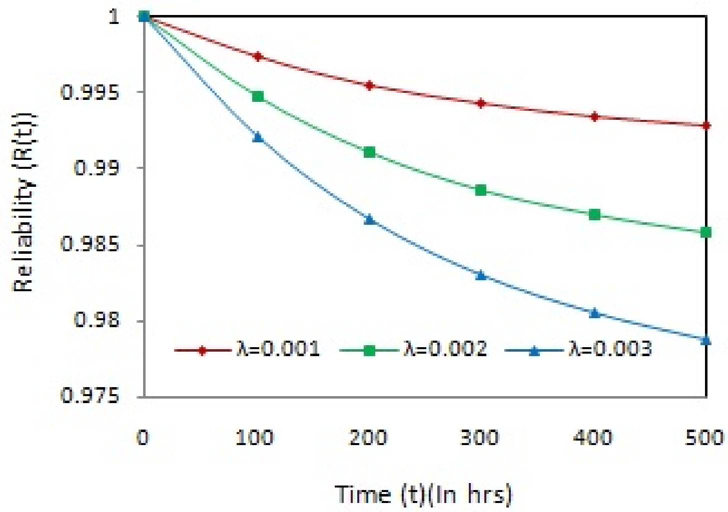


Figure 2: *Reliability {R(t)} w.r.t time (t)*

10.4. Trend of MTSF and Availability (AV_∞) for varying rates

The numerical values of MTSF and availability (AV_∞) are obtained for (λ, β) and (θ, γ) respectively. The other parameters are kept fixed as assumed in subsection 10.1. The results are tabulated as in Table 2 and 3 respectively. It is noted that,
(i) MTSF decreases as λ increases. However, it increases as β increases.
(ii) Availability (AV_∞) increases with the increase in both the parameters θ as well as γ.

Table 2: *MTSF w.r.t. λ for varied β*

λ	MTSF		
	β=0.55	β=0.65	β=0.75
0.0010	29215.21	31146.55	32740.61
0.0011	26563.96	28319.28	29768.06
0.0012	24354.58	25963.23	27290.93
0.0013	22485.11	23969.64	25194.91
0.0014	20882.71	22260.85	23398.31
0.0015	19493.96	20779.90	21841.26
0.0016	18278.80	19484.07	20478.85

Table 3: *AV_∞ w.r.t. θ for varied γ*

θ	AV_∞		
	γ=3	γ=5	γ=10
0.0020	0.9999343	0.9999346	0.9999348
0.0024	0.9999354	0.9999358	0.9999360
0.0028	0.9999364	0.9999369	0.9999371
0.0032	0.9999372	0.9999377	0.9999380
0.0036	0.9999379	0.9999384	0.9999387
0.0040	0.9999384	0.9999390	0.9999394

10.5. Trend of Profit function (P_∞) for varying rates/costs

The trend of profit function (P_∞) with respect to R_0 for varied β and C_B for varied R_0 is revealed by Fig.3 and Fig.4 respectively.Evidently,
(i)With the increase in R_0 and β, P_∞ increases.
(ii)With the increase in C_B, P_∞ decreases but increasing trend of P_∞ is observed with increase in R_0.

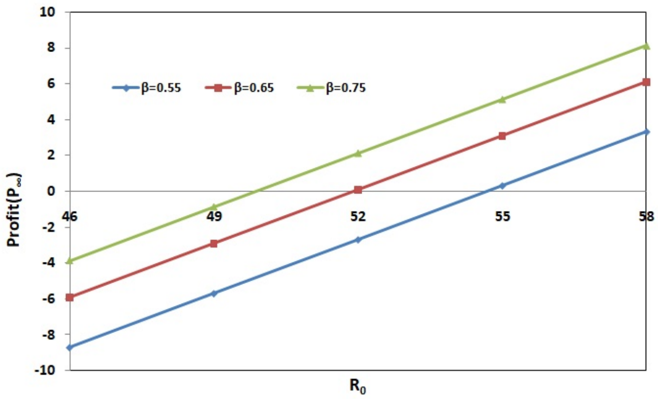


Figure 3: *P_∞ versus R_0 for varied β*

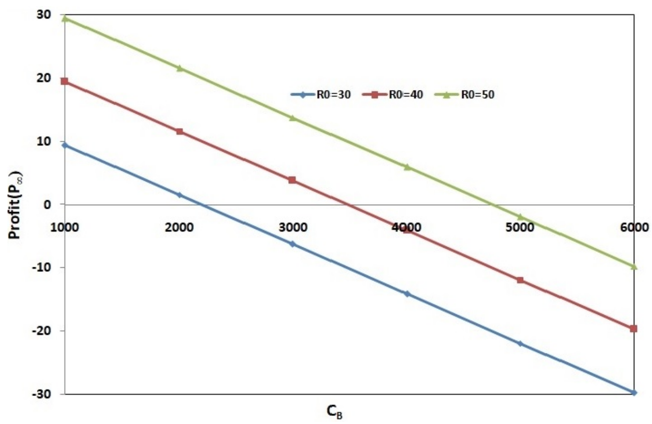


Figure 4: *P_∞ versus C_B for varied R_0*

Bearing economic viability of the system in mind, the bounds obtained for R_0 and C_B are shown in Table 4.

Table 4: *Bounds for revenue/cost*

Revenue/Cost	Varied Parameter	Bounds For Profitability($P_\infty > 0$)
R_0	β=0.55	R_0>54.69
	β=0.65	R_0>51.90
	β=0.75	R_0>49.86
C_B	R_0=30	C_B<2202.7
	R_0=40	C_B<3479.5
	R_0=50	C_B<4756.8

10.6. Numerical calculations for sensitivity analysis

Using the values of incoming variables (as considered in subsection 10.1) Table 5 and 6 represents the values of sensitivity and relative sensitivity functions (defined in section 9) for MTSF and AV_∞ respectively.

Table 5: *Sensitivity and Relative sensitivity of MTSF w.r.t. different rates*

Variable	MTSF	
(k)	π_k=$\frac{\partial MTSF}{\partial k}$	$\delta_k = \pi_k(\frac{k}{MTSF})$
λ	-8174507	-1.048
α	-685533	-0.352
β	8710.824	0.363
β_1	129.51	0.007
γ	0.567	0.0004
θ	-28955	-0.0074

Table 6: *Sensitivity and Relative sensitivity of availability (AV_∞) w.r.t. different rates*

Variable	Availability (AV_∞)	
(k)	ρ_k=$\frac{\partial A_\infty}{\partial k}$	$\tau_k = \rho_k(\frac{k}{AV_\infty})$
λ	0.0012	2.4×10^{-6}
α	-0.0056	-4.5×10^{-5}
β	0.0002	1.4×10^{-4}
β_1	-0.0001	-1.06×10^{-4}
γ	2.71×10^{-8}	1.4×10^{-7}
θ	0.0014	5.6×10^{-6}

Considering the absolute values of defined functions, Table 5 and Table 6 reveals that the MTSF is more sensitive with respect to failure rate of main unit λ whereas AV_∞ is impacted more by failure rate of standby unit α. However, the order of incoming variables in which they influence the MTSF and AV_∞ is:
MTSF: $\lambda > \alpha > \theta > \beta > \beta_1 > \gamma$.
$AV_\infty : \beta > \beta_1 > \alpha > \theta > \lambda > \gamma$.

11. Conclusion

This article proposes a probabilistic model for two non-identical units' standby system in which standby unit may be inspected randomly to ensure its operability. Various performability indices are derived. Keeping the cost factor in mind, bounds (lower/upper) for various costs are obtained to account for economic and budgetary constraints. The numerical study has been carried out for exponential case. Sensitivity analysis is performed for MTSF and steady-state availability of the system. The developed model is quite lucrative for any commercial/industrial establishment using such systems, in their production and operational commitments.

Funding

This research received no specific grant.

Disclosure statement

The authors declare that they have no conflict of interest.

References

[1] Singh, S.K. and Singh, R.P. (1989). Two unit standby system with imperfect switching device and maximum activation time. Microelectronics Reliability, 29(5): 717-720.
[2] Mokaddis, G.S., Labib, S.W. and El-Said, KH.M. (1994). Two models for two-dissimilar-unit standby redundant system with three types of repair facilities and perfect or imperfect switch. Microelectronics Reliability, 34(7): 1239-1247.
[3] Gupta, R., Tyagi, V. and Tyagi, P. (1997). Cost-benefit analysis of a two-unit standby system with post-repair,activation time and correlated failures and repairs. Journal of Quality in Maintenance Engineering, 3(1): 55-63.
[4] El-Said, K.M. and El-Sherbeny, M.S. (2005). Profit analysis of a two unit cold standby system with preventive maintenance and random change in units. Journal of Mathematics and Statistics, 1(1): 71-91.
[5] Parashar, B. and Taneja, G. (2007). Reliability and profit evaluation of a PLC hot standby system based on a master- slave concept and two types of repair facilities. IEEE Transactions on Reliability, 56(3): 534-539.
[6] Rashad, A.M., El-Sherbeny, M.S. and Hussien, Z.M. (2009). Cost analysis of a two-unit cold standby system with imperfect switch, patience time and two type of repair. Journal of the Egyptian Mathematical Society, 17(1): 65-81.
[7] Mahmoud, M.A.W. and Moshref, M.E. (2010). On a two-unit cold standby system considering hardware, human error failures and preventive maintenance. Mathematical and Computer Modelling, 51: 736-245.
[8] Mathews, A.G., Rizwan, S.M., Majumdar, M., Ramchandarann, K.P. and Taneja, G. (2011). Reliability analysis of identical two unit parallel CC plant system operative with full installed capacity. International Journal of Performability Engineering, 7(2): 179-187.
[9] Jain, M. and Rani, S. (2013). Availability analysis for repairable system with warm standby, switching failure and reboot delay. International Journal of Mathematics in Operational Research, 5(1): 19-39.
[10] 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.
[11] Jia, X., Chen, H., Cheng, Z. and Guo, B. (2016). A comparison between two switching policies for two-unit standby system. Reliability Engineering and System Safety, 148: 109-118.

[12] Barak, M.S., Yadav, D. and Kumari, S. (2018). Stochastic analysis of a two-unit system with standby and server failure subject to inspection. Life Cycle Reliability and Safety Engineering, 7:23-32.

[13] Wang, J., Xie, N. and Yang, N. (2019). Reliability analysis of a two dissimilar unit warm standby repairable system with priority in use. Communications in Statistics- Theory and Methods, 50(4): 792-814.

[14] El-Sherbeny, M.S. (2017). Stochastic behavior of a two-unit cold standby redundant system under poisson shocks. Arabian Journal for Science and Engineering, 42: 3043-3053.

[15] Branson, M.H. and Shah, B. (1971). Reliability analysis of systems comprised of units with arbitrary repair time distributions. IEEE Transactions on Reliability, 20(4): 217-223.

[16] Srinivasan, S.K. and Gopalan, M.N. (1973). Probabilistic analysis of a two-unit system with a warm standby and a single repair facility. Operations Research, 21(3): 748-754.

RT&A, No 2 (68)
Volume 17, June 2022

THE NEGATIVE BINOMIAL-AKASH DISTRIBUTION AND ITS APPLICATIONS

Rajitha.C.S and Ashly Regi

•

Department of Mathematics, Amrita School of Engineering, Coimbatore,
Amrita Vishwa Vidyapeetham, India.
rajitha.sugun@gmail.com,cs_rajitha@cb.amrita.edu

Abstract

A new two-parameter negative binomial mixture distribution named as negative binomial-Akash distribution is introduced in this paper. The proposed distribution is attained by compounding the negative binomial distribution with the Akash distribution. Some of its special characteristics are also derived, including factorial moments, mean, variance, index of dispersion etc. Furthermore, the behaviour of mean, variance and index of dispersion are discussed. The parameters of the proposed distribution are estimated using the maximum likelihood estimation method. This distribution can be used for modeling overdispersed count data. The usefulness and application of the proposed distribution are illustrated using two actual count data sets.

Keywords: Akash distribution, AIC, BIC, count data, method of maximum likelihood, mixture distribution.

1. Introduction

Count data modeling plays a vital role in the statistical literature. Usually Poisson distribution(PD) is used for analyzing the count data. It is a discrete probability distribution for modeling the number of occurrences of an event in a given period of time. Equi-dispersion is a main feature of PD which means that variance and mean are equal. However, in practice, the count data observed often shows overdispersion with mean smaller than variance (or underdispersion with mean greater than variance). Even in many fields, the count data shows the nature of the overdispersion. When this happens, the PD cannot handle overdispersed count data. To overcome this problem, an extension of PD is applicable. Therefore, the negative binomial distribution (NBD) was used in modeling over dispersed count data. The NBD is in fact a Poisson mixture distribution in which the distribution's parameter itself is considered as a random variable which follows gamma distribution. The NBD is a discrete failure distribution in a Bernoulli test sequence before a predetermined success occurs. The application of NBD can be found in various sectors, such as bio-statistics, accident statistics, actuarial science and economics. Although the NBD allows excessive dispersion, if the count data shows an excessive number of zeros, the NBD also does not work well. As a result, many studies have been conducted to find new distributions which provide better fit for overdispersed count data. Experiments show that mixed distribution, such as Poisson mixture and NB mixture distributions, provides better fit to count data than traditional count distributions. Numerous studies show that mixing PD and NBD with some lifetime probability distributions, such as the exponential distribution, provide better fit to overdispersed count data with an excessive number of zeros. However, studies show that Lindley distribution(LD) is a better model than one based on an exponential distribution. A detailed study of various mathematical properties, parameter estimation and application of LD was conducted by Ghitany et al., [3] and showed that LD is better than the exponential distribution. And this lifetime distribution was introduced by Lindley [5] . However, there are many cases where the LD is not adequate. Therefore, Shanker [14] introduced a new lifetime distribution with a parameter called Akash distribution to model lifetime data, which is more flexible than LD. He derived some important statistical properties of the proposed distribution and the usefulness and applicability of this distribution were also discussed and illustrated with two sets of real life data. In addition to these models many more mixture distributions have been proposed and studied in the literature.

The Poisson mixture distribution was proposed by Sankaran[9] as a mixture of the PD with the LD named Poisson Lindley distribution(PLD). Later Mishra [13] achieved a two-parameter PLD by mixing the PD and the two-parameter Lindley distribution(TP-LD). The two-parameter distribution of Lindley was proposed by Shanker and Mishra [10]. Further, Shanker and Tekie [12] introduced a new quasi PLD by mixing the Poisson distribution with a new quasi Lindley distribution(QLD)[11]. After that Zamani et al., [17] introduced a new mixed PD called the Poisson-weighted Exponential distribution. Moreover, literature shows that NB mixtures provide better fit for count data than the Poisson mixture distributions. A new distribution of NB mixture by combining the NB with LD was proposed by Zamani and Ismail [16] which perform better than PD and NBD for count data. Subsequently Lord and Geedipally [6] analyzed the crash data containing excess number of zeroes using the NB-Lindley distribution(NB-LD) and compared their performance with PD and the NBD and found that the NB-LD works better than PD and the NBD. A new mixture distribution named NB TP-LD was proposed by Denthet et al.,[2]. Later a new three parameter mixed NBD called NB-Erlang distribution was introduced, and the applications of this distribution were performed using two sets of actual count data [4]. Saengthong and Bodhisuwan[8] studied four parameters NB-Crack distribution and estimated the parameters for the NB-Crack distribution using the MLE method and the moment method and these methods were illustrated with an application to accident data. The NB-generalized exponential distribution was introduced by Aryuyuen and Bodhisuwan [1]. Recently a new NB mixture distribution named NB-Sushila distribution was proposed by Yamrubboon et al., [15]. The NB-LD is a special case of this distribution.

In this article we present a new mixed NBD attained by compounding the NBD with the parameters s and $\theta = e^{-\gamma}$ and Akash distribution with parameter α. Furthermore, we derive various properties of the negative binomial-Akash distribution, including factorial moments (FM), mean and second order moment. The parameters of the NB-Akash distribution are derived by moment method and MLE method. We present the performance of the PD, the NBD, NB-LD and NB-Akash distribution using two sets of real data in terms of chi-square test of goodness of fit, log-likelihood, p-value, AIC(Akaike Information Criteria) and BIC(Bayesian Information Criteria) In Section 2, subsection 1 deals with the NBD, subsection 2 discusses the Akash distribution, subsection 3 discusses the proposed distribution called negative binomial-Akash(NB-Akash) distribution and derives its probability mass function. In Section 3 we discuss distributional characteristics such as FM, the mean and the second order moment and the behaviour of mean, variance and index of dispersion(ID). The estimation of the parameters is reported in Section 4. The usefulness and application of the NB-Akash distribution is discussed in Section 5. Finally, the conclusion is discussed in Section 6.

