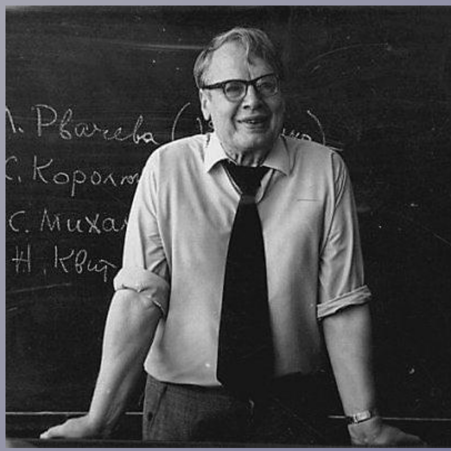
ELECTRONIC JOURNAL
OF INTERNATIONAL
GROUP ON RELIABILITY

Gnedenko Forum Publications



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

•

ISSN 1932-2321

•

VOL.13 NO.4 (51)
DECEMBER, 2018

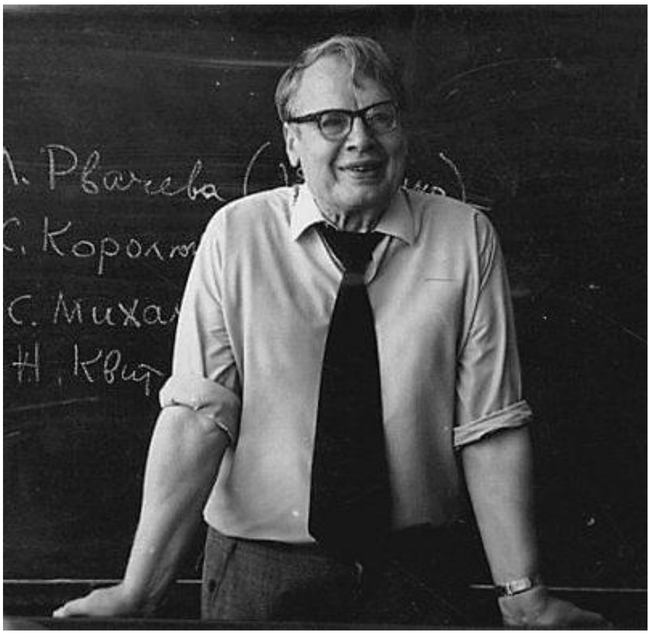
RELIABILITY:
THEORY&APPLICATIONS

San Diego

ISSN 1932-2321

Journal of International Group on Reliability





RELIABILITY:

THEORY & APPLICATIONS

Vol.13 No.4 (51),
December, 2018

San Diego
2018

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

Weibull-Lindley Distribution: A Bathtub Shaped Failure Rate Model

V.M.Chacko, Deepthi K S, Beenu Thomas, Rajitha C

•

Department of Statistics
St.Thomas College(Autonomous)
Thrissur-680001, Kerala, India
chackovm@gmail.com

Abstract

The Lindley and Weibull are the two most commonly used distributions for analyzing lifetime data. These distributions have several desirable properties and nice physical interpretations. This paper introduces a new distribution, which generalizes the well-known Lindley and Weibull distribution, having Bathtub shaped failure rate. The Statistical properties of this distribution are discussed in this paper. Applications in reliability study are discussed. A real data set is analyzed and it is observed that the present distribution can provide a better than some other very well known distributions.

Keywords: Reliability, Bathtub shaped failure rate, Weibull distribution, Lindley distribution.

I. Introduction

In order to apply suitable maintenance activities to a system or to apply reliability improvement procedures, one should know the dynamic behaviors of system reliability [2]. Increasing, decreasing and Bathtub curves are usually adopted to represent the failure rate of the system. Many statistical distributions are proposed in literature to model the Bathtub behavior of failure rate. The problem of getting optimal burn in time for the industrial burn in process is the major concern of industrial engineers. The failure rate of some engineering systems over time follows what is called the "bathtub" curve. There is a high rate of infant mortality initial failures. Then the failure rate drops, only to increase at the end of life due to wear out failures [3]. The reliability of a part can be enhanced by providing a burn-in at elevated temperatures prior to usage. This burn-in is typically done at pre specified time. It is also good to monitor the part performance during burning, so that the time point of failures can be detected. That data can be used to set the optimum burn-in length. A continuous distribution with a bathtub-shaped failure rate function with desirable characteristics is quite appropriate in this context, [9,7].

In analyzing lifetime data one often uses the Exponential, Generalized Lindley and Weibull distributions. It is well known that Exponential can have only constant hazard function, Generalized Lindley has a bathtub shape hazard function whereas Weibull can have constant or monotone (increasing/decreasing) hazard functions. Unfortunately, in practice often one needs to consider non-monotonic function such as bathtub shaped hazard function also. In this paper we present a new simple distribution which may have bathtub shaped hazard function, with high initial failure rate, which decreases rapidly and then slowly increases.

In this paper, we propose a new distribution whose failure rate function has monotone (increasing/decreasing) or bathtub shape. Section II discussed the definition of the Weibull-Lindley distribution (WLD). Section III discussed the statistical behaviours of the distribution. Section IV discussed the distribution of maximum and minimum. The maximum likelihood estimation of the parameters determined in section V. Section VI discussed three parameter Weibull-Lindley distribution (3WLD) and real data sets are analyzed in Section VII and the results are compared with existing distributions. Conclusions are given in Section VIII.

II. The Weibull-Lindley Distribution

Let X be a random variable with the following cumulative distribution function (CDF) for $\alpha, \beta, \lambda > 0$ as follows;

$$F(x;\alpha,\beta,\lambda) = 1 - e^{-\alpha\left((1+\lambda x)e^{(\lambda x)^{\beta}} - 1\right)}, x > 0 \qquad (2.1)$$

Assume λ=1 and $\alpha, \beta > 0$. Then, the probability density function (PDF) corresponding to Eq. (2.1) is given by

$$f(x;\alpha,\beta) = \alpha\left(\beta x^{\beta-1}(1+x)e^{x^{\beta}} + e^{x^{\beta}}\right)e^{-\alpha\left((1+x)e^{x^{\beta}} - 1\right)}, x > 0,\ \alpha, \beta > 0 \qquad (2.2)$$

Here β is shape parameter. The distribution with PDF of form (2.2) is said to be Weibull-Lindley distribution with parameters α, β and will be denoted by $WLD(\alpha,\beta)$.

$$f(x;\alpha,\beta) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{(-1)^{i+j}\alpha^{i+1}\beta(i-j+1)}{j!k!(i-j-k+1)!} e^{x^{\beta(i-j+1)}} x^{\beta+k-1} + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty} \frac{(-1)^{i+j}\alpha^{i+1}}{j!m!(i-j-m)!} e^{x^{\beta(i-j+1)}} x^{m} \qquad (2.3)$$

Figure 1 provide the PDFs of $WLD(\alpha,\beta)$ for different parameter values. From the below figures it is immediate that the PDFs are unimodal.

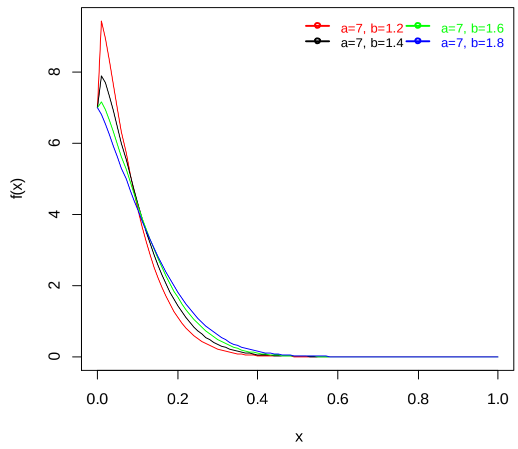


Figure 1: Probability density function of the $WLD(\alpha,\beta)$.

The survival function $S(x)$, reversed failure rate function $r(x)$ and cumulative failure rate function $H(x)$ of X are

$$S(x;\alpha,\beta)=1-F(x;\alpha,\beta)=e^{-\alpha\left((1+x)e^{x^{\beta}}-1\right)}, x>0 \tag{2.4}$$

$$r(x;\alpha,\beta)=\frac{\alpha e^{x^{\beta}}\left(1+\beta x^{\beta-1}(1+x)\right)e^{-\alpha\left((1+x)e^{x^{\beta}}-1\right)}}{1-e^{-\alpha\left((1+x)e^{x^{\beta}}-1\right)}}, x>0 \tag{2.5}$$

and

$$H(x;\alpha,\beta)=\int_0^x h(t;\alpha,\beta)dt=\alpha(1+x)e^{t^{\beta}} \tag{2.6}$$

As a result, the hazard rate function of the WL distribution can exhibit monotonically increasing, monotonically decreasing and bathtub shapes. We can see from that

$$\lim_{x\to 0} h(x) = \begin{cases} \infty, & \beta < 1 \\ 2\alpha, & \beta = 1 \\ \alpha, & \beta > 1 \end{cases}$$

Figure 2 provide the failure rate functions of $WLD(\alpha,\beta)$ for different parameter values. From the below figures it is immediate that the failure rate function can be increasing, decreasing or bathtub shaped.

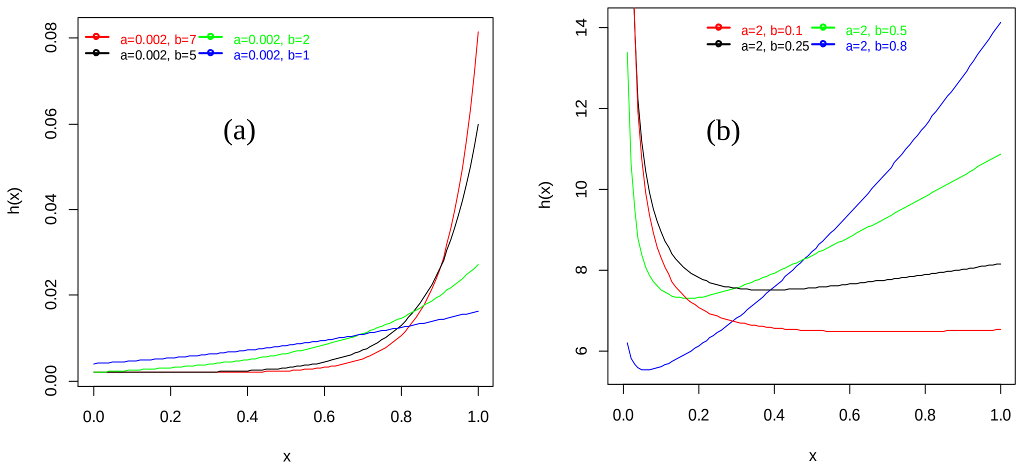


Figure 2: Failure rate function of the $WLD(\alpha,\beta)$.

It is clear that the PDF and the failure rate function have many different shapes, which allows this distribution to fit different types of lifetime data. For fixed α, the failure rate function is (a) non-decreasing function if $\beta>1$, and (b) non-increasing and bathtub function if $\beta<1$.

III. Statistical Properties

In this section, we study the statistical properties for the Weibull-Lindley distribution, specially Quantile function and Median, Mode, Moments etc

Quantile and median: We obtain the 100 p^{th} percentile,

$$(1+x)e^{x^{\beta}}=-\frac{1}{\alpha}\log(1-p)+1 \tag{3.1}$$

Setting $p = 0.5$ in Eq. (3.1), we get the median of WLD as follows.

$$(1+x)e^{x^\beta} = \frac{1}{\alpha}\log\left(\frac{1}{1-0.5}\right)+1$$

x_p is the solution of above monotone increasing function. Software can be used to obtain the Quantiles/Percentiles

Mode: Mode can be obtained as solution of

$$\frac{\partial}{\partial x}\left(\alpha\left(\beta x^{\beta-1}(1+x)e^{x^\beta}+e^{x^\beta}\right)e^{-\alpha\left((1+x)e^{x^\beta}-1\right)}\right)=0$$

$$\frac{\partial}{\partial x}\left(h(x;\alpha,\beta)\cdot S(x;\alpha,\beta)\right)=0$$

$$h^{'}(x;\alpha,\beta)\cdot S(x;\alpha,\beta)+h(x;\alpha,\beta)\cdot S^{'}(x;\alpha,\beta)=0$$

Then
$$[h^{'}(x;\alpha,\beta)-\left(h(x;\alpha,\beta)\right)^2]\cdot S(x;\alpha,\beta)=0 \tag{3.2}$$

It is not possible to get an analytic solution in x to Eq. (3.3) in the general case. It has to be obtained numerically by using methods such as fixed-point or bisection method.

Moments: If X has WLD, we obtain the r[th] moment of WLD in the form

$$\mu_r^{'}=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+\frac{\beta+r+k}{\beta(i-j+1)}}\alpha^{i+1}}{j!k!(i-j-k+1)!}\Gamma\left(\frac{\beta+r+k}{\beta(i-j+1)}\right)+\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{i+j+\frac{m+r+1}{\beta(i-j+1)}}\alpha^{i+1}}{j!m!(i-j-m)!}\frac{\Gamma\left(\frac{m+r+1}{\beta(i-j+1)}\right)}{\beta(i-j+1)} \tag{3.3}$$

If (3.3) is a convergent series for any $r \geq 0$, therefore all the moments exist and for integer values of α and β (3.4) can be represented as a finite series representation. Therefore putting $r=1$, we obtain the mean as

$$E(X)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+\frac{\beta+k+1}{\beta(i-j+1)}}\alpha^{i+1}}{j!k!(i-j-k+1)!}\Gamma\left(\frac{\beta+k+1}{\beta(i-j+1)}\right)+\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{i+j+\frac{m+2}{\beta(i-j+1)}}\alpha^{i+1}}{j!m!(i-j-m)!}\frac{\Gamma\left(\frac{m+2}{\beta(i-j+1)}\right)}{\beta(i-j+1)}$$

and putting $r=2$, we obtain the second moment as

$$E(X^2)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+\frac{\beta+k+2}{\beta(i-j+1)}}\alpha^{i+1}}{j!k!(i-j-k+1)!}\Gamma\left(\frac{\beta+k+2}{\beta(i-j+1)}\right)+\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{i+j+\frac{m+3}{\beta(i-j+1)}}\alpha^{i+1}}{j!m!(i-j-m)!}\frac{\Gamma\left(\frac{m+3}{\beta(i-j+1)}\right)}{\beta(i-j+1)}$$

which in turn can be used to obtained the higher central moments and variance.

Moment Generating Function and Characteristic function

The moment generating function, $M_X(t)$, is

$$M_X(t)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{r=0}^{\infty}\frac{t^r}{r!}\frac{(-1)^{i+j+\frac{\beta+r+k}{\beta(i-j+1)}}\alpha^{i+1}}{j!k!(i-j-k+1)!}\Gamma\left(\frac{\beta+r+k}{\beta(i-j+1)}\right)+\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}\frac{t^r}{r!}\frac{(-1)^{i+j+\frac{m+r+1}{\beta(i-j+1)}}\alpha^{i+1}}{j!m!(i-j-m)!}\frac{\Gamma\left(\frac{m+r+1}{\beta(i-j+1)}\right)}{\beta(i-j+1)}$$

The characteristic function is

$$\phi_X(t)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{r=0}^{\infty}\frac{(it)^r}{r!}\frac{(-1)^{i+j+\frac{\beta+r+k}{\beta(i-j+1)}}\alpha^{i+1}}{j!k!(i-j-k+1)!}\Gamma\left(\frac{\beta+r+k}{\beta(i-j+1)}\right)+\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}\frac{(it)^r}{r!}\frac{(-1)^{i+j+\frac{m+r+1}{\beta(i-j+1)}}\alpha^{i+1}}{j!m!(i-j-m)!}\frac{\Gamma\left(\frac{m+r+1}{\beta(i-j+1)}\right)}{\beta(i-j+1)}$$

IV. Distribution of Maximum and Minimum

Series, Parallel, Series-Parallel and Parallel-Series systems are general system structure of many engineering systems. The theory of order statistics provides a use-full tool for analysing life time data of such systems. Let $X_1, X_2, ..., X_n$ be a simple random sample from WLD with CDF and PDF as in (2.1) and (2.2), respectively. Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the order statistics obtained from this sample. The CDF of $X_{(r)}$ is given by,

$$F_{r:n}(x) = \sum_{j=r}^{n} \binom{n}{j} \left[1 - e^{-\alpha\left((1+x)e^{x^\beta}-1\right)}\right]^j \left[e^{-\alpha\left((1+x)e^{x^\beta}-1\right)}\right]^{n-j} \tag{4.1}$$

Reliability of a series system having n components with independent and identically distributed (iid) WLD distribution is $R(x) = \left[e^{-\alpha\left((1+x)e^{x^\beta}-1\right)}\right]^n$. Reliability of a parallel system having n components with iid WLD distribution is $R(x) = 1 - \left[1 - e^{-\alpha\left((1+x)e^{x^\beta}-1\right)}\right]^n$

V. Parameter Estimation

In this section, point estimation of the unknown parameters of the WLD are derived by using the method of maximum likelihood based on a complete sample data. First partial derivatives of the log-likelihood function with respect to the two-parameters are

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} (1 + x_i) e^{x_i^\beta} + n \tag{5.1}$$

and

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^{n} x_i^\beta \log x_i - \alpha \sum_{i=1}^{n} \left((1 + x_i) e^{x_i^\beta} x_i^\beta \log x_i\right) + \sum_{i=1}^{n} \frac{(1 + x_i)\left[x_i^{\beta-1} + \beta x_i^{\beta-1} \log x_i\right]}{\beta x_i^{\beta-1}(1 + x_i) + 1} \tag{5.2}$$

Setting the left side of the above two equations to zero, we get the likelihood equations as a system of two nonlinear equations in α and β. Solving this system in α and β gives the maximum likelihood estimates (MLE) of α and β. It is very easy to obtain estimates using R software by numerical methods.

Asymptotic Confidence bounds

In this section, we derive the asymptotic confidence intervals of these parameters when $\alpha > 0$ and $\beta > 0$ as the MLEs of the unknown parameters $\alpha > 0$ and $\beta > 0$ cannot be obtained in closed forms, by using variance covariance matrix I^{-1}, where I^{-1} is the inverse of the observed information matrix which defined as follows

$$I^{-1} = \begin{pmatrix} -\frac{\partial^2 L}{\partial \alpha^2} & -\frac{\partial^2 L}{\partial \alpha \partial \beta} \\ -\frac{\partial^2 L}{\partial \beta \partial \alpha} & -\frac{\partial^2 L}{\partial \beta^2} \end{pmatrix}^{-1} = \begin{pmatrix} Var(\hat{\alpha}) & Cov(\hat{\alpha}, \hat{\beta}) \\ Cov(\hat{\beta}, \hat{\alpha}) & Var(\hat{\beta}) \end{pmatrix} \tag{5.3}$$

The second partial derivatives as follows

$$\frac{\partial^2 L}{\partial \alpha^2} = -\frac{n}{\alpha}$$

$$\frac{\partial^2 L}{\partial \beta^2}=\sum_{i=1}^{n} x_i^{\beta}(\log x_i)^2-\alpha\sum_{i=1}^{n}(1+x_i)\left[e^{x_i^{\beta}}\left(x_i^{\beta}\log x_i\right)^2+e^{x_i^{\beta}}x_i^{\beta}(\log x_i)^2\right]+$$

$$\sum_{i=1}^{n}\frac{\left(\beta x_i^{\beta-1}(1+x_i)+1\right)\left((1+x_i)\beta x_i^{\beta-1}(\log x_i)^2+2(1+x_i)x_i^{\beta-1}\right)-\left((1+x_i)\left[x_i^{\beta-1}+\beta x_i^{\beta-1}\log x_i\right]\right)^2}{\left(\beta x_i^{\beta-1}(1+x_i)+1\right)^2}$$

$$\frac{\partial^2 L}{\partial \alpha \partial \beta}=\sum_{i=1}^{n}(1+x_i)e^{x_i^{\beta}}x_i^{\beta}\log x_i$$

We can derive the $(1-\delta)100\%$ confidence intervals of the parameters α and β by using variance matrix as in the following forms

$$\hat{\alpha}\pm Z_{\frac{\delta}{2}}\sqrt{Var(\hat{\alpha})},\ \hat{\beta}\pm Z_{\frac{\delta}{2}}\sqrt{Var(\hat{\beta})}$$

where $Z_{\frac{\delta}{2}}$ is the upper $\left(\frac{\delta}{2}\right)^{th}$ percentile of the standard normal distribution.

VI. Three parameter Weibull-Lindley Distribution

In order to address scaling problem, as given in (2.1), this section considered the CDF of Three parameter Weibull-Lindley Distribution (3WLD), for $\alpha,\beta,\lambda>0$ as follows;

$$F(x;\alpha,\beta,\lambda)=1-e^{-\alpha\left((1+\lambda x)e^{(\lambda x)^{\beta}}-1\right)},x>0 \tag{6.1}$$

The probability density function (PDF) corresponding to Eq. (6.1) is given by

$$f(x:\alpha,\beta,\lambda)=\alpha\left(\beta\lambda(\lambda x)^{\beta-1}(1+\lambda x)e^{(\lambda x)^{\beta}}+\lambda e^{(\lambda x)^{\beta}}\right)e^{-\alpha\left((1+\lambda x)e^{(\lambda x)^{\beta}}-1\right)},\ x>0,\ \alpha,\beta,\lambda>0 \tag{6.2}$$

Here β is shape parameter and λ is scale parameter. The distribution of this form with parameters $\alpha,\beta,$ and λ and will be denoted by $3WLD(\alpha,\beta,\lambda)$.

The survival function $S(x;\alpha,\beta,\lambda)$, failure rate function $h(x;\alpha,\beta,\lambda)$, reversed failure rate function $r(x;\alpha,\beta,\lambda)$ and cumulative failure rate function $H(x;\alpha,\beta,\lambda)$ of X are

$$S(x;\alpha,\beta,\lambda)=1-F(x;\alpha,\beta,\lambda)=e^{-\alpha\left((1+\lambda x)e^{(\lambda x)^{\beta}}-1\right)},x>0 \tag{6.3}$$

$$h(x;\alpha,\beta,\lambda)=\alpha\left(\beta\lambda(\lambda x)^{\beta-1}(1+\lambda x)e^{(\lambda x)^{\beta}}+\lambda e^{(\lambda x)^{\beta}}\right),\ x>0 \tag{6.4}$$

$$r(x;\alpha,\beta)=\frac{\alpha\left(\beta\lambda(\lambda x)^{\beta-1}(1+\lambda x)e^{(\lambda x)^{\beta}}+\lambda e^{(\lambda x)^{\beta}}\right)e^{-\alpha\left((1+\lambda x)e^{(\lambda x)^{\beta}}-1\right)}}{1-e^{-\alpha\left((1+\lambda x)e^{(\lambda x)^{\beta}}-1\right)}},x>0 \tag{6.5}$$

and

$$H(x;\alpha,\beta,\lambda)=\int_0^x h(t;\alpha,\beta,\lambda)dt=\alpha(1+\lambda x)e^{(\lambda x)^{\beta}} \tag{6.6}$$

respectively.

Figure 3 and Figure 4 provide the PDFs and the failure rate functions of $GoED(\alpha,\beta,\lambda)$ for different parameter values. From the below figures it is immediate that the PDFs can be unimodel and the failure rate function can be increasing, decreasing or bathtub shaped

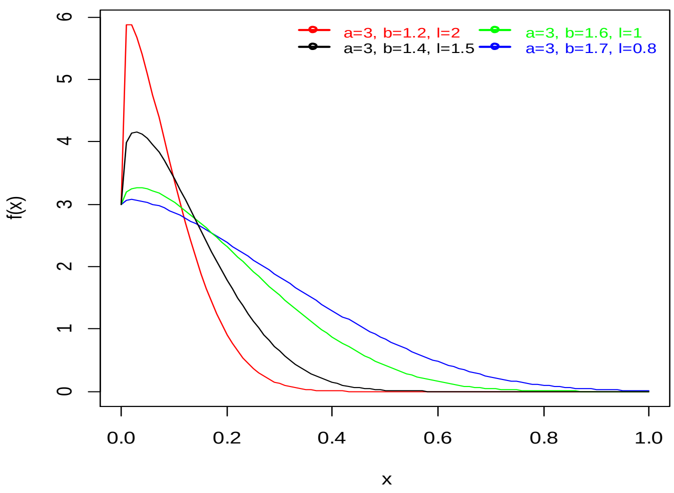


Figure 3: Probability density function of the $3WLD(\alpha, \beta, \lambda)$.

It is clear that the PDF and the failure rate function have many different shapes, which allows this distribution to fit different types of lifetime data.

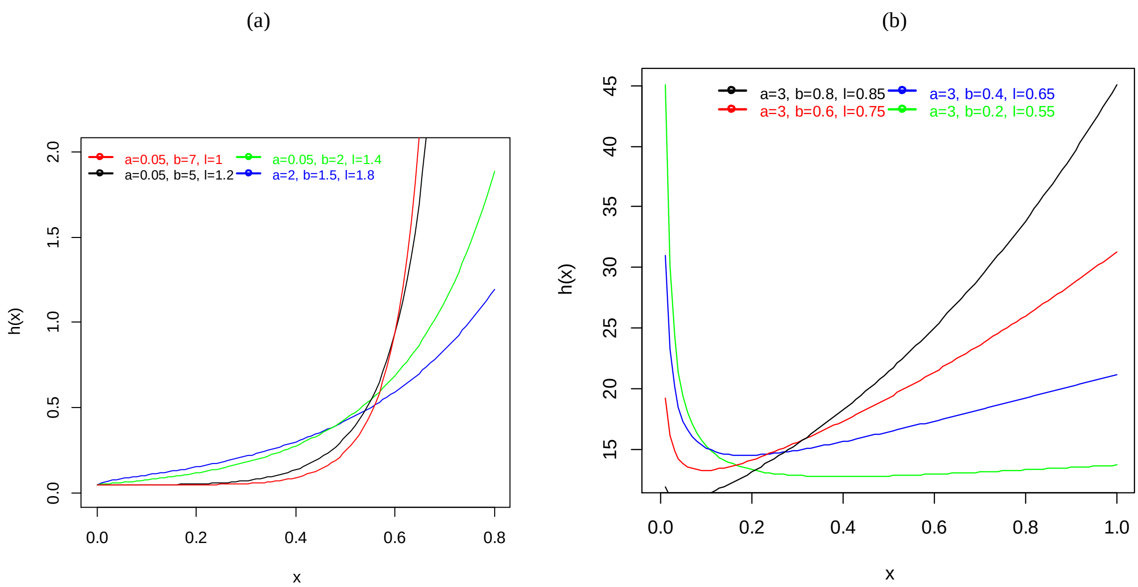


Figure 4: Failure rate function of the $3WLD(\alpha, \beta, \lambda)$.

For fixed α , the failure rate function is (a) non-decreasing function (IFR) if $\beta > 1$ and $\lambda > 1$, and (b) non-increasing (DFR) and bathtub function if $\beta < 1$ and $\lambda < 1$.

Parameter Estimation

In this section, point estimation of the unknown parameters of the 3WLD are derived by using the method of maximum likelihood based on a complete sample data. The first partial derivatives of the log-likelihood function with respect to the three-parameters are

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} (1 + \lambda x_i) e^{(\lambda x_i)^\beta} + n \tag{6.7}$$

and

$$\frac{\partial L}{\partial \beta}=\sum_{i=1}^{n}(\lambda x_i)^{\beta}\log(\lambda x_i)-\alpha\sum_{i=1}^{n}\left((1+\lambda x_i)e^{(\lambda x_i)^{\beta}}(\lambda x_i)^{\beta}\log(\lambda x_i)\right)+\sum_{i=1}^{n}\frac{(1+\lambda x_i)\left((\lambda x_i)^{\beta-1}+\beta(\lambda x_i)^{\beta-1}\log(\lambda x_i)\right)}{\beta(\lambda x_i)^{\beta-1}(1+\lambda x_i)+1} \tag{6.8}$$

$$\frac{\partial L}{\partial \lambda}=\frac{n}{\lambda}+\sum_{i=1}^{n}x_i^{\beta}\beta\lambda^{\beta-1}-\alpha\sum_{i=1}^{n}\left((1+\lambda x_i)e^{(\lambda x_i)^{\beta}}x_i\beta(\lambda x_i)^{\beta}+x_ie^{(\lambda x_i)^{\beta}}\right)+\sum_{i=1}^{n}\frac{\beta x_i^{\beta}\left((\beta-1)\lambda^{(\beta-2)}+x_i\beta\lambda^{\beta-1}\right)}{\left(\beta(\lambda x_i)^{\beta-1}(1+\lambda x_i)+1\right)} \tag{6.9}$$

Setting the left side of the above three equations to zero, we get the likelihood equations as a system of three nonlinear equations in α, β and λ. Solving this system in α, β and λ gives the maximum likelihood estimates (MLE) of α, β and λ. It is very easy to obtain estimates using R software by numerical methods.

VII. Application

In this section, we present the analysis of a real data set using the $WLD(\alpha, \beta)$ and $3WLD(\alpha, \beta)$ model and compare it with the other bathtub models such as Generalized Lindley distributions (GLD), [7], Exponentiated Weibull distribution (EW), [9], using Kolmogorov-Smirnov (K-S) statistic. We considered the data sets are obtained strengths of 1.5 cm glass fibres data [10] and infection for AIDS data [4] to estimate the parameter values.

Data Set 1: The data are the strengths of 1.5 cm glass fibres [10], measured at the National Physical Laboratory, England. Unfortunately, the units of measurement are not given in the paper. This data set 1 is in Table 1.

0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.5, 1.54, 1.6, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.5, 1.55, 1.61, 1.62, 1.66, 1.7, 1.77, 1.84, 0.84, 1.24, 1.3, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.7, 1.78 and 1.89.

Table 2 gives MLEs of parameters of the WLD, GLD, EW and 3WLD and goodness of fit statistics.

Table 2: MLEs of parameters, Log-likelihood.

Model	MLEs of parameters	log L	K-S	p-value
WLD	$\hat{\alpha}$ = 0.02852 $\hat{\beta}$= 1.8927	-16.63882	0.13681	0.189
GLD	$\hat{\alpha}$ = 26.17181 $\hat{\lambda}$ = 2.990087	-30.61986	0.22639	0.003136
EW	$\hat{\alpha}$ = 7.2847 $\hat{\beta}$ = 0.67122 $\hat{\lambda}$ =0.58203	-14.67552	0.14623	0.1352
3WLD	$\hat{\alpha}$ = 0.000212 $\hat{\beta}$=0.83783 $\hat{\lambda}$ =5.32574	-14.42277	0.12564	0.273

3WLD gives the largest Log-likelihood value and largest p value based on the KS statistic. The second largest Log-likelihood value and p value based on the KS statistic is given by the EW distribution. The third largest Log-likelihood value and p value based on the KS statistic is given by the WL distribution.

Figure 5 gives the form of the failure rate for the WLD and 3WLD which are used to fit the data after replacing the unknown parameters.

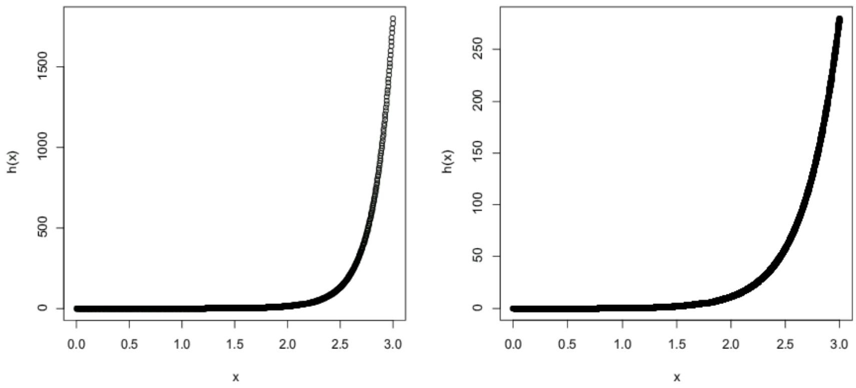


Figure 5: Failure rate function for WLD and 3WLD

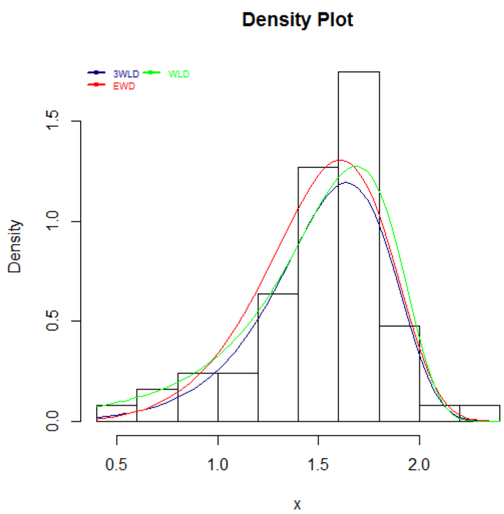


Figure 6: Fitted pdfs of the three best fitting distributions for data set 1.