2. Proposed Model

2.1. Negative Binomial Distribution

The NBD is used in cases where the data is overdispersed, that is, the mean smaller than the variance, to model count data. A discrete random variable Z is said to be a NBD with the parameters s and θ if its probability mass function (pmf) is

$$p(z) = \binom{z+s-1}{z}\theta^s(1-\theta)^z; \quad z = 0, 1, 2, ..., s > 0, 0 < \theta < 1 \tag{1}$$

θ is the probability of success and the experiment is repeated many times to obtain s successes. The FM of order m, the mean and the second order moment of NBD are

$$\mu_{[m]}(Z) = \frac{\Gamma(s+m)}{\Gamma(s)} \frac{(1-\theta)^m}{\theta^m}; \qquad m = 0, 1, 2, ... \tag{2}$$

$$E(Z) = \frac{s(1-\theta)}{\theta}$$

$$E(Z^2) = \frac{s(1-\theta)[1+s(1-\theta)]}{\theta^2}$$

The likelihood function (LF) of the NBD is given by

$$\mathcal{L}(s,\theta) = \prod_{j=1}^{n} \binom{s+z_j-1}{z_j} \theta^{ns} (1-\theta)^{\sum_{j=1}^{n} z_j}$$

The log-likelihood(LL) function is

$$\ell(s,\theta) = \sum_{j=1}^{n} \log\binom{s+z_j-1}{z_j} + ns\log(\theta) + \sum_{j=1}^{n} z_j \log(1-\theta)$$

2.2. Akash Distribution

The Akash distribution was proposed by [14]. It is a modified version of Lindley distribution. The probability density function(pdf) of the one parameter Akash distribution with parameter α is

$$f(z;\alpha) = \frac{\alpha^3}{\alpha^2+2}(1+z^2)\mathrm{e}^{-\alpha z}; \qquad z > 0, \alpha > 0 \tag{3}$$

The mean, variance and the MGF of the Akash distribution are

$$E(Z) = \frac{\alpha^2+6}{\alpha(\alpha^2+2)}$$

$$V(Z) = \frac{\alpha^4+16\alpha+12}{\alpha^2(\alpha^2+2)^2}$$

$$M_Z(t) = \frac{\alpha^3}{\alpha^2+2}\left[\frac{1}{\alpha-t} + \frac{2}{(\alpha-t)^3}\right] \tag{4}$$

2.3. Construction of NB-Akash Distribution

Definition 2.1. If the NBD has parameters $s > 0$ and $\theta = \mathrm{e}^{-\gamma}$, $Z|\gamma \sim NB(s, \theta = \mathrm{e}^{-\gamma})$, where γ is distributed as Akash distribution with parameter α, $\gamma \sim Akash(\alpha)$, then the r.v Z follows NB-Akash distribution with parameters s and α, $Z \sim NB - Akash(s, \alpha)$.

Theorem 1. Let $Z \sim NB - Akash(s, \alpha)$ be a NB-Akash distribution as defined in Definition 2.1, then the pmf of Z is

$$p(z) = \frac{\alpha^3}{\alpha^2+2}\binom{z+s-1}{z}\sum_{k=0}^{z}\binom{z}{k}(-1)^k\left[\frac{1}{\alpha+s+k} + \frac{2}{(\alpha+s+k)^3}\right]; \tag{5}$$

$z = 0, 1, 2, ..., s, \alpha > 0$

Proof. Since $Z|\gamma \sim NB(s, \theta = \mathrm{e}^{-\gamma})$ and $\gamma \sim Akash(\alpha)$, then pmf of Z can be attained by

$$p(z) = \int_0^\infty p(z|\gamma) f(\gamma;\alpha) \mathrm{d}\gamma \tag{6}$$

where

$$p(z|\gamma) = \binom{z+s-1}{z} e^{-\gamma s}(1-e^{-\gamma})^z = \binom{z+s-1}{z} \sum_{k=0}^{z} \binom{z}{k} (-1)^k e^{-\gamma(s+k)} \tag{7}$$

and $f(\gamma;\alpha)$ is the pdf of Akash distribution.
Substituting (7) in (6) we get

$$\begin{aligned} p(z) &= \binom{z+s-1}{z} \sum_{k=0}^{z} \binom{z}{k} (-1)^k \int_0^\infty \mathrm{e}^{-\gamma(s+k)} f(\gamma;\alpha)\mathrm{d}\gamma \\ &= \binom{z+s-1}{z} \sum_{k=0}^{z} \binom{z}{k} (-1)^k M_\gamma(-(s+k)) \\ &= \frac{\alpha^3}{\alpha^2+2} \binom{z+s-1}{z} \sum_{k=0}^{z} \binom{z}{k} (-1)^k \left[\frac{1}{\alpha+s+k} + \frac{2}{(\alpha+s+k)^3}\right] \end{aligned}$$

■

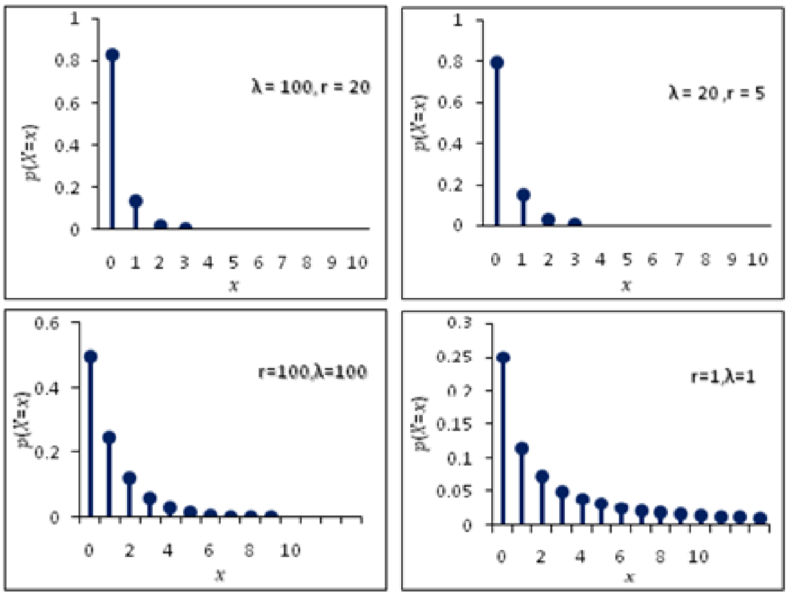


Figure 1: *The pmf of the NB-Akash distribution for various values of parameters*

3. Some Distributional Characteristics

This section is devoted to the discussion of FM, mean, variance and Index of Dispersion of NB-Akash distribution.

Theorem 2. If $Z \sim NB-Akash(s,\alpha)$, then the FM of order m of Z is

$$\mu_{[m]}(Z) = \frac{\alpha^3}{\alpha^2+2} \frac{\Gamma(s+m)}{\Gamma(s)} \sum_{k=0}^{m} \binom{m}{k} (-1)^k \left[\frac{1}{\alpha-m+k} + \frac{2}{(\alpha-m+k)^3}\right] \tag{8}$$

Proof. If $Z|\gamma \sim NB(s,\theta = \mathrm{e}^{-\gamma})$ and $\gamma \sim Akash(\alpha)$, then the FM of the order m of Z can be attained by

$$\mu_{[m]}(Z) = \mathrm{E}_\gamma[\mu_{[m]}(Z|\gamma)]$$

The FM of order m of a NB mixture distribution where $\theta = \mathrm{e}^{-\gamma}$ is

$$\begin{aligned} \mu_{[m]}(Z) &= \mathrm{E}_\gamma\left[\frac{\Gamma(s+m)}{\Gamma(s)} \frac{(1-\mathrm{e}^{-\gamma})^m}{\mathrm{e}^{-\gamma m}}\right] \\ &= \frac{\Gamma(s+m)}{\Gamma(s)} \mathrm{E}_\gamma(\mathrm{e}^\gamma - 1)^m \end{aligned}$$

Using binomial expansion for the term $(e^{\gamma}-1)^i$ we can write

$$\begin{aligned}\mu_{[m]}(Z) &= \frac{\Gamma(s+m)}{\Gamma(s)}\sum_{k=0}^{m}\binom{m}{k}(-1)^k \mathrm{E}(e^{\gamma(m-k)}) \\ &= \frac{\Gamma(s+m)}{\Gamma(s)}\sum_{k=0}^{m}\binom{m}{k}(-1)^k \mathrm{M}_{\gamma}(m-k) \\ &= \frac{\alpha^3}{\alpha^2+2}\frac{\Gamma(s+m)}{\Gamma(s)}\sum_{k=0}^{m}\binom{m}{k}(-1)^k\left[\frac{1}{\alpha-(m-k)}+\frac{2}{(\alpha-(m-k))^3}\right]\end{aligned}$$

■

The mean and variance are derived from (8) are given by

$$E(Z) = s\left[\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)}-1\right] \tag{9}$$

$$E(Z^2) = s(s+1)\frac{\alpha^3(\alpha^2-4\alpha+6)}{(\alpha-2)^3(\alpha^2+2)} - (2s^2+s)\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)} + s^2 \tag{10}$$

$$V(Z) = s(s+1)\frac{\alpha^3(\alpha^2-4\alpha+6)}{(\alpha-2)^3(\alpha^2+2)} - (3s^2+s)\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)} + 2s^2 \tag{11}$$

Index of Dispersion (ID) is defined as the ratio of the variance to the mean denoted by D=V(Z)/E(Z)

$$D = \frac{s(s+1)\frac{\alpha^3(\alpha^2-4\alpha+6)}{(\alpha-2)^3(\alpha^2+2)} - (3s^2+s)\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)} + 2s^2}{s\left[\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)}-1\right]} \tag{12}$$

Table 1: *Mean, variance and ID of NB-Akash distribution for various parameter values*

Mean

s \ α	3	4	5	6	7	8
3	2.5227	1.3457	0.9063	0.6834	0.5496	0.4604
4	3.3636	1.7942	1.2083	0.9112	0.7328	0.6138
5	4.2045	2.2428	1.5104	1.1389	0.9159	0.7673
6	5.0455	2.6914	1.8125	1.3667	1.0991	0.9208
7	5.8864	3.14	2.1146	1.5945	1.2823	1.0742
8	6.7273	3.5885	2.4167	1.8223	1.4655	1.2277

Variance

s \ α	3	4	5	6	7	8
3	52.340	8.7693	3.4687	1.9336	1.2835	7.838
4	85.685	13.965	5.3878	2.943	1.9224	1.3938
5	126.98	20.299	7.688	4.1349	2.6668	1.9137
6	176.22	27.7689	10.370	5.5091	3.5168	2.5022
7	233.41	36.3754	13.433	7.0659	4.4724	3.1593
8	298.56	46.1187	16.878	8.805	5.5335	3.8850

Index of Dispersion

s \ α	3	4	5	6	7	8
3	20.748	6.517	3.828	2.83	2.335	2.047
4	25.474	7.734	4.459	3.23	2.624	2.27
5	30.2	9.051	5.09	3.63	2.912	2.494
6	34.928	10.318	5.722	4.031	3.2	2.718
7	39.654	11.585	6.353	4.431	3.488	2.941
8	44.381	12.852	6.984	4.832	3.776	1.6055

Table 1 sum up the behavior of mean, variance and ID of NB-Akash distribution for selected parameter values, where ID is defined as the ratio of the variance to the mean denoted by D=V(Z)/E(Z)

Figure 2 shows the behaviour of mean, variance and ID for different values of the parameters. Since the ID is greater than 1, the distribution is suitable for overdispersed count data.

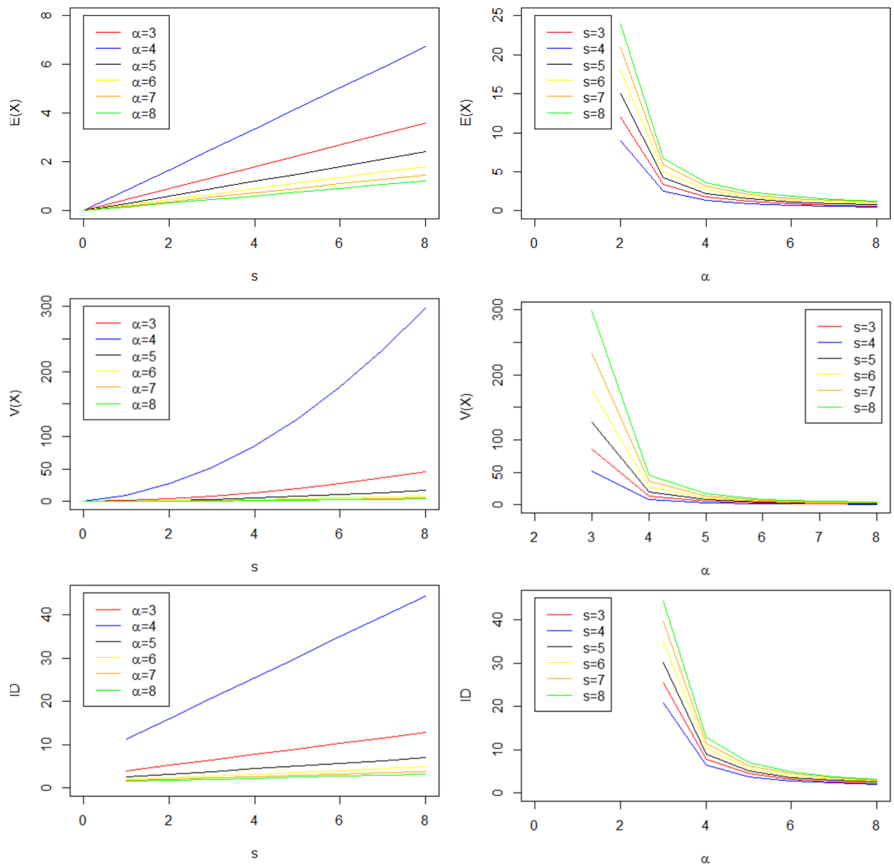


Figure 2: *Behavior of mean, variance and ID for various values of parameters*

4. Estimation of Parameters

The NB-Akash distribution parameters are evaluated using the moments method and the MLE procedure.

4.1. Method of Moments

The NB-Akash distribution has two parameters to estimate, which can be estimated using the first two moments about zero. For moment method, the parameters, s and α are evaluated by equating the moments of the sample and population.

$$m_1 = s\left[\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)}-1\right] \tag{13}$$

$$m_2 = s(s+1)\frac{\alpha^3(\alpha^2-4\alpha+6)}{(\alpha-2)^3(\alpha^2+2)}-(2s^2+s)\frac{\alpha^3(\alpha^2-2\alpha+3)}{(\alpha-1)^3(\alpha^2+2)}+s^2 \tag{14}$$

By equating equations 13 and 14 to the first two sample moments, the moment estimates of the two parameters s and α can be obtained.

4.2. Method of Maximum Likelihood

The LF of the $NB-Akash(s,\alpha)$ is

$$\begin{aligned}\mathcal{L}(z;s,\alpha) &= \frac{\alpha^{3n}}{(\alpha^2+2)^n}\prod_{j=1}^{n}\binom{s+z_j-1}{z_j}\sum_{k=0}^{z_j}\binom{z_j}{k}(-1)^k \\ &\quad\times\left[\frac{1}{\alpha+s+k}+\frac{2}{(\alpha+s+k)^3}\right]\end{aligned}$$

Hence the LL function is

$$\begin{aligned}\ell(z;s,\alpha) &= 3n\log(\alpha)-n\log(\alpha^2+2) \\ &\quad+\sum_{j=1}^{n}\log(s+z_j-1)!-\log z_j!-\log(s-1)! \\ &\quad+\sum_{j=1}^{n}\log\sum_{k=0}^{z_j}\binom{z_j}{k}(-1)^k\left[\frac{1}{\alpha+s+k}+\frac{2}{(\alpha+s+k)^3}\right]\end{aligned}$$

The optimal estimates of the parameters are obtained by partially differentiating this equation with respect to s and α.

$$\begin{aligned}\frac{\partial}{\partial s}\ell(z;s,\alpha) &= \sum_{j=1}^{n}(\Psi(s+z_j)-\Psi(s)) \\ &\quad+\sum_{j=1}^{n}\frac{\sum_{k=0}^{z_j}\binom{z_j}{k}(-1)^k\left(\frac{1}{(\alpha+s+k)^2}+\frac{6}{(\alpha+s+k)^5}\right)}{\sum_{k=0}^{z_j}\binom{z_j}{k}(-1)^k\left(\frac{1}{\alpha+s+k}+\frac{2}{(\alpha+s+k)^3}\right)}\end{aligned} \tag{15}$$

where $\Psi(k)=\frac{\Gamma'(k)}{\Gamma(k)}$ is a digamma function

$$\begin{aligned}\frac{\partial}{\partial \alpha}\ell(z;s,\alpha) &= \frac{3n}{\alpha}-\frac{2n\alpha}{\alpha^2+2} \\ &\quad+\sum_{j=1}^{n}\frac{\sum_{k=0}^{z_j}\binom{z_j}{j}(-1)^k\left(\frac{1}{(\alpha+s+k)^2}+\frac{6}{(\alpha+s+k)^5}\right)}{\sum_{k=0}^{z_j}\binom{z_j}{k}(-1)^k\left(\frac{1}{\alpha+s+k}+\frac{2}{(\alpha+s+k)^3}\right)}\end{aligned} \tag{16}$$

Maximum likelihood estimates are obtained by equating Eq.(15) and Eq.(16) to zero. But solving these equations is complicated and difficult. So these equations are solved numerically using Newton Raphson method.

5. Result and Discussion

This section explains the application and usefulness of the NB-Akash distribution. The proposed distribution is compared with PD, NBD, NB-LD using two real-time data sets. The distributions used for comparison are:

(a)Poisson distribution(PD): The pmf of Poisson distribution for the random variable Z can be written as

$$P(Z=z)=\frac{e^{-\lambda}\lambda^z}{z!};\qquad z=0,1,2,...,\qquad \lambda>0 \tag{17}$$

(b)Negative Binomial Distribution(NBD): If Z denotes a random variable which follows negative binomial distribution with parameters s and θ, then

$$\begin{aligned} P(Z=z) &= [\textit{probability of having } (s-1) \textit{ successes in } (z+s-1) \\ &\quad \textit{trial}] \times [\textit{probability of achieving } s^{th} \textit{ success}] \\ &= \left[\binom{z+s-1}{s-1}\theta^{s-1}(1-\theta)^{(z+s-1)-(s-1)}\right] \times \theta \\ &= \binom{z+s-1}{z}\theta^{s}(1-\theta)^{z}; \quad z=0,1,2,...,s>0, 0<\theta<1 \end{aligned} \tag{18}$$

(c)Negative Binomial Lindley Distribution (NB-LD): The pmf of NB-LD can be written as:

$$P(Z=z) = \frac{\theta^2}{\theta+1}\binom{s+z-1}{z}\sum_{j=0}^{z}\binom{z}{j}(-1)^j\frac{\theta+s+j+1}{(\theta+s+j)^2}; \tag{19}$$

$z = 0,1,2,...,s,\theta > 0$

(d)Negative Binomial Akash Distribution (NB-Akash):The pmf of NB-Akash distribution is given by

$$p(Z=z) = \frac{\alpha^3}{\alpha^2+2}\binom{z+s-1}{z}\sum_{k=0}^{z}\binom{z}{k}(-1)^k\left[\frac{1}{\alpha+s+k}+\frac{2}{(\alpha+s+k)^3}\right]; \tag{20}$$

$z = 0,1,2,...,s,\alpha > 0$

Example 5.1. The data for this example is taken from the article [16] which provides information on 9,461 motor insurance policies according to which the number of accidents of each policy is recorded. The data set is overdispersed because the variance of the data is greater than its mean. The result of the proposed distribution is compared with Poisson, negative binomial, negative binomial-Lindley distributions. The parameter estimation and goodness-of-fit analysis are done through R software. For model comparison, measures such as chi-square test, the p-value, LL, AIC and BIC are used. Based on these measurements, the table 2 shows the NB-Akash distribution performs better than the PD, NBD and NB-LD.

Table 2: *Observed and expected frequencies of Example 5.1*

No. of	No. of	Fitting of distribution			
claims	drivers	Poisson	NB	NB-LD	NB-Akash
0	7840	7638.3	7843.3	7853.6	7852.1
1	1317	1634.6	1290.2	1287.4	1288.4
2	239	174.9	257.7	247.6	247.9
3	42	12.5	54.5	54.2	54.3
4	14	0.7	11.8	13.2	13.2
5	4	0	2.6	3.5	3.5
6	4	0	.6	1	1
7	1	0	0.2	0.3	0.3
8	0	0	0.1	0.2	0.3
Estimated parameter		$\hat{\lambda} = 0.214$	$\hat{s} = 0.7$ $\hat{\theta} = 0.765$	$\hat{s} = 4.63$ $\hat{\theta} = 23.55$	$\hat{s} = 4.7477$ $\hat{\alpha} = 23.2$
degrees of freedom		2	3	4	4
Chi-square		293.8	8.66	6.997	6.79
p-value		≤ 0.01	0.01	0.072	0.079
LL		-5490.78	-5348.00	-5344.7	-5344.678
AIC		10983.56	10700	10693.4	10693.36
BIC		10982.95	10699.22	10692.98	10692.94

Example 5.2. The data are taken from the article [6]. The data contain fatal accidents at the exit of the single vehicle highway on horizontal two-lane rural curves between 2003 and 2008. The parameters are estimated using the MLE method and the Poisson, negative binomial, negative binomial-Lindley distribution are fitted to the data. The performances of these distributions are compared in terms of chi-square tests of goodness of fit, p-value, LL, AIC and BIC. Table 3 shows the NB-Akash distribution performs better than the PD, NBD and NB-LD.

Table 3: *Observed and expected frequencies of Example 5.2*

No. of	No. of	Fitting of distribution			
claims	drivers	Poisson	NB	NB-L	NB-Akash
0	29087	28471.6	29204.8	29133.6	29099.6
1	2952	3918	2706	2855.5	2906.1
2	464	269.6	567.4	503.1	498
3	108	12.4	141.1	120.9	116.2
4	40	0.4	37.8	35.9	33.4
5	9	0	10.6	13.1	11.2
6	5	0	3	3.3	4.2
7	2	0	.9	3.3	1.7
8	3	0	0.3	0	0.8
9	1	0	0.1	0	0.4
10+	1	0	0	3.3	0.4
Estimated parameter		$\hat{\lambda} = 0.138$	$\hat{s} = 0.138$ $\hat{\theta} = 0.2584$	$\hat{s} = 1.018$ $\hat{\theta} = 9.212$	$\hat{s} = 1.1881$ $\hat{\alpha} = 10.364$
degrees of freedom		2	3	4	4
Chi-square		2297.31	57.47	11.68	8.0666
p-value		≤ 0.01	≤ 0.01	0.02	0.089
LL		-14,208.1	-13,557.7	-13,529.8	-13528.43
AIC		28418.2	27119.4	27063.6	27060.86
BIC		28417.59	27118.98	27063.49	27060.75

6. Conclusion

In this paper a new mixed NB distribution named as negative binomial-Akash distribution is proposed by mixing the negative binomial distribution and Akash distribution. Some of the important characteristics such as FM, mean, variance and ID are studied, . The MLE method is used to evaluate the parameters of the NB-Akash distribution. The utility of NB-Akash distribution was illustrated using two real data sets. From the result it can be inferred that the NB-Akash distribution provides better fit than the PD, NBD and NB-LD.

References

[1] Aryuyuen, S. and Bodhisuwan, W. (2013). The negative binomial-generalized exponential (NB-GE) distribution. *Applied Mathematical Sciences*, 7:1093–1105.

[2] Denthet, Sunthree and Thongteeraparp, Ampai and Bodhisuwan, Winai. Mixed distribution of negative binomial and two-parameter Lindley distributions, 104–107, 2016.

[3] Ghitany, M. and Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78:493–506.

[4] Kongrod, S. and Bodhisuwan, W. and Payakkapong, P. (2014). The negative binomial-Erlang distribution with applications. *International Journal of Pure and Applied Mathematics*, 92:389–401.

[5] Lindley, D.V (1958). Fiducial Distributions and Bayes' Theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20:102–107.

[6] Lord, D and Geedipally, S (2011). The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. *Accident; analysis and prevention*, 43:1738–1742.

[7] Saengthong, P. and Bodhisuwan, W. (2013). Parameter estimation for negative binomial-Crack distribution and its application. *Journal of Science and Technology*, 33:125–130.

[8] Saengthong, P. and Bodhisuwan, W. (2013). Negative binomial-crack (NB-CR) distribution. *International Journal of Pure and Apllied Mathematics*, 84:213–230.

[9] Sankaran, M. (1970). The discrete poisson-lindley distribution. *Biometrics*, 26:145–149.

[10] Shanker, R and Mishra, A. (2013). A two-parameter Lindley distribution. *Statistics in Transition new series*, 1:45–56.

[11] Shanker, R and Amanuel, G.H. (2013). A New Quasi Lindley Distribution. *International Journal of Statistics and Systems*, 8:143–156.

[12] Shanker, R. and Tekie, A.L. (2014). A new quasi Poisson-Lindley distribution. *International Journal of Statistics and Systems*, 9:87–94.

[13] Mishra, A. (2014). A two-parameter Poisson - Lindley distribution. *International Journal of Statistics and Systems*, 9:79–85.

[14] Shanker, R. (2015). Akash Distribution and Its Applications. *International Journal of Probability and Statistics*, 4:65–75.

[15] Yamrubboon, D. and Bodhisuwan, W. and Saothayanun, L. and Sharma, S. (2017). The Negative Binomial-Sushila Distribution with Application in Count Data Analysis. *Thailand Statistician*, 15:69–77.

[16] Zamani, H. and Ismail, N. (2010). Negative Binomial-Lindley Distribution and Its Application. *Journal of Mathematics and Statistics*, 6:4–9.

[17] Zamani, H. and Ismail, N. and Faroughi, P. (2014). Poisson-weighted exponential: Univariate version and regression model with applications. *Journal of Mathematics and Statistics*, 10:148–154.

Generalized Transmuted Exponential-Exponential Distribution and its Applications

A. S. MOHAMMED[1,2*], F. I. UGWUOWO[2γ], T. S. PATRICE[3], H. MUHAMMAD[4]

•

[1]Department of Statistics, Ahmadu Bello University, Zaria
[2]Department of Statistics, University of Nigeria, Nsukka
[γ]Email: fidelis.ugwuowo@unn.edu.ng
[3]Department of Mathematics, University of Yaounde I, Cameroon
Email: patricetakam@gmail.com
[4]Division of Agriculture, Ahmadu Bello University, Zaria
Email: hamzamuhammad@abu.edu.ng
*Corresponding author: Email: mohammedas@abu.edu.ng

Abstract

Modeling of datasets requires knowledge of their appropriate distributional assumptions. In this research, we generalized the transmuted exponential-exponential distribution, and it was observed that the addition of the shape parameter to the model proved to be helpful in improving the flexibility of the model. Different characteristics, as well as structural properties of the model, were investigated and presented in an explicit form. The probability density function of the order statistics and numerical results for some descriptive statistics were obtained. A 95% confidence interval and interval widths, together with biases and mean square errors (MSEs) of the mean estimates, were equally evaluated using the Monte-Carlo simulation approach. To validate the flexibility of the model, we used real datasets and the generalized transmuted exponential-exponential distribution (GTE-ED) outperformed the competing distributions.

Keywords: Transmuted Exponential-Exponential distribution, descriptive statistics, Order statistics, Confidence Interval

1. Introduction

The procedure of parameter(s) induction to a parent distribution has fascinated the attention of numerous researchers in the recent years [1]. The addition of one or more shape parameter(s) to a given baseline model strengthens the distribution, especially when studying its tail characteristics. The parameter induction method has proved useful for boosting the fitness of a proposed model [2].

In statistics, the modelling of datasets requires knowledge of appropriate distributional assumptions about the datasets. In theory, the tractability of a probability distribution can be helpful since it is easier to manipulate when modelling a dataset. The concept of generalizing distributions was proposed by [3] which concern basically with raising the distribution function of the baseline distribution say $A(y)$ to the power of an arbitrary parameter $c > 0$ which give rise to a new model or distribution of the form $B(y) = (A(y))^c \; for\, c > 0$. The parameter (c) plays an important role in adding skewness to the function $A(y)$. In the early 1990s, generalized models were found to be useful in numerous areas of statistics and medical sciences due to their ability to model different forms of data. These distributions were proposed by statisticians from various fields. The concept of generating generalized distributions was used by [4] to developed a new

Weibull distribution named the exponentiated Weibull distribution. Furthermore, [5] studied the general characteristics of the exponentiated Weibull distribution. A new two-parameter model called exponentiated exponential distribution which outperformed other competing distribution in the study when applied to a real dataset was studied by [6]. Notable authors like [7], [8] and [9] applied the same methodology and developed the exponentiated type distributions, and exponentiated generalized inverse Gaussian distribution respectively. The properties of exponentiated transmuted generalized Rayleigh distribution was proposed and studied by [10] and [11] studied the exponentiated generalized class of distributions. Consequently, the properties and MLEs of generalized odd generalized exponential- exponential distribution was presented in an explicit form by [12]. The properties and applications of the transmuted exponential-exponential distribution (TE-ED) which has two scale parameters and a transmuted parameter. In practice, to find the distribution that captures the sensitive part of a given dataset, there are many possibilities was studied by [13]. We can either estimate non-parametrically the density function as well as the distribution function and compare them with the existing distributions to see which one is closest to the empirical distribution. However, in some situations for which we are obliged to consider some characteristics such as hazard rate, many of the existing distributions cannot adequately model a dataset with non-monotone hazard rates, and as such, these distributions are limited in applications.
The current kinds of literature in mathematical statistics as highlighted by [14] pay more attention to proposing more flexible distributions but give less concern to the hazard function of the distributions. It is critical to generate distributions with varying failure rates because the hazard rate function guides model selection [11]. Furthermore, many of the existing exponential extended distributions cannot adequately describe some of the existing datasets, particularly the ones with monotone and non-monotone hazard rates. For example, exponentiated exponential, transmuted exponential-exponential, and Weibull exponential distribution, among others. This has opened the room for more research that can account for monotone and non-monotone hazard rate data.

In this research, we are motivated by the above-mentioned rationale to develop a new exponential extended distribution called the generalized transmuted exponential-exponential distribution (GTE-ED). As compared to the existing exponential extended distributions, the GTE-ED is more flexible and can model both monotone and non-monotone hazard rate data.

Table 1: *Hazard rates behaviour for GTE-E and the competing distributions*

Distribution	Constant	Increasing	Decreasing	Unimodal
GTE-E	Yes	Yes	Yes	Yes
TE-E	Yes	Yes	No	No
EE	Yes	Yes	Yes	No
E	Yes	No	No	No

From table 1, we can deduce that GTE-ED has more advantages over the competing distributions in the study and, as such, it will be more robust in analysing data with different hazard rates.