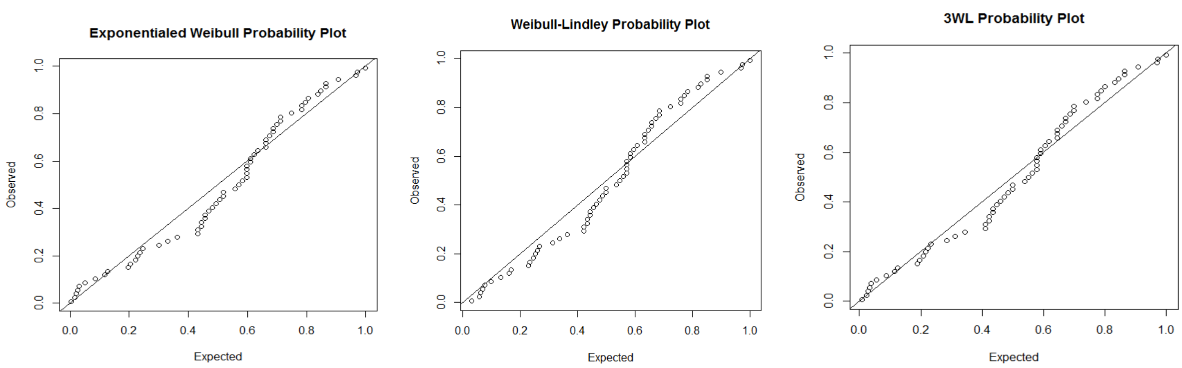


Figure 7: PP plots of the three best fitting distributions for data set 1.

Data Set 2: The second data set are times to infection for AIDS for two hundred and ninety five patients. The data were taken from Section 1.19 of Klein and Moeschberger [4]. The two distributions were fitted to this data. The parameter estimates and the goodness of fit statistics are given in Table 3. Table 3 gives MLEs of parameters of the WLD, GLD, EW and 3WLD and goodness of fit statistics.

Table 3: MLEs of parameters, Log-likelihood

Model	MLEs of parameters	log **L**	K-S	p-value
WLD	$\hat{\alpha}$ = 0.03552611 $\hat{\beta}$= 0.57122324	-457.3015	0.077618	0.08931
GLD	$\hat{\alpha}$ = 2.4144951 $\hat{\lambda}$ = 0.8924887	-453.523	0.71652	2.22×10^{-16}
EW	$\hat{\alpha}$ = 1.9565778 $\hat{\beta}$ = 0.9598033 $\hat{\lambda}$ = 0.3212501	-450.1305	0.063912	0.2426
3WLD	$\hat{\alpha}$ = 8.751896×10^{-04} $\hat{\beta}$ = 0.2994 $\hat{\lambda}$ = 15.0999	-451.8749	0.061941	0.2755

Here, EW gives the largest Log-likelihood value and largest p value based on the KS statistic. The second largest Log-likelihood value and p value based on the KS statistic is given by the 3WL distribution. The third largest Log-likelihood value and p value based on the KS statistic is given by the WL distribution.

Figure 8 and 9 gives the form of the failure rate for the WLD and 3WLD which are used to fit the data after replacing the unknown parameters.

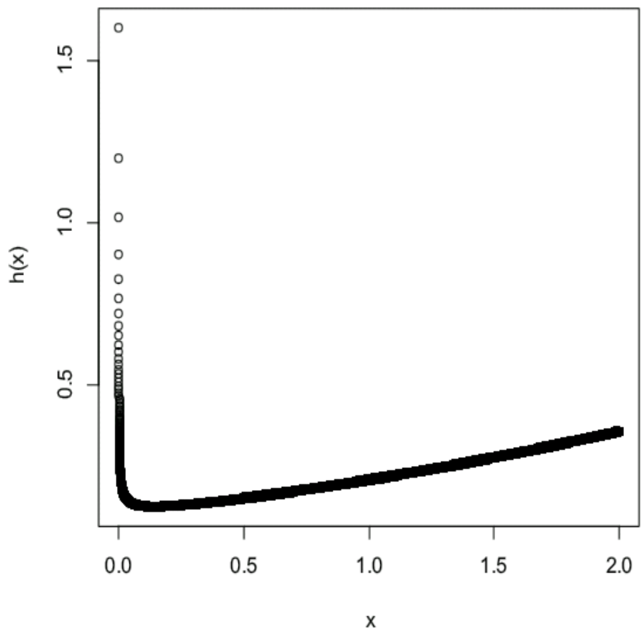


Figure 8: Failure rate function for WLD

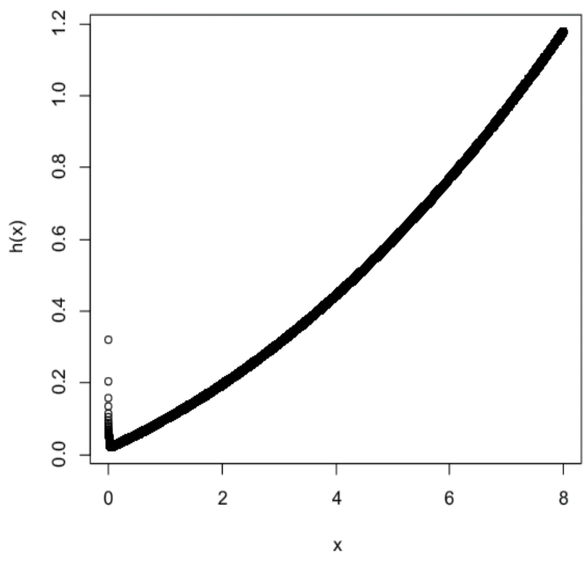


Figure 9: Failure rate function for 3WLD

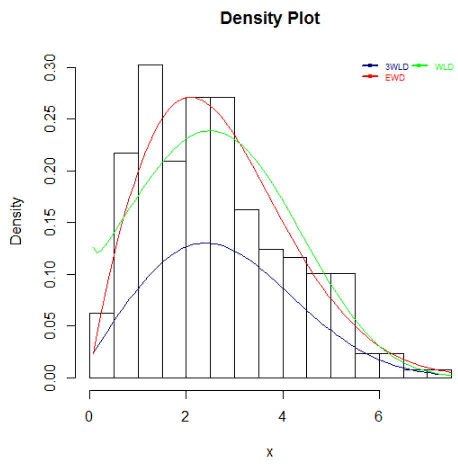


Figure 10: Fitted pdfs of the three best fitting distributions for data set 2.

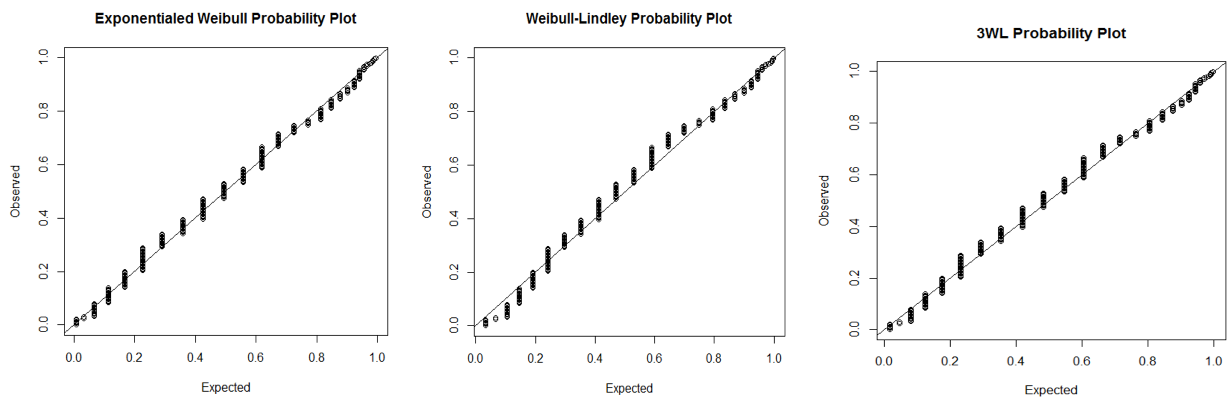


Figure 11: PP plots of the three best fitting distributions for data set 2.

It is observed that 3WLD fits the best in the first data set whereas EW fits the best in the second data in terms of likelihood and in terms of KS Statistic. Therefore, it is not guaranteed the 3WLD will behave always better than WLD or EW or GLD but at least it can be said in certain circumstances 3WLD might work better than WLD or EW or GLD.

VIII. Conclusion

A new distribution, Weibull-Lindley distribution (WLD), has been proposed and its properties studied. Three parameter Weibull-Lindley distribution (3WLD) is introduced for avoid scale problem. We have studied maximum likelihood estimators and the parameters estimation is carried out in the presence of real data. We present two real life data sets, where in one data set it is observed that 3WLD has a better fit compare to EW or WLD or GLD but in the other the EW has a better fit than 3WLD or WLD or GLD.

References

[1] Ammar M. Sarhan.: A two-parameter discrete distribution with a bathtub hazard shape. Com. for Stat. Appl. & Meth. 24(1), 15–27 (2017).

[2] Chacko, V.M.: A new bathtub shaped failure rate model, Reliability: Theo. & Appl. 3(42), Vol.11 (2016).

[3] Chacko, V.M.: X-Exponential bathtub failure rate model. Reliability: Theo & Appl. 4 (43), 55-66 (2016).

[4] Klein, J.P., Moesch Berger, M.L. (ed.): Survival Analysis Techniques for Censored and Truncated Data. Springer Verlag, New York (1997).

[5] Pal, M., Ali, M.M., Woo, J.: A Generalization of the Weibull Distribution with Application to the Analysis of Survival Data. STATISTICA. Anno LXVI, n. 2 (2006).

[6] Mustafa.A., Beih. S. El-Desouky, AL-Garash, S.: Weibull Generalized Exponential Distribution.

[7] Nadarajah, S., Bakouch, H.S., Tahmasbi, R.: A generalized Lindley distribution. Sankhya B, 73, 331-359 (2011).

[8] Nadarajah, S., Gupta, A.K.: The exponentiated gamma distribution with application to drought data. Calcutta Stat. Asso. Bulletin. 59, 233–234 (2007).

[9] Pal, M., Ali, M.M., Woo, J.: Exponentiated Weibull Distribution, Statistica, Anno LXVI, n.2 (2006).

[10] Smith, R. L. and Naylor, J. C.: A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. Appli. Stat. 36, 358–369 (1987).

A Practical Approach for Performing Multi-response Optimization for Advanced Process Control

Prasath Ganesan[1],
Sachin Patil[2], Vikas Pandey[3], Tony M Shaju[4], Srinivas Pothula[5]

•

Eaton Corporation
[1]PrasathGanesan@Eaton.com
[2]SachinPatil@Eaton.com
[3]VikasPandey@Eaton.com
[4]TonyMShaju@eaton.com
[5]SrinivasPothula@Eaton.com

Abstract

In account of the statistical methods used in advanced manufacturing process optimization, multi-response optimization is one of the key areas of focus. Previously multi-response optimization problems were solved by past experiences and engineering judgment by many industries which lead to uncertainty in the decision making & and less confidence in getting optimized process parameters to produce robust products. For identifying the optimal process parameters for a manufacturing a robust product in which multiple CTQ (Critical-to-Quality) characteristics need to be achieved, a systematic statistical optimization approach is required. This paper presents one of the practical systematic approaches for multi-response optimization of advanced manufacturing processes. This statistical methodology uses Taguchi DoE (Design of Experiment) based approach to optimize the process parameters for individual CTQ followed by a multi-response optimization using composite desirability functions to achieve multiple CTQs.

Keywords: Multi-response optimization, Design of Experiments, Critical-to-Quality, Taguchi, Regression

I. Introduction

In general, advanced polymer manufacturing processes require extreme control over multiple process parameters (control factors) to achieve desired quality in the final product. Quality of the product manufactured by different processes like injection molding, blow molding, compression molding, thermoforming, extrusion etc., drastically varies with respect to change in set process parameters like temperature, cooling time, cooling rate, pressure, material flow rate, etc. Therefore, it is important to choose the right settings for each control factor to achieve the right critical-to-quality (CTQ) parameter/properties. As the number of CTQs to be satisfied increases, the processing window becomes narrow. Hence, it is a challenging task for the engineers to arrive at the right process window when multiple CTQs need to be satisfied to achieve the desired quality.

Various Design of Experiment (DoE) based approaches are used in the industry to identify the right process window which would help to achieve the CTQs of the final product. In reality, the overall quality of the product is determined by multiple CTQs. It is almost impossible to achieve multiple CTQs using only one set of control factor values. Therefore, multiple control factors have to be optimized to get a balanced trade-off between all the CTQs, without compromising the overall performance.

Taguchi is one of the popular DoE approaches, where multiple process parameters can be controlled to manufacture a robust product with minimum number of experimental runs [1]. Although Taguchi method helps to optimize multiple process parameters (control factors) to satisfy single CTQ (single response) at a time, it may not optimize the process parameters when multiple CTQs (multiple responses) are to be satisfied simultaneously. Traditionally, Analytical Hierarchical Process (AHP) is used to obtain balanced trade-off when the importance level of multiple CTQs is already known based on engineering judgement. This method might not be a recommended solution for multi-response optimization when the manufacturing process is complex and is very sensitive to minor changes in control factors. Other methods include fuzzy logic [2] which employs grey relational ranking analysis, ridge analysis [3] where single response is maximized while keeping the other responses constrained within certain targeted values, loss function based approach etc. In the present work and alternative approach, where regression functions in combination with Taguchi responses was used to perform multi-response optimization. In this approach a regression functions was generated using Ordinary Least Squares regression (OLS) [4], Generalized Least Squares regression (GLS) [5], or Multivariate regression (MVR) [6] from the responses obtained from the minimum number of runs suggested by Taguchi method. Subsequently multi-response optimization is performed using the output functions obtained from the regression analysis through desirability function approach.

II. Multi-response Optimization Procedure

As described earlier, multi-response optimization is used specifically when there is a need to optimize the control factors in order to satisfy more than one CTQ at a time. It is all about determining a point or range in design space that helps to meet all the CTQ requirements. The system of equations become even more complex when there is interaction between the control factors.

The step-by-step approach employed in this work for multi-response optimization is depicted in Figure 1 using Minitab statistical software. The procedure starts with obtaining responses for every CTQ using Taguchi runs. These multiple responses are fitted into regression models and are fed into desirability functions which perform multiple iterations to arrive at a set of desired control factor values. The desirability function provides a maximum possible desirable value for every CTQ with one set of optimized control factors.

I. Identifications of Control factors

It is very important to understand the process of advanced manufacturing to recognize the effect of control factors on typical CTQs of the product. The parameters which can lead to variations on multiple CTQ characteristics are identified and are used as the control factors in the DoE.

II. Taguchi DoE

Taguchi method is commonly used for optimizing the design parameters with less number of experimental runs [1] as compared to factorial designs. Taguchi DoE uses orthogonal arrays to organize the parameters affecting (control factors) the process and the levels at which they should be varied. Based on the number of parameters and number of levels, appropriate orthogonal arrays can be selected, Table 1.

III. Optimal Process Parameters for Individual CTQs using Taguchi DoE

For example, if we consider a case with 4 control factors to be varied in 3 levels, an L9 orthogonal array with 9 runs is suggested by Taguchi method, Table 2. For each response (CTQ), the optimal control factor setting is obtained from the maximum S/N ratio (Signal–to-Noise ratio) value, while analyzing the DoE using Minitab. Depending upon whether to maximize or minimize the response, S/N ratio value was chosen as either smaller-the-better or larger-the-better or nominal-the-better option, as represented in Figure 2.

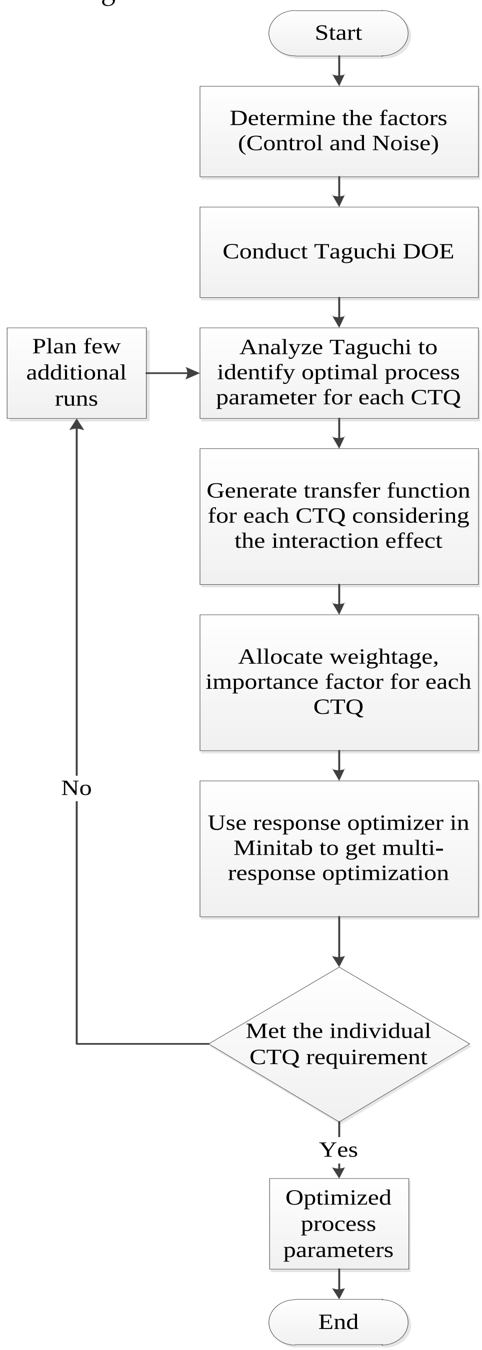


Figure 1: *Multi-Response Optimization Procedure*

IV. Optimal Process Parameters for Individual CTQs using Taguchi DoE

For example, if we consider a case with 4 control factors to be varied in 3 levels, an L9 orthogonal array with 9 runs is suggested by Taguchi method, Table 2. For each response (CTQ), the optimal control factor setting is obtained from the maximum S/N ratio (signal–to-Noise ratio) value, while analyzing the DoE using Minitab. Depending upon whether to maximize or minimize the response, S/N ratio value was chosen as either smaller-the-better or larger-the-better or nominal-the-better option, as represented in Figure 2.

Table 1: *Orthogonal Array selector for Taguchi*

Parameter# Level#	2	3	4	5	6	7	8	9	10	11	12	13
2	L4	L4	L8	L8	L8	L8	L12	L12	L12	L12		
3	L9	L9	L9	L18	L18	L18	L18	L27	L27	L27	L27	L27
4	L16	L16	L16	L16	L32	L32	L32	L32	L32			
5	L25	L25	L25	L25	L25	L50	L50	L50	L50	L50	L50	

Table 2: *L9 (32) Orthogonal Arrays*

	Independent Variables			
Exp#	Var 1	Var 2	Var 3	Var 4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

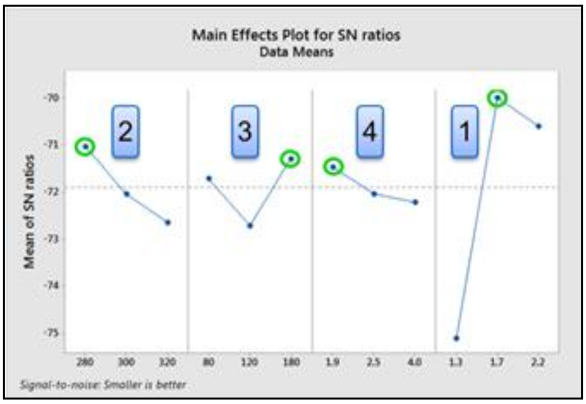


Figure 2: *Signal-to-Noise (S/N) Ratio (Smaller the Better)*

V. Transfer Function for Each CTQs using Regression

In general the transfer function generated from the Taguchi DoE analysis is a Taylor Series approximation. In order to obtain an analytical transfer function, the optimal control factor value obtained for individual CTQ responses from Taguchi method is used to generate transfer function by regression analysis, using Minitab. Each regression based transfer function is validated with the parameter effects and responses of the CTQs.

VI. Allocation of the Weightage and Importance to Each CTQ

The regression based transfer functions are allocated with weightages and importance based on the CTQs criticality. This will help in providing priority to certain CTQs over the other during the optimization process.

VII. Multi-Response Optimization using Minitab

Multi-response optimization is performed using the optimizer function inbuilt in the Minitab. The importance level of each CTQ is fed into the optimizer function [7]. This optimizer function performs iteration over the control factors and identifies an appropriate most favorable range which satisfies desired multiple responses using desirability function in the Minitab. Desirability in the response optimizer suggests the best combination of control factors which will satisfy the goals that are defined for the multiple CTQs. Individual desirability indicates how well single CTQ is satisfied whereas Composite desirability indicates how the requirements for multiple CTQs are satisfied simultaneously. Desirability has a range of 0 to 1 where 1 is the most favorable case and 0 indicates that one or more CTQ's are outside acceptable limits. The desirability function depends upon the weightage and important index given for each CTQ. The optimizer has a very important feature of visualizing the effect of each parameters on the CTQs which can be varied and cross verified; the below Figure 3 shows the response optimizer in Minitab as an example.

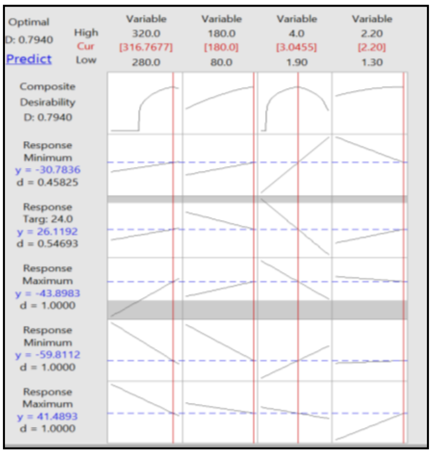


Figure 2: *Sample Response Optimizer in Minitab*

IV. Conclusion

The approach described in the work was found to be more practical and simpler one which can be adopted for optimizing new manufacturing processes, where the process history is not fully known. Especially when the process is very complex, where more process parameters to controlled, which involves higher lead time and considerable budget for identifying the right process setting, this approach would be simpler to adapt.

References

[1] Wysk, R. A., Niebel, B. W., Cohen, P. H., and Simpson, T. W. Manufacturing Processes: Integrated Product and Process Design, McGraw Hill, New York, 2000

[2] Kim, Kwang-Jae & Lin, Dennis K. J. (1998). Dual response Surface optimization: A Fuzzy

Modeling Approach. Journal of Quality Technology, 30 (1), 1-10.
[3] Myers, R. H. & Carter, W. H. (1973). Response Surface Techniques for Dual Response Surfaces. Technometrics, 15, 301-317.
[4] Draper, N. & Smith, H. (1981). Applied Regression Analysis. 2nd Ed. John Wiley, New York.
[5] Myers, R. H. (1986). Classical and Modern Regression with Applications. Duxbury Press, Boston.
[6] Seber, G. A. F. (1984). Multivariate Observations. John Wiley, New York.
[7] Minitab, Response Optimizer Factorial Designs.

Om Distribution With Properties And Applications

Rama Shanker & Kamlesh Kumar Shukla*
●
Department of Statistics
College of Science, EIT, Asmara, Eritrea
kkshukla22@gmail.com , shankerrama2009@gmail.com
* Corresponding Author

Abstract

A new one parameter lifetime distribution named, 'Om distribution' has been proposed and studied. Its various statistical properties including shapes for probability density, moments based measures, hazard rate function, mean residual life function, stochastic ordering, mean deviations, Bonferroni and Lorenz curves, distribution of order statistics, and stress-strength reliability have been discussed. Estimation of parameter has been discussed with the method of maximum likelihood. Applications of the distribution have been explained through two examples of real lifetime data from engineering and the goodness of fit found to be quite satisfactory over several one parameter lifetime distributions.

Keywords: Lifetime distributions, Statistical Properties, Maximum likelihood estimation, Applications

I. Introduction

In the present world, the time to the occurrence of some event is of interest for some populations of individuals in almost every field of knowledge. The event may be death of a person or any living creature, failure of a piece of electronic equipment, development (or remission) of symptoms. In reliability analysis, the time to the occurrences of events are known as "lifetimes" or "survival times" or "failure times" according to the event of interest in the fields of study. The modeling and statistical analysis of lifetime data has been a topic of considerable interest to statisticians and research workers in engineering, biomedical science, insurance, finance, amongst others. Applications of lifetime distributions range from investigations into the endurance of manufactured items in engineering to research involving human diseases in biomedical sciences.

During recent decades, a number of one parameter and two-parameter lifetime distributions for modeling lifetime data have been introduced by different researchers in statistics. The popular one parameter lifetime distributions available in statistics literature are exponential distribution and Lindley distribution introduced by Lindley (1958). Recently Shanker (2015 a, 2015 b, 2016 a, 2016 b, 2017 a, 2017 b, 2017 c, 2017 d) has proposed several one parameter lifetime distributions, namely Shanker, Akash, Sujatha, Aradhana, Rama, Akshaya, Amarendra and Devya, and it has been showed by Shanker that these distributions have advantages and disadvantages over the others. The probability density function (pdf) and the cumulative distribution function (cdf) of exponential, Lindley, Shanker, Akash, Sujatha, Aradhana, Rama, Akshaya, Amarendra and Devya distributions along with their introducers and year have been presented in table 1.

Table 1: pdf and cdf of one parameter lifetime distributions

Distributions	pdf and cdf	Introducer(Year)
Devya	$f(x;\theta)=\frac{\theta^5}{\theta^4+\theta^3+2\theta^2+6\theta+24}\left(1+x+x^2+x^3+x^4\right)e^{-\theta x}$	Shanker (2016 d)
	$F(x,\theta)=1-\left[1+\frac{\theta^4\left(x^4+x^3+x^2+x\right)+\theta^3\left(4x^3+3x^2+2x\right)+6\theta^2\left(2x^2+x\right)+24\theta x}{\theta^4+\theta^3+2\theta^2+6\theta+24}\right]e^{-\theta x}$	
Amarendra	$f(x;\theta)=\frac{\theta^4}{\theta^3+\theta^2+2\theta+6}\left(1+x+x^2+x^3\right)e^{-\theta x}$	Shanker (2016 c)
	$F(x,\theta)=1-\left[1+\frac{\theta^3x^3+\theta^2(\theta+3)x^2+\theta\left(\theta^2+2\theta+6\right)x}{\theta^3+\theta^2+2\theta+6}\right]e^{-\theta x}$	
Akshaya	$f(x;\theta)=\frac{\theta^4}{\theta^3+3\theta^2+6\theta+6}(1+x)^3e^{-\theta x}\quad ;x>0,\ \theta>0$	Shanker (2017 b)
	$F(x;\theta)=1-\left[1+\frac{\theta^3x^3+3\theta^2(\theta+1)x^2+3\theta\left(\theta^2+2\theta+2\right)x}{\theta^3+3\theta^2+6\theta+6}\right]e^{-\theta x}$	
Rama	$f(x;\theta)=\frac{\theta^4}{\theta^3+6}\left(1+x^3\right)e^{-\theta x}$	Shanker (2017 a)
	$F(x,\theta)=1-\left[1+\frac{\theta^3x^3+3\theta^2x^2+6\theta x}{\theta^3+6}\right]e^{-\theta x}$	
Aradhana	$f(x;\theta)=\frac{\theta^3}{\theta^2+2\theta+2}(1+x)^2e^{-\theta x}\quad ;x>0,\ \theta>0$	Shanker (2016 b)
	$F(x;\theta)=1-\left[1+\frac{\theta x(\theta x+2\theta+2)}{\theta^2+2\theta+2}\right]e^{-\theta x}\quad ;x>0,\theta>0$	
Sujatha	$f(x;\theta)=\frac{\theta^3}{\theta^2+\theta+2}\left(1+x+x^2\right)e^{-\theta x}\quad ;x>0,\ \theta>0$	Shanker (2016 a)
	$F(x,\theta)=1-\left[1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right]e^{-\theta x}$	
Akash	$f(x;\theta)=\frac{\theta^3}{\theta^2+2}\left(1+x^2\right)e^{-\theta x}\quad ;x>0,\ \theta>0$	Shanker (2015 b)
	$F(x;\theta)=1-\left[1+\frac{\theta x(\theta x+2)}{\theta^2+2}\right]e^{-\theta x}\quad ;x>0,\theta>0$	
Shanker	$f(x;\theta)=\frac{\theta^2}{\theta^2+1}(\theta+x)e^{-\theta x}\quad ;x>0,\ \theta>0$	Shanker (2015 a)
	$F(x,\theta)=1-\left[1+\frac{\theta x}{\theta^2+1}\right]e^{-\theta x}\quad ;x>0,\theta>0$	
Lindley	$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}\quad ;x>0,\ \theta>0$	Lindley (1958)
	$F(x;\theta)=1-\left[1+\frac{\theta x}{\theta+1}\right]e^{-\theta x}\quad ;x>0,\theta>0$	
Exponentia	$f(x;\theta)=\theta e^{-\theta x}\quad ;x>0,\ \theta>0$	

1	$F_1(x;\theta)=1-e^{-\theta x}\ ;x>0,\theta>0$	

Ghitany *et al* (2008) have discussed various statistical properties, estimation of parameter and application of Lindley distribution for modeling waiting time data in a bank. It has been observed that these lifetime distributions are not always suitable for modeling lifetime data from biomedical sciences and engineering. In the present paper an attempt has been made to propose a one parameter lifetime distribution named 'Om distribution' which gives better fit than all one parameter lifetime distributions. Its various statistical properties, estimations of parameter and applications for modeling two real lifetime data from engineering have been discussed.

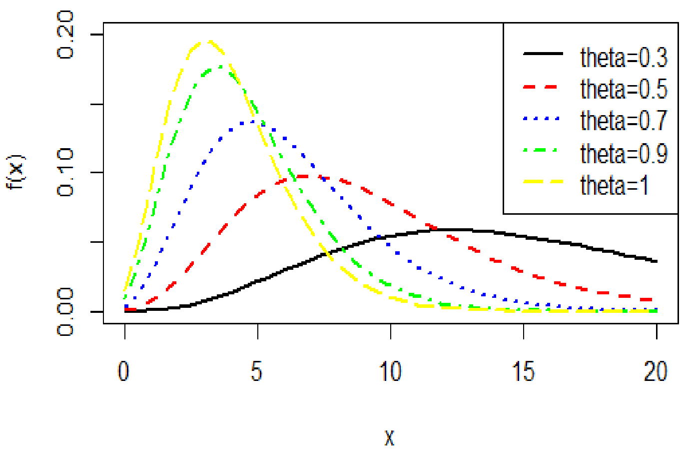
II. Om Distribution

A new one parameter lifetime distribution named Om distribution can be defined by its pdf and cdf

$$f(x;\theta)=\frac{\theta^5}{\theta^4+4\theta^3+12\theta^2+24\theta+24}(1+x)^4 e^{-\theta x}\quad ;x>0,\ \theta>0 \tag{2.1}$$

$$F(x;\theta)=1-\left[\frac{(1+x)^4\theta^4+4(1+x)^3\theta^3+12(1+x)^2\theta^2+24(1+x)\theta+24}{\theta^4+4\theta^3+12\theta^2+24\theta+24}\right]e^{-\theta x}\quad ;x>0,\theta>0 \tag{2.2}$$

The nature of the pdf and cdf of Om distribution for varying values of parameter θ have been shown graphically in figures 1 and 2 respectively.