2. The Generalized Transmuted Exponential-Exponential Distribution

Consider the density function $a(y;\lambda,\theta,\alpha) = \lambda\alpha\,(1-\theta)\,e^{-\alpha\lambda y} + 2\lambda\theta\alpha e^{-2\alpha\lambda y}$ and distribution function $A(y;\lambda,\theta,\alpha) = \left(1-e^{-\lambda\alpha y}\right)\left(1+\theta e^{-\lambda\alpha y}\right)$ of the transmuted exponential-exponential distribution with scale parameter $\alpha,\lambda > 0$, transmuted parameter $-1 \leq \theta \leq 1$ and $y \geq 0$ (Mohammed and Ugwuowo 2021). The cumulative distribution function and density of generalized transmuted exponential distribution (GTE-ED) are respectively derived from the following functions:

$$B(y) = (A(y))^c \ for\, c > 0 \tag{1}$$

and,

$$b(y) = ca(y)\,(A(y))^{c-1} \tag{2}$$

The GTE-ED is then defined as;

$$B(y) = \left[\left(1 - e^{-\lambda\alpha y}\right)\left(1 + \theta e^{-\lambda\alpha y}\right)\right]^{c} \tag{3}$$

and by taking the differential of $B(y)$, we have;

$$b(y) = c\left[\lambda\alpha\,(1-\theta)\,e^{-\alpha\,\lambda y} + 2\lambda\theta\alpha e^{-2\alpha\,\lambda y}\right]\left[\left(1 - e^{-\lambda\,\alpha\,y}\right)\left(1 + \theta e^{-\lambda\,\alpha\,y}\right)\right]^{c-1} \tag{4}$$

where, $y \geq 0,\ \alpha, \lambda, c > 0\, and\ -1 \leq \theta \leq 1$.

Equation (4) can be written in the following contracted form;

$$b(y) = me^{-\lambda\,\alpha(f+g+1)y}\left((1-\theta) + 2\theta\, e^{-\alpha\lambda y}\right) \tag{5}$$

where, $m = c\,\lambda\,\alpha \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g}$

Figure 1 displays some possible shapes of density and distribution function of the GTE-ED for chosen values of the parameters $a = \alpha,\ \ b = \lambda,\ d = \theta\, and\, c$. Moreover, the density changes in shape when the parameters take different values.

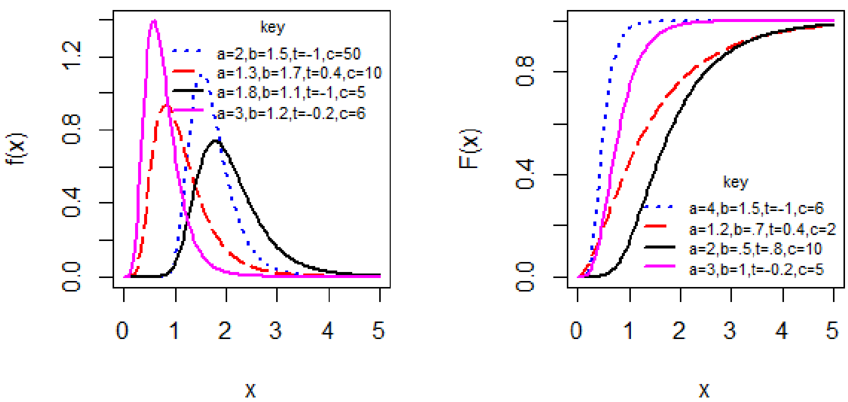


Figure 1: *density and distribution function of GTE-ED*

3. Statistical Properties of the model

Here, some statistical properties of GTE-ED including survival and hazard functions, quantile function, moments, moment generating function, limiting behaviour, and order statistics are considered and presented in an explicit form.

3.1. Survival and Hazard function

If Y has $GTE - E(\alpha, \lambda, c, \theta)$ model, then the survival and hazard function are respectively given as;

The survival function is defined mathematically as;

$S(x) = 1 - B(y)$

$$S(y) = 1 - \left(\left(1 - e^{-\lambda\alpha y}\right)\left(1 + \theta e^{-\lambda\alpha y}\right)\right)^{c} \tag{6}$$

The hazard function is defined mathematically as;

$h(y) = \frac{b(y)}{1-B(y)}$

$$h(y) = \frac{c\left(\lambda\alpha\left(1-\theta\right)e^{-\alpha\lambda y} + 2\lambda\theta\alpha e^{-2\alpha\lambda y}\right)\left(\left(1-e^{-\lambda\alpha y}\right)\left(1+\theta e^{-\lambda\alpha y}\right)\right)^{c-1}}{\left[1-\left(\left(1-e^{-\lambda\alpha y}\right)\left(1+\theta e^{-\lambda\alpha y}\right)\right)^{c}\right]} \tag{7}$$

Figures 3 and 4 displays some possible shapes of hazard and survival function (hf) of the GTE-ED for chosen values of the parameters $a = \alpha,\ \ b = \lambda,\ d = \theta\ and\ c$. The hf can take the shape of either increasing, decreasing, and unimodal as the parameter keep changing.

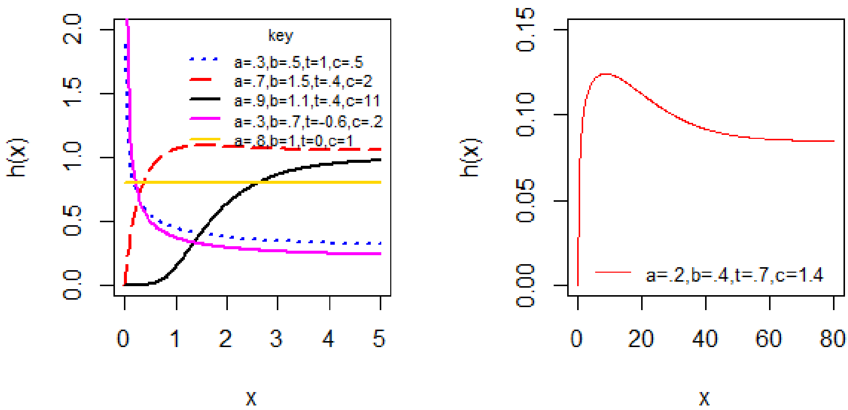


Figure 2: *Hazard function of GTE-ED*

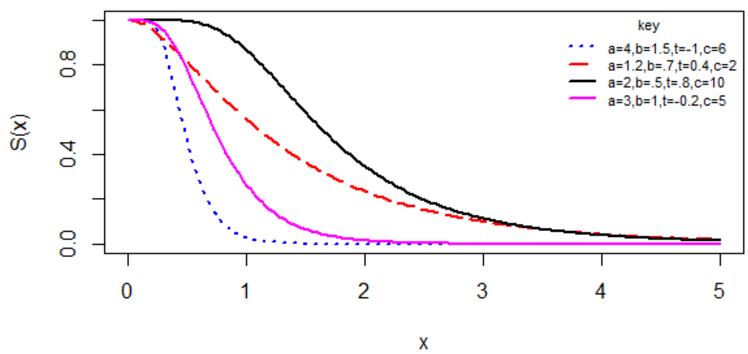


Figure 3: *Survival function of GTE-ED*

Limiting behaviour of the distribution

In this section, the asymptotic behaviour of the model is investigated by taking the limit $as\, y \to 0\, and\, y \to \infty$ of the distribution function.

$$\lim_{y\to 0} B(y) = \lim_{y\to 0} \left(\left(1-e^{-\lambda\alpha y}\right)\left(1+\theta e^{-\lambda\alpha y}\right)\right)^{c} = 0$$

and,

$$\lim_{y\to \infty} B(y) = \lim_{y\to \infty} \left(\left(1-e^{-\lambda\alpha y}\right)\left(1+\theta e^{-\lambda\alpha y}\right)\right)^{c} = 1$$

The results show that the GTE-ED is a valid distribution since $\lim_{y\to 0} B(y) = 0\, and \lim_{y\to\infty} B(y) = 1$.

3.2. The r^{th} moments and moment generating function

If $Y\, has\, GTEE(\alpha, \lambda, c, \theta)$ then, the r^{th} moments is given as;

$$E\left(y^{r}\right) = \frac{c\,\Gamma(r+1)}{(\lambda\,\alpha)^{r}} \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g} \left\{\frac{(1-\theta)}{(f+g+1)^{r+1}} + \frac{2\,\theta}{(f+g+2)^{r+1}}\right\} \tag{8}$$

By using (8) the first two moments about the origin are obtained which can pave way in finding the variance and the coefficient of variation.

When r =1,

$$E\left(y\right) = \frac{c}{(\lambda\,\alpha)} \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g} \left\{\frac{(1-\theta)}{(f+g+1)^{2}} + \frac{2\,\theta}{(f+g+2)^{2}}\right\}$$

When r =2,

$$E\left(y^{2}\right) = \frac{2\,c}{(\lambda\,\alpha)^{2}} \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g} \left\{\frac{(1-\theta)}{(f+g+1)^{3}} + \frac{2\,\theta}{(f+g+2)^{3}}\right\}$$

$$Let,\, A_1 = \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g} \left\{\frac{(1-\theta)}{(f+g+1)^{2}} + \frac{2\,\theta}{(f+g+2)^{2}}\right\}$$

$$A_2 = \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g} \left\{\frac{(1-\theta)}{(f+g+1)^{3}} + \frac{2\,\theta}{(f+g+2)^{3}}\right\}$$

Therefore, $E\left(y\right) = \frac{c}{(\lambda\,\alpha)} A_1\, and\, E\left(y^2\right) = \frac{2\,c}{(\lambda\,\alpha)^2} A_2$

If $Y\, has\, GTE - E(\alpha, \lambda, c, \theta)$ then, the variance and the coefficient of variation of GTE-ED are respectively given as;

$$Var(y) = \frac{2\,c}{(\lambda\,\alpha)^2} A_2 - \frac{c}{(\lambda\,\alpha)} A_1 \text{ and } C\,V = \frac{\sqrt{\frac{2\,c}{(\lambda\,\alpha)^2} A_2 - \frac{c}{(\lambda\,\alpha)} A_1}}{\frac{c}{(\lambda\,\alpha)} A_1}$$

Moment Generating Function

If $Y\, has\, GTE - E(\alpha, \lambda, c, \theta)$ distribution then, the moment generating function (MGF) is given as;

$$K_y(t) = c\,\lambda\,\alpha \sum_{f,g=0}^{\infty} (-1)^{f} \binom{c-1}{f}\binom{c-1}{g} \left\{\frac{(1-\theta)}{\lambda\,\alpha(f+g+1)-t} + \frac{2\,\theta}{\lambda\,\alpha(f+g+2)-t}\right\} \tag{9}$$

3.3. The r^{th} moments about the mean

If $Y\, has\, GTE-E(\alpha,\lambda,c,\theta)$ distribution then, the r^{th} moments about the mean is given;

$$E(y-\mu)^r = c\sum_{f,g=0}^{\infty}\sum_{h=0}^{r}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{r}{h}\mu^h\times \left\{\frac{(1-\theta)\Gamma(r-h+1)}{((f+g+1))^{r-h+1}(\lambda\alpha)^{r-h}}+\frac{2\theta\Gamma(r-h+1)}{((f+g+2))^{r-h+1}(\lambda\alpha)^{r-h}}\right\} \tag{10}$$

If $\mu=0$, the result will give us the moment about the origin.

$$E(y^r) = c\sum_{f,g=0}^{\infty}(-1)^{f}\binom{c-1}{f}\binom{c-1}{g}\left\{\frac{(1-\theta)\Gamma(r+1)}{((f+g+1))^{r+1}(\lambda\alpha)^{r}}+\frac{2\theta\Gamma(r+1)}{((f+g+2))^{r+1}(\lambda\alpha)^{r}}\right\}$$

In order to find the skewness and kurtosis, we have to find the expressions for r =1,2,3 and 4. The expressions are given below;

If r =1,

$$E(y-\mu) = c\sum_{f,g=0}^{\infty}\sum_{h=0}^{1}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\times \left\{\frac{(1-\theta)\Gamma(2-h)}{((f+g+1))^{2-h}(\lambda\alpha)^{1-h}}+\frac{2\theta\Gamma(2-h)}{((f+g+2))^{2-h}(\lambda\alpha)^{1-h}}\right\}$$

If r =2,

$$E(y-\mu)^2 = c\sum_{f,g=0}^{\infty}\sum_{h=0}^{2}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\times \left\{\frac{(1-\theta)\Gamma(3-h)}{((f+g+1))^{3-h}(\lambda\alpha)^{2-h}}+\frac{2\theta\Gamma(3-h)}{((f+g+2))^{3-h}(\lambda\alpha)^{2-h}}\right\}$$

If r =3,

$$E(y-\mu)^3 = c\sum_{f,g=0}^{\infty}\sum_{h=0}^{3}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\times \left\{\frac{(1-\theta)\Gamma(4-h)}{((f+g+1))^{4-h}(\lambda\alpha)^{3-h}}+\frac{2\theta\Gamma(4-h)}{((f+g+2))^{4-h}(\lambda\alpha)^{3-h}}\right\}$$

If r =4,

$$E(y-\mu)^4 = c\sum_{f,g=0}^{\infty}\sum_{h=0}^{4}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\times \left\{\frac{(1-\theta)\Gamma(5-h)}{((f+g+1))^{5-h}(\lambda\alpha)^{4-h}}+\frac{2\theta\Gamma(5-h)}{((f+g+2))^{5-h}(\lambda\alpha)^{4-h}}\right\}$$

The coefficient of skewness are kurtosis of the GTE-ED are respectively given as;

$$CS = \frac{E(y-\mu)^3}{\left(E(y-\mu)^2\right)^{\frac{3}{2}}}$$

$$CS = \frac{c\sum_{f,g=0}^{\infty}\sum_{h=0}^{3}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\left\{\frac{(1-\theta)\Gamma(4-h)}{((f+g+1))^{4-h}(\lambda\alpha)^{3-h}}+\frac{2\theta\Gamma(4-h)}{((f+g+2))^{4-h}(\lambda\alpha)^{3-h}}\right\}}{\left(c\sum_{f,g=0}^{\infty}\sum_{h=0}^{2}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\left\{\frac{(1-\theta)\Gamma(3-h)}{((f+g+1))^{3-h}(\lambda\alpha)^{2-h}}+\frac{2\theta\Gamma(3-h)}{((f+g+2))^{3-h}(\lambda\alpha)^{2-h}}\right\}\right)^{\frac{3}{2}}}$$

and

$$CK = \frac{E(y-\mu)^4}{\left(E(y-\mu)^2\right)^{2}}$$

$$CK = \frac{c\sum_{f,g=0}^{\infty}\sum_{h=0}^{4}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\left\{\frac{(1-\theta)\Gamma(5-h)}{((f+g+1))^{5-h}(\lambda\alpha)^{4-h}}+\frac{2\theta\Gamma(5-h)}{((f+g+2))^{5-h}(\lambda\alpha)^{4-h}}\right\}}{\left(c\sum_{f,g=0}^{\infty}\sum_{h=0}^{2}(-1)^{f+h}\binom{c-1}{f}\binom{c-1}{g}\binom{1}{h}\mu^h\left\{\frac{(1-\theta)\Gamma(3-h)}{((f+g+1))^{3-h}(\lambda\alpha)^{2-h}}+\frac{2\theta\Gamma(3-h)}{((f+g+2))^{3-h}(\lambda\alpha)^{2-h}}\right\}\right)^2}$$

Table 2: *Some selected measures of $Y \sim GTE - E$ for some chosen values of the c and $\alpha = 2$, $\lambda = 0.3$, $\theta = 0.5$. The standard errors (SEs) in bracket, where τ_1 and τ_2 stands for the mean deviation about mean and the mean deviation about the median*

	Parameter (c)				
	1	2	3	4	5
Mean	1.2704 (0.1022)	1.9317 (0.1168)	2.3875 (0.1239)	2.7374 (0.1284)	3.0223 (0.1315)
Variance	2.1170 (0.5922)	2.7815 (0.6601)	3.1413 (0.6927)	3.3761 (0.7127)	3.5445 (0.7265)
Skewness	2.4648 (0.7892)	2.0447 (0.6535)	1.8681 (0.5996)	1.7658 (0.5696)	1.6971 (0.5501)
Kurtosis	12.1973 (7.7423)	9.5389 (5.6921)	8.5507 (4.9328)	8.0148 (4.5232)	7.6712 (4.2619)
τ_1	0.4023 (0.0408)	0.4793 (0.0425)	0.5168 (0.0433)	0.5401 (0.0438)	0.5564 (0.0441)
τ_2	0.3666 (0.0349)	0.4490 (0.0376)	0.4888 (0.0388)	0.5134 (0.0395)	0.5306 (0.04)

Table 3: *Some selected measures of $Y \sim GTE - E$ for some chosen values of the c and $\alpha = 2$, $\lambda = 0.3$, $\theta = -0.5$ The standard errors (SEs) in bracket, where τ_1 and τ_2 stands for the mean deviation about mean and the mean deviation about the median*

	Parameter (c)				
	1	2	3	4	5
Mean	2.1051 (0.1317)	3.0448 (0.1414)	3.6465 (0.1450)	4.0886 (0.1468)	4.4381 (0.1479)
Variance	3.5522 (0.7368)	4.1168 (0.7727)	4.3324 (0.7859)	4.4453 (0.7927)	4.5147 (0.7969)
Skewness	1.7785 (0.5441)	1.5157 (0.4921)	1.4258 (0.4758	1.3808 (0.4679)	1.3539 (0.4633)
Kurtosis	7.8311 (4.3142)	6.7415 (3.5522)	6.4140 (3.3160)	6.2586 (3.0221)	6.1681 (3.1351)
τ_1	0.5588 (0.0451)	0.06278 (0.0454)	0.6278 (0.0456)	0.6370 (0.0457)	0.6425 (0.0458)
τ_2	0.5286 (0.0409)	0.5860 (0.0418)	0.6060 (0.0422)	0.6162 (0.0424)	0.6223 (0.0425)

From tables 2 and 3, it can be deduced that as the value of the parameter (c) increases, the mean, variance, mean deviation about mean and the mean deviation about the median also increases. While the skewness and kurtosis decrease.

3.4. The Quantile function of the model

The quantile function can be defined mathematically as; $Q(u) = Inf\{y \in \Re : u \leq F(y)\}$ for which $0 < u < 1$. Since the function $F(y)$ is continuous and monotonically increasing, then we have $Y = F^{-1}(u)$.

Corollary 1. The quantile function of the GTE-ED is given as;

$$Q(u) = -\frac{1}{\alpha\,\lambda} \ln\left(\frac{\theta - 1 + \sqrt{(1+\theta)^2 - 4\,\theta\,u^{\frac{1}{c}}}}{2\,\theta}\right), \qquad 0 < u < 1 \tag{11}$$

Note that, when u=0.5 (11) gives the median.

The effect of the shape and Transmuted parameter were examined on the skewness and kurtosis and it was evaluated by using the relationship of Bowley (BS) and Moors (MK). Figures (a) and (b) shows the plot of Bowley (BS) and Moors (MK) for GTE-ED for fixed parameters (α and λ) respectively. The plot for the skewness shows a steady decrease as the parameter (c)increases while for parameter θ shows a steady increase to a minimum point before decreasing as its value increases. Again, the Kurtosis shows a steady decrease to a certain point and decreases as the parameter (c) increases while for parameter θ, shows a steady decrease as its value increases.

$$Sk_B = \frac{Q(\frac{3}{4}) - 2Q(\frac{1}{2}) + Q(\frac{1}{4})}{Q(\frac{3}{4}) - Q(\frac{1}{4})} \text{ and } Ku_M = \frac{Q(\frac{7}{8}) - Q(\frac{5}{8}) + Q(\frac{3}{8}) - Q(\frac{1}{8})}{Q(\frac{6}{8}) - Q(\frac{2}{8})}$$

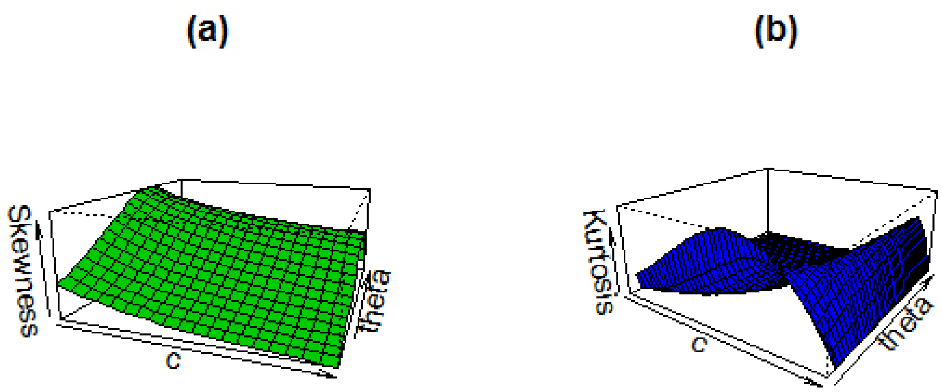


Figure 4: *3D diagram for Skewness and kurtosis*

3.5. Order Statistics of the GTE-ED

The general form of the density of the h^{th} order statistics for a given random samples $y_1, y_2, ..., y_n$ from the distribution function is obtained as; $b_{n,h}(y) = \frac{n!}{(h-1)!(n-h)!} b(y) B(y)^{h-1} (1 - B(y))^{n-h}$. Therefore, by substituting the resulting density as well as the distribution function of the GTE-ED we have;

$$\begin{aligned} b_{n,h}(y) = &\frac{n!}{(h-1)!(n-h)!} \left(c \left(\lambda\alpha (1-\theta) e^{-\alpha\lambda y} + 2\lambda\theta\alpha e^{-2\alpha\lambda y} \right) \left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c-1} \right) \\ &\times \left(\left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c} \right)^{h-1} \left(1 - \left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c} \right)^{n-h} \end{aligned} \tag{12}$$

The distribution of the minimum and maximum order statistics for the GTE-ED are respectively given as;

$$\begin{aligned} b_{n,1}(y) = n &\left(c \left(\lambda\alpha (1-\theta) e^{-\alpha\lambda y} + 2\lambda\theta\alpha e^{-2\alpha\lambda y} \right) \left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c-1} \right) \\ &\times \left(1 - \left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c} \right)^{n-1} \end{aligned}$$

and,

$$\begin{aligned} b_{n,n}(y) = n &\left(c \left(\lambda\alpha (1-\theta) e^{-\alpha\lambda y} + 2\lambda\theta\alpha e^{-2\alpha\lambda y} \right) \left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c-1} \right) \\ &\times \left(\left(\left(1 - e^{-\lambda\alpha y}\right) \left(1 + \theta e^{-\lambda\alpha y}\right) \right)^{c} \right)^{n-1} \end{aligned}$$

4. Estimation of the Parameters of GTE-ED

If the parameters of the GTE-ED are unknown, then the maximum likelihood estimates of the parameters are presented below, let $y_1, y_2, ..., y_n$ be the random sample of size (n) from the GTE-ED, then the log-likelihood function of (4) is obtained as;

$$\begin{aligned} ll(\Psi) = n \log \alpha + n \log \lambda + n \log c - \lambda\alpha \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} \log \left(1 - \theta + 2\theta e^{-\alpha\lambda y_i} \right) \\ + (c-1) \sum_{i=1}^{n} \log \left(1 + \theta e^{-\alpha\lambda y_i} - e^{-\alpha\lambda y_i} - \theta e^{-2\alpha\lambda y_i} \right) \end{aligned} \tag{13}$$

By differentiating the $ll(\Psi)$ with respect to the parameters. The following results were obtained;

$$\begin{aligned} \frac{\delta ll(\Psi)}{\delta\alpha} &= \frac{n}{\alpha} - \lambda \sum_{i=1}^{n} y_i - 2\lambda\theta \sum_{i=1}^{n} \frac{y_i e^{-\alpha\lambda y_i}}{\left(1 - \theta + 2\theta e^{-\alpha\lambda y_i}\right)} \\ &\quad + (c-1) \sum_{i=1}^{n} \frac{\lambda y_i e^{-\alpha\lambda y_i} - \theta\lambda y_i e^{-\alpha\lambda y_i} + 2\theta\lambda y_i e^{-2\alpha\lambda y_i}}{\left(1 + \theta e^{-\alpha\lambda y_i} - e^{-\alpha\lambda y_i} - \theta e^{-2\alpha\lambda y_i}\right)} \\ \frac{\delta ll(\Psi)}{\delta\lambda} &= \frac{n}{\lambda} - \alpha \sum_{i=1}^{n} y_i - 2\alpha\theta \sum_{i=1}^{n} \frac{y_i e^{-\alpha\lambda y_i}}{\left(1 - \theta + 2\theta e^{-\alpha\lambda y_i}\right)} \\ &\quad + (c-1) \sum_{i=1}^{n} \frac{\alpha y_i e^{-\alpha\lambda y_i} - \theta\alpha y_i e^{-\alpha\lambda y_i} + 2\theta\alpha y_i e^{-2\alpha\lambda y_i}}{\left(1 + \theta e^{-\alpha\lambda y_i} - e^{-\alpha\lambda y_i} - \theta e^{-2\alpha\lambda y_i}\right)} \\ \frac{\delta ll(\Psi)}{\delta\theta} &= \sum_{i=1}^{n} \frac{2e^{-\alpha\lambda y_i} - 1}{\left(1 - \theta + 2\theta e^{-\alpha\lambda y_i}\right)} + (c-1) \sum_{i=1}^{n} \frac{e^{-\alpha\lambda y_i} - e^{-2\alpha\lambda y_i}}{\left(1 + \theta e^{-\alpha\lambda y_i} - e^{-\alpha\lambda y_i} - \theta e^{-2\alpha\lambda y_i}\right)} \\ \frac{\delta ll(\Psi)}{\delta c} &= \frac{n}{c} + \sum_{i=1}^{n} \log \left(1 + \theta e^{-\alpha\lambda y} - e^{-\alpha\lambda y} - \theta e^{-2\alpha\lambda y} \right) \end{aligned}$$

The ML Estimator $\widehat{\Phi} = \left(\widehat{\alpha}, \widehat{\lambda}, \widehat{c}, \widehat{\theta} \right)^T$ of the parameter vector is gotten by finding the solution of the set of nonlinear system of equations. The results will give the MLEs $\widehat{\alpha}, \widehat{\lambda}, \widehat{c}$ *and* $\widehat{\theta}$. We applied an optimization technique to numerically maximize the log- likelihood (LL) function given in (13).

5. Simulation Study

5.1. The Design

Here, a simulation study is conducted to assess the performance of the maximum likelihood estimation method as expressed above. So also, 10,000 random samples were generated for different sizes, n= 20, 50, 100, 200, 300, and 500 from GTE-ED. Furthermore, the estimates, Biases, MSEs, Confidence Interval (C. Is) at 95%, widths are evaluated. The steps are:

i Choose the initial values of the parameters and the sample size say (n).

ii Draw a random sample of size (n) from the GTE-ED using the quantile function given in (11).

iii Evaluate the estimates of the parameters using the approach of maximum likelihood.

iv Repeat steps i and ii for N=10,000 times to evaluate the bias, mean square error, 95% C.I and the interval width of the given estimates.

Table 4: *Results for the MLEs, Biases and MSEs, 95% C. Is, and Widths of the GTE-ED for* $\alpha = 3,\ \lambda = 2,\ c = 0.5, \theta = 0.5$

					C I		
Sample	Parameter	Esimate	Bias	MSE	LC	UC	width
	α	3.1657	0.1657	0.1334	2.9580	3.3734	0.4154
n=20	λ	2.2169	0.2169	0.2664	1.7868	2.6469	0.8601
	c	0.5379	0.0379	0.0345	0.4732	0.6026	0.1294
	θ	0.3583	-0.1417	0.3579	-0.3038	1.0203	1.13241
	α	3.1072	0.1072	0.0794	2.9741	3.2403	0.2662
n=50	λ	2.1433	0.1433	0.1680	1.8543	2.4324	0.5782
	c	0.5022	0.0022	0.0105	0.4817	0.5227	0.0410
	θ	0.3480	-0.1520	0.2873	-0.1697	0.8658	1.0355
	α	3.0888	0.0888	0.0640	2.9787	3.1989	0.2206
n=100	λ	2.1280	0.1280	0.1416	1.8827	2.3734	0.4907
	c	0.4952	-0.0048	0.0050	0.4854	0.5051	0.0196
	θ	0.3514	-0.1486	0.2227	-0.0419	0.7447	0.7866
	α	3.0670	0.0670	0.0510	2.9757	3.1582	0.1825
n=200	λ	2.0978	0.0978	0.1165	1.8883	2.3073	0.4190
	c	0.4943	-0.0057	0.0023	0.4898	0.4988	0.0090
	θ	0.3841	-0.1159	0.1552	0.1063	0.6620	0.5557
	α	3.0539	0.0539	0.0429	2.9754	3.1323	0.1568
n=300	λ	2.0787	0.0787	0.0962	1.9023	2.2551	0.3528
	c	0.4942	-0.0058	0.0015	0.4913	0.4972	0.0058
	θ	0.4035	-0.0965	0.1228	0.1811	0.6258	0.4448
	α	3.0430	0.0430	0.0346	2.9788	3.1072	0.1284
n=500	λ	2.0580	0.0580	0.0770	1.9135	2.2024	0.2888
	c	0.4.951	-0.0049	0.0008	0.4935	0.4967	0.0032
	θ	0.4267	-0.0733	0.0920	0.2569	0.5964	0.3395

The bias and MSE are respectively calculated as;

$B\left(\Psi_j\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\Psi}_j - \Psi_j\right)$ and $MSE\left(\Psi_j\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\Psi}_j - \Psi_j\right)^2$ where, $\widehat{\Psi}_j$ stands for the estimate of Ψ_j for $j = 1, ..., 4$.

Table 5: *Results for the MLEs, Biases and MSEs, 95% C. Is, and Widths of the GTE-ED for $\alpha = 0.3,\ \lambda = 0.5,\ c = 1,\ \theta = 0.7$*

Sample	Parameter	Esimate	Bias	MSE	C I LC	C I UC	width
	α	0.3361	0.0361	0.0108	0.3175	0.3547	0.0372
n=20	λ	0.5095	0.0095	0.0126	0.4850	0.5340	0.0490
	c	1.1427	0.1427	0.2028	0.7852	1.5002	0.7150
	θ	0.6781	-0.0219	0.1180	0.4478	0.9084	0.9084
	α	0.3225	0.0225	0.0057	0.0057	0.3327	0.0204
n=50	λ	0.5003	0.0003	0.0007	0.4867	0.5140	0.0273
	c	1.0390	0.0390	0.0486	0.9467	1.1314	0.1847
	θ	0.6606	-0.0394	0.1028	0.4622	0.8590	0.3968
	α	0.3168	0.0168	0.0038	0.3098	0.3237	0.0139
n=100	λ	0.4993	-0.0007	0.0046	0.4903	0.5082	0.0180
	c	1.0111	0.0111	0.0211	0.9701	1.0523	0.0822
	θ	0.6523	-0.0477	0.0905	0.4795	0.8252	0.3458
	α	0.3101	0.0101	0.0026	0.3051	0.3150	0.0099
n=200	λ	0.5007	0.0007	0.0028	0.4952	0.4952	0.0110
	c	0.9992	-0.0008	0.0099	0.9798	1.0186	0.0389
	θ	0.6608	-0.0392	0.0705	0.5257	0.7960	0.2703
	α	0.3083	0.0083	0.0023	0.3040	0.3126	0.0086
n=300	λ	0.5016	0.0016	0.0022	0.4972	0.5060	0.0088
	c	0.9948	-0.0052	0.0067	0.9817	1.0080	0.0263
	θ	0.6599	-0.0401	0.0626	0.5403	0.7794	0.2390
	α	0.3067	0.0067	0.0019	0.3031	0.3103	0.0071
n=500	λ	0.5015	0.0015	0.0018	0.4979	0.5051	0.0072
	c	0.9935	-0.0065	0.0042	0.9854	1.0015	0.0162
	θ	0.6649	-0.0351	0.0522	0.5651	0.7647	0.1997

Interpretation of the results for tables 3 and 4:

a . The difference between the true and the estimated values of the parameters are relatively small.

b . As the sample size increases the estimates converge toward the true values of the parameters.

c . The interval widths, biases and MSEs decrease with an increase in sample size.

6. Applications

6.1. Datasets

The first data is on the remission times (in months) of a randomly selected (128) bladder cancer patients, which can be found in [15]. The second data was used by [16] and it represents the number of million revolutions before failure for each of the twenty-three ball bearings in the life tests. These datasets are used in order to check the flexibility of the proposed distribution over the competing distribution in the study.