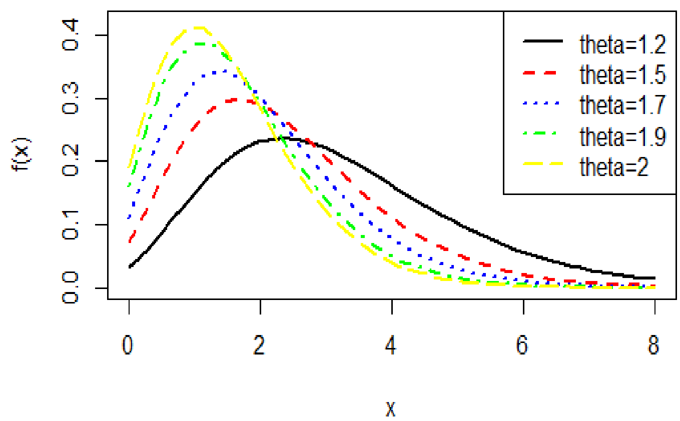
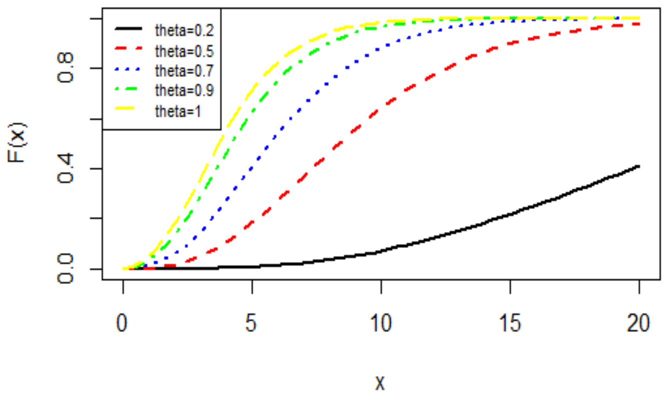


Fig. 1: Nature of the pdf of Om distribution for varying values of parameter θ



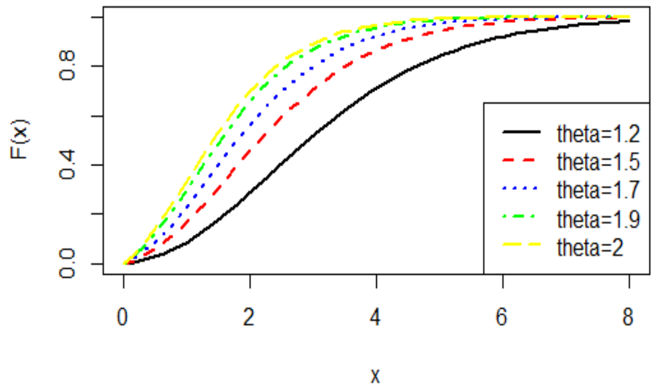


Fig. 2: Nature of the cdf of Om distribution for varying values of parameter θ

3. Moments and Associated measures

The r th moment about origin of Om distributon (2.1) can be obtained as

$$\mu_r' = \frac{r!\left\{\begin{matrix}\theta^4+4(r+1)\theta^3+6(r+1)(r+2)\theta^2+4(r+1)(r+2)(r+3)\theta \\ +(r+1)(r+2)(r+3)(r+4)\end{matrix}\right\}}{\theta^r\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)} \quad ;r=1,2,3,... \tag{3.1}$$

The first four moments about origin of Om distribution can be given as

$$\mu_1' = \frac{\theta^4+8\theta^3+36\theta^2+96\theta+120}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

$$\mu_2' = \frac{2\left(\theta^4+12\theta^3+72\theta^2+240\theta+360\right)}{\theta^2\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

$$\mu_3' = \frac{6\left(\theta^4+16\theta^3+120\theta^2+480\theta+840\right)}{\theta^3\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

$$\mu_4' = \frac{24\left(\theta^4+20\theta^3+180\theta^2+840\theta+1680\right)}{\theta^4\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

Thus the moments about mean of the Om distribution (2.1) are obtained as

$$\mu_2 = \frac{\theta^8+16\theta^7+128\theta^6+624\theta^5+1920\theta^4+3840\theta^3+5760\theta^2+5760\theta+2880}{\theta^2\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)^2}$$

$$\mu_3 = \frac{2\begin{pmatrix}\theta^{12}+24\theta^{11}+276\theta^{10}+1928\theta^9+8856\theta^8+28512\theta^7+70848\theta^6+141696\theta^5+233280\theta^4\\+311040\theta^3+311040\theta^2+207360\theta+69120\end{pmatrix}}{\theta^3\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)^3}$$

$$\mu_4 = \frac{3\begin{pmatrix}3\theta^{16}+96\theta^{15}+1472\theta^{14}+14048\theta^{13}+92672\theta^{12}+454656\theta^{11}+1767936\theta^{10}+5640960\theta^9\\+15034752\theta^8+33675264\theta^7+63148032\theta^6+98039808\theta^5+123863040\theta^4\\+123863040\theta^3+92897280\theta^2+46448640\theta+11612160\end{pmatrix}}{\theta^4\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)^4}$$

The coefficient of variation $\left(C.V\right)$, coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis $\left(\beta_2\right)$ and Index of dispersion $\left(\gamma\right)$ of Om distribution (2.1) are thus obtained as

$$C.V = \frac{\sigma}{\mu_1'} = \frac{\sqrt{\theta^8+16\theta^7+128\theta^6+624\theta^5+1920\theta^4+3840\theta^3+5760\theta^2+5760\theta+2880}}{\theta^4+8\theta^3+36\theta^2+96\theta+120}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\begin{pmatrix}\theta^{12}+24\theta^{11}+276\theta^{10}+1928\theta^9+8856\theta^8+28512\theta^7+70848\theta^6+141696\theta^5+233280\theta^4\\+311040\theta^3+311040\theta^2+207360\theta+69120\end{pmatrix}}{\left(\theta^8+16\theta^7+128\theta^6+624\theta^5+1920\theta^4+3840\theta^3+5760\theta^2+5760\theta+2880\right)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^{\,2}} = \frac{3\begin{pmatrix}3\theta^{16}+96\theta^{15}+1472\theta^{14}+14048\theta^{13}+92672\theta^{12}+454656\theta^{11}+1767936\theta^{10}+5640960\theta^9\\+15034752\theta^8+33675264\theta^7+63148032\theta^6+98039808\theta^5+123863040\theta^4\\+123863040\theta^3+92897280\theta^2+46448640\theta+11612160\end{pmatrix}}{\left(\theta^8+16\theta^7+128\theta^6+624\theta^5+1920\theta^4+3840\theta^3+5760\theta^2+5760\theta+2880\right)^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{\theta^8+16\theta^7+128\theta^6+624\theta^5+1920\theta^4+3840\theta^3+5760\theta^2+5760\theta+2880}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)\left(\theta^4+8\theta^3+36\theta^2+96\theta+120\right)}.$$

The behaviors of coefficient of variation, skewness, kurtosis and index of dispersion of Om distribution have been shown graphically for varying values of parameter θ in figure 3.

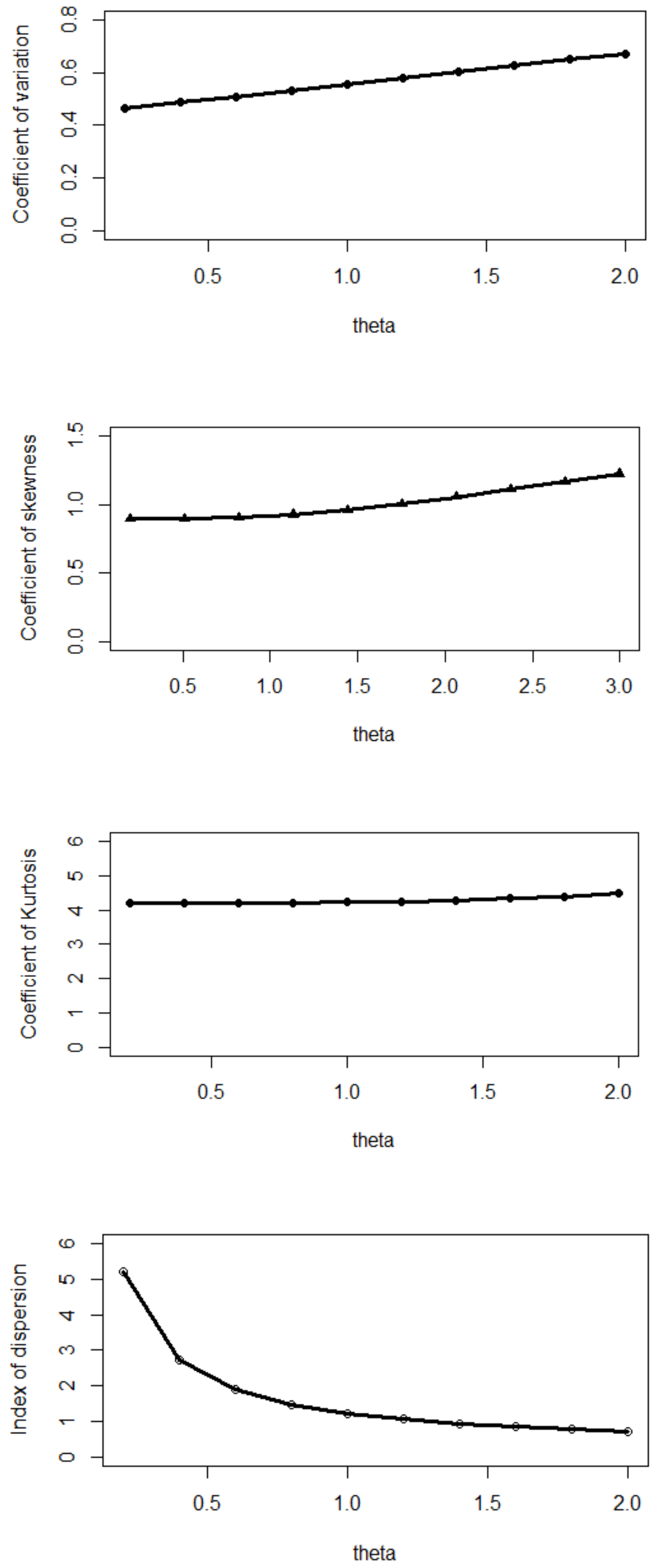


Fig. 3: Behavior of coefficient of variation, skewness, kurtosis and index of dispersion of Om distribution for varying values of parameter θ

The conditions of dispersion of Om distribution along with other one parameter lifetime distribution for values of the parameter θ have been presented in table 2.

Table 2. Over-dispersion, equi-dispersion and under-dispersion of Om distribution and other one parameter lifetime distributions for varying values of their parameter θ

Distributions	Over-dispersion $\left(\mu<\sigma^2\right)$	Equi-dispersion $\left(\mu=\sigma^2\right)$	Under-dispersion $\left(\mu>\sigma^2\right)$
Om	$\theta<1.306113562$	$\theta=1.306113562$	$\theta>1.306113562$
Devya	$\theta<1.451669994$	$\theta=1.451669994$	$\theta>1.451669994$
Amarendra	$\theta<1.525763580$	$\theta=1.525763580$	$\theta>1.525763580$
Akshaya	$\theta<1.327527885$	$\theta=1.327527885$	$\theta>1.327527885$
Rama	$\theta<1.950164618$	$\theta=1.950164618$	$\theta>1.950164618$
Aradhana	$\theta<1.283826505$	$\theta=1.283826505$	$\theta>1.283826505$
Sujatha	$\theta<1.364271174$	$\theta=1.364271174$	$\theta>1.364271174$
Akash	$\theta<1.515400063$	$\theta=1.515400063$	$\theta>1.515400063$
Shanker	$\theta<1.171535555$	$\theta=1.171535555$	$\theta>1.171535555$
Lindley	$\theta<1.170086487$	$\theta=1.170086487$	$\theta>1.170086487$
Exponential	$\theta<1$	$\theta=1$	$\theta>1$

IV. Statistical Properties

I. Survival function, Hazard rate function and Mean Residual life function

Suppose $f(x)$ and $F(x)$ be the pdf and cdf of a continuous random variable X. The survival function, $S(x)$, hazard rate function $h(x)$ (also known as the failure rate function) and the mean residual life function $m(x)$ of X are respectively defined as

$$S(x)=P(X>x)=1-F(x)$$

$$h(x)=\lim_{\Delta x\to 0}\frac{P\left(X<x+\Delta x \mid X>x\right)}{\Delta x}=\frac{f(x)}{1-F(x)}$$

and $m(x)=E\left[X-x \mid X>x\right]=\dfrac{1}{1-F(x)}\int_x^{\infty}\left[1-F(t)\right]\,dt$

The corresponding survival function $S(x)$, hazard rate function, $h(x)$ and the mean residual life function, $m(x)$ of Om distribution are thus obtained as

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

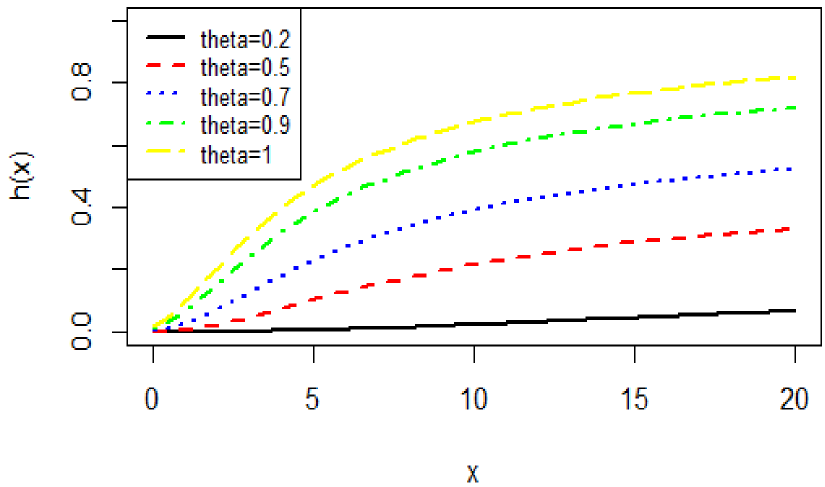
$$h(x)=\frac{\theta^5(1+x)^4}{(1+x)^4\theta^4+4(1+x)^3\theta^3+12(1+x)^2\theta^2+24(1+x)\theta+24}\quad ;x>0,\theta>0$$

and $$m(x)=\frac{\theta^4+4\theta^3+12\theta^2+24\theta+24}{\left[(1+x)^4\theta^4+4(1+x)^3\theta^3+12(1+x)^2\theta^2+24(1+x)\theta+24\right]e^{-\theta x}}\times$$

$$\int_x^{\infty}\left[(1+t)^4\theta^4+4(1+t)^3\theta^3+12(1+t)^2\theta^2+24(1+t)\theta+24\right]e^{-\theta t}dt$$

$$= \frac{(1+x)^4\theta^4+8(1+x)^3\theta^3+36(1+x)^2\theta^2+96(1+x)\theta+120}{\theta\left[(1+x)^4\theta^4+4(1+x)^3\theta^3+12(1+x)^2\theta^2+24(1+x)\theta+24\right]}$$

It can be easily verified that $h(0)=\frac{\theta^5}{\theta^4+4\theta^3+12\theta^2+24\theta+24}=f(0)$ and $m(0)=\frac{\theta^4+8\theta^3+36\theta^2+96\theta+120}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}=\mu_1'$. The behaviors of $h(x)$ and $m(x)$ of Om distribution have been shown in figures 4 and 5 respectively.



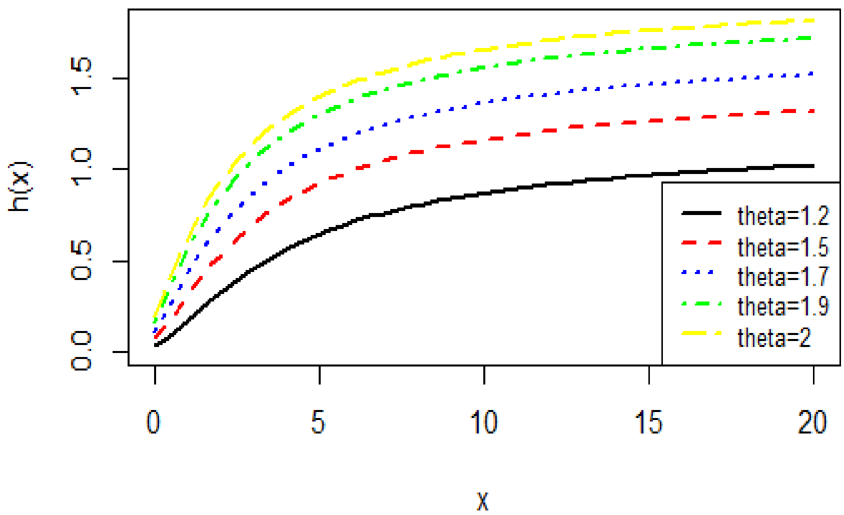
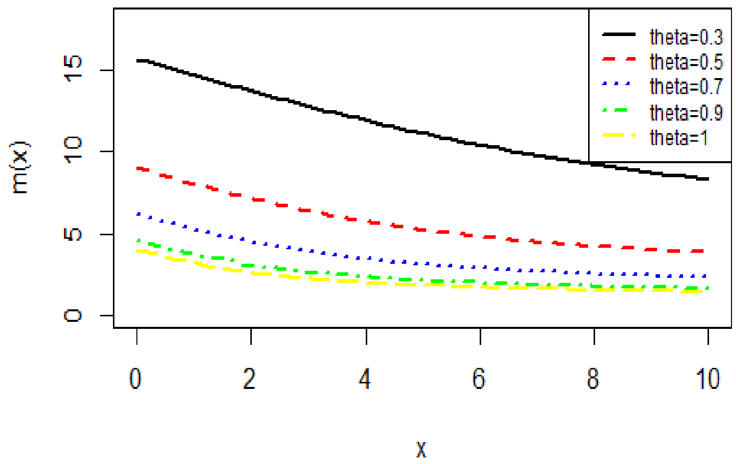


Fig. 4: Behavior of $h(x)$ of Om distribution for varying values of parameter θ



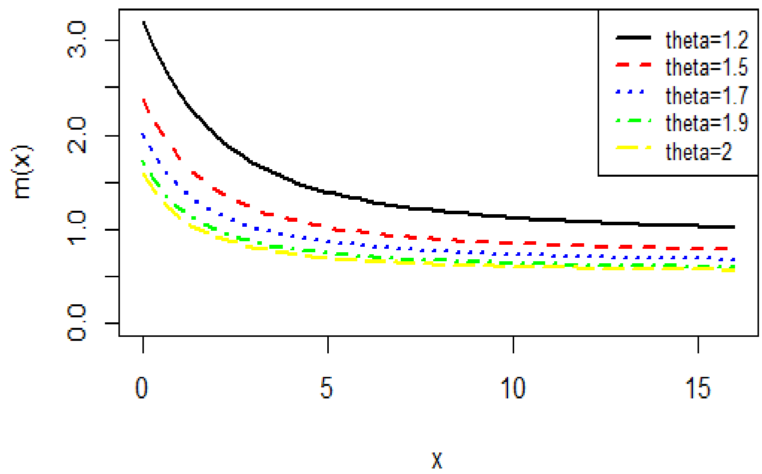


Fig. 5: Behavior of $m(x)$ of Om distribution for varying values of parameter θ

II. Mean deviations from the mean and the Median

The amount of scatter in a population is measured to some extent by the totality of deviations usually from their mean and median. These are known as the mean deviation about the mean and the mean deviation about the median and are defined as

$\delta_1(X)=\int_0^{\infty}|x-\mu|\, f(x)\,dx$ and $\delta_2(X)=\int_0^{\infty}|x-M|\, f(x)\,dx$, respectively, where $\mu=E(X)$ and $M=\text{Median}\,(X)$. The measures $\delta_1(X)$ and $\delta_2(X)$ can be computed using the following simplified relationships

$$\delta_1(X)=2\mu F(\mu)-2\int_0^{\mu} x\, f(x)\,dx \qquad (4.2.1)$$

and

$$\delta_2(X)=\mu-2\int_0^{M} x\, f(x)\,dx \qquad (4.2.2)$$

Using pdf (2.1) and expression for the mean of Om distribution, we get

$$\int_0^{\mu} x\, f(x;\theta)\,dx = \mu - \frac{\left\{\begin{array}{l}\left(\mu^5+4\mu^4+6\mu^3+4\mu^2+\mu\right)\theta^5+\left(5\mu^4+16\mu^3+18\mu^2+8\mu+1\right)\theta^4 \\ +\left(20\mu^3+48\mu^2+36\mu+8\right)\theta^3+\left(60\mu^2+96\mu+36\right)\theta^2+\left(120\mu+96\right)\theta+120\end{array}\right\}e^{-\theta\mu}}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

(4.2.3)

$$\int_0^{M} x\, f(x;\theta)\,dx = \mu - \frac{\left\{\begin{array}{l}\left(M^5+4M^4+6M^3+4M^2+M\right)\theta^5+\left(5M^4+16M^3+18M^2+8M+1\right)\theta^4 \\ +\left(20M^3+48M^2+36M+8\right)\theta^3+\left(60M^2+96M+36\right)\theta^2+\left(120M+96\right)\theta+120\end{array}\right\}e^{-\theta M}}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

(4.2.4)

Using expressions (4.2.1), (4.2.2), (4.2.3) and (4.2.4), the mean deviation about mean, $\delta_1(X)$ and the mean deviation about median, $\delta_2(X)$ of Om distribution (2.1),, after tedious algebraic simplification are obtained as

$$\delta_1(X) = 2\left[\frac{(1+\mu)^4\theta^4+8(1+\mu)^3\theta^3+36(1+\mu)^2\theta^2+96(1+\mu)\theta+120}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}\right]e^{-\theta\mu}$$

(4.2.5)

$$\partial_2(X) = \frac{\left\{\begin{array}{l}\left(M^5+4M^4+6M^3+4M^2+M\right)\theta^5+\left(5M^4+16M^3+18M^2+8M+1\right)\theta^4 \\ +\left(20M^3+48M^2+36M+8\right)\theta^3+\left(60M^2+96M+36\right)\theta^2+\left(120M+96\right)\theta+120\end{array}\right\}e^{-\theta M}}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)} - \mu$$

III. Bonferroni and Lorenz Curves

The Bonferroni and Lorenz curves (Bonferroni, 1930) and Bonferroni and Gini indices have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medicine. The Bonferroni and Lorenz curves are defined as

$$B(p) = \frac{1}{p\mu}\int_0^q x\, f(x)\,dx = \frac{1}{p\mu}\left[\int_0^\infty x\, f(x)\,dx - \int_q^\infty x\, f(x)\,dx\right] = \frac{1}{p\mu}\left[\mu - \int_q^\infty x\, f(x)\,dx\right]$$

(4.3.1)

and

$$L(p) = \frac{1}{\mu}\int_0^q x\, f(x)\,dx = \frac{1}{\mu}\left[\int_0^\infty x\, f(x)\,dx - \int_q^\infty x\, f(x)\,dx\right] = \frac{1}{\mu}\left[\mu - \int_q^\infty x\, f(x)\,dx\right] \quad (4.3.2)$$

respectively or equivalently

$$B(p) = \frac{1}{p\mu}\int_0^p F^{-1}(x)\,dx \quad (4.3.3)$$

and

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(x)\,dx \quad (4.3.4)$$

respectively, where $\mu = E(X)$ and $q = F^{-1}(p)$.

The Bonferroni and Gini indices are thus define

$$B = 1 - \int_0^1 B(p)\,dp \tag{4.3.5}$$

$$G = 1 - 2\int_0^1 L(p)\,dp \tag{4.3.6}$$

respectively

Using pdf (2.1), we get

$$\int_q^\infty x\, f(x;\theta)\,dx = \frac{\left\{\begin{array}{l}\left(q^5+4q^4+6q^3+4q^2+q\right)\theta^5+\left(5q^4+16q^3+18q^2+8q+1\right)\theta^4\\ +\left(20q^3+48q^2+36q+8\right)\theta^3+\left(60q^2+96q+36\right)\theta^2+\left(120q+96\right)\theta+120\end{array}\right\}e^{-\theta q}}{\theta\left(\theta^4+4\theta^3+12\theta^2+24\theta+24\right)}$$

(4.3.7)

Now using equation (4.3.7) in (4.3.1) and (4.3.2), we get

$$B(p) = \frac{1}{p}\left[1-\frac{\left\{\begin{array}{l}\left(q^5+4q^4+6q^3+4q^2+q\right)\theta^5+\left(5q^4+16q^3+18q^2+8q+1\right)\theta^4\\ +\left(20q^3+48q^2+36q+8\right)\theta^3+\left(60q^2+96q+36\right)\theta^2+\left(120q+96\right)\theta+120\end{array}\right\}e^{-\theta q}}{\theta^4+8\theta^3+36\theta^2+96\theta+120}\right]$$

(4.3.8)

$$L(p) = 1-\frac{\left\{\begin{array}{l}\left(q^5+4q^4+6q^3+4q^2+q\right)\theta^5+\left(5q^4+16q^3+18q^2+8q+1\right)\theta^4\\ +\left(20q^3+48q^2+36q+8\right)\theta^3+\left(60q^2+96q+36\right)\theta^2+\left(120q+96\right)\theta+120\end{array}\right\}e^{-\theta q}}{\theta^4+8\theta^3+36\theta^2+96\theta+120}$$

(4.3.9)

Now using equations (4.3.8) and (4.3.9) in (4.3.5) and (4.3.6), the Bonferroni and Gini indices of Om distribution (2.1) are obtained as

$$B = 1-\frac{\left\{\begin{array}{l}\left(q^5+4q^4+6q^3+4q^2+q\right)\theta^5+\left(5q^4+16q^3+18q^2+8q+1\right)\theta^4\\ +\left(20q^3+48q^2+36q+8\right)\theta^3+\left(60q^2+96q+36\right)\theta^2+\left(120q+96\right)\theta+120\end{array}\right\}e^{-\theta q}}{\theta^4+8\theta^3+36\theta^2+96\theta+120}$$

(4.3.9)

$$G = \frac{\left\{\begin{array}{l}\left(q^5+4q^4+6q^3+4q^2+q\right)\theta^5+\left(5q^4+16q^3+18q^2+8q+1\right)\theta^4\\ +\left(20q^3+48q^2+36q+8\right)\theta^3+\left(60q^2+96q+36\right)\theta^2+\left(120q+96\right)\theta+120\end{array}\right\}e^{-\theta q}}{\theta^4+8\theta^3+36\theta^2+96\theta+120}-1$$

(4.3.10)

IV. 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

(i) stochastic order $\left(X \leq_{st} Y\right)$ if $F_X\left(x\right) \geq F_Y\left(x\right)$ for all x

(ii) hazard rate order $\left(X \leq_{hr} Y\right)$ if $h_X\left(x\right) \geq h_Y\left(x\right)$ for all x

(iii) mean residual life order $\left(X \leq_{mrl} Y\right)$ if $m_X\left(x\right) \leq m_Y\left(x\right)$ for all x

(iv) likelihood ratio order $\left(X \leq_{lr} Y\right)$ if $\dfrac{f_X\left(x\right)}{f_Y\left(x\right)}$ decreases in x.

The following results due to Shaked and Shanthikumar (1994) are well known for establishing stochastic ordering of distributions

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y$$
$$\Downarrow$$
$$X \leq_{st} Y$$

The Om distribution is ordered with respect to the strongest 'likelihood ratio' ordering as shown in the following theorem:

Theorem: Let X ~ Om distribution$\left(\theta_1\right)$ and Y ~ Om distribution$\left(\theta_2\right)$. If $\theta_1 \geq \theta_2$, then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Proof: We have

$$\frac{f_X\left(x;\theta_1\right)}{f_Y\left(x;\theta_2\right)} = \frac{\theta_1^5\left(\theta_2^4+4\theta_2^3+12\theta_2^2+24\theta_2+24\right)}{\theta_2^5\left(\theta_1^4+4\theta_1^3+12\theta_1^2+24\theta_1+24\right)} e^{-\left(\theta_1-\theta_2\right)x}; x>0$$

Now

$$\ln \frac{f_X\left(x;\theta_1\right)}{f_Y\left(x;\theta_2\right)} = \ln\left[\frac{\theta_1^5\left(\theta_2^4+4\theta_2^3+12\theta_2^2+24\theta_2+24\right)}{\theta_2^5\left(\theta_1^4+4\theta_1^3+12\theta_1^2+24\theta_1+24\right)}\right] - \left(\theta_1-\theta_2\right)x$$

This gives $\dfrac{d}{dx}\ln \dfrac{f_X\left(x;\theta_1\right)}{f_Y\left(x;\theta_2\right)} = -\left(\theta_1-\theta_2\right)$. Thus for $\theta_1 \geq \theta_2$, $\dfrac{d}{dx}\ln \dfrac{f_X\left(x;\theta_1\right)}{f_Y\left(x;\theta_2\right)} \leq 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

V.Distribution of Order Statistics

Let $X_1, X_2, ..., X_n$ be a random sample of size n from Om distribution (2.1). Let $X_{(1)} < X_{(2)} < ... < X_{(n)}$ denote the corresponding order statistics. The pdf and the cdf of the k th order statistic, say $Y = X_{(k)}$ are given by

$$f_Y\left(y\right) = \frac{n!}{\left(k-1\right)!\left(n-k\right)!} F^{k-1}\left(y\right)\left\{1-F\left(y\right)\right\}^{n-k} f\left(y\right)$$

$$= \frac{n!}{\left(k-1\right)!\left(n-k\right)!}\sum_{l=0}^{n-k}\binom{n-k}{l}\left(-1\right)^l F^{k+l-1}\left(y\right) f\left(y\right)$$

and

$$F_Y\left(y\right) = \sum_{j=k}^{n}\binom{n}{j} F^j\left(y\right)\left\{1-F\left(y\right)\right\}^{n-j}$$

$$=\sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^{l}F^{j+l}(y),$$

respectively, for $k=1,2,3,...,n$.

Thus, the pdf and the cdf of k th order statistics of Om distribution (2.1) are obtained as

$$f_Y(y)=\frac{n!\theta^5(1+x)^4e^{-\theta x}}{(\theta^4+4\theta^3+12\theta^2+24\theta+24)(k-1)!(n-k)!}$$

$$\times\sum_{l=0}^{n-k}\binom{n-k}{l}(-1)^{l}\left[1-\left\{\frac{(1+x)^4\theta^4+4(1+x)^3\theta^3+12(1+x)^2\theta^2+24(1+x)\theta+24}{\theta^4+4\theta^3+12\theta^2+24\theta+24}\right\}e^{-\theta x}\right]^{k+l-1}$$

and

$$F_Y(y)=\sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^{l}\left[1-\left\{\frac{(1+x)^4\theta^4+4(1+x)^3\theta^3+12(1+x)^2\theta^2+24(1+x)\theta+24}{\theta^4+4\theta^3+12\theta^2+24\theta+24}\right\}e^{-\theta x}\right]^{j+l}$$

VI. Stress-Strength Reliability

The stress- strength reliability describes the life of a component which has random strength X that is subjected to a random stress Y. When the stress Y applied to it exceeds the strength X, the component fails instantly and the component will function satisfactorily till $X>Y$. Therefore, $R=P(Y<X)$ is a measure of component reliability and is known as stress-strength reliability in statistical literature. It has wide applications in almost all areas of knowledge especially in engineering such as structures, deterioration of rocket motors, static fatigue of ceramic components, aging of concrete pressure vessels etc.

Let X and Y be independent strength and stress random variables having Om distribution (2.1) with parameter θ_1 and θ_2 respectively. Then the stress-strength reliability R of Om distribution can be obtained as

$$R=P(Y<X)=\int_0^\infty P(Y<X\mid X=x)f_X(x)\,dx$$

$$=\int_0^\infty f(x;\theta_1)\,F(x;\theta_2)\,dx$$

$$=1-\frac{\theta_1^5\left[\begin{array}{l}(\theta_2^4+4\theta_2^3+12\theta_2^2+24\theta_2+24)(\theta_1+\theta_2)^8+(8\theta_2^4+28\theta_2^3+72\theta_2^2+120\theta_2+96)(\theta_1+\theta_2)^7\\+(56\theta_2^4+168\theta_2^3+360\theta_2^2+480\theta_2+288)(\theta_1+\theta_2)^6\\+(336\theta_2^4+840\theta_2^3+1440\theta_2^2+1440\theta_2+576)(\theta_1+\theta_2)^5\\+(1680\theta_2^4+3360\theta_2^3+4320\theta_2^2+2880\theta_2+576)(\theta_1+\theta_2)^4\\+(6720\theta_2^4+10080\theta_2^3+8640\theta_2^2+2880\theta_2)(\theta_1+\theta_2)^3\\+(20160\theta_2^4+20160\theta_2^3+8640\theta_2^2)(\theta_1+\theta_2)^2+(40320\theta_2^4+20160\theta_2^3)(\theta_1+\theta_2)+40320\theta_2^4\end{array}\right]}{(\theta_1^4+4\theta_1^3+12\theta_1^2+24\theta_1+24)(\theta_2^4+4\theta_2^3+12\theta_2^2+24\theta_2+24)(\theta_1+\theta_2)^9}$$

V. Maximum likelihood estimation

Let $(x_1, x_2, ..., x_n)$ be a random sample of size n from Om distribution (2.1). The likelihood function L of Om distribution can be expressed as

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

$$= \left(\frac{\theta^5}{\theta^4 + 4\theta^3 + 12\theta^2 + 24\theta + 24} \right)^n \prod_{i=1}^{n} \left(1 + 4x_i + 6x_i^{\,2} + 4x_i^{\,3} + x_i^{\,4}\right) e^{-n\theta \bar{x}}$$

The log likelihood function is thus given by

$$\log L = n \log \left(\frac{\theta^5}{\theta^4 + 4\theta^3 + 12\theta^2 + 24\theta + 24} \right) \sum_{i=1}^{n} \log\left(1 + 4x_i + 6x_i^{\,2} + 4x_i^{\,3} + x_i^{\,4}\right) - n\theta \bar{x} \, .$$

The maximum likelihood estimates (MLE) $\hat{\theta}$ of parameter θ is the solution of the log-likelihood equation $\frac{d \log L}{d\theta} = 0$ and is given by

$$\frac{d \log L}{d\theta} = \frac{5n}{\theta} - \frac{n\left(4\theta^3 + 12\theta^2 + 24\theta + 24\right)}{\theta^4 + 4\theta^3 + 12\theta^2 + 24\theta + 24} - n\,\bar{x} = 0 \, .$$

This gives a fifth degree polynomial equation in θ as

$$\bar{x}\,\theta^5 + \left(4\bar{x} - 1\right)\theta^4 + 4\left(3\bar{x} - 2\right)\theta^3 + 12\left(2\,\bar{x} - 3\right)\theta^2 + 24\left(\bar{x} - 4\right)\theta - 120 = 0 \, .$$

This equation can be easily solved using any numerical iterationmethod namely, Newton-Raphson method, Regula Falsi method or Bisection method. In this paper Newton-Raphson method has been used to estimate the parameter θ from above equation. It should be noted that equating the population mean to the corresponding sample mean, the method of moment estimate is the same as method of maximum likelihood.

VI. Data analysis

In this section the goodness of fit of Om distribution has been discussed with following two real lifetime datasets from engineering.

Data Set 1: The data is given by Birnbaum and Saunders (1969) on the fatigue life of 6061 – T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second. The data set consists of 101 observations with maximum stress per cycle 31,000 psi. The data ($\times 10^{-3}$) are presented below (after subtracting 65**).**

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	55	56	56	56	58	59	59	59	59	59
63	63	64	64	65	65	65	66	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	94	97	98	98	99	101	103
105	109	136	147								

Data Set 2: This data set is the strength data of glass of the aircraft window reported by Fuller *et al* (1994)

18.83	20.8	21.657	23.03	23.23	24.05	24.321	25.5	25.52	25.8	26.69	26.77
26.78	27.05	27.67	29.9	31.11	33.2	33.73	33.76	33.89	34.76	35.75	35.91
36.98	37.08	37.09	39.58	44.045	45.29	45.381					

For these two datasets, Om distribution has been fitted along with other one parameter lifetime distributions. The ML estimate, value of $-2\ln L$, Akaike Information criteria (AIC), K-S statistics and p-value of the fitted distributions are presented in tables 3 and 4.. The AIC and K-S Statistics are computed using the following formulae: $AIC = -2\ln L + 2k$ and $\text{K-S} = \sup_x \left| F_n(x) - F_0(x) \right|$, where k = the number of parameters, n = the sample size , $F_n(x)$ is the empirical (sample) cumulative distribution function, and $F_0(x)$ is the theoretical cumulative distribution function. The best distribution is the distribution corresponding to lower values of $-2\ln L$, AIC, and K-S statistics and higher p-value

Table 3: MLE's, - 2ln L, AIC, K-S and p-values of the fitted distributions for dataset 1

Distributions	$\text{MLE}(\hat{\theta})$	$\text{S.E}(\hat{\theta})$	$-2\log L$	AIC	K-S	P-Value
Om	0.07211	0.00322	924.64	926.64	0.138	0.043
Shambhu	0.08755	0.00357	918.61	920.61	0.117	0.131
Devya	0.07289	0.00326	924.26	926.26	0.333	0.000
Amarendra	0.05824	0.00213	934.38	936.38	0.163	0.010
Suja	0.07317	0.00327	924.21	926.21	0.136	0.049
Akshaya	0.05769	0.00288	935.11	937.11	0.164	0.008
Rama	0.05854	0.00293	934.05	934.05	0.162	0.012
Aradhana	0.04327	0.00249	952.58	954.58	0.196	0.001
Sujatha	0.04356	0.00251	951.78	953.78	0.195	0.001
Akash	0.04387	0.00253	950.97	952.97	0.194	0.001
Shanker	0.02925	0.00206	980.97	982.97	0.248	0.000
Lindley	0.02887	0.00204	983.11	985.11	0.252	0.000
Exponential	0.01463	0.00145	1044.87	1046.87	0.366	0.000

Table 4: MLE's, - 2ln L, AIC, K-S and p-values of the fitted distributions for dataset 2

Distributions	$\text{MLE}(\hat{\theta})$	$\text{S.E}(\hat{\theta})$	$-2\log L$	AIC	K-S	P-Value
Om	0.15718	0.01262	228.81	230.81	0.230	0.061
Shambhu	0.19339	0.01417	223.40	225.40	0.199	0.148
Devya	0.16087	0.01292	227.68	229.68	0.422	0.000
Amarendra	0.12829	0.01210	233.41	235.41	0.257	0.027
Suja	0.16227	0.01303	227.25	229.25	0.223	0.077
Akshaya	0.12574	0.01129	234.44	236.44	0.263	0.022
Rama	0.12978	0.01165	232.79	234.79	0.253	0.030
Aradhana	0.09432	0.00978	242.22	244.22	0.306	0.004
Sujatha	0.09561	0.00990	241.50	243.50	0.303	0.005
Akash	0.09706	0.01005	240.68	242.68	0.298	0.006
Shanker	0.64716	0.00820	252.35	254.35	0.358	0.000
Lindley	0.06299	0.00800	253.98	255.98	0.365	0.000
Exponential	0.03245	0.00582	274.53	276.53	0.458	0.000

VII. Concluding remarks and future works

This paper proposes a new one parameter lifetime distribution named, 'Om distribution'. Statistical properties including shapes for probability density, moments based measures, hazard rate function, mean residual life function, stochastic ordering, mean deviations, Bonferroni and Lorenz curves, distribution of order statistics, and stress-strength reliability have been discussed. Method of maximum likelihood has been discussed for estimating the parameter of the distribution. Applications of the distribution have been explained through two examples of real lifetime data from engineering and the goodness of fit has been found to be quite satisfactory over several one parameter lifetime distributions.

Since the present distribution is a new distribution in statistics literature, a lot of works can be done on the distribution. The future works to be done on the distribution includes Poisson mixture of the distribution, weighted version of the distribution, Power version of the distribution, discretization of the distribution using infinite series and survival function approach, exponentiation of the distribution, Bayesian method of estimation, some among others. All these works will appear in statistics literature with passage of time.

References

[1] Birnbaum, Z.W. and Saunders, S.C. (1969): Estimation for a family of life distributions with applications to fatigue, *Journal of Applied Probability*, 6, 328-347.

[2] Bonferroni, C.E. (1930): Elementi di Statistca generale, Seeber, Firenze

[3] Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008): Lindley distribution and its Application, *Mathematics Computing and Simulation*, 78, 493 – 506

[4] Fuller, E.J., Frieman, S., Quinn, J., Quinn, G., and Carter, W.(1994): Fracture mechanics approach to the design of glass aircraft windows: A case study, SPIE Proc 2286, 419-430.

[5] Lindley, D.V. (1958): Fiducial distributions and Bayes' theorem, *Journal of the Royal Statistical Society, Series B*, 20, 102- 107.

[6] Shaked, M. and Shanthikumar, J.G. (1994): Stochastic Orders and Their Applications, Academic Press, New York.

[7] Shanker, R. (2015 a): Shanker Distribution and Its Applications, *International Journal of Statistics and Applications,* 5 (6), 338 – 348

[8] Shanker, R. (2015 b): Akash Distribution and Its Applications, *International Journal of Probability and Statistics,* 4 (3), 65 – 75

[9] Shanker, R. (2016 a): Sujatha Distribution and Its Applications, *Statistics in Transition-New series*, 17(3), 1- 20.

[10] Shanker, R. (2016 b): Aradhana Distribution and Its Applications, *International Journal of Statistics and Application*s, 6(1), 23 – 34.

[11] Shanker, R. (2017 c): Amarendra Distribution and Its Applications, *American Journal of Mathematics and Statistics*, 6(1), 44 – 56

[12] Shanker, R. (2016 d): Devya Distribution and Its Applications, *International Journal of Statistics and Applications*, 6(4), 189 – 202

[13] Shanker, R. (2017 a): Rama Distribution and its Application, *International Journal of Statistics and Applications*, 7(1), 26 – 35

[14] Shanker, R. (2017 b): Akshaya Distribution and its Application, *American Journal of Mathematics and Statistics*, 7(2), 51 – 59

Imperfect Production Model for Price Sensitive Demand with Shortage

Uttam Kumar Khedlekar and Ram Kumar Tiwari

•

Department of Mathematics and Statistics
Dr. Harisingh Gour Vishwavidyalaya, Sagar M.P. India
(A Central University)
mailto: uvkkcm@yahoo.co.in

Abstract

In this paper, we have presented an economic production inventory model considering non-linear demand depanding on selling price. Here, all imperfect quality items are reworked after the regular production process and the reworked items are considered as similar as good quality items. Rework is important in those businesses where last product is expensive and raw materials are insuficient. Now, our objective is to find out the optimal ordering lot size, optimal selling price and shortage, for which the profit of the system is maximum. A numerical example is presented to illustrate the validity of the model. Manageral implications has been presented based on sensitivity analysis.

Keywords: Dynamic pricing, Non linear demand, Optimal price settings, Imperfect item, Rework, Partial backlogging

AMS Subject Classification: 90B05, 90B30, 90B50

1 Introduction

Inventory control is an important part of business because it ensure quality control in business. Inventory management secure the business and help to smooth runing of business affair. Today, pricing and production strategies are two fundamental components of the daily operations for manufacturers, particularly in the presence of imperfect production system. Production system is one of the most important ascepts of company's business strategy. To avoid the off overeges and shortages of products, firm should carefully design the production process to enrich the business.

As a consequence of this paper, the topic of pricing with production system has recently been the focus of acadmic research increase diverse as economics, marketing and operation managemant. There has been several studies analysing condition under which different pricing strategys optimize the compnies profitability Bose *et al.* (1995) desinged an economoc order quantity inventory model for deteriorating products with linear demand and positive trend under allowable shortage and backlogging. Chakrabarti and chauduri (1997) presented an inventory model for perishable items. In this model the demend was taken as linear function and shortage in each cycle. Wee (1999) developed an inventory model for deteriorating items. In which, shortage was partially backlogged at constant rate and demand was taken to be linear function of selling price.

The production is an essential part of inventory system and not produced hundred persent perfect items. Many researchers designed a production inventory model under backlogging situation such as Chern *et al.* (2008); Dye *et al.* (2007); Lodree (2007); Leung (2008) Goyal and Imran (2008); Thannyam and Uthayakumar (2008); Cardenas and Berson (2009); Taleizadeh (2011) Roy *et al.* (2011). Das *et al.* (2011) presented an economic order quantity model for imperfect quality items with partialy backlogging. In this model, they also considered the cost of lost sale. Taleizadeh *et al.* (2012) proposed an EOQ model in which they considered a special sale price along with partial backlogging, and customer may take the advantage of discount in price. Lee and Dye (2012) formulated an economic order quantity model with shortage, in that model demand was taken stock dependent. They also considered the optimal ordering and preservation policies to maximize the total profit.

Several inventory model considered demand dependence on other factors such as product selling price and quality. Datta (2013) investigated an inventory model assuming that the demand depends on both the selling price and quality. Kumar *et al.* (2013) proposed EOQ model under the consideration of price-dependent demand, where the carrying cost is a function of the trade credit for deteriorating products. Sana (2010) designed an economic order quantity (EOQ) model in that model, the demand was considered as function of selling price and they also assumed the deterioration rate of defective item is time proportional. Sana (2011) suggested an inventory model in which, they taken the demand function as quadratic function and the selling price increases in each cycle, but demand decreases quadraticly with selling price.

By use of item preservation concept for deteriorating items, Khedlekar *et al.* (2016) conceptualized an EOQ in that model the demand was considered as function of selling price and linearly decreases. They considered as the profit is the concave function of the optimal selling price, also calculated the optimal selling price, the length of the replenishment cycle and the optimal preservation concept investment simultaneously. Mishra (2016) proposed a single-manufacturer single-retailer inventory model by incorporating preservation technology cost for defective items and determined optimal retail price, replenishment cycle and the cost of preservation lechnology.

Taleizadeh and Noori-daryan (2016) studid a production inventory model with a three-level decentralised supply chain with price sensitive demand. Haider *et al.* (2016) proposed an economic production quantity (EPQ) model from this they reveal if we make the discount in defective item and apply rework process then we get maximum profit. Teksan and Geunes (2016) reported an economic order quantity model for finished goods. In this model they they assumed that the demand rate was more price sensitive for supplier and customer both. Taleizadeh *et al.* (2017) outlined an imperfect production inventory model without shortages. Pal and Adhikari (2017) conceptualized an imperfect production inventory model with exponential partially backlogging with rework, in that model they assumed that all imperfect quality products are reworked after the regular production process and demand rate was price sensitive and it was monotonic decreasing function selling price. Among other researcher in the exposure, the notworthy contribution of Sarkar, Sana and Chaudhuri (2011); Yu, and Chen (2007); Wee, and Kuo (2013); Pal, Sana, and Chaudhuri (2014); Sarkar (2012, 2013); Haider, Salameh, and Nasr (2016); Tyyab and Sarkar (2016) should be mentioned.

We have considered an imperfect production model which depend on the selling price. We assumed, all the defective products are reworked just after the regular production process and no any scarp product is produced during production as well as reworking run time. Shortage occurs at the beginning of the cycle and production starts after backorder time and backlogging rate is variable. The price of goods is definitely shown to the customer at the beginning of time cycle in many situations. So it is very difficult to take the different price within same inventory cycle. In this paper we deal with the three issues: first, what will be selling price for the items, second one how much inventory should be produced and third one what time period shortage would be allowed in order to optimum profit.

2 Assumptions & Notations

2.1 Assumptions

The model is designed for infinite time horizon, This model is developed for single item, Production rate if perfect item p is constant and production rate of defective items is $p_d = xp$, where x is continuous randam variable, In this model the shortages occur at the beginning of the cycle and during the shortage time interval a fraction of the demand varying with waiting time is backlogged for the clients, who have patience to wait, assume that customers impatient function by

$B(\tau) = e^{-\alpha\tau}$, $\alpha > 0$, After the continue production process all imperfect items are reworked, The holding cost for both type (perfect and imperfect) items is the same, Every constant costs as inspection cost and purchasing cost are included within the production cost of the items, The demand function of the product is $D(s) = \varphi s^{-\eta}$; $\eta > 0$.

2.2 Notations

$[D(s)]$ – Demand function for good products,

$[I(t)]$ – On-hand inventory of product at time t in j^{th} cycle,

$[p]$ – Production rate for perfect item per units per unit time, [p_d] – Production rate for imperfect quantity items unit per unit time, $[x]$ – Percentage of produced imperfect quality items which is randam variable, $[f(x)]$ – Probability density function of x, $[r]$ – Rework rate of imperfect quality item per unit per unit time, $[\omega]$ – Backorder level, $[B(\tau)]$ – Customers impatient function, where τ is the waiting time for customer, $[c_h]$ – Holdig cost per item per unit time, $[c_{h_1}]$ – Holdig cost of reworked item per item per unit time, $[c_p]$– Production cost per unit of item, $[c_b]$– Backorder cost per item, $[c_k]$– Per production set-up cost, $[c_l]$– Lost sale cost per item, $[s]$– Selling price per item, $[\varphi]$– Stock dependent parameter, [Π]– The total profit,

1. – Average total profit,
2. – Excepted average total profit.

3 The Mathematical Model

Suppose a business start with shortage of products which are partially backlogged. The backlogging rate is a function of customer waiting time as $B(\tau) = e^{-a\tau}$, $a > 0$, where τ is waiting time $\tau = t_1 - t$. Suppose the production start at time t_1 and it continue up to time t_3. Due to production run, all the products which are backlogged, during time period $[0, t_2]$ are provide at the time t_2. The production rate is considered constant. The qx amount of defective item is produced by the total production. The rework rate of defective products is r, and these are reworked after the regular production process. $\frac{qx}{r}$ is the amount of time required for reworking of defective items, where qx is total items produced and r is rework rate. There is the same price of good products and reworked product and demand rate is depend on selling price and defined as,

$$D(s) = \varphi s^{-\eta} \tag{1}$$

We take $T_i = t_i - t_{i-1}$.

For the Time period $0 \le t \le t_1$, the differential equation governing the inventory level is

$$\frac{dI}{dt} = -D(s)B(\tau) \tag{2}$$

with the boundary condition$I(0) = 0$ and $I(t_1) = -\omega$ where $\tau = t_1 - t$.

The solution of above differential equation by using the boundary condition is

$$I(t) = \frac{D(s)e^{-}at - e^{a(t_1 - t)}}{a} \tag{3}$$

and using the boundary condition $I(t_1) = -\omega$, we get

$$\omega = \frac{D(s)(1-e^{-at_1})}{a} \tag{4}$$

The backorder cost during $0 \leq t \leq t_1$ is

$$c_b \int_0^{t_1} (I(t))dt = \frac{c_b D(s)\{1-at_1e^{-at_1}-e^{-at_1}\}}{a^2} \tag{5}$$

The demand rate is D(s), out of this only $D(s)e^{-a(t_1-t)}$ is fullfilld during $[0, t_1]$ and $D(s) - D(s)e^{-a(t_1-t)}$ wich not fullfilld. Then the cost of lost sale is given by

$$c_l \int_0^{t_1} D(s)\{1 - e^{-a(t_1-t)}\}dt = \frac{c_l D(s)(at_1-1+e^{-at_1})}{a} \tag{6}$$

For the time interval $t_1 \leq t \leq t_2$, the governing differential equation of inventory level is

$$\frac{dI}{dt} = p - p_d - D(s) \tag{7}$$

with boundary condition $I(t_1) = -\omega$, $I(t_2) = 0$

Then the solution of above differential equation is

$$I(t) = \{(1-x)p - D(s)\}(t - t_2) \tag{8}$$

using the condition $I(t) = -\omega$, we have

$$\omega = \{(1-x)p - D(s)\}T_2, \tag{9}$$

where $T_2 = t_2 - t_1$

The cost of backorder in time interval $t_1 \leq t \leq t_2$ is

$$c_b \int_0^{t_1} (I(t))dt = \frac{c_b \omega T_2}{2} \tag{10}$$

Eq. (9) & Eq. (10) leads the back order cost during $t_1 \leq t \leq t_2$

$$= \frac{c_b \omega^2}{2\{(1-x)p-D(s)\}} \tag{11}$$

For the time interval $t_2 \leq t \leq t_3$, the governing differential equation of inventory level is

$$\frac{dI}{dt} = p - p_d - D(s) \tag{12}$$

with boundary condition $I(t_2) = 0$, $I(t_3) = z_3$ where z_3, is inventory level of good product.

Then the solution of above differential equation is

$$I(t) = \{(1-x)p - D(s)\}(t - t_2) \tag{13}$$

using $I(t_3) = z_3$, we get

$$z_3 = \{(1-x)p - D(s)\}T_3 \tag{14}$$

The holding cost for good items in time period $t_2 \leq t \leq t_3$ is

$$c_h \int_{t_2}^{t_3} (I(t))dt = \frac{c_h z_3 T_3}{2} \tag{15}$$

Now $T_2 + T_3 = \frac{q}{p}$, using the Eq. (9) & Eq. (14) the holding cost is

$$= \frac{c_h}{2}\{(1-x)p - D(s)\}\frac{q^2}{p^2} - \frac{c_h q \omega}{p} + \frac{c_h \omega^2}{2\{(1-x)p-D(s)\}} \tag{16}$$

The differential equation for time period $t_3 \leq t \leq t_4$, is

$$\frac{dI}{dt} = r - D(s) \tag{17}$$

with boundary condition $I(t_3) = z_3$, $I(t_4) = z_4$, where z_4 is the highest inventory level of good items

$$I(t) = z_3 + \{r - D(s)\}(t - t_3) \tag{18}$$

by using the condition $I(t_4) = z_4$

$$z_4 - z_3 = \{r - D(s)\}T_4 \tag{19}$$

After some simplification and putting $T_4 = \frac{qx}{r}$, we get

$$z_4 = q\{1 - \frac{D(s)(r+x)}{pr}\} - \omega \tag{20}$$

Holding cost for good poducts for the time interval $t_3 \leq t \leq t_4$ is given by

$$c_h \int_{t_2}^{t_3} (I(t))dt = \frac{c_h}{2}(z_3 + z_4)T_4 \tag{21}$$

Putting the value from Eq. (19) then holding cost

$$\begin{aligned}
&= \frac{c_h T_4}{2}\{z_3 + z_3 + \{r - D(s)\}T_4\} \\
&= c_h T_4 z_3 + \frac{c_h\{r-D(s)\}T_4{}^2}{2} \\
&= c_h\{(1-x)p - D(s)\}T_3T_4 \quad .2cm by Eq.(3.14) \\
&= c_h\{(1-x)p - D(s)\}(\frac{q}{p} - \frac{\omega}{\{(1-x)p-D(s)\}})\frac{qx}{r} + \{r - D(s)\}\frac{q^2x^2}{r^2} \\
&= c_h\{(1-x)p - D(s)\}\frac{q^2x}{pr} - \frac{c_h\omega qx}{r} + \frac{c_h}{2}\{r - D(s)\}\frac{q^2x^2}{r^2}
\end{aligned} \tag{22}$$

Now it can be seen that the difective products produced during the time interval $t_1 \le t \le t_3$ at rate p_d. The defective products are reworked perfectly during the time interval $[t_3, t_4]$ by the rework rate r. In this system there is no defective items after time $t = t_4$.

The differential equation for time period $t_4 \le t \le t_5$, that show inventory level is

$$\frac{dI}{dt} = -D(s) \tag{23}$$

with boundary conditions $I(t_4) = z_4$ and $I(t_5) = 0$

Then the solution of this differential equation

$$I(t) = D(s)(t_5 - t) \tag{24}$$

$$By using \quad .2cm I(t) = z_4, \quad .5cm z_4 = D(s)T_5 \tag{25}$$

Holding cost for the time interval $t_4 \le t \le t_5$ is given by

$$\begin{aligned}
c_h \int_{t_4}^{t_4} (I(t))dt = &\ \frac{c_h}{2} z_4 T_5 \\
= &\ \frac{c_h z_4{}^2}{2D(s)} \\
= &\ \frac{c_h}{2D(s)}\left[q\{1 - \frac{\beta(r+x)}{pr}\} - \omega\right]^2
\end{aligned} \tag{26}$$

The inventory of defective products is given figure(2) then the differential equation for time period $t_1 \le t \le t_3$

$$\frac{dI_d}{dt} = p_d, \quad .2cm with boudary condition \quad .2cm I_d(t_1) = 0, \quad .2cm I_d(t_3) = qx \tag{27}$$

Then the solution is

$$I_d(t) = p_d(t - t_1) \tag{28}$$

Holding cost for the defective products is

$$c_h \int_{t_1}^{t_3} (I_d(t))dt = \frac{c_h q^2 x}{2p} \tag{29}$$

For time interval $t_3 \le t \le t_4$ the governing differential equation inventory level of the defective item, is given by

$$\frac{dI_d}{dt} = -r, \quad .2cm with boudary condition \quad .2cm I_d(t_3) = qx, \quad .2cm I_d(t_4) = 0 \tag{30}$$

Then the solution is

$$I_d(t) = r(t_4 - t) \tag{31}$$

The holding cost of reworked items

$$c_{h_r} \int_{t_3}^{t_4} (I_d(t))dt = \frac{c_{h_1} q^2 x^2}{2r} \tag{32}$$

The total profit = Revenue - total cost

= Revenue - (backorder cost + cost of lost sale + holding cost for good and defective products + holding cost for reworked items + purchase cost + repairing cost for defective items + set-up cost)

$$\begin{aligned}\Pi(q,t_1,s) = \quad & sq-\frac{c_bD(s)\{1-at_1e^{-at_1}-e^{-at_1}\}}{a^2}-\frac{c_lD(s)(at_1-1+e^{-at_1})}{a}\\ &-\frac{c_b\omega^2}{2\{(1-x)p-D(s)\}}-\frac{c_h}{2}(1-x)\frac{q^2}{p}+\frac{c_h}{2}\frac{D(s)q^2}{p^2}+\frac{c_hq\omega}{p}\\ &-\frac{c_h\omega^2}{2\{(1-x)p-D(s)\}}-c_h\{(1-x)p-D(s)\}\frac{q^2x}{pr}+\frac{c_h\omega qx}{r}\\ &-\frac{c_h}{2}\{r-D(s)\}\frac{q^2x^2}{r^2}-\frac{c_h}{2D(s)}\left[q\{1-\frac{\beta(r+x)}{pr}\}-\omega\right]^2\\ &-\frac{c_hq^2x}{2p}-\frac{c_{h_1}q^2x^2}{2r}-c_pq-c_rqx-k\end{aligned} \tag{33}$$

The total average profit of the model

$$\begin{aligned}\Pi_{\tilde{a}\tilde{t}\tilde{p}} = \quad & \frac{D(s)}{q}\Pi(q,t_1,s)\\ = \quad & \frac{D(s)}{q}[sq-\frac{c_bD(s)\{1-at_1e^{-at_1}-e^{-at_1}\}}{a^2}-\frac{c_lD(s)(at_1-1+e^{-at_1})}{a}\\ &-\frac{c_b\omega^2}{2\{(1-x)p-D(s)\}}-\frac{c_h}{2}(1-x)\frac{q^2}{p}+\frac{c_h}{2}\frac{D(s)q^2}{p^2}+\frac{c_hq\omega}{p}\\ &-\frac{c_h\omega^2}{2\{(1-x)p-D(s)\}}-c_h\{(1-x)p-D(s)\}\frac{q^2x}{pr}+\frac{c_h\omega qx}{r}\\ &-\frac{c_h}{2}\{r-D(s)\}\frac{q^2x^2}{r^2}-\frac{c_h}{2D(s)}\left[q\{1-\frac{\beta(r+x)}{pr}\}-\omega\right]^2\\ &-\frac{c_hq^2x}{2p}-\frac{c_{h_1}q^2x^2}{2r}-c_pq-c_rqx-k]\end{aligned} \tag{34}$$

The total expected average profit of the model

$$\begin{aligned}\Pi_{\tilde{a}\tilde{t}\tilde{p}} = \quad & \frac{D(s)}{q}[sq-\frac{c_bD(s)\{1-at_1e^{-at_1}-e^{-at_1}\}}{a^2}-\frac{c_lD(s)(at_1-1+e^{-at_1})}{a}\\ &-\frac{c_b\omega^2}{2\{(1-m)p-D(s)\}}-\frac{c_h}{2}(1-m)\frac{q^2}{p}+\frac{c_h}{2}\frac{D(s)q^2}{p^2}+\frac{c_hq\omega}{p}\\ &-\frac{c_h\omega^2}{2\{(1-m)p-D(s)\}}-c_h\{(1-m)p-D(s)\}\frac{q^2m}{pr}+\frac{c_h\omega qm}{r}\\ &-\frac{c_h}{2}\{r-D(s)\}\frac{q^2(m^2+\sigma^2)}{r^2}-\frac{c_h}{2D(s)}\left[q\{1-\frac{\beta(r+m)}{pr}\}-\omega\right]^2\\ &-\frac{c_hq^2m}{2p}-\frac{c_{h_1}q^2(m^2+\sigma^2)}{2r}-c_pq-c_rqm-k]\end{aligned} \tag{35}$$

Eq. (4) & Eq. (35) leads to

$$\Pi_{\tilde{e}\tilde{a}\tilde{t}\tilde{p}} = f_1(q,s,t_1) = u_0(s)+u_1(s,t_1)+\frac{u_2(s,t_1)}{\Psi(s)q} \tag{36}$$

where

$$\begin{aligned}
u_0(s) &= x_{00}+x_{01}D(s)+x_{02}D(s)^2\\
u_1(s,t_1) &= w_1(s)+w_2(s)e^{-at_1}\\
u_2(s,t_1) &= v_1(s)e^{-2at_1}+\{v_2(s)+t_1v_3(s)\}e^{-at_1}+v_4(s)t_1+v_5(s)\\
\Psi(s) &= 2a^2\{(1-m)p-D(s)\}\\
v_1(s) &= \lambda_{11}D(s)^2+\lambda_{12}D(s)^3\\
v_2(s) &= \lambda_{21}D(s)+\lambda_{22}D(s)^2+\lambda_{22}D(s)^3\\
v_3(s) &= \lambda_{31}D(s)^2+\lambda_{32}D(s)^3\\
v_4(s) &= \lambda_{41}D(s)^2+\lambda_{42}D(s)^3\\
v_5(s) &= \lambda_{51}D(s)+\lambda_{52}D(s)^2+\lambda_{53}D(s)^3\\
w_1(s) &= x_{11}D(s)+x_{12}D(s)^2\\
w_2(s) &= x_{21}D(s)+x_{22}D(s)^2\\
\lambda_{11} &= -c_hp(1-m);\lambda_{12}=-c_b;\lambda_{21}=0\\
\lambda_{22} &= 2c_bp(1-m)+2\{c_hp(1-m)-c_lp(1-m)a\};\lambda_{23}=2c_la\\
\lambda_{31} &= 2c_bp(1-m)a;\lambda_{32}=-2c_ba;\lambda_{41}=-2c_lp(1-m)a^2\\
\lambda_{51} &= 2kp(1-m)a^2;\lambda_{52}=2c_bp(1-m)+\{-c_hp(1-m)+c_lp(1-m)a\}\\
\lambda_{53} &= c_b-2c_la;x_{00}=\frac{-c_h}{2};\\
x_{01} &= -\frac{c_h(1-m)}{2p}-\frac{c_hm}{2p}-\frac{c_h(1-m)m}{r}-\frac{c_h(m^2+\sigma^2)}{2r}-\frac{c_{h_1}(m^2+\sigma^2)}{2r}+\frac{c_h(m+r)}{pr}\\
x_{02} &= \frac{c_h}{2p^2}+\frac{c_hm}{pr}+\frac{c_h(m^2+\sigma^2)}{2r^2}-\frac{c_h(m+r)^2}{2p^2r^2};x_{11}=-c_p+s-c_rm+\frac{c_h}{a}\\
x_{12} &= \frac{c_hm}{ra}-\frac{c_hm}{pra};x_{21}=-\frac{c_h}{a};x_{22}=-\frac{c_hm}{ra}+\frac{c_hm}{pra}
\end{aligned}$$

Proposition. The profit function $f_1(q,s,t_1)$ is concave if the corresponding Hessian matrix H of expected profit function is negative definite. where

$$H=\begin{pmatrix} \frac{\partial^2 f_1}{\partial q^2} & \frac{\partial^2 f_1}{\partial s\,\partial q} & \frac{\partial^2 f_1}{\partial q\,\partial t_1} \\ \frac{\partial^2 f_1}{\partial s\,\partial q} & \frac{\partial^2 f_1}{\partial s^2} & \frac{\partial^2 f_1}{\partial t_1\,\partial s} \\ \frac{\partial^2 f_1}{\partial q\,\partial t_1} & \frac{\partial^2 f_1}{\partial t_1\,\partial s} & \frac{\partial^2 f_1}{\partial {t_1}^2} \end{pmatrix}$$

Proof: We have

$$\Pi_{\tilde{e}\tilde{a}\tilde{t}\tilde{p}} = f_1(q,s,t_1) = u_0(s) + u_1(s,t_1) + \frac{u_2(s,t_1)}{\Psi(s)q}$$

$$\frac{\partial f_1}{\partial q} = x_{00} + x_{01}D(s) + x_{02}D(s)^2 - \frac{v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)}{q^2\Psi(s)}$$

$$\frac{\partial f_1}{\partial s} = w_1'(s) + w_2'(s)e^{-at_1} + q\{x_{01}D'(s) + 2x_{02}D'(s)D(s)\} - \frac{\{v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)\}\Psi'(s)}{q\Psi(s)^2} + \frac{v_1'(s)e^{-2at_1} + \{v_2'(s)+t_1v_3'(s)\}e^{-at_1} + v_4'(s)t_1 + v_5'(s)}{q\Psi(s)}$$

$$\frac{\partial f_1}{\partial t_1} = -aw_2(s)e^{-at_1} + \frac{-2av_1(s)e^{-2at_1} + e^{-at_1}v_3(s) - \{v_2(s)+t_1v_3(s)\}ae^{-at_1} + v_4(s)}{q\Psi(s)}$$

Solve above equations by puting

$$\frac{\partial f_1}{\partial q} = 0, \frac{\partial f_1}{\partial s} = 0, \frac{\partial f_1}{\partial t_1} = 0$$

and get the values of variable q, s, t_1

$$x_{00} + x_{01}D(s) + x_{02}D(s)^2 - \frac{v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)}{q^2\Psi(s)} = 0$$

Then

$$q = \sqrt{\frac{v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)}{\{x_{00}+x_{01}D(s)+x_{02}D(s)^2\}\Psi(s)}} \tag{37}$$

Substituting the value of q in the Eq. $\frac{\partial f_1}{\partial s} = 0$ & $\frac{\partial f_1}{\partial t_1} = 0$ and solving them, we get the solution of decision variable q, s, t_1 of the model.