Table 6: *Summary Statistics for the first dataset*

Min.	Mean	Variance	Max.	Skewness	Kurtosis
0.080	9.366	110.425	79.050	3.2866	18.4831

Table 7: *Summary Statistics for the second dataset*

Min.	Mean	Variance	Max.	Skewness	Kurtosis
17.88	72.23	1404.783	173.40	0.9419	3.4889

6.2. The Criteria

The generalized transmuted exponential-exponential (GTE-ED), transmuted exponential-exponential distribution (TE-ED), exponentiated exponential distribution (EED) and exponential distribution (ED) are compared using some goodness-of-fit statistics, including Akaike Information Criterion (AIC), Cramer-von Mises Criterion (W), Anderson-Darling Criterion (AD) and Kolmogorov Smirnov (KS). Furthermore, the model with the smallest values of these criteria indicates better fit. The R software (AdequacyModel package) is employed to evaluate these statistics.

Table 8: *Estimated parameters for the first data*

Model	$\widehat{\alpha}$	$\widehat{\lambda}$	$\widehat{\theta}$	$\widehat{c}$
GTE-E	0.8313	0.1007	0.7914	1.3506
TE-E	0.6401	0.0922	0.8898	-
EE	0.1199	-	-	1.2222
E	0.1066	-	-	-

Table 9: *Goodness-of-fit statistics for the first dataset*

Model	-LL	AIC	AD	W	KS
GTE-E	**410.8999**	**829.7998**	**0.3314**	**0.0523**	**0.0523**
TE-E	413.5223	833.0447	0.4945	0.0824	0.0725
EE	413.0901	830.1802	0.6705	0.1116	0.0778
E	414.3419	830.6841	0.7156	0.1192	0.0844

Tables 8 and 10 give the estimates of the parameters for the GTE-ED and the competing models in the study. The values of the computed goodness-of-fits statistics are given in tables 9 and 11. It was observed that GTE-ED has the lowest values of these statistics among the competing distributions in this study. Hence, the GTE-ED provides a better fit to the datasets.

Figures 5 and 6 show that the GTE-ED fits both the datasets well compared to the competing distributions.

Table 10: *Estimated parameters for the second data*

Model	$\widehat{\alpha}$	$\widehat{\lambda}$	$\widehat{\theta}$	$\widehat{c}$
GTE-E	0.1516	0.2095	0.1407	5.8194
TE-E	0.1053	0.2037	-0.9996	-
EE	0.0150	-	-	1.3535
E	0.0139	-	-	-

Table 11: *Goodness-of-fit statistics for the second dataset*

Model	-LL	AIC	AD	W	KS
GTE-E	**112.9714**	**233.9428**	**0.1868**	**0.0315**	**0.1037**
TE-E	116.0201	238.0402	0.2064	0.0367	0.2172
EE	118.8677	241.7355	0.2110	0.0377	0.2226
E	121.4366	244.8731	0.2157	0.0386	0.3072

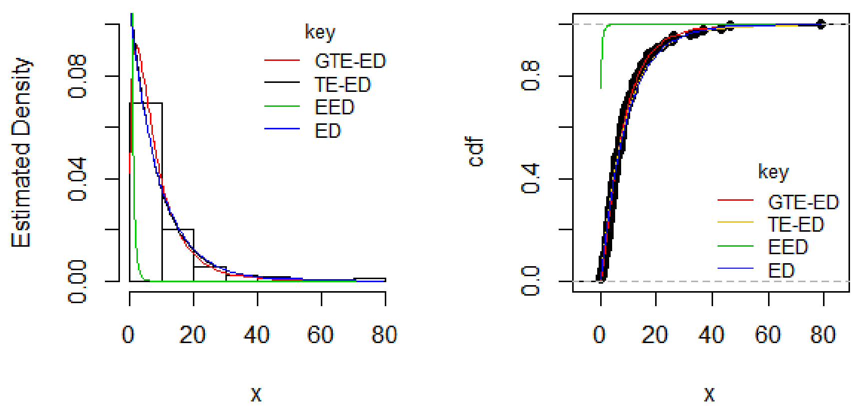


Figure 5: *Shows the estimated densities and ecdf for first data*

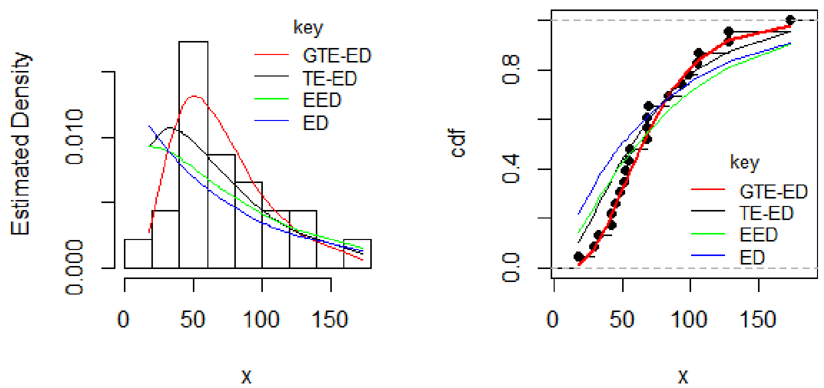


Figure 6: *Shows the estimated densities and ecdf for second data*

7. Conclusion

This research proposed an extension of the transmuted exponential-exponential distribution named the generalized transmuted exponential-exponential distribution. Expressions for some of its statistical properties, including the moments, moment generating function, limiting behaviour, and quantile function, were explicitly derived. A simulation study was conducted, and numerical values for some of the descriptive statistics were obtained and presented. The method of the maximum likelihood is adopted in estimating the unknown parameters of GTE-ED. A 95% confidence interval and interval widths together with biases, mean square errors (MSEs) of the mean estimates were equally presented on a table for different parameter values. An application to real datasets proved that the GTE-ED outperformed the competing distributions with lower values of the goodness-of-fit statistics used in this research.

Disclosure statement
On behalf of The authors, I declare that no potential conflict of interest was reported.
Funding
No funding was provided for the work.

References

[1] Tahir, M. H., & Cordeiro, G. M. (2016). *Journal of Statistical Distributions and Applications,* 3(1), 1-35.

[2] Bourguignon, M., Ghosh, I., & Cordeiro, G. M. (2016). General results for the transmuted family of distributions and new models. *Journal of Probability and Statistics,* Article ID 7208425, http://dx.doi.org/10.1155/2016/7208425

[3] Gupta, R. C., Gupta, P. L., & Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and methods,* 27(4), 887-904.

[4] Mudholkar, G. S., Srivastava, D. K., & Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data.

[5] Mudholkar, G. S., & Hutson, A. D. (1996). The exponentiated Weibull family: some properties and a flood data application. *Communications in Statistics-Theory and Methods,* 25(12), 3059-3083.

[6] Gupta, R. D., & Kundu, D. (1999). Theory and methods: Generalized exponential distributions. *Australian and New Zealand Journal of Statistics,* 41(2), 173-188.

[7] Nadarajah, S., & Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae,* 92(2), 97-111.

[8] Silva, R. B., Barreto-Souza, W., & Cordeiro, G. M. (2010). A new distribution with decreasing, increasing and upside-down bathtub failure rate. *Computational Statistics and Data Analysis,* 54(4), 935-944.

[9] Lemonte, A. J., & Cordeiro, G. M. (2011). . The exponentiated generalized inverse Gaussian distribution. *Statistics and Probability Letters,* 81(4), 506-517.

[10] Afify, A. Z., Nofal, Z. M. & Ebraheim, A. E. D. (2015).Exponentiated transmuted generalized Rayleigh Distribution. *Pakistan Journal of Statistics and Operation Research,* 11(1), 115-134.

[11] Cordeiro, G. M., Ortega, E. M., & da Cunha, D. C. (2013). The exponentiated generalized class of distributions. *Journal of data science,* 11(1), 1-27.

[12] Abba, B., Mohammed, A. S., & Musa, A. G. (2020). On the properties and mles of generalized odd generalized exponential- exponential distribution. *Fudma Journal of Sciences,*4(4), 155 - 165.

[13] Mohammed, A. S., & Ugwuowo, F. I. (2021). The Transmuted Exponential-Exponential Distribution with Application to Breast Cancer Data. *Pakistan Journal of Statistics,* 37(1), 1-18.

[14] Mohammed, A. S., & Ugwuowo, F. I. (2021).On Transmuted Exponential-Topp Leon Distribution with Monotonic and Non-Monotonic Hazard Rates and its Applications. *Reliability: Theory & Applications,* 16(4 (65)), 197-209.

[15] Lee, E. T. & 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.
[16] Lawless, J.F (1982) Statistical models and methods for lifetime data. John Wiley and Sons, New York, USA.

ANALYSIS OF MARKOVIAN BATCH SERVICE QUEUE WITH FEEDBACK AND SECOND OPTIONAL SERVICE

P. Vijaya Laxmi*[,1], Hasan A.B.D. Qrewi[2] and Andwilile A. George[3]

Department of Applied Mathematics, Andhra University,
Visakhapatnam, India.
Email: [1]vijayalaxmiau@gmail.com, [2]hasanqrewi@gmail.com,
[3]gandwilile@gmail.com
*Corresponding author.

Abstract

The aim of this paper is to analyze a single server batch service queue model with feedback and second optional service under a transient and steady state environment. The server provides the first essential service (FES) to all customers who arrive at the system and the second optional service (SOS) to those who need it. After completion of FES, if the customer is not satisfied with the service, he may rejoin the queue or may opt for SOS or exit the system with a particular probability. The service times of both FES and SOS follow exponential distribution. We use the probability generating function and the Laplace transform expression to obtain probabilities in the transient state after inverting Laplace transforms into the time domain. Also, we apply the Tauberian property in the Laplace transform expression to get the steady state probabilities. Finally, some performance measures and numerical results are provided.

Keywords: Batch service queue; Transient state; Feedback; First essential service; Second optional service.

I. Introduction

In queue theory, the customers are served one by one or served in batches whose sizes are fixed or variable in size. When the customers are served, they will depart, but when the customer does not seem satisfied they will return to the queue again to be served; this situation is called feedback. Queueing model with feedback have been studied by many researchers such as [1] who is the first to introduce the concept of feedback mechanism in queues which includes the probability of the customer to back the counter and take the service. Later, [2] investigated a single server with feedback wherein the queue is formed in two categories, one is formed in a waiting room and the other is formed in the service room and obtained the queue size, waiting time, and total time spent in the system. [3] presented a single server model with limiting behavior of the waiting time process and having a certain feedback property. Lemoine [3] classified the queue into two types: the primary queue where customers receive a maximum of time units of service. The secondary queue is formed from the customers who are not satisfied with the service in the primary queue. Those customers are attended to only when all customers who enter the system before them have departed and when the primary waiting room is empty.

[4] discussed $M/G/1/\infty$ queueing system with instantaneous Bernoulli feedback. They obtained the time-stationary distribution of the number of customers in the system at an arbitrary epoch. The queueing system with delayed feedback has been studied by [5]. They considered two servers; a lower server with a general service and an upper server with an exponential service. The decision to feedback or not depends on the queue length at the two servers. Other existing works on feedback are found in [7], [10], [9], [11], [12], [14], [15], [16], [17], etc.

In many real service systems, customers want both essential and optional services provided by the server. More precisely, we may consider a system where a single server's service is classified into two phases. The first phase is required for all customers and only some of them are routed to the second phase service. [6] studied a single server batch arrival and arbitrary service time distribution queue system with second multi-optional service and a finite number of immediate Bernoulli feedback. They provided a steady state analysis of the model, including the asymptotic behavior under a high rate of retrials.

[8] studied the priority retrial queue with immediate Bernoulli feedback and second multi-optional service. In this case, the customers' feedback after completing both phases of services. They use the embedded Markov chain technique and probability generating function to obtain the queue size and orbit size. [13] and [17] presented the queue system with batch arrival and multi-optional service with feedback. They consider two kinds of independent customers' arrival: positive and negative arrival. Positive arrivals in batches with the Poisson process and negative arrived singly with the Poisson process. The arrival of a negative customer removes the positive customer in service from the system and makes down the server.

The batch service queues have potential applications in various areas, including manufacturing, production, computer networks, cargo loading and unloading at a harbor, etc. In this situation, the number of items is processed in batches with a limit on the number of items taken at a time for processing. A number of researchers have made substantial contributions to batch service models. [19] investigated a batch service retrial system with feedback in a steady state. They considered two kinds of the arrival of the customers; the positive customer who is served in batch and the negative customer who arrives in the system and removes the positive customer. Some additional works on queuing with batch service are found in [20], [21], etc. One can study a queue system in two states viz., : transient and steady state. However, most of the literature works involved steady-state only. Steady state results are inappropriate in cases where the time horizon of operation is finite. In this situation, we need time-dependent (transient state) to analyze the system behavior by tracking down the system operation at any instant of time. The transient state has been widely studied in [23], [22], and the references therein.

This paper extends the work of [18] by including the feedback policy. The inclusion of feedback, batch service and SOS makes the model more adaptable. This motivates us to explore the Markovian batch service queue with feedback and SOS under transient and steady state environments. The rest of the paper is organized in the following way: Section 2 contains a discussion of the model as well as the mathematical formulation of differential difference equations. We use the Laplace transform and probability generating functions with Rouche's theorem in Section 3 to develop the system transient state equations. Section 4 presents the steady-state analysis, followed by Section 5 with performance measures. Section 6 presents the numerical investigation and conclusions are made in Section 7.

II. Mathematical Formulation and Description of the Model

Consider a feedback of single batch service $M/M^{[b]}/1$ queue system with SOS, in which the arrival of customers follows a Poison process with parameter λ. During FES and SOS the service times are distributed exponentially with rate μ_1 and μ_2, respectively. The services are rendered in batches that does not exceed the maximum capacity b. This means if the server finds the units that are equal or fewer than b in the waiting line, then he serves them all in the batch. Otherwise, if he sees in excess of b units in the waiting line, he takes a batch of b on a first-come, first-served, and the other units remain waiting in the queue. All arriving units required FES, and after a batch of units complete the FES, the batch (on the same server) may opt for SOS with the probability r_0 or leave the system with probability r_1. Further, if the batch is unsatisfied with the service in FES, they rejoin the queue (feedback) with probability ($r_2 = 1 - r_0 - r_1$).