If the second order condition of of optimization method will be satisfied then above solution will be optimal.

Now the second order derivatives

$$\frac{\partial^2 f_1}{\partial q^2} = \frac{2[v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)]}{q^2\Psi(s)} \tag{38}$$

$$\frac{\partial^2 f_1}{\partial q\,\partial t_1} = -\frac{-2v_1(s)e^{-2at_1} + v_3(s)e^{-at_1} + v_4(s) - \{v_2(s)+t_1v_3(s)\}ae^{-at_1}}{q^2\Psi(s)} \tag{39}$$

$$\frac{\partial^2 f_1}{\partial s^2} = -\frac{2\{v_1'(s)e^{-2at_1} + \{v_2'(s)+t_1v_3'(s)\}e^{-at_1} + v_4'(s)t_1 + v_5'(s)\}\Psi'(s)}{q\Psi(s)^2} + \frac{2\{v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)\}\{\Psi'(s)\}^2}{q\Psi(s)^3} + \frac{v_1''(s)e^{-2at_1} + \{v_2''(s)+t_1v_3''(s)\}e^{-at_1} + v_4''(s)t_1 + v_5''(s)}{q\Psi(s)} + w_1''(s) + w_2''(s)e^{-2at_1} + q[x_{01}D''(s) + 2x_{02}D''(s)D(s) + 2x_{02}\{D'(s)\}^2] \tag{40}$$

$$\frac{\partial^2 f_1}{\partial s\,\partial q} = -\frac{\{v_1'(s)e^{-2at_1} + \{v_2'(s)+t_1v_3'(s)\}e^{-at_1} + v_4'(s)t_1 + v_5'(s)\}}{q^2\Psi(s)} + \frac{\{v_1(s)e^{-2at_1} + \{v_2(s)+t_1v_3(s)\}e^{-at_1} + v_4(s)t_1 + v_5(s)\}\{\Psi'(s)\}}{q\Psi(s)^2} + 2x_{01}x_{02}D'(s) + D'(s)D(s) \tag{41}$$

$$\frac{\partial^2 f_1}{\partial {t_1}^2} = a^2 w_2(s) e^{-at_1} + \frac{4a^2 v_1(s) e^{-2at_1} - 2e^{-at_1} v_3(s) + \{v_2(s) + t_1 v_3(s)\} a^2 e^{-at_1}}{q\Psi(s)} \quad (42)$$

$$\frac{\partial^2 f_1}{\partial s\, \partial t_1} = \frac{-2av_1'(s) e^{-2at_1} + v_3'(s) e^{-at_1} + v_4'(s) - \{v_2'(s) + t_1 v_3'(s)\} a e^{-at_1}}{q\Psi(s)} - aw_2'(s) e^{-at_1} - \frac{[-2av_1(s) e^{-2at_1} + e^{-at_1} v_3(s) - \{v_2(s) + t_1 v_3(s)\} a e^{-at_1} + v_4(s)]\Psi'(s)}{q\Psi(s)^2} \quad (43)$$

putting all values of second derivatives in Hessian matrix

$$\mathrm{H}= \begin{pmatrix} \frac{\partial^2 f_1}{\partial q^2} & \frac{\partial^2 f_1}{\partial s\, \partial q} & \frac{\partial^2 f_1}{\partial q\, \partial t_1} \\ \frac{\partial^2 f_1}{\partial s\, \partial q} & \frac{\partial^2 f_1}{\partial s^2} & \frac{\partial^2 f_1}{\partial t_1\, \partial s} \\ \frac{\partial^2 f_1}{\partial q\, \partial t_1} & \frac{\partial^2 f_1}{\partial t_1\, \partial s} & \frac{\partial^2 f_1}{\partial {t_1}^2} \end{pmatrix}$$

If all eigen values are nagetive i.e Hessian matrix H of expected profit function is negative definite, then the profit function is concave.

4 Numerical Example & Sensitivity Analysis

Consider a numerical example taking the demand function as given in Eq. (1)

4.1 Example

We consider the demand function $D(s) = \varphi s^{-\eta}$ and the value of the parameter in appropriate units are as follows $\eta = 1.2$, $c_l = 2$ per unit per unit time, $c_b = 1.5$ per unit per unit time, $k = 500$, $c_h = 1$ per unit per unit time, $c_{h_1} = 1$ per unit per unit time, $c_r = 1.5$ per unit, $c_p = 4$ per unit, $\varphi = 3000$, $r = 1200$ units per unit time, $\alpha = 1.6$, $m = 0.05$, $\sigma^2 = \frac{1}{1200}$, $p = 800$ units per unit time, and randam variable follows uniform distribution in the interval (0,0.1). Then the optimal values for the model are $f_1^* = 1107.4$, $s^* = 39.15$, $q^* = 206$, $t_1^* = 0.69$. These values are optimal as the eigen value of the Hessian matrix $\begin{pmatrix} \frac{\partial^2 f_1}{\partial q^2} & \frac{\partial^2 f_1}{\partial s\, \partial q} & \frac{\partial^2 f_1}{\partial q\, \partial t_1} \\ \frac{\partial^2 f_1}{\partial s\, \partial q} & \frac{\partial^2 f_1}{\partial s^2} & \frac{\partial^2 f_1}{\partial t_1\, \partial s} \\ \frac{\partial^2 f_1}{\partial q\, \partial t_1} & \frac{\partial^2 f_1}{\partial t_1\, \partial s} & \frac{\partial^2 f_1}{\partial {t_1}^2} \end{pmatrix}$ are negative. i.e -31.88, -0.39, -0.002. So the profit function is concave.

4.2 Sensitive Analysis

We observed the sensitiveness of the key parameters which help the decision makers to take appropriate decision on their marketing strategy.

From Table 1, we observed that, with the increasing values of holding cost of products there is a minor change in the optimal lot size and selling price, but the expected average profit decreases shortlly and there is negligible changes in the period of shortage. It is clear that higher holding cost reduce the lot size. So smaller commodity causes the increas in shortage period. In this situation the expected average total profit in decreasing order.

From Table 2, we noticed that, the optomal lot size, shortage period and selling price are increasing with increasing production cost and we also fund that expected profit decreases with increasing the production cost.

From Table 3, we observed that, with the increasing values of backorder cost there is a minor changes in the optimal lot size and selling price, and there is negligible changes in the expected profit and shortage period.

We observed that, with the increasing values of parameter η there is a major change in the optimal lot size and selling price, the expected average profit decreases and there is negligible changes in the period of shortage (table 5). With the changes of parameter a, ther are minor change in optimal lot size, selling price and expected average profit with the increasing values of parameter a shortage period decreases. (table 4). If the demand function parameter φ increases, the expected average profit, and lot size increases highly while the selling price and shortage period decreases (from table 6).

Now we have followed graphical analysis method three-dimensional (3D) plots for the profit function $\Pi_{\tilde{e}\tilde{a}\tilde{t}\tilde{p}}$, Figure 1 and 2 present the piecewise 3D plots for the profit function, $\Pi_{\tilde{e}\tilde{a}\tilde{t}\tilde{p}}$, versus the two corresponding variables subsequently out of the three variables, s, q and t_1 . In each Figure 1 and 2, 3D plot of function, t_1 using the other two variables, s and q at a fixed shortage time period t_1 and 3D plot of function, $\Pi_{\tilde{e}\tilde{a}\tilde{t}\tilde{p}}$, using the other two variables, s and t_1 at a fixed lot-size q.

Table 1: Changges in c_h

c_h	s	t_1	q	f_1
1	39.15	0.69	206	1107.4
1.1	39.49	0.72	195	1099.26
.2	37.30	0.72	197	1090.8
.3	25.69	0.70	275	1052.24

Table 2: Changges in c_p

c_p	s	t_1	q	f_1
3	31.30	0.64	240	1144.37
4	39.15	0.69	106	1107.4
	46.83	0.74	183	1074.41
	54.39	0.78	166	1047.31

Table 3: Changges in c_b

c_b	s	t_1	q	f_1
.5	19.88	0.73	446	1027.81
1	39.06	0.72	207	1107.82
.5	39.15	0.64	206	1107.4
	39.24	0.66	206	1107.01

Table 4: Changges in a

a	s	t_1	q	f_1
1.3	38.90	0.82	209	1109.22
1.4	31.78	0.73	245	1103.37
.5	32.62	0.70	240	1103.95
.6	39.15	0.69	206	1107.4

Table 5: Changges in η

η	s	t_1	q	f_1
1.1	77.56.15	0.78	167	1687.86
1.2	39.15	0.69	206	1107.4
.3	22.29	0.68	289	744.67
.4	22.71	0.69	210	519.93

Table 6: Changges in φ

φ	s	t_1	q	f_1
3000	39.15	0.69	206	1107.4
3500	37.77	0.65	235	1308.24
	36.69	0.62	254	1510.38
	35.83	0.59	277	1713.58

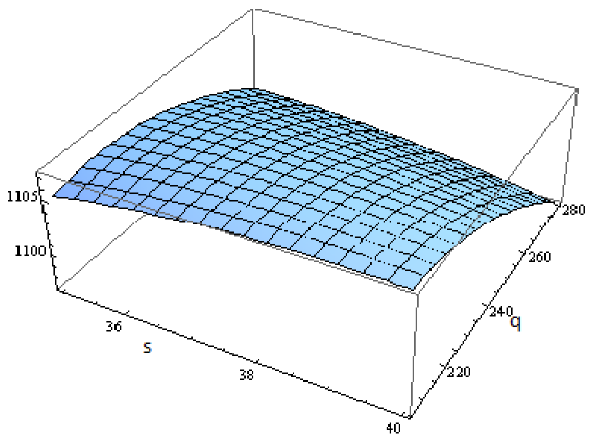


Fig.1. Expected average total profit versus quantity and price

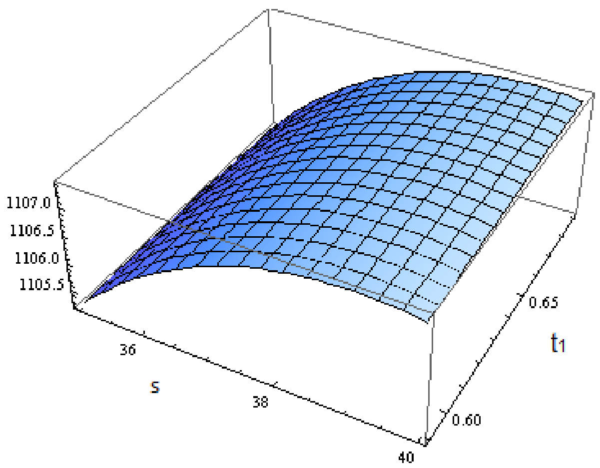


Fig.2. Expected average total profit versus shortage time and price

5 Conclusion

Several manufacturers have to call back their items after use and rework on them to make protect. satisfy the demands with new ones in recent years. This type of remanufacturing system may prevent disposal cost and reduce environment dilemmas. To overcome this problem, an economic production quantity model has been portrayed for imperfect items with rework and production.

We have presented an imperfect production inventory model by considering demand as nagetive power function of selling price. The shortage occurs in begning bears the more cost for inventory manager, but it helps to project the product and optimize the selling price also. We have also illustrated the model numerically for demand depending on selling price. In the sensitivity of parameters of the model, we observed that the optimal expected average profit decreases with higher holding cost of items and optimal expected average profit increases with higher value of parameter φ.

References

[1] Bose, S., Goswami, A., and Chaudhuri, K. S. (1995). An EOQ model for deteriorating items with linear time-dependent demand rate and shortage under inflation and time discount. *The Journal of Operational Research Society*, 46(6),771-782.

[2] Chakrabarti, T., Chaudhuri, K. S. (1997). An inventory model for deteriorating items with alinear trend *International Journal of Production Economics*, 49(3), 205-213.

[3] Chern, M.S., Yang, H.L., Teng, J.T., and Papachristos, S. (2008). Partial backlogging inventory lot-size models for deteriorating items with fluctuating demand under inflation, European Journal of Operational Research, 191(1), 127-141.

[4] Cárdenas Barrón, L.E. (2009). Economic production quantity with rework process at a single-stage manufacturing system with planned backorders. *Computers and Industrial Engineering*, 57(3), 1105-1113.

[5] Datta, T.K. (2013). An inventory model with price and quality dependent demand where some items produced are defective. *Advances in Operations Research*.

[6] Dye, C.Y., Hsieh, T.P., and ouyang, L.Y. (2007). Determining optimal selling price and lot-size with a varying rate of deteriorating and exponential partial backlogging. *European Journal of Operational Research*, 181(2), 668-678.

[7] Das Roy, M., Sana, S.S., and Chauduri, K.S. (2011). An economic order quantitymodel of imperfect quality items with partial backlogging. *International Journal of Systems Science*, 42(8), 1409-1419.

[8] Haider, M.L., Salameh, M., and Nasr, W. (2016). Production lot sizing with quality screening and rework. *Applied Mathematical Modelling*, 40(4), 3242-3256.

[9] Jaber, M.Y., Goyal, S.K.,and Imran, M. (2008). Economic production quantitymodel for items with imperfect quality subject to learning effects. *International Journal of Production Economics*, 115(1), 143-150.

[10] Khedlekar, U. K., Shukla, D. and Namdeo, A. (2016). Pricing policy for declining demand using item preservation technology. *SpringerPlus*, 5(1):19-57.

[11] Konstantaras I., Goyal, S.K., Papachristos, S. (2007). Economic ordering policy for an item with imperfect quality subject to the in-house inspection. *International Journal of Systems Science*, 38(6), 473-482.

[12] Kumar, M., Chauhan, A. and Kumar, R. (2013). A Deterministic inventory model for deteriorating items with price dependent demand and time varyring holding cost under trade credit. *International Journal of Soft Computing and Engineering*, 2231-2307.

[13] Lee, Y.P., Dye, C.Y. (2012). An inventory model for deteriorating item under stock dependent demand and controllable deterioration rate. *Computers & Industrial Engineering*, 63(2), 474-82

[14] Leung, K.N.F. (2008). Using the complete squares method to analyse a lot – size model when the quantity backordered and the quantity received. *European Journal of Operational Research*, 187, 9-30.

[15] Lodree, E.J. (2007). Advanced supply chain planning with mixtures of backorders, lost sales, and lose contract. *European Journal of Operational Research*, 181(1), 168-183.

[16] Mishra, V. K. (2016). Inventory model of deteriorating item with revenue sharing on preservation technology investment under price sensitive stock dependent demand. *International Journal of Mathematical Modelling & Computations*, 37-48.

[17] Pal, B., Adhikari, S. (2017). Price-sensitive imperfect production inventory model withexponential partial backlogging. *International Journal of System Science: Operations & Logistics*, 2330-2674.

[18] Roy, T. and Chaudhuri, K. S. (2008). An EPLS model for a variable production rate with stock price sensitive demand and deterioration. *Yugoslav Journal of Operational Research*, 21.

[19] Sarkar, B., Sana, S.S. and Chaudhuri, K.S. (2011). An imperfect process for time varying demand with inflation and time value of money An EMQ model. *Expert Systems with Applications*, 38, 13543-3548.

[21] Sarkar, B. (2012). An inventory model with reliability in an imperfect production process. *Applied Mathematics and Computation*, 218, 4881-4891.

[21] Sarkar, B. (2013). A production-inventory model with probabilistic deterioration in two-echelon supply chain management. *Applied Mathematical Modelling*, 37, 3138-3151.

[22] Sajadieh, M. S. and Akbari, M. and Jokar, R. (2009). Optimizing Shipment ordering and pricing policies in a two stage supply chain with price sensitive demand. *Transportation Research Part E*, 45, 564-571.

[24] Sana, S. S. (2010). Optimal selling price and lotsize with time varying deterioration and partial backlogging. *Applied Mathematics and Computation*, 217, 185-194.

[24] Sana, S. S. (2011). Price sensitive demand for perishable item an EOQ model, Applied Mathematics and Computation, 217, 6248-6259.

[25] Taleizadeh, A.A., Sajadi, S.J., and Niaki, S. T. A. (2011). Multiproduct EPQ model with single machine, backordering and immediate rework proces. *European Journal Industrial Engineering*, 5(4), 388-411.

[26] Taleizadeh, A.A., Jalali-Naini, S. Gh., Wee, H.M., and Kuo, T.C. (2013). An imperfect multi-product production system with rework. *Scientia Iranica E*, 20(3), 811-823.

[27] Taleizadeh, A.A., Noori-daryan, M. (2016a). Pricing, manufacturing and inventory policies for rawmaterial in a threelevel supply chain. *International Journal of System Science*, 47(4), 919-931.

[28] Taleizadeh, A.A., Samimi, H., Sarkar, B., Mohammadi, B. (2017). Stochastic machine breakdown and discrete delivery in an imperfect inventory production system. *Journal of Industrial and Management Optimization*, doi:10.3934/jimo.2017005

[29] Tayyab, M., and Sarkar, B. (2016). Optimal batch quantity in a cleaner multi-stage lean production system with random defective rate. *Journal of Cleaner Production*, 139, 922-934.

[30] Teksan, Z.M., Geunes, J. (2016). An EOQ model with pricedependent supply and demand, *International Journal of Production Economics*, 178, 22-33.

[31] Thangam, A., and Uthayakumar, R.(2008). A two level supply chain with partial backordering and approximate Poisson demand. *European Journal of Operational Research*, 187, 228-242.

[32] Wee, H.M., Yu, J., and Chen, M.C. (2007). Optimal inventory model for items with imperfect quality and shortage backordering. *Omega*, 35(1), 7-11

A Bayes Analysis and Comparison of Weibull and Lognormal Based Accelerated Test Models with Actual Lifetimes Unknown

Prof. S.K. Upadhyay

•

Department of Statistics, Banaras Hindu University, Varanasi- 221005, India;
DST- Centre for Interdisciplinary Mathematical Sciences, Banaras Hindu University, Varanasi- 221005, India
Email id: skupadhyay@gmail.com

Miss. Reema Sharma

•

Department of Statistics, Banaras Hindu University, Varanasi- 221005, India
Email id: reema.stat@gmail.com

Abstract

The paper considers an accelerated test situation where the actual lifetimes of the items are not directly observable rather their status are known in the form of binary outcomes. By assuming two widely entertained models, namely the Weibull and the lognormal distributions, for the actual lifetimes, the paper provides full Bayesian analysis of the entertained models when both scale and shape parameters of the models are allowed to vary over the covariates involved in the study, thus giving rise to corresponding accelerated test models. The Bayes implementation is based on sample based approaches, namely the Metropolis algorithm and the Gibbs sampler using proper priors of the parameters where the prior elicitation is based on the expert testimonies. The situation involving missing items where actual status is also unknown is additionally entertained using the same modelling assumption. A comparison between the two entertained models is carried out using some standard Bayesian model comparison tools. Finally, numerical illustration is provided based on a given set of current status data and some relevant findings are reported.

Keywords: Binary outcomes, Missing items, Accelerated testing, Weibull distribution, Lognormal distribution, Log-linear link function, Metropolis algorithm, Gibbs sampler, Model comparison.

1 Introduction

Generally, in life testing experiments, the items or equipments are put on test to observe their exact failure times and, based on the same; various reliability characteristics of the items under consideration are studied. There are situations, however, where exact lifetimes are not observable and the experimenter only happens to know the status of the items with regard to their failure or survival. That is the item is either surviving at the time of observation or found in the failed state. Thus the resulting data, often in the form of binary outcomes, may represent one of the two states of the items at the time of observation and generally referred to as the current status data. An important example includes time to occurrence of tumour in animal carcinogenicity experiments, where one might not be able to observe the exact time to appearance of tumour in the subject rather only observes status of the tumour at a particular time, that is, whether the tumour is present or not. Other examples include testing of electro explosive devices, missiles, rocket motors, air bags in cars, etc. (see Balakrishnan and Ling (2012, 2013)) where the items are found either in working state or in failed state at the time of observation. Thus in all these situations the actual lifetimes are unknown and the experimenter is only observing the status of the items or subjects at the time of observation. Also, while dealing with such life testing experiments, there is a possibility of getting some of the experimental units missing during experimentation due to some known or unknown reasons and, as such, the experimenter is even not in a position to know exactly if such items were surviving or already failed at the time of observation. Say, for example, in animal carcinogenicity experiments involving mice, some of the experimental units (mice) might not be available at their expected places at a specified point of time and, as such, it is not possible to know exactly about their current status even. A similar kind of situation was also studied by Sharma and Upadhyay (2018a) with regard to engineering experiments when the actual lifetimes and the status of some of the items are both unknown.

Fan et al. (2009), Balakrishnan and Ling (2012) and Sharma and Upadhyay (2018a) are some of the important references on the analyses of current status data. Whereas Fan et al. (2009) and Sharma and Upadhyay (2018a) provide Bayesian analysis of such datasets, Balakrishnan and Ling (2012) deals mainly with the classical inferences. Other important references on the analysis of current status data include Balakrishnan and Ling (2013) and Balakrishnan and Ling (2014) where the authors used different lifetime models in their work on classical maximum likelihood (ML) estimation and observed that Weibull distribution stands better than other considered models.

Before we come across some other relevant concepts, we need to consider the appropriate lifetime distributions that are capable of representing the actual lifetimes, which are not exactly known in the present scenario. It may be noted here that we do not have the actual lifetimes of items in the present situation rather only have information on the status of the units at the time of observation, that is, failed or surviving. Thus, if the time of observation is T, the actual lifetime either falls below T or goes beyond T. Since the lifetimes are continuous variates, it is almost unlikely that the failure occurs exactly at T.

Among the various lifetime models, the two-parameter Weibull distribution and the two-parameter lognormal distribution, specified by their scale and shape parameters, are widely used lifetime models in the literature. The two-parameter Weibull distribution is a quite flexible and a rich family that has the capability of accommodating all three hazard rate shapes, that is, increasing, decreasing and constant. This is perhaps the reason that the model is highly explored model and used in a wide variety of situations (see, for example, Lawless (2002), Upadhyay (2010)). Similarly, the two-parameter lognormal distribution is known for its non-monotone hazard rate shape that initially increases and attains maxima. It then decreases and finally approaches to zero for large lifetimes and also at the initial lifetimes (see Lawless (2002)). It is often proclaimed that this decreasing nature of lognormal hazard rate with large lifetimes makes lognormal a less popular lifetime distribution regardless of its versatile hazard rate. However, in spite of this discouraging fact, the distribution receives the attention of a number of reliability practitioners, especially in situations where very large lifetimes are not of interest. We do not go into the details of various inferential developments related to these models due to paucity of space. The interested readers may, however, refer to Mann et al. (1974), Lawless (2002) and Singpurwalla (2006), among others, where the last reference primarily concerns with Bayesian developments.

The Weibull and lognormal distributions differ in their tails and both may fit a dataset equally well in their middle ranges. In fact, when both the distributions are fitted to a lifetime data, the Weibull distribution has an earlier lower tail than the corresponding lognormal distribution. In other words, we can say that a low Weibull percentile is always below the corresponding lognormal percentile making the Weibull distribution more pessimistic (see, for example, Nelson 1990). In spite of several such comparative remarks, the two models are used simultaneously by a number of authors for a variety of lifetime data sets (see, for example, Dumonceaux et al. (1973), Wang (1999) and Upadhyay and Peshwani (2003)). The authors have concluded that the two models appear to be good contenders to each other and, therefore, each one can be used as an alternative to other in a variety of situations. The important classical references on the model comparison include Dumonceaux et al. (1973), Meeker (1984), Kim and Yum (2008), etc. Meeker (1984), however, extended the task of model comparison by focussing on accelerated test plans involving censored data for the two models. The Bayesian contributions on model comparison between Weibull and lognormal models include Kirn et al. (2000), Upadhyay and Peshwani (2003) and Araújo and Pereira (2007). It may be noted that some of these references provide extensive treatment on model comparison and conclude their findings based on various model comparison tools of Bayesian paradigm.

Generally, the experimental units used for the considered situations are highly reliable and, therefore, laboratory based experimentation may result in a very few failures or even no failures in normal operating environment. As a matter of fact, the outcomes of such experimentation may provide one-sided information, that is, all the experimental units are surviving at the time of observation and none have failed. The problem can be resolved to a large extent if the experiment is conducted in an accelerated environment where we allow the items under test to operate at the accelerated levels of the stress(es) or covariates to induce early failures. Say, for instance, this might be the accelerated level of dose of the chemical responsible to induce the tumour with reference to the animal carcinogenicity experiment or the accelerated levels of the stresses such as temperature, humidity and voltage, etc., with reference to the testing of electro explosive devices, missiles, rocket motors, airbags in cars, etc. Moreover, although such experiments are performed in an accelerated environment, the primary objective remains drawing the inferences based on the observed data in the normal operating environment. This can, of course, be achieved by means of a suitable life-stress relationship used to relate the lifetimes with the applied covariate(s)/stress(es). Such relationships are generally decided based on several biological considerations in animal carcinogenicity experiment or physical considerations in engineering experiments (see, for example, Nelson (1990) and Lawless (2002)).

A number of life-stress relationships are suggested in the literature of accelerated testing. Important among these are Arrhenius, Eyring, inverse power, exponential, exponential-power, quadratic and polynomial relationships, etc. These relationships are generally used when the characteristic life is assumed to be influenced by only one covariate or stress variable involved in the study. When multiple stresses or covariates are involved in the process, the most commonly used relationship is the log-linear relationship, which is formed under the assumption that the characteristic life has a log-linear relation with the stress(es). This relationship offers a generalized version although it can be used for the situation where one or two stresses affect the process. Another apparent advantage associated with the log-linear relationship is mathematical convenience in its use. These are some of the reasons that led to maximum usage of log-linear relationship in a variety of situations. A detailed discussion on life-stress relationships and the related issues can be had from Meeker and Escobar (1998), Wang (1999), Nelson (1990) and Sen (2016), etc.

Most of the accelerated tests work under the assumption that only the scale parameter is influenced by the covariate(s)/stress(es) involved in the study whereas the shape parameter remains constant over the covariate(s)/stress(es) (see, for example, Nelson (1990)). However, this is not true in practice for all kinds of datasets and, therefore, the assumption of constant shape parameter with respect to the covariates may hide some important features of the units involved in the study. Meeter and Meeker (1994) is a good reference where applicability of non-constant shape parameter has been given by means of some examples (see also Balakrishnan and Ling (2013)).

The present paper provides Bayes analysis of both Weibull and lognormal based accelerated test models with an assumption that both the parameters of the two models are likely to be affected by the considered stress variates. The paper also considers missing data situation by assuming a hypothetical scenario in the assumed current status data although the hypothetical missing data scenario appears to be quite realistic in practice. Finally, the two models are compared using some standard Bayesian tools such as the deviance information criterion (DIC) and the expected posterior predictive loss (EPPL) criterion.

The plan of the paper is as under. The next section provides the formulation of the likelihood function corresponding to two considered models for a general form of current status data when the experiment is subject to accelerated testing and both scale and shape parameters of the model are affected by the stress variables. Section 3 details the Bayesian model formulation and also comments on the implementation of the Metropolis and the Gibbs sampler algorithms to get the desired posterior based inferences in both non-missing and missing data situations. In section 4, model selection criteria, namely the DIC and the EPPL are discussed in brief. This section is given for completeness only. Section 5 provides numerical illustration based on a real dataset. Finally, the paper ends with a brief conclusion given in the last section.

2 The Models and the Likelihood Functions

Let us consider a life testing experiment that involves I experimental groups where i^{th} group consists of K_i experimental units, $i = 1, 2, \dots, I$. Thus, in all, the experiment involves testing of $\sum_{i=1}^{I} K_i$ experimental units. Besides, we also assume that in i^{th} experimental group, the K_i units are observed for their status in terms of either failure or survival at time T_i where the lifetime of each unit is affected by J kinds of covariates, say x_{ij}; $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$. Accordingly, the observed number of failures or survivors is recorded. Obviously, the resulting outcomes are available in the form of binary data where binary zero is used to represent the failed state and binary one for the state of survival. Let n_i and $r_i = (K_i - n_i)$ denote the number of failures (count of binary zeros) and number of survivals (count of binary ones), respectively, observed at time T_i in the i^{th} experimental group when each unit in the group subject to J covariates or stresses, $i = 1, 2, \dots, I$. The complete structure of the data is shown in Table 1. We have also considered missing data case but the same will be discussed separately.

Table 1: Current status data observed at different points of time under different stresses or covariates (the values in parentheses correspond to missing data situation)

Experimental group	Number of experimental units	Observation time	Number of failures	Number of survivals	Number of missing units	Covariates		
1	K_1	T_1	$n_1(n'_1)$	$r_1(r'_1)$	(m_1)	x_{11}	...	x_{1J}
2	K_2	T_2	$n_2(n'_2)$	$r_2(r'_2)$	(m_2)	x_{21}	...	x_{2J}
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮
I	K_I	T_I	$n_I(n'_I)$	$r_I(r'_I)$	(m_I)	x_{I1}	...	x_{IJ}

Now suppose t_{ik} denote the lifetime for the k^{th} experimental unit in the i^{th} experimental group, where $i = 1, 2, \dots, I$ and $k = 1, 2, \dots, K$. If the lifetimes t_{ik}; $i = 1, 2, \dots, I$, $k = 1, 2, \dots, K$, are assumed to be the independent and identically distributed (iid) Weibull variates then the associated probability density function (pdf) can be written as

$$f_W(t_{ik}) = \left(\frac{\beta_i}{\theta_i^{\beta_i}}\right) t_{ik}^{\beta_i - 1} \exp\left\{-\left(\frac{t_{ik}}{\theta_i}\right)^{\beta_i}\right\};\ t_{ik} > 0,\ \theta_i,\ \beta_i > 0, \forall\, i, k, \quad (1)$$

where θ_i and β_i denote the scale and the shape parameters, respectively, associated with the Weibull model corresponding to i^{th} experimental group. Let us assume that both θ_i and β_i are related to the covariates x_{ij}; $i = 1, 2, \ldots, I$, $J = 1, 2, \ldots, J$, via means of log-linear relationship given as

$$\theta_i = \exp\left(a_0 + \sum_{j=1}^{J} a_j x_{ij}\right) \text{ and } \beta_i = \exp\left(b_0 + \sum_{j=1}^{J} b_j x_{ij}\right);\ i = 1, 2, \ldots, I. \quad (2)$$

The parameters $\underline{a} = (a_0, a_1, \ldots, a_J)$ and $\underline{b} = (b_0, b_1, \ldots, b_J)$ are the parameters associated with the log-linear relationships corresponding to the Weibull scale parameter θ_i and the shape parameter β_i, respectively, $i = 1, 2, \ldots, I$. Obviously, these parameters contribute in the model due to the involvement of covariates or stress variables in the study and the resulting Weibull model can be referred to as the accelerated Weibull model. More specifically, a_0 corresponds to the constant effect of covariates on the scale parameter θ_i whereas the parameter a_j gives the effect of covariate x_{ij} on the same, $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$. A similar interpretation can be given for the components of $\underline{b}$ associated with β_i. Moreover, the components of $\underline{a}$ and $\underline{b}$ are assumed to be real on their support, an assumption that appears justified as well.

To proceed further, it may be noted that in the experiment considered here, we do not observe the actual lifetime data rather only get the information regarding the fact that if the actual lifetimes are either less than the observation time (that is, $t_{ik} \leq T_i$) or exceed it (that is, $t_{ik} > T_i$), $i = 1,2,\ldots, I$ and $k = 1,2,\ldots, K$. The probabilities corresponding to these units under the assumption of Weibul lifetime distribution can be given as

$$P_W(t_{ik} \leq T_i) = 1 - \exp\left\{-\left(\frac{T_i}{\exp(a_0 + \sum_{j=1}^{J} a_j x_{ij})}\right)^{\exp(b_0 + \sum_{j=1}^{J} b_j x_{ij})}\right\}, \quad (3)$$

and $P_W(t_{ik} > T_i) = 1 - P_W(t_{ik} \leq T_i)$. It may be further noted that the equality sign in (3) is used to avoid the discontinuity at time T_i. Of course, this does not make any difference as $t_{ik}; i = 1,2,\ldots, I, k = 1,2,\ldots, K$, are the continuous variates. The expression given in (3) is nothing but the cumulative distribution function (cdf) of the corresponding Weibull distribution and its complimentary probability $P_W(t_{ik} > T_i)$ is the corresponding reliability or survival probability.

On the other hand, if we assume lognormal distribution for the iid random variates $t_{ik}; i = 1,2,\ldots, I$, $k = 1,2,\ldots, K$, the corresponding pdf can be written as

$$f_{LN}(t_{ik}) = \frac{1}{t_{ik}\sigma_i\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_i^2}(\log t_{ik} - \mu_i)^2\right];\ t_{ik} > 0, -\infty < \mu_i < \infty, \sigma_i > 0, \forall\, i, k, \quad (4)$$

where $\exp(\mu_i)$, σ_i denote the scale and shape parameters, respectively, of the lognormal model for the i^{th} experimental group, $i = 1,2,\ldots, I$, and the script LN is used to distinguish the lognormal density with that of Weibull density. Analogous to (2), the parameters of lognormal distribution can be written as

$$\mu_i = a'_0 + \sum_{j=1}^{J} a'_j x_{ij} \text{ and } \sigma_i = \exp(b'_0 + \sum_{j=1}^{J} b'_j x_{ij});\ i = 1,2,\ldots, I, \quad (5)$$

where $\underline{a}' = (a'_0, a'_1, \ldots, a'_J)$ and $\underline{b}' = (b'_0, b'_1, \ldots, b'_J)$ are the parameters associated with the log-linear relationship corresponding to lognormal parameters μ_i and σ_i, respectively, $i = 1,2,\ldots, I$. A detailed interpretation of such parameters is already given while discussing Weibull distribution and, therefore, we presume that the components of $\underline{a}'$ and $\underline{b}'$ can be similarly dealt. Obviously, substituting μ_i and σ_i from (5) into (4) results into accelerated lognormal lifetime model.

The probability associated with the event $t_{ik} \leq T_i$; $i = 1,2,\ldots, I$, $k = 1,2,\ldots, K$, under the assumption of lognormal distribution can be written analogous to (3) as

$$P_{LN}(t_{ik} \leq T_i) = \phi\left\{\frac{\log T_i - a_0 - \sum_{j=1}^{J} a_j x_{ij}}{\exp(b_0 + \sum_{j=1}^{J} b_j x_{ij})}\right\}, \tag{6}$$

where ϕ is the standard normal cdf, $\phi\{z\} = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\xi^2}{2}\right] d\xi$. The other probability $P_{LN}(t_{ik} > T_i)$ is simply $1 - P_{LN}(t_{ik} \leq T_i)$.

Once the model formulation is done as given above, the likelihood function for the model parameters $\underline{a}$ and $\underline{b}$ based on the counts of binary outcomes n_i and r_i under the assumption of Weibull lifetimes can be written as

$$\begin{aligned} L_W(\underline{a}, \underline{b}|n, r, T, x) &= \prod_{i=1}^{I} [P_W(t_{ik} \leq T_i)]^{n_i} [P_W(t_{ik} > T_i)]^{r_i} \\ &= \prod_{i=1}^{I} \left[1 - \exp\left\{-\left(\frac{T_i}{\exp(a_0 + \sum_{j=1}^{J} a_j x_{ij})}\right)^{\exp(b_0 + \sum_{j=1}^{J} b_j x_{ij})}\right\}\right]^{n_i} \\ &\quad \times \left[\exp\left\{-\left(\frac{T_i}{\exp(a_0 + \sum_{j=1}^{J} a_j x_{ij})}\right)^{\exp(b_0 + \sum_{j=1}^{J} b_j x_{ij})}\right\}\right]^{r_i}. \end{aligned} \tag{7}$$

It may be noted that the binary outcomes n_i and r_i are observed at time T_i with corresponding covariates x_{ij}; $j = 1,2,\ldots,J$, for the experimental groups $i = 1,2,\ldots,I$. In (7), we have used the notations n for $(n_1, n_2, \ldots, n_I)$, r for $(r_1, r_2, \ldots, r_I)$, T for $(T_1, T_2, \ldots, T_I)$ and x for $(x_{.1}, x_{.2}, \ldots, x_{.J})$ where $x_{.j} = (x_{1j}, x_{2j}, \ldots, x_{Ij})'$ for $j = 1,2,\ldots,J$.

Analogous to (7), the likelihood function for the model parameters $\underline{a}'$ and $\underline{b}'$ under the assumption of lognormal lifetimes can be written as

$$\begin{aligned} L_{LN}(\underline{a}', \underline{b}'|n, r, T, x) &= \prod_{i=1}^{I} [P_{LN}(t_{ik} \leq T_i)]^{n_i} [P_{LN}(t_{ik} > T_i)]^{r_i} \\ &= \prod_{i=1}^{I} \left[\phi\left\{\frac{\log T_i - a'_0 - \sum_{j=1}^{J} a'_j x_{ij}}{\exp(b'_0 + \sum_{j=1}^{J} b'_j x_{ij})}\right\}\right]^{n_i} \\ &\quad \times \left[1 - \phi\left\{\frac{\log T_i - a'_0 - \sum_{j=1}^{J} a'_j x_{ij}}{\exp(b'_0 + \sum_{j=1}^{J} b'_j x_{ij})}\right\}\right]^{r_i}, \end{aligned} \tag{8}$$

where the various notations used in (8) are already defined while discussing Weibull based likelihood.

2.1 Missing data case

To formalise the missing data situation informally introduced in Section 1, let us assume that m_i is the observed number of missing units out of K_i experimental units tested in i^{th} experimental group. As mentioned earlier, these K_i units are scheduled to be observed at time T_i for their status (failed or surviving) when operated under J different types of covariates or stresses x_{ij}; $i = 1,2,\ldots,I$, $J = 1,2,\ldots,J$ (see Table 1) but m_i missing units are, of course, not observable. Obviously, m_i missing units will consist of two different kinds of units, failed or surviving, if they were continued on experimentation and not found missing. Suppose m'_i ($(m_i - m'_i)$) corresponds to number of failed (surviving) units out of m_i missing units where obviously m'_i is unknown and so is $(m_i - m'_i)$, $i = 1,2,\ldots,I$. Let n'_i and $r'_i = (K_i - n'_i - m_i)$ denote the observed number of failures (count of binary zeros) and survivals (count of binary ones), respectively, out of K_i experimental units when tested in i^{th} experimental group and observed at the time T_i. As usual, the covariates or stresses x_{ij}; $i = 1,2,\ldots,I$, $J = 1,2,\ldots,J$, are assumed to have their effects in missing data case as well.

Now the likelihood function for the model parameters $\underline{a}$ and $\underline{b}$ based on the observed counts n'_i, r'_i and m_i under the assumption of Weibull lifetimes can be written as

$$\begin{aligned}L_{W_m}(\underline{a},\underline{b}|n',r',m,T,x) &= \prod_{i=1}^{I} \; [P_W(t_{ik} \le T_i)]^{n'_i} \; [P_W(t_{ik} > T_i)]^{r'_i} \\ &\times [P_W(t_{ik} \le T_i)]^{m'_i} \; [P_W(t_{ik} > T_i)]^{m_i - m'_i} \\ &= \prod_{i=1}^{I} \left[1 - \exp\left\{-\left(\frac{T_i}{\exp(a_0+\sum_{j=1}^{J} a_j x_{ij})}\right)^{\exp(b_0+\sum_{j=1}^{J} b_j x_{ij})}\right\}\right]^{n'_i+m'_i} \\ &\times \left[\exp\left\{-\left(\frac{T_i}{\exp(a_0+\sum_{j=1}^{J} a_j x_{ij})}\right)^{\exp(b_0+\sum_{j=1}^{J} b_j x_{ij})}\right\}\right]^{r'_i+m_i-m'_i},\end{aligned} \tag{9}$$

where the counts n'_i, r'_i and m_i are observed at the time T_i when the items under the test are exposed to covariates or stresses $x_{ij}; j = 1,2,\ldots,J$, for the experimental groups $i = 1,2,\ldots,I$. In (9), we have used the notations n' for $(n'_1, n'_2, \ldots, n'_I)$, r' for $(r'_1, r'_2, \ldots, r'_I)$, m for $(m_1, m_2, \ldots, m_I)$ and the script m with W stands for Weibull likelihood corresponding to missing data case.

Similarly, if we use the same notations as described for the Weibull case, the likelihood function for the model parameters $\underline{a}'$ and $\underline{b}'$ under the assumption of lognormal lifetimes and missing data situation can be written as

$$\begin{aligned}L_{LN_m}(\underline{a}',\underline{b}'|n',r',m,T,x) &= \prod_{i=1}^{I} \; [P_{LN}(t_{ik} \le T_i)]^{n'_i} \; [P_{LN}(t_{ik} > T_i)]^{r'_i} \\ &\times [P_{LN}(t_{ik} \le T_i)]^{m'_i} \; [P_{LN}(t_{ik} > T_i)]^{m_i - m'_i} \\ &= \prod_{i=1}^{I} \left[\phi\left\{\frac{\log T_i - a'_0 - \sum_{j=1}^{J} a'_j x_{ij}}{\exp(b'_0+\sum_{j=1}^{J} b'_j x_{ij})}\right\}\right]^{n'_i+m'_i} \\ &\times \left[1 - \phi\left\{\frac{\log T_i - a'_0 - \sum_{j=1}^{J} a'_j x_{ij}}{\exp(b'_0+\sum_{j=1}^{J} b'_j x_{ij})}\right\}\right]^{r'_i+m_i-m'_i},\end{aligned} \tag{10}$$

where as usual the script m with LN stands for lognormal likelihood corresponding to missing data case.

The likelihoods given in (9) and (10) are incompletely specified in the sense that they involve unknown components $m' = (m'_1, m'_2, \ldots, m'_I)$ in m but this is not a deterrent issue with regard to Bayesian implementation if attempted using sample based approaches (see, for example, Sharma and Upadhyay (2018a)). For this, we only need to generate the binary response data corresponding to the observed missing units and this can be easily done by generating the iid Bernoulli variates with success probabilities $P_W(t_{ik} > T_i)$ and $P_{LN}(t_{ik} > T_i)$, for accelerated Weibull and lognormal lifetimes, respectively. Once the binary response data corresponding to the missing units are made available, availability of the unknown component m' is obvious (see Sharma and Upadhyay (2018a)). The implementation has been briefed in the next section. It may be noted that the situation may not be straightforward if tried using the tools of classical paradigm. The details of Bayesian implementation for missing data case will be discussed in the next section.

3 Bayesian Model Formulation

As mentioned, the parameters a_j and b_j; $j = 0,1,\ldots,J$, associated with the accelerated Weibull model are assumed to be real on their support and, therefore, we consider normal priors for these model parameters given as

$$g_W(a_j|\mu_{a_j},\sigma_{a_j}) = \frac{1}{\sqrt{2\pi}\sigma_{a_j}} \exp\left\{-\frac{1}{2}\left(\frac{a_j-\mu_{a_j}}{\sigma_{a_j}}\right)^2\right\};\ j = 0,1,\ldots,J, \tag{11}$$

$$g_W(b_j|\mu_{b_j},\sigma_{b_j}) = \frac{1}{\sqrt{2\pi}\sigma_{b_j}} \exp\left\{-\frac{1}{2}\left(\frac{b_j-\mu_{b_j}}{\sigma_{b_j}}\right)^2\right\};\ j = 0,1,\ldots,J, \tag{12}$$

where (μ_{a_j},σ_{a_j}) and (μ_{b_j},σ_{b_j}) are the hyperparameters associated with the prior distributions of a_j and $b_j; j = 0,1,\ldots,J$, respectively. Now assuming that the parameters a_j and b_j; $j = 0,1,\ldots,J$, are *a priori* independent, the joint prior distribution can be written as

$$g(\underline{a},\underline{b}|\mu_a,\sigma_a,\mu_b,\sigma_b) = \prod_{j=0}^{J} g(a_j|\mu_{a_j},\sigma_{a_j}) \times g(b_j|\mu_{b_j},\sigma_{b_j}), \tag{13}$$

where $\mu_a = (\mu_{a_0},\mu_{a_1},\ldots,\mu_{a_J})$, $\sigma_a = (\sigma_{a_0},\sigma_{a_1},\ldots,\sigma_{a_J})$, $\mu_b = (\mu_{b_0},\mu_{b_1},\ldots,\mu_{b_J})$ and $\sigma_{b_0} = (\sigma_{b_0},\sigma_{b_1},\ldots,\sigma_{b_J})$.

For successful implementation of Bayesian techniques, the choice of hyperparameters becomes a significant issue. In case, the experimenter fails to have enough information to define an appropriate prior distribution, it is often recommended going for such choices that make the prior distributions essentially vague. Alternatively, one can attempt eliciting the prior hyperparameters based on the subjective opinion of the experts if the same are made available. To clarify, suppose the expert suggests that the parameters a_j and b_j are bounded within the finite intervals $[l_{a_j},u_{a_j}]$ and $[l_{b_j},u_{b_j}]$, respectively, $j = 0,1,\ldots,J$, and due to non-availability of any other significant information, it is assumed that each value within the intervals is equally likely (see Bousquet et al. (2006)). Thus presuming a_j and b_j to be uniformly distributed in the intervals $[l_{a_j},u_{a_j}]$ and $[l_{b_j},u_{b_j}]$, respectively, $j = 0,1,\ldots,J$, one can equate the means and variances of the assumed normal priors with those of the assumed uniform distributions over the corresponding intervals. As a result, the hyperparameters (μ_{a_j},σ_{a_j}) and (μ_{b_j},σ_{b_j}) associated with the prior distributions of a_j and b_j; $j = 0,1,\ldots,J$, can be obtained as

$$\mu_{a_j} = \frac{l_{a_j}+u_{a_j}}{2}, \qquad \sigma_{a_j} = \sqrt{\frac{(u_{a_j}-l_{a_j})^2}{12}};\ j = 0,1,\ldots,J, \tag{14}$$

and

$$\mu_{b_j} = \frac{l_{b_j}+u_{b_j}}{2}, \qquad \sigma_{b_j} = \sqrt{\frac{(u_{b_j}-l_{b_j})^2}{12}};\ j = 0,1,\ldots,J. \tag{15}$$

The prior information formalized in equations (13) gets updated by the medium of likelihood (7) in order to give the joint posterior, which up to proportionality under the accelerated Weibull distribution is given as

$$\begin{aligned} &p_W(\underline{a},\underline{b}|\mu_a,\sigma_a,\mu_b,\sigma_b,n,r,T,x) \\ &\propto \prod_{i=1}^{I}\left[1-\exp\left\{-\left(\frac{T_i}{\exp(a_0+\sum_{j=1}^{J}a_jx_{ij})}\right)^{\exp(b_0+\sum_{j=1}^{J}b_jx_{ij})}\right\}\right]^{n_i} \\ &\times\left[\exp\left\{-\left(\frac{T_i}{\exp(a_0+\sum_{j=1}^{J}a_jx_{ij})}\right)^{\exp(b_0+\sum_{j=1}^{J}b_jx_{ij})}\right\}\right]^{r_i} \\ &\times\prod_{j=0}^{J}\frac{1}{\sigma_{a_j}}\exp\left\{-\frac{1}{2}\left(\frac{a_j-\mu_{a_j}}{\sigma_{a_j}}\right)^2\right\}\times\frac{1}{\sigma_{b_j}}\exp\left\{-\frac{1}{2}\left(\frac{b_j-\mu_{b_j}}{\sigma_{b_j}}\right)^2\right\}. \end{aligned} \tag{16}$$

One can similarly obtain the joint posterior distribution, up to proportionality, under the assumption of accelerated lognormal distribution and the same can be given as

$$p_{LN}(\underline{a}', \underline{b}' | \mu_{a'}, \sigma_{a'}, \mu_{b'}, \sigma_{b'}, n, r, T, x)$$
$$\propto \prod_{i=1}^{I} \left[\phi \left\{ \frac{\log T_i - a'_0 - \sum_{j=1}^{J} a'_j x_{ij}}{\exp(b'_0 + \sum_{j=1}^{J} b'_j x_{ij})} \right\} \right]^{n_i} \times \left[1 - \phi \left\{ \frac{\log T_i - a'_0 - \sum_{j=1}^{J} a'_j x_{ij}}{\exp(b'_0 + \sum_{j=1}^{J} b'_j x_{ij})} \right\} \right]^{r_i}$$
$$\times \prod_{j=0}^{J} \frac{1}{\sigma_{a'_j}} \exp\left\{ -\frac{1}{2} \left(\frac{a'_j - \mu_{a'_j}}{\sigma_{a'_j}} \right)^2 \right\} \times \frac{1}{\sigma_{b'_j}} \exp\left\{ -\frac{1}{2} \left(\frac{b'_j - \mu_{b'_j}}{\sigma_{b'_j}} \right)^2 \right\}, \quad (17)$$

where the prior distributions for the parameters a'_j and b'_j; $j = 0,1,\ldots,J$, are formed in a manner similar to those for the accelerated Weibull parameters but with hyperparameters $({\mu_{a_j}}', {\sigma_{a_j}}')$ and $({\mu_{b_j}}', {\sigma_{b_j}}')$, respectively. Likewise, we define ${\mu_a}' = ({\mu_{a_0}}', {\mu_{a_1}}', \ldots, {\mu_{a_J}}')$, ${\sigma_a}' = ({\sigma_{a_0}}', {\sigma_{a_1}}', \ldots, {\sigma_{a_J}}')$, ${\mu_b}' = ({\mu_{b_0}}', {\mu_{b_1}}', \ldots, {\mu_{b_J}}')$ and ${\sigma_b}' = ({\sigma_{b_0}}', {\sigma_{b_1}}', \ldots, {\sigma_{b_J}}')$.

Both (16) and (17) do not appear to offer closed form solutions. Therefore, we propose to consider sample based approaches to get the desired posterior based inferences (see, for example, Upadhyay et al. (2001) and the references cited therein). No doubt, the sample based approaches to Bayesian computation are beneficial in the sense that they are automatic and simultaneously capable of providing variety of inferential aspects in a routine manner. The commonly used sample based approaches are the Metropolis and the Gibbs sampler algorithms where the former is a more generalised version in the sense that it also includes latter as a special case. The implementation of Gibbs sampler algorithm requires the specification of full conditional distributions for all the generating variates and simultaneously necessitates the availability of such full conditionals from the viewpoint of easy generation of samples. On the other hand, the Metropolis algorithm does not require any such specification of full conditionals rather seeks alternative randomization mechanism for sample generation often in a multidimensional framework.

The joint posterior given in (16) and (17) are the $((J+1)+(J+1))$ dimensional distributions. If we look carefully on various associated full conditionals, it is obvious that the corresponding full conditionals are not easy from the viewpoint of sample generation and, therefore, Metropolis algorithm appears to be an obvious choice for simulating the samples from joint posterior (16) and (17). The Metropolis algorithm is a Markovian updating scheme that uses a suitably chosen candidate generating density, say $q(\omega, \nu)$, for sample generation where ω and ν are the realizations from $q(\omega, \nu)$ at s^{th} and $(s+1)^{th}$ stage, respectively. If the chosen kernel is symmetric, the value generated at each step from $q(\omega, \nu)$ follows a randomization step based on the probability of acceptance $\alpha(\omega, \nu) = \min\left\{1, \frac{p(\nu)}{p(\omega)}\right\}$ where $p(.)$ is the corresponding posterior distribution of possibly vector valued parameter $(.)$. If ν is accepted with probability $\alpha(\omega, \nu)$, it is retained as the candidate point from the target posterior at s^{th} stage otherwise ω is retained as the candidate point from the target posterior at s^{th} stage. The commonly used candidate generating densities are multivariate normal, rectangular and t distributions, etc. (see also Upadhyay et al. (2001)). One can also use the non-symmetric candidate generating density and define a generalized version, known as the Metropolis-Hastings algorithm. We shall not discuss this version as it is beyond the scope of the present work.

For the implementation of the algorithm on the posteriors under consideration, we consider a multivariate normal distribution $(((J+1)+(J+1))$ dimensional) as a candidate generating density. The multivariate normal distribution is chosen because it can be easily specified by its mean vector and covariance matrix, approximate values for the same can be guessed on the basis of ML estimates and the corresponding Hessian based approximation obtained at ML estimates. The calculation of ML estimates and the corresponding Hessian based approximations for the models under consideration cannot be directly obtained rather one requires going for an efficient optimisation technique. Balakrishan and Ling (2013) is an important reference that considers expectation maximization (EM) algorithm for the current status data under the assumed lifetime models. We, however, consider using simulated annealing (SA) optimisation technique (see Robert and Casella (2010)) to obtain the ML estimates of the model parameters for both the accelerated Weibull and the accelerated lognormal distributions and accordingly obtain Hessian based approximations at corresponding ML estimates. Thus using the initial values in the form of ML estimates and corresponding Hessian based approximations, the implementation of the Metropolis algorithm can be done routinely through successive generations from (16) and (17) to get an iterative sequence of states. This sequence, after a long run, converges in distribution to a random sample from the true posterior distribution. The process is to be implanted separately on the two posteriors (see also Robert and Casella (2013) and Smith and Roberts (1993)).

Next considering the aforesaid missing data situation, the joint posterior can be obtained by updating the prior distribution (13) with the likelihood (9) under the accelerated Weibull distribution. Let $p_{W_m}(\underline{a},\underline{b}|\mu_a,\sigma_a,\mu_b,\sigma_b,n',r',m,T,x)$ denotes the posterior in this case. Obviously, this posterior does not behave similar to that for non-missing case given in (16) due to the involvement of an unknown component m' in m. The situation can, however, be easily dealt with by constructing a cycle of Gibbs sampler algorithm for these additional unknowns. It may be noted here that the Gibbs sampler algorithm is also a Markovian updating scheme that requires successive generation from each of its full conditional distributions by using the most recent values of rest of the conditioning variates and the algorithm ultimately proceeds in a cyclic manner (see also Sharma and Upadhyay (2018a)). To implement the algorithm in the present case, we need to form two full conditional distributions. One of these full conditionals focuses on generation of count m' given all other variates $(\underline{a},\underline{b},\mu_a,\sigma_a,\mu_b,\sigma_b,n',r',m,T,x)$. The other full conditional is the joint full conditional of the model parameters $(\underline{a},\underline{b})$ given all other variates including m' that appears in m. The generation of count m' (binary zeros) may be done by Bernoulli variate generation with success probability $P_W(t_{ik} > T_i)$, conditioned on the recently generated values of the model parameters. The generation from other full conditional can be achieved by implementation of the Metropolis algorithm that has already been discussed earlier for non-missing case. As a result, the algorithm can routinely proceed to get the desired posterior samples for missing data case as well. The situation can be similarly detailed for the posterior arising from the accelerated lognormal distribution involving missing data.

4 Bayesian Model Selection

A number of model comparison tools are available in the literature. The logically appealing tools among these are those which take into accountability both goodness of fit and model complexity while deciding the model. The goodness of fit can obviously be summarized on the basis of deviance statistic whereas the model complexity can be based on the number of free parameters in the model. A few such tools are Akaike information criterion (AIC), Bayesian information criterion (BIC), DIC and the posterior predictive loss (PPL) criterion, etc. We, however, restrict ourselves to two most exploited and justified criteria, namely the DIC and EPPL where latter is nothing but PPL with slight change in the definition and nomenclature.

4.1 The deviance information criterion

Spiegelhalter et al. (2002) proposed a measure based on the posterior distribution of deviance that deals with both Bayesian measure of fit and complexity. This measure, known as DIC, obviously consists of two terms. The first term gives the measure of fit and the second term is used to measure the complexity involved in the entertained models. The DIC can be defined as

$$\mathrm{DIC} = E[-2\log(L(\eta|(.)))] + p_D, \tag{18}$$

where $p_D = E[-2\log(L(\eta|(.)))] + 2\log(L(\overline{\eta}|(.)))$, η is a vector of parameters and $\overline{\eta}$ is the corresponding posterior estimates that are usually taken as the posterior mean but posterior mode has also been recommended (see, for example, Upadhyay et al. (2013)). The term $-2\log(L(\eta|(.)))$, considered as a function of η, is classical deviance and is defined as a measure of fit in classical modelling framework. In this light, the first term of (18) is defined to give the measure of fit, which is nothing but the posterior mean of classical deviance. On the other hand, the second terms p_D is known as the effective number of parameters and is used to measure the complexity involved in the modelling. The model providing the least value of DIC is preferred over all other models.

4.2 The expected posterior predictive loss

Initially, the EPPL criterion was developed by Buck and Sahu (2000) for the multinomial cell frequencies and later on Sharma and Upadhyay (2018b) derived the same for the binomial counts. We shall use the same criterion as discussed by Sharma and Upadhyay (2018b) keeping in view the dataset given in Table 1. To discuss the criterion briefly, let r_i and r_i^p denote the number of successes corresponding to observed and predicted future observations, respectively, out of K_i experimental units in the i^{th} experimental group, where $i = 1,2,\ldots,I$. The EPPL criterion derived in Sharma and Upadhyay (2018b) can be rewritten as

$$\begin{aligned} E\{L(r,r^p)\} &= 2\sum_{i=1}^{I}\left[r_i\left\{\log\left(\frac{r_i}{K_i}\right)-\log(p_i^*)\right\}+(K_i-r_i)\left\{\log\left(\frac{K_i-r_i}{K_i}\right)-\log(1-p_i^*)\right\}\right] \\ &+2\sum_{i=1}^{I}\left[r_i\left\{\log(p_i^*)-E\left(\log\left(\frac{r_i^p}{K_i}\right)\right)\right\}+(K_i-r_i)\left\{\log(1-p_i^*)-E\left(\log\left(\frac{K-r_i^p}{K_i}\right)\right)\right\}\right], \end{aligned} \tag{19}$$

where $r^p = (r_1^p, r_2^p, \ldots, r_I^p)$, $p_i^* = E\left(\frac{r_i^p}{K_i}\right)$ and $L(r, r^p)$ is the aggregated loss function over the components of r and r_p. For the i^{th} experimental group, it is given by

$$L(r_i, r_i^p) = 2\left[r_i\log\left(\frac{r_i}{r_i^p}\right)+(K_i-r_i)\log\left(\frac{K_i-r_i}{K-r_i^p}\right)\right];\ i = 1,2,\ldots,I. \tag{20}$$

It is to be noted here that the loss function (20) is derived under the binomial modelling assumption in each of the i^{th} experimental group for r_i (r_i^p) number of successes corresponding to observed (predictive) dataset, in K_i Bernoulli trials (see Sharma and Upadhyay (2018b)), where $i = 1,2,\ldots,I$. A careful look on (19) reveals the situation when we observe exactly zero or K_i; $i = 1,2,\ldots,I$, counts. This problem can be resolved by adding (subtracting) $\frac{1}{2}$ when we observe exactly zero (K_i) counts in the i^{th} experimental group, $i = 1,2,\ldots,I$. Moreover, in (19), expectation is taken with respect the posterior predictive distribution of the future observation r_i^p; $i = 1,2,\ldots,I$, given the observed data. One may note that for the situation considered here, the posterior predictive distribution is not available in analytically closed form although it is manageable with the help of simulated posterior samples. Say, for instance, we can easily generate the predictive samples from the considered distributions once the corresponding posterior samples are made available and hence the expectation in (19) can be obtained accordingly.

The EPPL is a decision theoretic measure for model comparison that comprises of two types of losses. The first one being the loss due to fitting (LDF) also known as goodness of fit term and is given in terms of the likelihood ratio statistic (see the first term on RHS of (19)). The second one is the loss incurred due to complexity (LDC) of the model that may also be used to approximate the predictive variance (see the second term on RHS of (19)). Finally, the criterion recommends a model that has the least value of EPPL over all the entertained models.

5 Numerical Illustration

The numerical illustration is based on a real dataset obtained from a survival/sacrifice experiment conducted at National center for toxicological research (NCTR). The subjects involved in the experiment were the male and female mice from two strains of offspring. The first strain F1 consisted of offspring from mating of BALB/c males to C57BL/6 females. The second strain F2 consisted of offspring from non brother-sister mating of the Fl progeny. The considered dataset is taken from Balakrishnan and Ling (2013) which is a compact form of the original dataset reported in Kodell and Nelson (1980).

The dataset consisted of 823 male and female mice from two strains of offspring classified in $I = 51$ experimental groups with varied number of mice in each group. The mice in different groups were subject to four different doses, 60 ppm, 120 ppm, 200 ppm and 400 ppm, of chemical benzidine dihydrochloride, responsible to develop tumours in the mice, where ppm stands for parts per million. The number of mice with tumours, say, $n_1, n_2, \ldots, n_{51}$, was then recorded in different groups at times $T_1, T_2, \ldots, T_{51}$, respectively. Besides, we also recorded the number of mice without tumours, say, $r_1, r_2, \ldots, r_{51}$, in each group . Obviously, the mice with (without) tumours correspond to the observed number of failures (survivals) at the time of observation. One may also note that the experiment involves three covariates, namely the two strains of offspring (say, F1=0 and F2=1), sex of mice (say, F=0 for females and M=1 for males) and doses of chemical (60 ppm, 120 ppm, 200 ppm and 400 ppm). The distribution of mice according to these three covariates are reported in Table 2.