Model Formulation of Differential Difference Equations

The single batch service queue with SOS can be modeled by continuous time of two dimensional Markov process $\{(K(t), W(t)); t \geq 0\}$; where

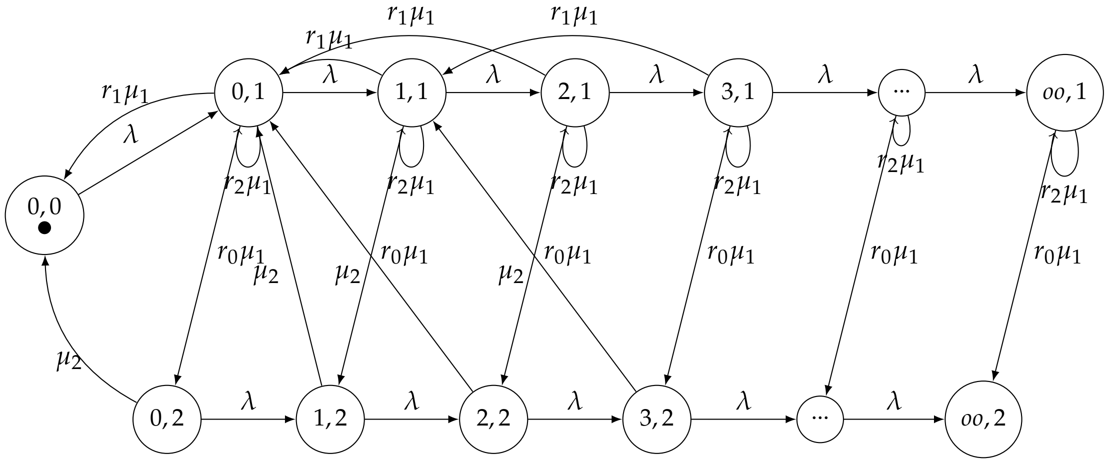


Figure 1: *Transition State diagram for b=2*

$K(t)$ is number of units in a line at time t,
$W(t)$ is server state at time t with

$$W(t) = \begin{cases} 1, & \text{FES providing by server,} \\ 2, & \text{SOS providing by server.} \end{cases}$$

The state-space of a Markov process is given as follows:

$$\Omega = \{(m, j); \quad m \geq 0;\ j = 1, 2\}.$$

The probabilities in transient state are given by

$$P_{m,j}(t) = Pr\{K(t) = m, \quad W(t) = j\}; \quad m \geq 0, \quad j = 1, 2$$

where the probability in transient state is $P_{m,j}(t)$ when m units in the waiting line at time t and the server is rendering FES ($j = 1$) or SOS ($j = 2$).
$O(t)$ is the probability in transient state when the waiting line is empty at time t and the server is idle. The differential difference equations for our model using Markov theory are as follows:

$$O'(t) = -\lambda O(t) + r_1\mu_1 P_{0,1}(t) + \mu_2 P_{0,2}(t), \tag{1}$$

$$P'_{0,1}(t) = -(\lambda + r_0\mu_1 + r_1\mu_1)P_{0,1}(t) + \lambda O(t) + r_1\mu_1 \sum_{i=1}^{b} P_{i,1}(t) + \mu_2 \sum_{i=1}^{b} P_{i,2}(t), \tag{2}$$

$$P'_{m,1}(t) = -(\lambda + r_0\mu_1 + r_1\mu_1)P_{m,1}(t) + \lambda P_{m-1,1}(t) + r_1\mu_1 P_{m+b,1}(t) + \mu_2 P_{m+b,2}(t),$$
$$m \geq 1, \tag{3}$$

$$P'_{0,2}(t) = -(\lambda + \mu_2)P_{0,2}(t) + r_0\mu_1 P_{0,1}(t), \tag{4}$$

$$P'_{m,2}(t) = -(\lambda + \mu_2)P_{m,2}(t) + \lambda P_{m-1,2}(t) + r_0\mu_1 P_{m,1}(t), \quad m \geq 1. \tag{5}$$

III. The Model's Transient Solution

In this part, the Laplace transform (L.T) and probability generating function are employed to get the transient probability of the number of units in the waiting line during server idle and busy. We consider that time has been calculated from the moment that the server has taken a batch for service, without leaving none in the queue. i.e, $P_{0,1}(0) = 1$. Let us denote the Laplace transform for $O(t), P_{m,1}(t)$ $P_{m,2}(t)$ as $O^*(s), P^*_{m,1}(s)$ and $P^*_{m,2}(s)$, respectively. Taking L.T of equations (1), (2), (3), (4), (5), we get

$$(s+\lambda)O^*(s) = r_1\mu_1 P^*_{0,1}(s) + \mu_2 P^*_{0,2}(s), \tag{6}$$

$$(s+\lambda+r_0\mu_1+r_1\mu_1)P^*_{0,1}(s) = 1 + \lambda O^*(s) + r_1\mu_1 \sum_{i=1}^{b} P^*_{i,1}(s) + \mu_2 \sum_{i=1}^{b} P^*_{i,2}(s), \tag{7}$$

$$(s+\lambda+r_0\mu_1+r_1\mu_1)P^*_{m,1}(s) = \lambda P^*_{m-1,1}(s) + r_1\mu_1 P^*_{m+b,1}(s) + \mu_2 P^*_{m+b,2}(s), \quad m \geq 1, \tag{8}$$

$$(s+\lambda+\mu_2)P^*_{0,2}(s) = r_0\mu_1 P^*_{0,1}(s), \tag{9}$$

$$(s+\lambda+\mu_2)P^*_{m,2}(s) = \lambda P^*_{m-1,2}(s) + r_0\mu_1 P^*_{m,1}(s), \quad m \geq 1. \tag{10}$$

The probability generating functions are defined as:

$$P_1(s,z) = \sum_{m=0}^{\infty} P^*_{m,1}(s) z^m, \quad P_2(s,z) = \sum_{m=0}^{\infty} P^*_{m,2}(s) z^m.$$

Multiplying equations (7) and (8) by z^m then summing over $m=0$ to $m=\infty$, adding to (6) and re-arranging the terms, we have

$$P_1(s,z) = \frac{z^b(sO^*(s)-1) + (1-z^b)\left(r_1\mu_1 \sum_{m=0}^{b-1} P^*_{m,1}(s)z^m + \mu_2 \sum_{m=0}^{b-1} P^*_{m,2}(s)z^m\right) - \mu_2 P_2(s,z)}{\lambda z^{b+1} - (s+\lambda+\mu_1)z^b + r_1\mu_1}. \tag{11}$$

Similarly, from equations (9) and (10), we get

$$P_2(s,z) = \frac{-r_0\mu_1 P_1(s,z)}{\lambda z - (s+\lambda+\mu_2)}. \tag{12}$$

Substituting equation (12) in equation (11), we have

$$P_1(s,z) = \frac{\left[z^b(sO^*(s)-1) + (1-z^b)\left[r_1\mu_1 \sum_{m=0}^{b-1} P^*_{m,1}(s)z^m + \mu_2 \sum_{m=0}^{b-1} P^*_{m,2}(s)z^m\right]\right](\lambda z - (s+\lambda+\mu_2))}{\lambda^2 z^{b+2} - \lambda(2s+2\lambda+r_0\mu_1+r_1\mu_2+\mu_2)z^{b+1} + (s+\lambda++r_0\mu_1+r_1\mu_2)(s+\lambda+\mu_2)z^b + \lambda r_1\mu_1 z - (s+\lambda+\mu_2)r_1\mu_1 - r_0\mu_1\mu_2}. \tag{13}$$

The expression for $P_1(s,z)$ has the characteristic of converging within the unit circle. We can see that the denominator of $P_1(s,z)$ has exactly $b+2$ zeros. Applying the Rouche's theorem for denominator of $P_1(s,z)$ to find the number of zeros on and within the unit circle of the analytic function, we observe that b of these roots lie inside or on the unit circle, one of zeros is $z=1$ and others $b-1$, lies inside the circle and must agree with zeros of the numerator of $P_1(s,z)$ to be converged. As a result, one zero in the denominator of $P_1(s,z)$ is canceled by $P_1(s,z)$ of the numerator, so that the two remaining zeros of the denominator are found outside the unit circle and let us consider them as z_0 and z_1. Since two polynomial differ by at most a multiplicative function (constant), let us call it $A(s)$ and is independent of z. Therefore, equation (13) can be written as

$$P_1(s,z) = \frac{(\lambda z - (s+\lambda+\mu_2))(1-z^b)A(s)}{(z-1)(z-z_0)(z-z_1)}. \tag{14}$$

Applying the rule of L'Hospital and setting $z=1$ in (14), we obtain

$$P_1(s,1) = \frac{(s+\mu_2)bA(s)}{(1-z_0)(1-z_1)}, \tag{15}$$

Taking $z=1$ in (12), we get

$$P_2(s,1) = \frac{r_0\mu_1 P_1(s,1)}{(s+\mu_2)}. \tag{16}$$

Applying the normalization condition

$$P_1(s,1) + P_2(s,1) + O^*(s) = \frac{1}{s},$$

we get

$$P_1(s,1) = \frac{(1 - sO^*(s))(s + \mu_2)}{s(s + r_0\mu_1 + \mu_2)}. \tag{17}$$

Using (15) and (17), we can determined the function $A(s)$ as

$$A(s) = \frac{(1 - sO^*(s))(1 - z_0)(1 - z_1)}{s(s + r_0\mu_1 + \mu_2)b}. \tag{18}$$

Substituting (18) into (14), we have

$$P_1(s,z) = \frac{(1 - sO^*(s))(1 - z_0)(1 - z_1)(\lambda z - (s + \lambda + \mu_2))(1 - z^b)}{s(s + r_0\mu_1 + \mu_2)b(z - 1)(z - z_0)(z - z_1)}. \tag{19}$$

When $z = 0$, equation (19) becomes

$$P^*_{0,1}(s) = \frac{(1 - sO^*(s))(q_0 - 1)(q_1 - 1)(s + \lambda + \mu_2)}{s(s + r_0\mu_1 + \mu_2)b}, \tag{20}$$

where $q_0 = \frac{1}{z_0}$, $q_1 = \frac{1}{z_1}$. Note that $P^*_{0,1}(s)$ is the L.T of the probability of empty queue and server providing FES.
From equation (12), when $z = 0$ and using (20), we have

$$P^*_{0,2}(s) = \frac{r_0\mu_1(1 - sO^*(s))(q_0 - 1)(q_1 - 1)}{s(s + r_0\mu_1 + \mu_2)b}. \tag{21}$$

$P^*_{0,2}(s)$ is the L.T of the probability of empty queue and server providing SOS.
Using (6), (20) and equation (21), we can determine $O^*(s)$ as below:

$$O^*(s) = \frac{r_1\mu_1(q_0 - 1)(q_1 - 1)(s + \lambda + \mu_2)r_0\mu_1\mu_2(q_0 - 1)(q_1 - 1)}{s\left[(s + \lambda)(s + r_0\mu_1 + \mu_2)b + (r_1\mu_1(s + \lambda + \mu_2) + r_0\mu_1\mu_2)(q_0 - 1)(q_1 - 1)\right]}. \tag{22}$$

Equation (22) indicates the Laplace transform of the state probability that the server is idle and the queue is empty.

IV. Steady State Analysis

Using the Tauberian property, we get the closed form expressions of the stationary probability for the number of items in the waiting line while the server is idle or active in FES and SOS. We define the stationary probabilities.

$$O = \lim_{t\to\infty} O(t) = \lim_{s\to 0} sO^*(s), \tag{23}$$

$$P_{m,1} = \lim_{t\to\infty} P_{m,1}(t) = \lim_{s\to 0} sP^*_{m,1}(s), \tag{24}$$

$$P_{m,2} = \lim_{t\to\infty} P_{m,2}(t) = \lim_{s\to 0} sP^*_{m,2}(s). \tag{25}$$

Assuming that the steady state probabilities exist, the equations (20), (21) and (22) are, respectively given by

$$P_{0,1} = \frac{(1 - O)(q_0 - 1)(q_1 - 1)(\lambda + \mu_2)}{(r_0\mu_1 + \mu_2)b},$$

$$P_{0,2} = \frac{r_0\mu_1(1 - O)(q_0 - 1)(q_1 - 1)}{(r_0\mu_1 + \mu_2)b},$$

$$O = \frac{r_1\mu_1(q_0 - 1)(q_1 - 1)(\lambda + \mu_2) + r_0\mu_1\mu_2(q_0 - 1)(q_1 - 1)}{\lambda(r_0\mu_1 + \mu_2)b + (r_1\mu_1(\lambda + \mu_2) + r_0\mu_1\mu_2)(q_0 - 1)(q_1 - 1)}.$$

As a special case, if $r_2 = 0$, the model reduces to $M/M^b/1$ queueing system with second optional service [18].

V. Performance Measures

The measures of performance are key aspects of queueing models as they demonstrate the models efficient and effective index. In this section, we present the probability when the server is active with FES or SOS and the anticipated number of units in the waiting line during the server busy with FES or SOS.

I. The Measures of Performance in Transient State

The probability while the server is active with FES

$$P[FES](s) = \sum_{m=0}^{\infty} P_{m,1}^{*}(s) = \frac{(1 - sO^{*}(s))(s+\mu_2)}{s(s+r_0\mu_1+\mu_2)}. \tag{26}$$

Similarly, the probability while the server is active with SOS

$$P[SOS](s) = \sum_{m=0}^{\infty} P_{m,2}^{*}(s) = \frac{r_0\mu_1(1 - sO^{*}(s))}{s(s+r_0\mu_1+\mu_2)}. \tag{27}$$

When the server is active with FES or SOS, the probability is

$$P_b(s) = \sum_{m=0}^{\infty} P_{m,1}^{*}(s) + \sum_{m=0}^{\infty} P_{m,2}^{*}(s). \tag{28}$$

When the server is active with FES the anticipated number of items in the waiting line is

$$E[FES](s) = \sum_{m=0}^{\infty} mP_{m,1}^{*}(s).$$

Taking derivative of (19) with respect to z, putting $z = 1$, then applying the rule of L'Hospital, the expected number of units in the waiting line when the server is active with FES is obtained as

$$\sum_{m=0}^{\infty} mP_{m,1}^{*}(s) = \frac{\begin{array}{r}[[((s+\mu_2)b(b-1) - 2\lambda b](q_0-1)(q_1-1) - [(s+\mu_2)b(4q_0q_1 - 2(q_0+q_1))]] \\ (1-sO^{*}(s))\end{array}}{2b\,(s(s+r_0\mu_1+\mu_2)(q_0-1)(q_1-1))}. \tag{29}$$

Similarly, from equations (12) and (29), the anticipated number of units in the waiting line when the server is active with SOS is given by

$$\begin{aligned} E[SOS](s) &= \sum_{m=0}^{\infty} mP_{m,2}^{*}(s) \\ &= \frac{\begin{array}{r}[[(b(b-1) - 2\lambda b](q_0-1)(q_1-1) - [b(4q_0q_1 - 2(q_0+q_1))]] \\ r_0\mu_1(1-sO^{*}(s))\end{array}}{2b\,(s(s+r_0\mu_1+\mu_2)(q_0-1)(q_1-1))} \\ &+ \frac{\lambda r_0\mu_1(1-sO^{*}(s))}{s(s+\mu_2)(s+r_0\mu_1+\mu_2)}. \end{aligned} \tag{30}$$

The overall queue length is given by

$$L_q(s) = \sum_{m=1}^{\infty} mP_{m,1}(s) + \sum_{m=1}^{\infty} mP_{m,2}(s) = E[FES](s) + E[SOS](s). \tag{31}$$

The overall amount of time a unit spends in the waiting line is determined by

$$W_q(s) = \frac{\sum_{m=1}^{\infty} mP_{m,1}(s) + \sum_{m=1}^{\infty} mP_{m,2}(s)}{\lambda} = \frac{E[FES](s) + E[SOS](s)}{\lambda}. \tag{32}$$

II. The Measures of Performance in Steady State

Suppose the limit of (23), (24) and (25) exist, the steady state quantities from (26) to (32), respectively are denoted by

$$
\begin{aligned}
P[FES] &= \sum_{n=0}^{\infty} P_{m,1} = \frac{(1-O)\mu_2}{r_0\mu_1+\mu_2}, \\
P[SOS] &= \sum_{m=0}^{\infty} P_{m,2} = \frac{r_0\mu_1(1-O)}{r_0\mu_1+\mu_2}, \\
P_b &= \sum_{m=0}^{\infty} P_{m,1} + \sum_{m=0}^{\infty} P_{m,2}, \\
\sum_{m=0}^{\infty} mP_{m,1} &= \frac{[[\mu_2 b(b-1)-2\lambda b](r_0-1)(r_1-1)-[\mu_2 b(4r_0r_1-2(r_0+r_1))]] \, (1-O)}{2b\,(r_0\mu_1+\mu_2)(r_0-1)(r_1-1))}, \\
\sum_{m=0}^{\infty} mP_{m,2} &= \frac{[[b(b-1)-2\lambda b](q_0-1)(q_1-1)-[b(4q_0q_1-2(q_0+q_1))]]\, r_0\mu_1(1-O)}{2b\,(\mu_2(r_0\mu_1+\mu_2)(q_0-1)(q_1-1))} \\
&\quad + \frac{\lambda r_0\mu_1(1-O)}{\mu_2(r_0\mu_1+\mu_2)}, \\
L_q &= \sum_{m=0}^{\infty} mP_{m,1} + \sum_{m=0}^{\infty} mP_{m,2} = E[FES]+E[SOS], \\
W_q &= \frac{\sum_{m=1}^{\infty} mP_{m,1} + \sum_{m=1}^{\infty} mP_{m,2}}{\lambda} = \frac{E[FES]+E[SOS]}{\lambda}.
\end{aligned}
\tag{33}
$$

VI. Numerical Results and Discussion

In this part, the numerical examples are presented after inverting L.T of equations (20), (21), (22) and (26) to (32) into time domain with the help of Mathematica software. Since an inverted L.T expressions are too long, therefore, we compute the model numerically by using the arbitrary model parameters as $\lambda = 2,\ \mu_1 = 4.5,\ \mu_2 = 3.5,\ r_0 = 0.36,\ b = 5,\ r_1 = 0.4$.
The effect of these parameters on probability and other performance measure with respect to time are shown in terms of graphs and tables.