Let us now consider the case of missing data situation in the considered dataset and assume that some of the mice are found missing when observed at different times $T_1, T_2, \ldots, T_{51}$. In this case, we therefore observe three different categories, that is, number of mice with tumours (n'_i), number of mice without tumours (r'_i) and number of mice found missing (m_i) at the time T_i in the i^{th} experimental group, where $i = 1,2,\ldots,I$. It may be noted that the number of mice considered in the i^{th} experimental group, K_i, is same as that for the non-missing case. Also, the mice in each group are subject to the same covariates as mentioned for the non-missing case. The corresponding dataset is shown in Table 2 where number of missing observations are taken hypothetically for illustration.

Table 2: Observed number of failures $n_i(n'_i)$, survivals $r_i(r'_i)$ and missing units (m_i) at time T_i in the presence of covariates x_{i1} (F1=0, F2=1), x_{i2} (F=0, M=1) and x_{i3} (dose of chemical in ppm), for the i^{th} experimental group, $i = 1,2,\ldots,I$ (the values in parentheses correspond to missing data cases)

Experimental group	Number of experimental units	Observation time (in months)	Number of failures	Number of survivals	Number of missing units	Covariates		
i	K_i	T_i	$n_i(n'_i)$	$r_i(r'_i)$	(m_i)	x_{i1}	x_{i2}	x_{i3}
1	48	9.27	1 (1)	47 (41)	(6)	0	0	60
2	24	9.37	0 (0)	24 (20)	(4)	0	0	60
3	24	13.97	1 (1)	23 (21)	(2)	0	0	60
4	24	9.37	0 (0)	24 (21)	(3)	0	0	120
5	24	13.97	5 (3)	19 (16)	(5)	0	0	120
6	23	14.03	9 (7)	14 (14)	(2)	0	0	120
7	26	18.67	25 (20)	1 (3)	(3)	0	0	120
8	48	9.27	0 (0)	48 (40)	(8)	0	1	120
9	44	14.00	7 (6)	37 (33)	(5)	0	1	120
10	22	18.73	7 (5)	15 (14)	(3)	0	1	120

Experimental group	Number of experimental units	Observation time (in months)	Number of failures	Number of survivals	Number of missing units	Covariates		
11	20	19.30	4 (4)	16 (14)	(2)	0	1	120
12	24	9.27	0 (0)	24 (22)	(2)	1	1	60
13	23	9.30	0 (0)	23 (20)	(3)	1	1	60
14	21	9.37	0 (0)	21 (19)	(2)	1	1	60
15	44	14.00	3 (3)	41 (34)	(7)	1	1	60
16	18	18.67	2 (2)	16 (15)	(1)	1	1	60
17	20	18.70	2 (2)	18 (16)	(2)	1	1	60
18	1	16.53	1 (1)	0 (0)	(0)	0	0	120
19	1	16.57	1 (1)	0 (0)	(0)	0	0	120
20	1	16.90	1 (1)	0 (0)	(0)	0	0	120
21	1	15.13	1 (1)	0 (0)	(0)	0	0	120
22	1	15.40	1 (1)	0 (0)	(0)	0	0	120
23	47	9.33	4 (4)	43 (38)	(5)	0	0	200
24	45	14.00	38 (32)	7 (10)	(3)	0	0	200
25	22	14.00	11 (8)	11 (13)	(1)	0	1	400
26	15	18.70	11 (7)	4 (6)	(2)	0	1	400
27	1	7.87	1 (1)	0 (0)	(0)	0	1	400
28	1	14.73	1 (1)	0 (0)	(0)	0	1	400
29	18	18.70	5 (5)	13 (12)	(1)	1	1	120
30	1	9.57	1 (1)	0 (0)	(0)	1	1	120
31	1	14.43	1 (1)	0 (0)	(0)	1	1	120
32	1	17.87	1 (1)	0 (0)	(0)	1	1	120
33	1	18.03	1 (1)	0 (0)	(0)	1	1	120
34	1	5.13	0 (0)	1 (1)	(0)	1	1	120
35	1	13.53	1 (1)	0 (0)	(0)	0	0	200
36	1	14.03	1 (1)	0 (0)	(0)	0	0	200
37	1	14.23	1 (1)	0 (0)	(0)	0	0	200
38	1	18.67	1 (1)	0 (0)	(0)	0	0	200
39	24	9.33	16 (11)	8 (10)	(3)	0	0	400
40	10	14.00	9 (7)	1 (2)	(1)	0	0	400
41	1	9.87	1 (1)	0 (0)	(0)	0	0	400
42	1	17.13	0 (0)	1 (1)	(0)	1	0	60
43	24	9.27	2 (2)	22 (19)	(3)	1	0	120
44	22	9.37	0 (0)	22 (20)	(2)	1	0	120
45	41	14.00	15 (12)	26 (25)	(4)	1	0	120
46	1	15.43	1 (1)	0 (0)	(0)	1	1	200
47	24	9.30	1 (1)	23 (20)	(3)	1	1	400
48	21	14.00	4 (4)	17 (16)	(1)	1	1	400
49	12	18.67	6 (5)	6 (6)	(1)	1	1	400
50	1	11.90	1 (1)	0 (0)	(0)	1	1	400
51	1	14.77	1 (1)	0 (0)	(0)	1	1	400

Now, going back to Section 2, we can easily see that the present dataset involves $J = 3$ covariates and, therefore, the total number of associated parameters are $((J + 1) + (J + 1)) = 8$. Out of these eight parameters, the first four, that is, a_0, a_1, a_2, a_3 are associated with the Weibull scale parameter and the remaining four, that is, b_0, b_1, b_2, b_3 are associated with the Weibull shape parameter. Similarly, when one considers the analysis of lognormal distribution, the associated parameters are a'_0, a'_1, a'_2, a'_3 and b'_0, b'_1, b'_2, b'_3, respectively, with its scale and shape parameters. One may also note that the parameters a_1, a_2, a_3 and a'_1, a'_2, a'_3 affecting the scale parameters of the corresponding distributions are due to the covariates considered as strain of offspring, sex of mice and dose of the chemical, respectively. Similarly, the parameters b_1, b_2,b_3 and b'_1, b'_2, b'_3 causing the affect on shape parameters of the corresponding distributions enter into the modelling formulation due to the three covariates entertained in the analysis.

To implement the Bayesian modelling formulation as given in Section 3, we first begin with the exact specification of the considered prior distributions for the model parameters of the two entertained models. The prior distribution given in (13) was evaluated based on the interval [-25, 25] specified by the experts for $a_j$and b_j; $j = 0,1,2,3$, under the assumption of accelerated Weibull distribution. The hyperparameters $(\mu_{a_j}, \sigma_{a_j})$ and $(\mu_{b_j}, \sigma_{b_j})$ associated with a_j and b_j can be obvious from (14) and (15) for all j. The prior distributions associated with the accelerated lognormal parameters were also managed in an identical manner.

We next considered simulation of posterior samples from (16) and (17) using Metropolis algorithm as per details given in Section 3, considering a single long run of the chain in each case. For the implementation of the Metropolis algorithm, we considered an eight-dimensional normal distribution with mean vector η and covariance matrix Σ as a candidate generating density. We, however, used a scaling constant $c_s = 0.5$ with Σ to minimize number of rejections of generated values from the candidate generating density (see also Upadhyay et al. (2001)). As initial values for starting the chain, we used $\hat{\eta}$ as a vector of ML estimates of η assuming $\eta = (\underline{a}, \underline{b})$ for the posterior based on accelerated Weibull distribution and $\eta = (\underline{a'}, \underline{b'})$ for the posterior based on accelerated lognormal distribution. Thus ML estimates were obtained using the likelihood functions given in (7) and (8) corresponding to Weibull and lognormal based accelerated models, respectively. Similarly, the initial values for Σ was obtained as $\hat{\Sigma}$ in each case where $\hat{\Sigma}$ is Hessian based approximation corresponding to (7) and (8) evaluated at corresponding ML estimates $\hat{\eta}$. The ML estimates of the model parameters are also reported in Tables 3 and 4 for the accelerated Weibull and the accelerated lognormal distributions, respectively.

Once the simulated chain started producing, we monitored the convergence based on ergodic averages. This was achieved at about 50K iterations for both the models. Once the convergence was achieved, we took a random sample of size 4K separately from the two posteriors by choosing every 10th observation. This gap among the generating variates made serial correlations almost negligible. Thus we finally obtain samples from the joint posteriors (16) and (17), each component of which will form samples from the corresponding marginal posterior. Once these samples are obtained, the desired sample based posterior inferences can be easily drawn (see also Upadhyay et al. (2001)).

The simulation of posterior samples corresponding to missing data situation can be a routine extension of non-missing case described above. As mentioned in Section 3 (see also subsection 2.1), we need to create hybridization with the help of idea inherent in the Gibbs sampler algorithm where the corresponding full conditional distributions can be managed as discussed in Section 3 (see also Sharma and Upadhyay (2018a)).

The estimated posterior modes for various parameters in the two models and the corresponding highest posterior density (HPD) intervals with coverage probability 0.95 are reported in Tables 3-4, where the values in parentheses correspond to the same for the missing data cases. It can be seen that the estimated posterior modes are, in general, close enough to the corresponding ML estimates, a conclusion that is normally appreciated by the classical statisticians. Besides, it conveys the fact that the subjective opinion provided by the experts does not lead to strong prior distributions and the inferences are mostly data dependent. We also observe that the estimates obtained under the missing data cases are, in general, close enough to those corresponding to non-missing cases for both the models. This, in turn, ensures that we do not loose enough if some of the observations are missing during the experimentation. Among other important conclusions, it can be seen that the estimated 0.95 HPD intervals are mostly narrow reflecting small variability among the different variates. We do not go into the details of other conclusions although the same can be easily drawn once the posterior samples are made available.

Table 3: ML estimates and estimated posterior characteristics for the parameters of accelerated Weibull distribution (the values in parentheses correspond to missing data case)

Parameter	ML estimate	Posterior mode	0.95 HPD interval
a_0	2.944	2.950 (2.945)	2.878, 3.033 (2.877, 3.035)
a_1	0.049	0.041 (0.039)	-0.065, 0.256 (-0.072, 0.262)
a_2	0.622	0.594 (0.608)	0.398, 1.017 (0.412, 1.046)
a_3	-0.002	-0.002 (-0.002)	-2.210e-03, -0.001 (-2.221e-03, -0.001)
b_0	2.205	2.214 (2.207)	1.880, 2.572 (1.889, 2.573)
b_1	-0.088	-0.128 (-0.116)	-0.620, 0.285 (-0.622, 0.279)
b_2	-0.816	-0.755 (-0.836)	-1.422, -0.320 (-1.481, -0.318)
b_3	-0.002	-0.002 (-0.002)	-0.004, -0.001 (-0.004, -0.001)

Table 4: ML estimates and estimated posterior characteristics for the parameters of accelerated lognormal distribution (the values in parenthesis correspond to missing data case)

Parameter	ML estimate	Posterior mode	0.95 HPD interval
a'_0	2.900	2.906 (2.932)	2.829, 3.005 (2.824, 3.035)
a'_1	0.096	0.095 (0.087)	-0.037, 0.285 (-0.052, 0.344)
a'_2	0.552	0.540 (0.567)	0.390, 0.781 (0.368, 0.957)
a'_3	-0.002	-0.002 (-0.002)	-2.534e-03, -1.589e-03 (-2.485e-03, -0.001)
b'_0	-1.720	-1.751 (-1.649)	-2.113, -1.406 (-2.088, -1.325)
b'_1	0.283	0.307 (0.245)	-0.065, 0.771 (-0.128, 0.803)
b'_2	0.787	0.768 (0.812)	0.407, 1.235 (0.362, 1.376)
b'_3	0.001	0.001 (0.002)	-0.001, 0.003 (-8.421e-05, 0.004)

The simulated marginal posterior samples corresponding to the model parameters can be further utilized to obtain various survival characteristics such as the survival probability at a specified point of time, hazard rate and mean time to appearance of tumours (MTAT), etc. associated with the mice involved in testing. We have, however, worked out for MTAT only based on the final simulated posterior samples. The MTAT under the accelerated Weibull distribution is $\theta\Gamma\left(\frac{1}{\beta}+1\right)$ where the scale parameter $\theta = \exp(a_0 + \sum_{j=1}^{3} a_j x_j)$ and the shape parameter $\beta = \exp(b_0 + \sum_{j=1}^{3} b_j x_j)$. Similarly, the MTAT under the accelerated lognormal distribution is $\exp\left(\mu + \frac{\sigma^2}{2}\right)$ where $\mu = a'_0 + \sum_{j=1}^{J} a'_j x_j$ and $\sigma = \exp(b'_0 + \sum_{j=1}^{J} b'_j x_j)$. It may be noted here that while writing the expression for MTAT, we have simplified the notations considerably. As a result, the term x_j associated with the model parameters corresponds to the j^{th} covariate, where $j = 1,2,3$. Thus x_1 corresponds to two levels of strain (F_1=0 and F_2=1), x_2 corresponds to sex of mice (0 for females and 1 for males) and x_3 corresponds to four doses of chemical (60 ppm, 120 ppm, 200 ppm and 400 ppm).

The estimated posterior characteristics for MTAT at different levels of three covariates are reported in Table 5 under the assumption of accelerated Weibull distribution. The values in parentheses represent the results corresponding to missing data case. We have also reported the corresponding ML estimates for MTAT in order to have a comparison of our results with the classical ones (see Table 5). It can be seen that there is appreciable difference between the MTAT estimates corresponding to two sexes and the female mice are more susceptible to the chemically developed tumours than the male mice for both the levels of strains and all the four doses of the chemical. The results given in Table 5, however, do not stipulate appreciable difference between the two strains of offspring for both the sexes at all the four doses of chemical. There is yet another important finding that can be seen from Table 5. The mice receiving the higher dose of the chemical are more susceptible to the chemically developed tumour, a conclusion that is absolutely in accordance with dose-response relationship. Besides, we also obtained the estimates for MTAT under the assumption of accelerated lognormal distribution. More or less similar results were observed in this case as well except a few marginal differences in the estimates for male mice at two levels of strains of offspring. We do not report these results presuming that these are not going to offer any additional benefit rather unnecessarily increase the length of paper.

We next focus on the estimates of MTAT in missing data case. We can see that both point and interval estimates are close enough to the corresponding estimates in non-missing data case when the considered distribution is accelerated Weibull (see Table 5). More or less similar observations were marked when the considered distribution was accelerated lognormal except for some of the estimated HPD intervals for male mice. These HPD intervals were found to be wider, in general, than the corresponding HPD intervals for non-missing data case. Obviously, this finding is important in the sense that it provides large variability among the MTAT estimates associated with the male mice in missing data case. In this very sense, the accelerated Weibull distribution can be visualized to be a better candidate than the accelerated lognormal distribution simply because the former distribution offers more or less stable estimates for MTAT almost in every considered situation.

Before we end the section, we compare the two accelerated distributions formally using DIC and EPPL measures discussed in Section 4. The DIC is evaluated on the basis of 4K posterior samples from each of the two posteriors (16) and (17) associated with the accelerated Weibull and the accelerated lognormal distributions, respectively. For the evaluation of EPPL, we generated 4K predictive data sets exactly in the same form and size as the observed data. It may be noted that these 4K predictive samples were obtained with the help of simulated posterior samples corresponding to each of the two considered distributions. DIC and EPPL were similarly evaluated for missing data case as well. The evaluated values of the two measures are reported in Table 6.

Table 5: ML estimates and estimated posterior characteristics for MTAT under the accelerated Weibull distribution (the values in parentheses correspond to missing data case)

Strain	Sex	Dose of chemical (in ppm)	ML estimate	Posterior mode	0.95 HPD interval
F1	F	60	16.122	16.189 (16.237)	15.387, 17.198 (15.305, 17.146)
		120	14.448	14.437 (14.495)	13.968, 15.085 (13.898, 15.031)
		200	12.476	12.443 (12.429)	11.939, 12.919 (11.941, 12.915)
		400	8.632	8.474 (8.545)	7.539, 9.435 (7.490, 9.360)
F1	M	60	28.722	27.742 (27.851)	23.241, 42.981 (23.206, 44.484)
		120	25.752	24.723 (25.005)	20.978, 38.108 (20.415, 38.900)
		200	22.286	21.256 (21.681)	17.824, 32.919 (17.747, 33.388)
		400	15.686	17.036 (15.692)	12.063, 27.602 (12.074, 28.903)
F2	F	60	16.885	16.622 (16.665)	15.499, 20.234 (15.387, 20.385)
		120	15.139	14.997 (14.932)	13.787, 17.981 (13.763, 18.152)
		200	13.085	12.985 (12.854)	11.813, 15.706 (11.655,15.672)
		400	9.077	8.901 (8.953)	7.535, 11.343 (7.365, 11.208)
F2	M	60	30.052	30.757 (31.216)	25.030, 44.570 (24.330, 45.438)
		120	26.958	27.639 (27.597)	22.464, 39.396 (22.039, 40.562)
		200	23.354	23.829 (23.699)	19.539, 34.449 (19.515, 35.736)
		400	16.531	17.563 (16.883)	13.652, 31.258 (13.692, 33.687)

Table 6: DIC and EPPL values under the accelerated Weibull and the accelerated lognormal distributions (the values in parentheses correspond to missing data case)

Distribution	DIC	Under EPPL criterion		
		LDF	LDC	EPPL
Accelerated Weibull	599.943 (600.012)	38.066 (38.009)	28.991 (29.110)	67.057 (67.119)
Accelerated lognormal	600.276 (611.988)	38.315 (49.368)	28.306 (29.987)	66.621 (79.355)

As regards the results, it can be seen that the values of DIC and those of EPPL under the two distributions are almost close to each other and, therefore, one may consider either of the two distributions for the considered dataset. It may, however, be seen that the accelerated Weibull distribution performs better than the accelerated lognormal distribution even in case some of the observations are missing. So we prefer to conclude in favour of the accelerated Weibull distribution although it offers slight poor loss due to complexity and hence the overall loss for non-missing data case. The difference is, however, marginal only when compared to the corresponding values obtained under the accelerated lognormal distribution.

6 Conclusion

The paper is a successful attempt of analyzing current status data when exact lifetimes are not observable rather the information is available only in the form of failure status or surviving. The other novel feature of the paper is the use of accelerated lifetime models, namely the two-parameter Weibull and the two-parameter lognormal distributions when both the parameters are allowed to be affected by the covariates or the stress variables occurring in the experimentation. Normally, in such situations only the scale parameter of the model is allowed to be varied in accordance with the covariates and the other shape parameter is kept constant. Of course, the resulting distributions are complex but not a deterrent issue when allowed to be dealt by sample based approaches to Bayesian computation. The results on model comparison considered in the paper finally recommend the accelerated Weibull model when both missing and non-missing datasets are allowed to be entertained. If, however, there is no missing data in the experimentation, one can consider either of the two models for drawing the necessary inferences.

References

[1] M. I. Araújo and B. d. B Pereira. A comparison of Bayes factors for separated models: some simulation results. *Communications in Statistics- Simulation and Computation*, 36(2):297–309, 2007.

[2] N. Balakrishnan and M. H. Ling. EM algorithm for one-shot device testing under the exponential distribution. *Computational Statistics and Data Analysis*, 56(3):502–509, 2012.

[3] N. Balakrishnan and M. H. Ling. Expectation maximization algorithm for one shot device accelerated life testing with Weibull lifetimes, and variable parameters over stress. *IEEE Transactions on Reliability*, 62(2):537–551, 2013.

[4] N. Balakrishnan and M. H. Ling. Gamma lifetimes and one-shot device testing analysis. *Reliability Engineering & System Safety*, 126:54–64, 2014.

[5] N. Bousquet, H. Bertholon, and G. Celeux. An alternative competing risk model to the Weibull distribution for modelling aging in lifetime data analysis. *Lifetime Data Analysis*, 12(4):481, 2006.

[6] C. E. Buck and S. K. Sahu. Bayesian models for relative archaeological chronology building. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 49(4):423–440, 2000.

[7] R. Dumonceaux, C. E. Antle, and G. Haas. Likelihood ratio test for discrimination between two models with unknown location and scale parameters. *Technometrics*, 15(1):19–27, 1973.

[8] T.-H. Fan, N. Balakrishnan, and C.-C. Chang. The Bayesian approach for highly reliable electro-explosive devices using one-shot device testing. *Journal of Statistical Computation and Simulation*, 79(9):1143–1154, 2009.

[9] J. S. Kim and B.-J. Yum. Selection between Weibull and lognormal distributions: a comparative simulation study. *Computational Statistics & Data Analysis*, 53(2):477–485, 2008.

[10] D. H. Kirn, W. D. Lee, and S. G. Kang. Bayesian model selection for life time data under type II censoring. *Communications in Statistics- Theory and Methods*, 29(12):2865–2878, 2000.

[11] R. L. Kodell and C. J. Nelson. An illness-death model for the study of the carcinogenic process using survival/sacrifice data. *Biometrics*, 36(2):267–277, 1980.

[12] J. F. Lawless. *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc., Hoboken, New Jersey., 2002.

[13] N. R. Mann, N. D. Singpurwalla, and R. E. Schafer. *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley & Sons, Inc., New York, 1974.

[14] W. Q. Meeker and L. A. Escobar. *Statistical Methods for Reliability Data*. John Wiley & Sons, 1998.

[15] W. Q. Meeker. A comparison of accelerated life test plans for Weibull and lognormal distributions and type I censoring. *Technometrics*, 26(2):157–171, 1984.

[16] C. A. Meeter and W. Q. Meeker. Optimum accelerated life tests wth a nonconstant scale parameter. *Technometrics*, 36(1):71–83, 1994.

[17] W. Nelson. *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey, 1990.

[18] C. Robert and G. Casella. *Introducing Monte Carlo Methods with R*. Springer Science & Business Media, 2010.

[19] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Science & Business Media, 2013.

[20] R. Sen. *Bayesian Analysis of Some Non-Regular Families in Accelerated Life Testing*. PhD thesis, Department of Statistics, Banaras Hindu University, Varanasi, India, 2016.

[21] R. Sharma and S. K. Upadhyay. A hierarchical Bayes analysis for one-shot device testing experiment under the assumption of exponentiality. *Communications in Statistics- Simulation and Computation*, 47(5):1297–1314, 2018a.

[22] R. Sharma and S. K. Upadhyay. A hierarchical Bayes analysis for one-shot device accelerated life test under the assumption of Weibull and exponential models. *Submitted for Publication*, 2018b.

[23] N. D. Singpurwalla. *Reliability and Risk: a Bayesian Perspective*. John Wiley & Sons, 2006.

[24] A. F. M. Smith and G. O. Roberts. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(1):3–23, 1993.

[25] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. v. d. Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.

[26] S. K. Upadhyay. Common failure distributions. *Wiley Encyclopedia of Operations Research and Management Science*, 2010.

[27] S. K. Upadhyay, A. Gupta, B. Mukherjee, and A. Khokhar. A Baye comparison of Weibull extension and modified Weibull models for data showing bathtub hazard rate. *Journal of Statistical Computation and Simulation*, 83(1):68–81, 2013.

[28] S. K. Upadhyay, N. Vasishta, and A. F. M. Smith. Bayes inference in life testing and reliability via Markov chain Monte Carlo simulation. *Sankhyā: The Indian Journal of Statistics, Series A*, 63(1):15–40, 2001.

[29] S. K. Upadhyay and M. Peshwani. Choice between Weibull and lognormal models: a simulation based Bayesian study. *Communications in Statistics-Theory and Methods*, 32(2):381–405, 2003.

[30] W. Wang. *Fitting the Weibull and Lognormal Log-linear Models to Accelerated Life Test Data*. PhD thesis, The University of Arizona, Tucson, Arizona, 1999.

Continuous Multistate System Universal Generating Function

V. M. Chacko
•
Department of Statistics
St.Thomas College (Autonomous), Thrissur
Kerala-680001, India
chackovm@gmail.com

Abstract

Usually, systems and components are described as being in one of two modes, "on" or "off." Such systems are described using binary structure functions. In multistate systems (MSS), components can be in more than two states—for example, there can be partially failed or partially operating modes. The system state can be described by continuously many values. A system that can have different task performance levels is named multi-state system (MSS). In this paper, we present a technique for solving a family of Continuous MSS reliability problems. A universal generating function (UGF) method is proposed for fast reliability estimation of continuous MSSs. The UGF method provides the ability to estimate relatively quickly different MSS reliability indices for series-parallel, parallel-series and bridge structures. It can be applied to MSS with different physical nature of system performance measure.

Keywords: multi-state system, universal generating function, reliability

I. Introduction

Many technical systems are distinguished by their structural complexity. They can perform their task at several different levels like the system failure can lead to decreased ability to perform the given task without moving to complete failure. Such a system element can also perform its task with some different levels in between perfect functioning to complete failure. Such systems are named as Multistate Systems (MSS). For example, the generating unit in power systems has its nominal generating capacity, which is fully available if there are no failures. Some types of failures can cause complete unit outage while other types of failures can cause a unit to work with reduced capacity. For example, in a power generation system, the generator can produce 100MW, later on, due to some technical problem, it may produce 80MW and so on. The physical characteristics of the performance depend on physical nature of the system outcome. One need to choose reliability procedures according to the physical performance of system such as productivity, capacity etc. Therefore, it is important to measure performance rates of system components by their contribution into the entire MSS output performance. Continuous materials or energy transmission systems or oil transportation systems, power generation systems etc are examples of MSS. Billinton and Allan (1996), Aven (1990) discussed the flow network problem, which provide the desired throughput or transmission capacity for continuous energy, material or information flow. There may be components in series or parallel in the system. In redundancy optimization problems, data processing speed can also be considered as a performance measure and the main task of system is to complete the task within the desired time, see Levitin et.al (1998) and Lisnianski et.al (2000). Several type of MSS were considered in Gnedenko and Ushakov (1995). Several properties of MSSs and importance measures in MSSs are considered in Chacko and Manoharan (2008, 2011).

A rigorous work on Binary state system can be seen in Barlow and Proschan (1975). Many standard works on reliability theory adopt a framework in which systems and components can be in only one of two modes 'on' or 'off'. Consequently, the system structure function is a binary function of binary variables. Most of the standard results in Barlow and Proschan (1975) are set in this framework. These structure functions fail to model important situations when systems have redundant standby components. Furthermore, if the components or systems can be in intermediate modes besides the two extremes of completely functioning or completely failed, the above framework does not suffice. To remedy this situation, authors such as Barlow and Wu (1978), El. Neweihi et al. (1978), and Griffith (1980) have considered situations in which components and systems can assume a finite number of values. In these works, the basic concepts of MSS reliability were formulated and the system structure function was defined for coherent MSS. The research on such systems – called 'multi-state systems' is still continuing. The aim of this research work is to advance the state-of-the art of the highly promising multi-state reliability theory so that it can be applied to design and maintenance of practical engineering systems. In Griffith (1980), the coherence definition was generalized and three types of coherence were studied. The reliability importance was extended to MSS from the binary state system in Butler (1979). Concepts of MSS importance are also discussed in Block and Savits (1982).

The steady state behavior of Markov systems is very useful in reliability analysis. The steady-state behavior of multi-state monotone systems was considered by applying the theory for stationary and synchronous processes with embedded point process in Natvig and Streler (1984). The modeling technique was suggested by Wood (1985), which allows existing binary algorithms for block diagrams and fault trees to be applied to multi-state system. The concept of equivalent behavior was introduced in Garriba et.al (1980) which provide the analysis of multiple-valued logic tree aimed at eliciting prime implicants. These prime implicants are the multiple-valued logic analogue of minimal cut sets encountered in binary fault trees. The prime implicants were also successfully used in dependability analysis of software controlled systems Yau (1998). As in the Binary state system reliability analysis availability and unavailability plays a very major role in system maintenance through corrective maintenance of MSSs. A method for the two-sided estimation of MSS unavailability was proposed based on the binary model, Pouret et.al (1999). Large system analysis using extreme value theory is important in the MSS theory. An asymptotic approach to the MSS reliability evaluation was presented in Kolowrocki (2000). Chacko and Manoharan (2009), Chacko et. al. (2018) considered MSSs reliability problems like ageing properties with semi-Markov modeling.

MSS reliability assessment are based on three different approaches Aven (1993): the structure function approach - where Boolean models are extended for the multi-valued case, the stochastic process (mainly Markov) approach, and Monte-Carlo simulation. The structure function approach is also extremely time consuming and difficult to deal with. The stochastic process method can be applied only to relatively small MSS, because the number of system states increases drastically with the increase in number of system components. A Monte-Carlo simulation model may be a fairly true representation of the real world, but the main disadvantage of the simulation technique is the time and expense involved in the development and execution of the model Aven (1993). This is an especially important drawback when the optimization problems are solved. In spite of these limitations, the above mentioned methods are often used by practitioners, for example in the field of power systems reliability analysis Pouret et.al. (1999).

MSSs reliability analysis is more complex in reality. In real-world problems of MSS reliability analysis, the great number of system states that need to be evaluated makes it difficult to use traditional techniques in various optimization problems. The universal generating function (UGF) technique is fast enough to be used in these problems in discrete state MSSs, Ushakov (1986) and Ushakov (1988). In addition, this technique allows practitioners to find the entire MSS performance distribution based on the performance distributions of its components. An engineer can find it by using the same procedures for MSS with different physical nature of performance. In the following sections the application of the UGF to MSS reliability analysis is considered especially for continuum state systems. The discretization of continuous systems make variations in reliability analysis. The results of measure theory and probability theory will become applicable while using continuous performance variables. So it is necessary to study continuous MSSs and introduce analysis tools.

Section II provided the UGF for continuous MSSs. Performance measure evaluation is given in section III. Numerical example is given in section IV. Conclusions are given in final section.

II. Performance Measures of Continuous MSSs

Consider a system consisting of n units. We suppose that any system unit i can have continuous states: from complete failure up to perfect functioning. The entire MSS system has continuous states as determined by the states of its units. Denote a MSS state at instance t as Y(t) $\in [0, b]$, where Y(t)=0 corresponds to the worst state and Y(t)=b corresponds to the best state. The performance level G_y is associated with each state $y \in [0, b]$ and $G_y \geq G_s$ if y>s. The MSS behavior is characterized by its evolution in the space of states. To characterize numerically this evolution process, one has to determine the MSS reliability indices. These indices can be considered as extensions of the corresponding reliability indices for a binary-state system.

The Continuum MSS reliability measures were systematically studied Brunelle and Kapur (1999). In this paper, we consider three measures which are most commonly used by engineers, namely MSS availability, MSS expected performance, and MSS expected unsupplied demand (lost throughput).

MSS availability A(t) is the probability that the MSS will be in the states with performance level greater than or equal to W at a specified moment t>0, where the MSS initial state at the instance t=0 is the best state K or some other predetermined state M (G(y)>W). For large t the initial state has practically no influence on the availability. Therefore, the index A is usually used for the steady state case and is called the stationary availability coefficient, or simply, the MSS availability. MSS availability is the function of required demand W. It may be defined as

$$A(W) = \int_W^\infty f(y)dy \tag{1}$$

Where f(y) is the probability density function of MSS performance state y. The resulting integral is taken only for the states where MSS performance is greater than or equal to the specified demand W.

In practice, the system operation period T is often partitioned into M intervals, T_m ($1 \leq m \leq M$) and each T_m has its own demand level W_m. The following generalization of the availability index as in Levitin et.al. (1998) is used in these cases:

$$E_A = \sum_{m=1}^{M} A(W_m) \cdot q_m,. \tag{2}$$

where

$$q_m = T_m / \sum_{m=1}^{M} T_m \tag{3}$$

is the steady state probability of demand level m.