Table 1: *The effect of r_0 on r_1, L_qFES, L_qSOS and L_q with $\lambda = 2,\ \mu_1 = 4.5,\ \mu_2 = 3.5,\ r_2 = 0.1,\ b = 5$.*

r_0	r_1	$LqFES$	$LqSOS$	L_q
0.1	0.8	1.61413	0.136298	1.75042
0.2	0.7	1.52905	0.258306	1.78736
0.3	0.6	1.44568	0.366441	1.81213
0.4	0.5	1.36723	0.462234	1.82947
0.5	0.4	1.29436	0.54716	1.84152
0.6	0.3	1.22737	0.622792	1.85016
0.7	0.2	1.16602	0.69048	1.8565
0.8	0.1	1.10997	0.751454	1.86143

Table 1 depicts the effect of the probability of opting for SOS (r_0) on the expected queue length (L_q). We observe that for fixed feedback probability (r_2), an increase in r_0 leads to a decrease in the probability of the departure (r_1), resulting in decrease of L_qFES while L_qSOS and overall L_q increases. This is because as r_0 increases, more customers attend for SOS, resulting in increasing the L_q in SOS and overall queue size, as we expect.

Table 2: *Effect of feedback probability (r_2) on r_0, L_qFES, L_qSOS and L_q with $\lambda = 2$, $\mu_1 = 4.5$, $\mu_2 = 3.5$, $r_1 = 0.2$, $b = 5$*

r_2	r_0	$LqFES$	$LqSOS$	L_q
0.2	0.6	1.29082	0.663747	1.95457
0.3	0.5	1.44771	0.630166	2.07788
0.4	0.4	1.65249	0.586778	2.23927
0.5	0.3	1.93281	0.527648	2.46045
0.6	0.2	2.34909	0.441877	2.79096
0.7	0.1	3.05593	0.30109	3.35702

Table 2 shows the effect of feedback probability (r_2) on the expected queue length. For the fixed probability of the departure (r_1), as r_2 increases, the probability of opting for SOS (r_0) decreases, which leads to decrease L_qSOS and opposite trend observed in L_qFES and overall L_q. This implies that as r_2 increases more customers are not satisfied with service (FES). Therefore, the customers feedback by joining the queue tends to increase L_qFES and overall queue size.

Table 3: *The effect of r_1 on feedback probability r_2, L_qFES, L_qSOS and L_q with $\lambda = 2$, $\mu_1 = 4.5$, $\mu_2 = 3.5$, $r_0 = 0.1$, $b = 5$*

r_1	r_2	$LqFES$	$LqSOS$	L_q
0.1	0.8	1.61413	0.136298	1.75042
0.2	0.7	1.52905	0.258306	1.78736
0.3	0.6	1.44568	0.366441	1.81213
0.4	0.5	1.36723	0.462234	1.82947
0.5	0.4	1.29436	0.54716	1.84152
0.6	0.3	1.22737	0.622792	1.85016
0.7	0.2	1.16602	0.69048	1.8565
0.8	0.1	1.10997	0.751454	1.86143

From Table 3, for the fixed probability of opting for SOS r_0, as r_1 increase, r_2 decrease, resulting in decreasing L_qFES, L_qSOS and overall L_q, which agrees with our intuition.

Table 4: *The effect of μ_1 on the expected queue length with different values of μ_2 and $r_0 = 0.36, r_1 = 0.4$, $\lambda = 3.8$.*

μ_1	$\mu_2 = 1.8$	$\mu_2 = 2$	$\mu_2 = 2.2$
2.4	6.54206	6.09142	5.7638
2.8	5.71485	5.31505	5.02568
3.2	5.23358	4.85613	4.57894
3.6	4.92135	4.55096	4.27954
4.0	4.7030 8	4.33623	4.06484
4.4	4.54338	4.17542	3.90165

The impact of μ_1 on the expected queue length (L_q) with different values of μ_2 is shown in Table 4. Here L_q deceases as μ_1 increases, as we expect. Furthermore, for fixed μ_1, as μ_2 increases, L_q decreases. The reason is that, by increasing μ_1(μ_2), customers are serviced faster and results in reducing the length of the queue.

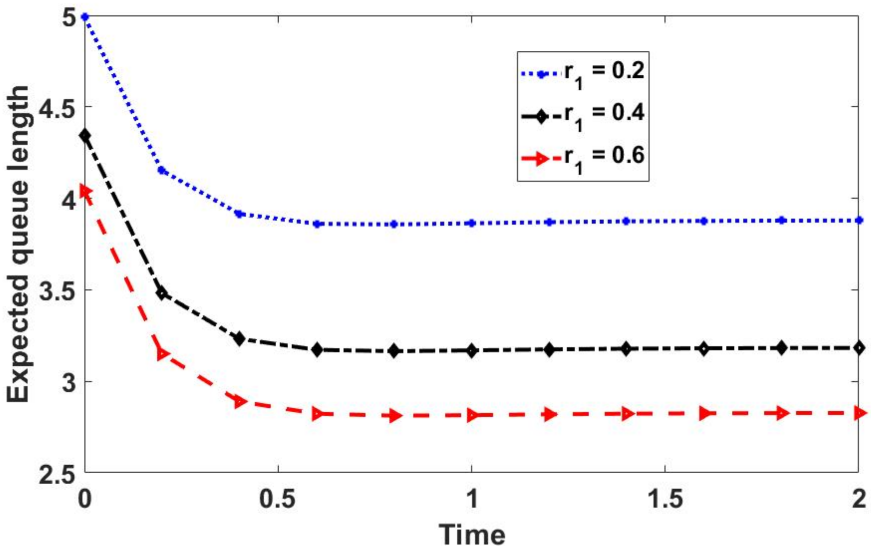


Figure 2: *The effect of departure probability r_1 on the L_q*

Figure 2 shows the effect of departure probability r_1 after completing FES on the expected queue (L_q). We observe that L_q decreases with time until it reaches a steady state. Further, we notice that as r_1 increases , L_q decreases. The reason is that more customers depart from the system after completion of FES, hence reducing L_q.

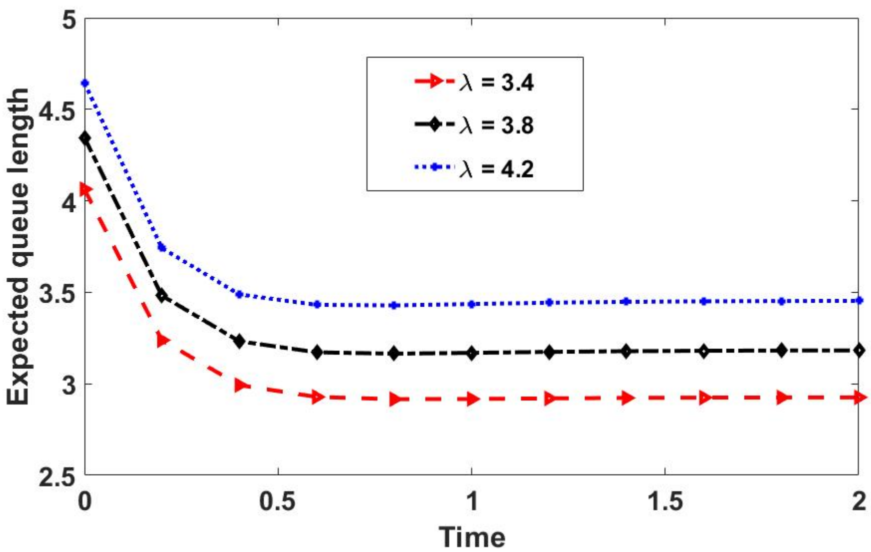


Figure 3: *The effect of λ on the expected L_q as time progresses*

.

Figure 3 display the effect of the rate of arrival (λ) on the expected queue length (L_q). As time progresses, L_q decreases until it attains steady state. Moreover, as λ increases, more customers enter the queue, resulting in increase in L_q.

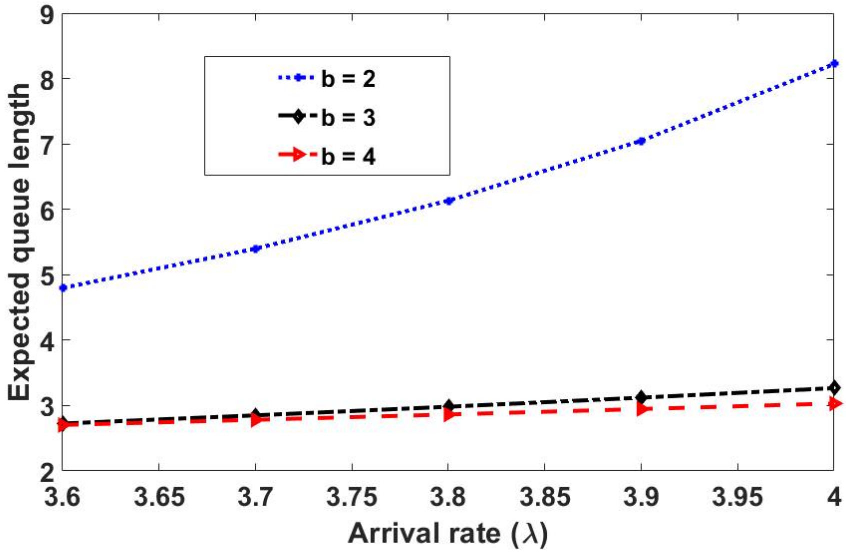


Figure 4: *The effect of arrival rate on L_q with different batch size*

.

Figure 4 displays the effect of λ on the expected length of queue (L_q) with different batch size. For fixed λ, we observe that as b increases, L_q decreases. This is because, more customers are served in batch at a time, which results in reducing the queue length. Further, as b keeps constant, L_q increases as λ increases, as intuitively expected.

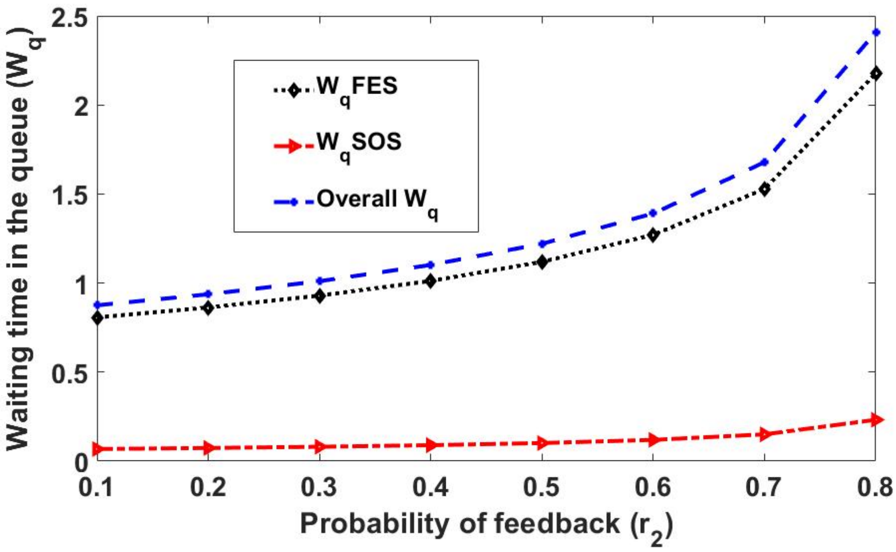


Figure 5: *The effect of r_2 on W_q*

.

The effect of the probability of feedback (r_2) on the expected waiting time in the queue (W_q) is presented in Figure 5 . We can observe from the graph that as r_2 increases, W_qFES, W_qSOS, and overall W_q increase. Because as r_2 increases, more customers rejoin the queue (feedback), resulting in increasing W_qFES, W_qSOS and overall W_q.

VII. Conclusion

We investigate the transient and steady state behavior of a single server batch service queue with SOS and feedback. We use the probability generating function, Laplace transforms, and Rouche's theorem to obtain the transient state probabilities after inverting Laplace transform into the time domain. Also, the Tauberian theorem is applied in Laplace transform expression to obtain the steady state probabilities. Finally, we present numerical results as tables and figures to show the effects of various parameters on the model performance measures.

References

[1] Takacs, L. (1963). A single server queue with feed back. *Bell System Technical journal*, 42:505–519.

[2] Takacs, L. (1977). A queuing model with feedback. *RAIRO Operations Research*, 4:345–354.

[3] Lemoine, A. J. (1974). Some limiting results for a queueing model with feedback. *The Indian Journal of Statistics, Series*, 36(3):293–304.

[4] Disney, R. L., KÃűnig, D. and Schmidt, V. (1984). Stationary queue-length and waiting-time distributions in single-server feedback queues. *Applied Probability Trust*, 16(2): 437–446.

[5] Foley, R. D. and Disney, R. L. (1983). Queues with delayed feedback. *Advances in Applied Probability*, 15(1):162–184.

[6] Al-Jirarha, J. and Madan, C. K. (2003). An $M/G/1$ queue with second optional service with general service time distribution. *Information and Management Sciences*, 14(2):47–56.

[7] Wang, W. (2008). The well-posedness and regularity of an $M^x/G/1$ queue with feedback and optional server vacations based on a single vacation policy. *International Journal of Information and Management Sciences*, 20:205–216.

[8] Kalidass, K. and Ramanath, K. (2011). A priority retrial queue with second multi optional service and m immediate Bernoulli feedback, Association for Computing Machinery. *Proceedings of the 6th International Conference on Queuing Theory and Network Applications*, 67–76.

[9] Arivudainambi, D. and Godhandaraman, P.(2012). A batch arrival retrial queue with two phases of service,feedback and K optional vacations. *Applied Mathematical Sciences*, 6(22): 1071–1087.

[10] Kalyanaraman, R. (2012). A feedback retrial queueing system with two types of batch arrivals. *Hindawi Publishing Corporation International Journal of Stochastic Analysis*, 2012.

[11] Kumar, R., Jain, N. K. and Som, B. K. (2014). Optimization of an $M/M/1/N$ feedback queuing with retention of reneged customers. *Operations Research and Decisions*, 24(3):45–58.

[12] Kumar, K., Som, B. K., and Jain, S. (2015). An $M/M/1/N$ feedback queuing system with reverse balking. *Journal of Reliability and Statistical Studies*, 8(1):31–38.

[13] Kirupa, K., and Chandrika, K. U. (2015). Batch arrival retrial queue with negative customers multi-optional service and feedback. *Communications on Applied Electronics (CAE)*, 2(4).

[14] Varalakshmi, M., Chandrasekaran, V. M. and Saravanarajan, M. C. (2018). A single server queue with immediate feedback working vacation and server breakdown. *International Journal of Engineering and Technology*, 7(4.10) Special Issue 10:476–479.

[15] Kumar, S. and Taneja, G. (2018). A feedback queuing model with chances of revisit of customer at most twice to any of the three servers. *International Journal of Applied Engineering Research*, 13(17):3093–13102.

[16] 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):1025âĂŞ-1028.

[17] Sangeetha, N. and Chandrika, K. U. (2020). Multi stage and multi optional retrial G-queue with feedback and starting failure. *AIP publishing*, 2261.

[18] Vijaya Laxmi, P. and George, A. A. (2021). Transient and steady state analysis of $M/M([b])/1$ queue with second optional service. *Journal of Industrial and Production Engineering*.

[19] Ayyappan, G., Devipriya, G. and Subramanian, A. M. G. (2014). Analysis of single Server queueing system with batch service under multiple vacations with loss and feedback. *Mathematical Theory and Modeling*, 14(11):78–89.

[20] Goswami, V. and Samanta, S. K. (2009). Discrete time bulk service queue with two heterogeneous servers. *Computers and Industrial Engineering*, 56(2):1348âĂŞ1356.

[21] Krishnamoorth, A. and Ushakumari, P. V. (2000). A queueing system with single arrival bulk service and single departure. *Mathematical and Computer Modelling*, 31:99âĂŞ108.

[22] Shanmugasundaram, S. and Chitra, S. (2016). Study on $M/M/2$ transient queue with feedback under catastrophic effect. *International Journal of Computational Engineering Research*, 6(10):2250–3005.

[23] Chandrasekaran, V. M. and Saravanarajan, M. C. (2012). Transient and reliability analysis of $M/M/1$ feedback queue subject to catastrophes, server failures and repairs. *International Journal of Pure and Applied Mathematics*, 77(5):605–625.