For example, in power system reliability analysis, the index (1-E_A) is often used and treated as loss of load probability, see Billinton and Alen (1996). This measure is commonly used in power system reliability analysis. The MSS performance in this case is interpreted as power system generating capacity.

The value of MSS expected performance could be determined as

$$EG = \int_o^b G(y)f(y)dy. \tag{4}$$

One can note that expected MSS performance does not depend on demand W. EG defines the average productivity (capacity) or processing speed of the system.

When penalty expenses are proportional to the unsupplied demand, the expected unsupplied demand EU may be used as a measure of system output performance. This index may be presented by the following expression:

$$EU = \sum_{m=1}^{M} \int_0^b \max(W_m - G(y) \ \ ,0)\, f(y)dy.\, q_m, \tag{5}$$

Examples of the EU measure are the unsupplied power in power distribution systems and expected output tardiness in information processing systems. In this case EU may be interpreted as expected electric power unsupplied to consumers.

In the following section we consider MSS reliability assessment based on MSS reliability indices based on the UGF technique.

III. Universal Generating Function of Continuous MSSs

The UGF was introduced in Ushakov (1986) and principles of its application were formulated in Ushakov (1987) and Ushakov (1988). The most systematical description of mathematical aspects of the method can be found in Ushakov (2000), where the method is referred to as generalized generating sequences approach. A brief overview of the method with respect to its applications for MSS reliability assessment can be seen in Levitin et.al (1998). Chacko and Manoharan (2011) discussed application of UGF in finding joint importance measures of MSSs. The method was first applied to the real power system reliability assessment and optimization in Lisnianski et.al (1994,1996)

For MSS which continuous states, there can be different levels of output performance at each time t: G(t)∈**G**={G} and the system output performance distribution (OPD) can be defined by two sets **G** and f(g(t)) .

The u-function of a continuous random variable Y is defined as

$$u(z) = \int_a^b z^y f(y)dy, \tag{6}$$

where the variable G lies between a and b and f(g) is the is the probability density function of G. u-function $u_j(z)$ can be introduced to represent the performance distribution of element j by relating the probabilities of each state y_j, $0 \le y \le b_j$, to the corresponding performance G_{y_j} of the element in that state:

$$u_j(z) = \int_0^{b_j} z^{y_j} f(y_j)dy_j.$$

To obtain the u-function of a subsystem containing two elements, composition operators are introduced. All the composition operators take the form

$$u_i(z) * u_j(z) = \int_0^{b_i} z^{y_i} f_i(y_i) dy_i * \int_0^{b_j} z^{y_j} f_j(y_j) dy_j = \int_0^{b_i}\int_0^{b_j} z^{\omega(y_i, y_j)} f_i(y_i) f_j(y_j) dy_i dy_j$$

The definition of the function $\omega(\cdot)$ strictly depends on the physical nature of the system performance measure and on the nature of the interaction of the system elements, for example for a series system, $\omega(.) = \min(.,.)$, and for a parallel system, $\omega(.) = sum(.,.)$ or $max(.,.)$. Because, the total performance of a pair of elements connected in parallel is equal to the sum of the performance rates (e.g. productivity and capacity) of the individual and when several elements are connected in series, the element with the lowest performance becomes the bottleneck of the subsystem: in other words, the performance of the subsystem is equal to the minimum of the performances of the individual.

Consecutively applying the operators to all elements and replacing pairs of macro-elements by equivalent elements one can obtain the u-function representing the performance distribution of the entire MSS. Obtain the u-functions of all of the system elements. If the system contains a pair of elements connected in parallel or in series, replace this pair with an equivalent macro-element with u-function obtained by 'sum' or 'min' operator for $\omega(\cdot)$. If the system contains more than one element, do it again and again. Then, determine the u-function of the entire series-parallel system as the u-function of the remaining single equivalent macro-element. The system probability and performance density functions f(.), g are represented by the coefficients and exponents of this u-function.

In order to use the UGF in evaluation in various reliability measures, we consider the following approach. Let g_{jy_j} be the output performance of multistate system when element j is in state y_j while the rest of the elements evolve stochastically among their corresponding states with performance distributions. $f_i(y_i)$, $0 \le y_i \le b_i, 1 \le i \le n$. Assume that the element j is in one of its states y_j with performance not greater than α. We denote by $y_{j\alpha}$ the state in the ordered set of states of element j whose performance $g_{jy_{j\alpha}}$ is equal or immediately below α, i. e., $g_{jy_{j\alpha}} \le \alpha \le g_{jy_{j\alpha+}}$. The conditional probability of the element j being in a generic state k characterized by a performance $g_{jy_{j\alpha}}$ not greater than a pre-specified level threshold α is

$$f_j\left(Y_j = y_j \middle| G_j \le g_{jy_{j\alpha}}\right) = \frac{f_j(Y_j = y_j)}{F(g_{jy_{j\alpha}})}$$

.

Similarly, the conditional probability of the element j being in a generic state k characterized by a performance $g_{jy_{j\alpha}}$ greater than a pre-specified level threshold α is

$$f_j\left(Y_j = y_j \middle| G_j > g_{jy_{j\alpha}}\right) = \frac{f_j(Y_j = y_j)}{1 - F(g_{jy_{j\alpha}})}$$

In this model we get $OPM^{\le\alpha}{}_j$:

$$OPM^{\le\alpha}{}_j = \int_0^{g_{jy_{j\alpha}}} Y_j\ f_j\left(Y_j = y_j \middle| G_j \le g_{jy_{j\alpha}}\right)\ dy_j = \int_0^{g_{jy_{j\alpha}}} Y_j \frac{f_j(Y_j = y_j)}{F(g_{jy_{j\alpha}})} dy_j$$

,

Similarly, we define as $OPM^{>\alpha}{}_j$:

.

$$OPM^{>\alpha}{}_j = \int_{g_{jy_{j\alpha}}}^{b_j} Y_j\ f_j\left(Y_j = y_j \middle| G_j > g_{jy_{j\alpha}}\right)\ dy_j = \int_{g_{jy_{j\alpha}}}^{b_j} Y_j \frac{f_j(Y_j = y_j)}{1 - F(g_{jy_{j\alpha}})} dy_j$$

In order to obtain the state space restricted measures, one has to modify the UGF of elements as follows,

$$U^{\le\alpha}{}_j = \int_0^{g_{jy_{j\alpha}}} z^{y_j}\ f_j\left(Y_j = y_j \middle| G_j \le g_{jy_{j\alpha}}\right)\ dy_j = \int_0^{g_{jy_{j\alpha}}} z^{y_j} \frac{f_j(Y_j = y_j)}{F(g_{jy_{j\alpha}})} dy_j$$

$$U^{>\alpha}{}_j = \int_{g_{jy_{j\alpha}}}^{b_j} z^{y_j}\, f_j\left(Y_j = y_j \middle| G_j > g_{jy_{j\alpha}}\right)\, dy_j = \int_{g_{jy_{j\alpha}}}^{b_j} z^{y_j} \frac{f_j(Y_j = y_j)}{1 - F(g_{jy_{j\alpha}})} dy_j$$

$$U_{j,k}{}^{\leq\alpha,\leq\beta} = \int_0^{g_{jy_{j\alpha}}} z^{y_j} \frac{f_j(Y_j = y_j)}{F(g_{jy_{j\alpha}})} dy_j \int_0^{g_{ky_{k\beta}}} z^{y_k} \frac{f_k(Y_k = y_k)}{F(g_{ky_{k\beta}})} dy_k$$

$$U_{j,k}{}^{>\alpha,>\beta} = \int_{g_{jy_{j\alpha}}}^{b_j} z^{y_j} \frac{f_j(Y_j = y_j)}{1 - F(g_{jy_{j\alpha}})} dy_j \int_{g_{ky_{k\beta}}}^{b_k} z^{y_k} \frac{f_k(Y_k = y_k)}{1 - F(g_{ky_{k\beta}})} dy_k$$

$$U_{j,k}{}^{\leq\alpha,>\beta} = \int_0^{g_{jy_{j\alpha}}} z^{y_j} \frac{f_j(Y_j = y_j)}{F(g_{jy_{j\alpha}})} dy_j \int_{g_{ky_{k\beta}}}^{b_k} z^{y_k} \frac{f_k(Y_k = y_k)}{1 - F(g_{ky_{k\beta}})} dy_k$$

$$U_{j,k}{}^{>\alpha,\leq\beta} = \int_{g_{jy_{j\alpha}}}^{b_j} z^{y_j} \frac{f_j(Y_j = y_j)}{1 - F(g_{jy_{j\alpha}})} dy_j \int_0^{g_{ky_{k\beta}}} z^{y_k} \frac{f_k(Y_k = y_k)}{F(g_{ky_{k\beta}})} dy_k$$

when evaluating UGF of

$$OPM^{\leq\alpha}{}_i, OPM^{>\alpha}{}_i, OPM^{\leq\beta}{}_j, OPM^{>\beta}{}_j, OPM^{\leq\beta,\leq\alpha}{}_{ij}, OPM^{>\beta,>\alpha}{}_{ij}, OPM_{ij}^{\leq\alpha,>\beta}\, and\, OPM_{ij}^{>\alpha,\leq\beta}.$$

Having MSS OPD, one can obtain the system availability for the arbitrary t and W using the following operator δ_A:

$$A(t,W) = \delta_A(U_{MSS}(t,z),W) = \delta_A(\int_0^b Z^y f(y(t)dy, W) = \int_0^b f(y(t))\alpha(y(t) - W)dy \tag{7}$$

where

$$\alpha(x) = \begin{cases} 1, x \geq 0, \\ 0, x < 0. \end{cases} \tag{8}$$

The expected system output performance value during the operating time can be obtained for given $U_{MSS}(z)$ using the following δ_G operator:

$$E_G(t) = \delta_G(U_{MSS}(t,z)) = \delta_G(\int_0^b Z^y f(y(t)dy) = \int_0^b y(t) f(y(t)dy \tag{9}$$

In order to obtain the expected unsupplied demand EU for the given $U_{MSS}(z)$ and constant demand W according to (4), the following δ_U operator should be used:

$$E\ U(t) = \delta_U(U_{MSS}(t,z)) = \delta_U(\int_0^b Z^y f(y(t)dy, W) = \int_0^b \max(0, W - y(t)) f(y(t)dy$$

$$EU(\mathbf{W}) = \sum_{m=1}^{M} q_m \delta_U(U_{MSS}(z), W_m), \tag{10}$$

where

$$EU(W_m) = \delta_U(U_{MSS}(z), W_m) = \delta_U(\int_0^b Z^y f(y(t)dy, W_m) = \int_0^b \max(0, W_m - y(t)) f(y(t))dy$$

Consider, for example, two power system generators with nominal capacity 100 MW as two separate MSS, Billonton (1996). In the first generator some types of failures require the capacity to be reduced to 60 MW and some types lead to the complete generator outage. In the second one some types of failures require the capacity to be reduced to 80 MW, some types lead to capacity reduction to 40 MW and some types lead to the complete generator outage. So, there are three possible relative capacity levels that characterize the performance of the first generator:

Universal Generating function (UGF) is found to be a useful toll in determining the system performance for the MSSs. Real world MSS are often very complex and consist of a large number of elements connected in different ways. To obtain the MSS OPD and the corresponding u-function, for the continuum state MSSs, we must develop some rules to determine the system u-function based on the individual u-function of its elements.

In order to obtain the u-function of a subsystem (component) containing a number of elements, composition operators are introduced. These operators determine the subsystem u-function expressed as integral for a group of elements using simple algebraic operations over individual u-functions of elements. All the composition operators for two different elements take the form

$$\Omega_\omega(u_1(t,z),u_2(t,z)) = \Omega_\omega[\int_0^{b_1} Z^{y_1(t)} f_1(y_1(t))dy_1, \int_0^{b_2} Z^{y_2(t)} f_2(y_2(t))dy_2] = \int_0^{b_1}\int_0^{b_2} Z^{\omega(y_1(t),y_2(t))} f_1(y_1(t)) f_2(y_2(t))du_1du_2 , \tag{11}$$

where $u_1(t,z)$, $u_2(t,z)$ are individual U-function of elements and $\omega(\cdot)$ is a function that is defined according to the physical nature of the MSS performance and the interactions between MSS elements. The function $\omega(\cdot)$ in composition operators expresses the entire performance of a subsystem consisting of different elements in terms of the individual performance of the elements. The definition of the function $\omega(\cdot)$ strictly depends on the type of connection between the elements in the reliability diagram sense, i.e. on the topology of the subsystem structure. It also depends on the physical nature of system performance measure.

For example in MSS, where performance measure is defined as capacity or productivity (MSSc), the total capacity of a pair of elements connected in parallel is equal to the sum of the capacities of elements. Therefore, the function $\omega(\cdot)$ in composition operator takes the form:

$$\omega(g_1,g_2)=g_1+g_2. \tag{12}$$

For a pair of elements connected in series the element with the least capacity becomes the bottleneck of the system. In this case, the function $\omega(\cdot)$ takes the form:

$$\omega(g_1,g_2)=\min(g_1,g_2). \tag{13}$$

In MSS where the performances of elements are characterized by their processing speed (MSSs) and parallel elements cannot share their work, the task is assumed to be completed by the group of parallel elements when it is completed by at least one of elements. The entire group processing speed is defined by the maximum element processing speed:

$$\omega(g_1,g_2)=\max(g_1,g_2). \tag{14}$$

If a system contains two elements connected in series, the total processing time is equal to the sum of processing times t_1 and t_2 of individual elements: $T=t_1+t_2=g^{-1}_1+g^{-1}_2$.

Therefore, the total processing speed of the system can be obtained as $T^{-1}=g_1g_2/(g_1+g_2)$ and the $\omega(\cdot)$ function for a pair of elements is defined as follows:

$$\omega(g_1,g_2)=g_1g_2/(g_1+g_2). \tag{15}$$

Ω operators were determined in Levitin et.al (1998), Lisnianski et.al (2000) for several important types of series-parallel systems MSS. Some additional composition operators were also derived for bridge structures Levitin and Lianianski (1998).

Applying the Ω operators in sequence, one can obtain the u-function representing the system performance distribution for an arbitrary number of elements connected in series, in parallel, or forming bridge structure.

If Y_1 follows Exp(θ) and Y_2 follows Exp(μ), then

$$\Omega_\omega(u_1(t,z),u_2(t,z)) = \int_0^\infty\int_0^\infty Z^{y_1(t)+y_2(t)} \theta\mu e^{-\theta y_1(t)-\mu y_2(t)} dy_1dy_2$$

for parallel structure

$$\Omega_\omega(u_1(t,z),u_2(t,z)) = \int_0^\infty\int_0^\infty Z^{\min(y_1(t),y_2(t))}\theta\mu e^{-\theta y_1(t)-\mu y_2(t)}dy_1dy_2$$

for series structure

$$\Omega_\omega(u_1(t,z),u_2(t,z)) = \int_0^\infty\int_0^\infty e^{[y_1(t)+y_2(t)]\log Z}\theta\mu e^{-\theta y_1(t)-\mu y_2(t)}dy_1dy_2$$

$$\Omega_\omega(u_1(t,z),u_2(t,z)) = \int_0^\infty\int_0^\infty \theta\mu e^{-(\theta-\log Z)y_1(t)-(\mu-\log Z)y_2(t)}dy_1dy_2 = \frac{\theta\mu}{(\theta-\log Z)(\mu-\log Z)}$$

Putting z=1, we get output performance for parallel structure. Similarly for series system.

$$A_{Series}(u_1(t,z),u_2(t,z)) = I_{(z=1,g_1(t)\geq W\ ,g_2(t)\geq W)}\int_0^\infty\int_{0_2}^\infty \theta\mu e^{-(\theta)y_1(t)-(\mu)y_2(t)}dy_1dy_2 = \frac{\theta\mu e^{-\theta y_w\ -\mu y_w}}{(\theta)(\mu)} = e^{-\theta y_w-\mu y_w}$$

$$A_{Parallel}(u_1(t,z),u_2(t,z),W) = 1 - I_{(z=1,g_1(t)+g_2(t)\leq W)}\int_0^\infty\int_0^\infty \theta\mu e^{-(\theta)y_1(t)-(\mu)(y_2(t))}dy_2dy_1$$

$$A_{Parallel}(u_1(t,z),u_2(t,z),W) = 1 - \int_0^\infty\int_0^{W-u_1(t)} \theta\mu e^{-(\theta)u_1(t)-(\mu)(u_2(t))}du_2du_1 = 1 - \int_0^\infty \theta e^{-\theta u_1(t)}(1-e^{-\mu(w-u_2(t))})du_1$$

$$A_{Parallel}(u_1(t,z),u_2(t,z),W) = 1 - \int_0^\infty \theta e^{-\theta u_1(t)}(1-e^{-\mu(w-u_1(t))})du_1 = \frac{\theta}{\theta-\mu}e^{-\mu w}$$

$$A_{Parallel}(u_1(t,z),u_2(t,z),W) = 1 - I_{(z=1,y_1(t)\leq w,y_2(t)\leq W)}\int_0^\infty\int_0^\infty \theta\mu e^{-(\theta)y_1(t)-(\mu)(y_2(t))}dy_2dy_1$$

$$A_{Parallel}(u_1(t,z),u_2(t,z),W) = 1 - \int_0^W\int_0^W \theta\mu e^{-(\theta)y_1(t)-(\mu)(y_2(t))}dy_2dy_1 = 1-(1-e^{-\theta W})(1-e^{-\mu W})$$

$$F(w) = 1 - \frac{\theta}{\theta-\mu}e^{-\mu w}$$

$$\Omega_\omega(u_1(t,z),u_2(t,z)) = \int_0^\infty\int_0^\infty Z^{\min(y_1(t),y_2(t))}\theta\mu e^{-\theta y_1(t)-\mu y_2(t)}dy_1dy_1$$

$$A_{Series}(u_1(t,z),u_2(t,z)) = I_{(z=1,y_1(t)0\geq W,y_2(t)\geq W)}\int_0^\infty\int_0^\infty \theta\mu e^{-(\theta)y_1(t)-(\mu)y_2(t)}du_1du = \frac{\theta\mu e^{-\theta w-\mu w}}{(\theta)(\mu)} = e^{-(\theta+\mu)w}$$

IV. Numerical Example

Time to failure of two components an a system is given in table 1. The availability is estimated if the components are connected in series and in parallel. Both of the data follows exponential distribution, since the failures are due to shocks occurred during operation. The parameters are estimated and availability formula is obtained.

The data is found to be follows Exponential distribution with mean 19.22 and 27.54 respectively.

If the components are connected in series, the availability would be, at w,

$$A_{Series}(u_1(t,z),u_2(t,z)) = e^{-46.76w}$$

If the components are connected in parallel, the availability would be, at w,

$$A_{Parallel}(u_1(t,z),u_2(t,z),W) = 1 - \int_0^W\int_0^W \theta\mu e^{-(\theta)y_1(t)-(\mu)(y_2(t))}dy_2dy_1 = 1-(1-e^{-19.22W})(1-e^{-27.54W})$$

The availability for various w values van be easily obtained.

Table 1: Time to failure of two components (1 and 2)

Time of failure (Component 1)	Time of failure (Component 2)
4.6	15.0
5.6	7.2
6.6	8.5
7.6	9.8
8.6	11.2
9.7	32.0
10.8	14.0
11.9	15.5
13.1	18.0
14.3	18.5
17.0	17.0
16.7	21.8
18.0	53.0
19.4	72.0
20.8	27.0
22.2	30.0
22.0	22.0
25.2	32.7
26.8	35.0
28.4	36.9
31.0	31.0
31.9	41.4
33.7	43.8
36.0	36.0
39.0	39.0

V. Conclusions

The Universal Generating Function for continuous performance distributions are introduced. Discretization of continuous system becomes sometimes more unrealistic inferences. Method for analyzing continuous MSSs is desired. In this paper, we have made an attempt to deal with continuous MSSs, which will guide to obtain performance measures of complex MSSs. More analysis has to be explored in future.

References

[1] Aven T. (1990) Availability evaluation of flow networks with varying throughput-demand and deferred repairs. *IEEE Transactions on Reliability*; **38**: 499-505.

[2] Billinton R, Allan R. *Reliability Evaluation of Power System*s, Plenum Press, 1996

[3] Chacko V. M. and Manoharan M (2008) Joint Importance measures for the Multistate systems,*Advances in Safety and Performance of Complex Systems*, Advances in Safety and

Performance of Complex Systems, A. K. Verma, P. K. Kapur, S. G. Gadge (Eds), Macmillan, India, 308-312, 2008

[4] Chacko V M and Manoharan M (2009) Ageing properties of first passage time distribution of multistate reliability system, *ProbStat Forum, An e-journal in Probability, Statistics and Optimizati*on, Vol.2, 22-30, Jan-2009

[5] Chacko V M and Manoharan M (2011) Joint Importance measures for multistate reliability system, Opsearch (Available online) Vol.48, No.3, 257-278, 2011

[6] Chacko V M, Deepa Paul and Rajitha C (2018) Ageing properties of First passage time of Semi-Markov system, Birth-Death models and Queuing Models, Advances in Mathematical Sciences, Ed. Dr.Ratnakar D B, Dr.Sr.Alphonsa Mathew, Appeared in International Journal of Mathematical Sciences, 2018

[7] Levitin G, Lisnianski A, Ben-Haim H, Elmakis D. (1998) Redundancy optimization for series-parallel multi-state systems. *IEEE Transactions on Reliability*; **47**:(2):165-172.

[8] Lisnianski A, Levitin G, Ben-Haim H. (2000) Structure optimization of multi-state system with time redundancy. *Reliability Engineering and System Safety*; 67: 103-112.

[9] Gnedenko B, Ushakov I.(1995) *Probabilistic Reliability Engineering*. John Wiley & Sons, Inc, NY/Chichester/Brisbane.

[10] Murchland J. (1975) Fundamental concepts and relations for reliability analysis of Multistate systems, Reliability and Fault Tree Analysis, in *Theoretical and Applied Aspects of System Reliability*; SIAM: 581-618.

[11] El-Neveihi E, Proschan F, Setharaman J. (1978) Multistate coherent systems, *J. Applied Probability*; **15**: 675-688.

[12] Barlow R, Wu A. (1978) Coherent systems with multistate components. *Math. Operation Research*; **3**: 275-281.

[13] Ross S. (1979) Multivalued state component systems. *Annals of Probability* ; **7**: 379-383.

[14] Griffith W. (1980) Multistate reliability models. *J. Applied Probability*; **17**: 735-744.

[15] Butler D. (1979) A complete importance ranking for components of binary coherent systems with extensions to multistate systems. *Naval Research Logistics Quarterly*; **26**: 556-578.

[16] Block H, Savits T.(1982) Decomposition of Multistate monotone system. *J. Applied Probability*; **19**: 391-402.

[17] Natvig B, Streller A. (1984) The steady-state behavior of multistate monotone systems. *J. Applied Probabilit*; **21**: 826-835.

[18] Wood A. (1985) Multistate Block Diagrams and Fault Trees. *IEEE Transactions on Reliability,* **34**: 236-240.

[19] Garriba S, Mussio P, Naldi F. (1980) Multivalued logic in the representation of engineering systems, *in Synthesis and analysis methods for safety and reliability studies,* Apostolakis G, Garriba S, Volta G. eds., Plenum Press; 183-197.

[20] Yau M, Apostolakis G, Guarro S. (1998) The use of prime imlicants in dependability analysis of software controlled systems. *Reliability Engineering and System Safety*; **62**: 23-32.

[21] Pouret O, Collet J, Bon J-L. (1999) Evaluation of the unavailability of a multistate-component system using a binary Model. *Reliability Engineering & System Safety*; **66**: 13-17.

[22] Kolowrocki K.(2000), An asymptotic approach to multistate systems reliability evaluation, in Recent advances in reliability theory, methodology, practice and inference, Limnios N, Nikulin M, eds., Birkhauser: 163-180.

[23] Aven T. (1993) On performance measures for multistate monotone systems, *Reliability Engineering and System Safety*: **41**: 259-266.

[24] Painton L, Campbell J. (1995) Genetic algorithm in optimization of system reliability, *IEEE Trans. Reliability*: **44**: 172-178.

[25] Coit D, Smith A (1996) . Reliability optimization of series-parallel systems using genetic algorithm, IEEE Trans. Reliability; **45**: 254-266.

[26] Yang J, Hwang M, Sung T, Jin Y. (1999) Application of genetic algorithm for reliability allocation in nuclear power plant, *Reliability Engineering & System Safety*; **65**: 229-238.

[27] Gen M, Kim J. (1999) GA-based reliability design: state-of-the-art survey, *Computers ind. Engn*; **37**: 151-155.

[28] Gen M, Cheng R. (1997) *Genetic Algorithms and engineering design*, John Wiley & Sons, New York.

[29] Kumar A, Pathak R, Gupta Y, (1995) Genetic algorithm-based reliability optimization for computer network expansion, *IEEE Trans. Reliability*; **44**: 63-72.

[30] Bartlett L, Andrews J. (1999) Efficient basic event ordering schemes for fault tree analysis, *Quality and reliability engineering International*; **15**: 95-101.

[31] Shao S, Murotsu Y. (1999) Approach to failure mode analysis of large structures, Probabilistic engineering Mechanics;**14**: 169-177.

[32] Brunelle R, Kapur K. (1999) Review and classification of reliability measures for multistate and continuum models. *IIE Transactions*; **31**: 1171-1180.

[33] Ushakov I. (1986) A universal generating function. *Soviet J. Comput. Systems Sci.* 1986; **24**: 37-49.

[34] Ushakov I.(1987) Optimal standby problem and a universal generating function. *Soviet J. Comput. Systems Sci.* 7; **25**: 61-73.

[35] Ushakov I. (1988) Reliability analysis of multi-state systems by means of a modified generating function. *J. Inform. Process. Cybernet.*; **34**: 24-29.

[36] Ushakov I. (2000)The method of generalized generating sequences, *European Journal of Operating Research*; **125**: 2: 316-323.

[37] Abouammoh, A. M. and Al-kadi (1991). On measures of importance for components in multi-state coherent system, Microel. Reliab., 31, 109-122.

[38] Abouammoh, A. M. and Al-kadi (1995). Multi-state coherent system of order k, Microel. Reliab., 35, 1415-1421.

[39] Barlow, R. E. and Proschan, F. . Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, New York, 1975

[40] Belzunce, F., Ortega, E. M. and Ruizon, J. M. (2002). Ageing properties of first- passage times of increasing Markov processes, Adv. Appl. Prob., 34, 241-259.

[41] Brown, M. and Chaganty, N. R. (1983). On the first passage time distribution for a class of Markov chains, Ann. Prob., 11, 1000-1008.

[42] Bueno, V. C. (1989). On the importance of components for multi-state monotone systems, Stat. Prob. Lett., 7, 51-59.

[43] Deshpande, J. V., Kochar, S. C. and Singh, H. (1986). Aspects of positive aging, J. Appl. Prob., 23, 748-758.

[44] El-Neweihi, E., Proschan, F., and Sethuraman, J. (1978). Multi-state coherent systems, J. Appl. Prob., 15, 675-688.

[45] Hjort, N. B., Natvig, B. and Funnemark, E. (1985). The association in time of a Markov process with applications to multi-state reliability theory, J. Appl. Prob., 22 473-479.

[46] Lisnianski, A. and Levitin, G. Multi-state System Reliability-Assessment, Optimization and Applications, World Scientific, Singapore (2003).

[47] Meng, F. C. (1993). Component relevancy and characterization results in multi-state systems, IEEE Trans. Reliab., 42, 478-483.

[48] Wu, S. (2005). Joint importance of multi-state systems, Comp. Indust. Eng., 49, 63-75.

[49] Zio, E. and Podofillini, L. (2003). Importance measures of multi-state components in multi-state systems, In. J. Reliab. Qual. Saft. Eng., 10, 289-310.

Mission-Based System Reliability Modeling for Establishing Testable Performance Requirements of a Distributed Network Monitoring System

Arthur Fries[1] and Garfield Jones[2]

•

[1]Institute for Defense Analyses, Alexandria, Virginia 22311, USA, afries@ida.org
[2]Morgan State University, Baltimore, Maryland, 21251, USA, Garfield.jones@morgan.edu

Abstract

Mission-based subsystem reliability requirements are derived for a parent distributed network monitoring system operating under circumstances that differ from standard analytical constructs in a number of ways. First, the system comprises a hierarchy of elements of different functionalities individually adhering to distinct operational profiles. Second, some constituent elements only need to perform during relatively small and non-predetermined portions of the overall system mission accomplishment window. Third, failed elements can be restored or replaced in time to enable additional opportunities for satisfying mission needs.

Keywords: Distributed Network Monitoring System, Subsystem reliability, Operational profile, Mean time between operational mission failures

I. Introduction

A distributed network monitoring system (DNMS) is to be integrated into the current architecture of an existing computer network supporting operations across an extensively dispersed organization. The DNMS will provide the capability to regularly check and report on the security posture of the devices on the parent network. A challenge is to establish credible performance requirements for the constituent elements of the DNMS – to aid design and implementation planning, and to enable reliability demonstration analyses that can accommodate historical DNMS element reliability data as well as dedicated DNMS test results at both the element and system levels. To that end, this paper formulates tractable analytical models that plausibly represent anticipated DNMS operational and maintenance profiles, which vary by DNMS element type, and link DNMS mission performance specifications, as prescribed by organization management, to reliability requirements for the individual classes of DNMS elements.

This setting deviates from standard calculations of system reliability requirements in three fundamental ways. First, the DNMS comprises a hierarchy of subordinate elements of different functionalities individually adhering to distinct mission-based operational profiles. Second, some constituent elements only need to perform during relatively small and non-predetermined portions of the overall DNMS mission accomplishment window. Third, a failed element can be restored or replaced in time to enable additional chances for satisfying DNMS mission needs, with the number of opportunities depending on sub-system and prevailing network support processes.

While much of the operational functionality of a DNMS element is software centric, the composition of DNMS elements includes both dedicated hardware and software whose configurations, both in numbers and design, contribute to the reliability ascribed to particular ensembles. For example, higher quality parts and/or redundancy of supporting integral equipment and operational processes can be built in to enhance system reliability. Accordingly, it is appropriate to pursue more traditional reliability formulations [1, 2] vice focusing on software engineering perspectives [3].

Section II sketches the general structure of a DNMS. Section III describes associated operational and maintenance profiles and translates them into tractable reliability modeling approaches. The discussions presented in Section IV elaborate on the analytical constructs and outline potential follow-on and related reliability analyses.

II. Architecture

The notional DNMS depicted in Figure 1 is composed conceptually of four constituent element types:

1. Individual automated sensors that scan network hardware and software objects for specified defects. Different types of sensors search for distinctive classes of network defects. For each type of sensor, multiple copies are needed to scan the entire network in a reasonably time-efficient manner.
2. A data interface and integration layer that standardizes, processes, and transmits information collected by the automated sensors to base-level dashboards.
3. Base dashboards that process local network scanning data and display aggregated statistics to attendant network security monitors and administrators.
4. A master dashboard that encapsulates summaries from lower level dashboards and enables top-level organization management to track the security posture across the entire network. A back-up master dashboard, operating in a warm standby mode, provides redundancy.

Note that the execution steps essential to DNMS performance are mutually independent across the four layers. Further, within any given layer there are no dependencies among individual elements.

The domain for a single base dashboard encompasses a natural subdivision of the network, e.g., a particular division, component, agency, sub-organization, or geographical location. In addition to receiving data from subordinate dashboards, the master dashboard supports communications down to lower level dashboards and the associated staff. The redundancy provided by the back-up master dashboard enhances organization leadership's access to DNMS information at any critical time point. Dashboards cannot continuously provide real-time status reports for the whole organization, as that would necessitate constant sensor scanning across the entire network. Some acceptable data latency period (e.g., less than a nominal number of prescribed business days) is tolerable and is reflected in DNMS operating profiles.

From the user viewpoint, the new DNMS, while adding modestly to the day-to-day operational mission workload of the parent organization, enhances existing network security processes. In particular, dashboard displays illuminate categories of detected network security defects and characterize their incidence and distribution across the network. These promote the development of mitigation strategies and prioritized implementations, both at the overall and localized network levels.

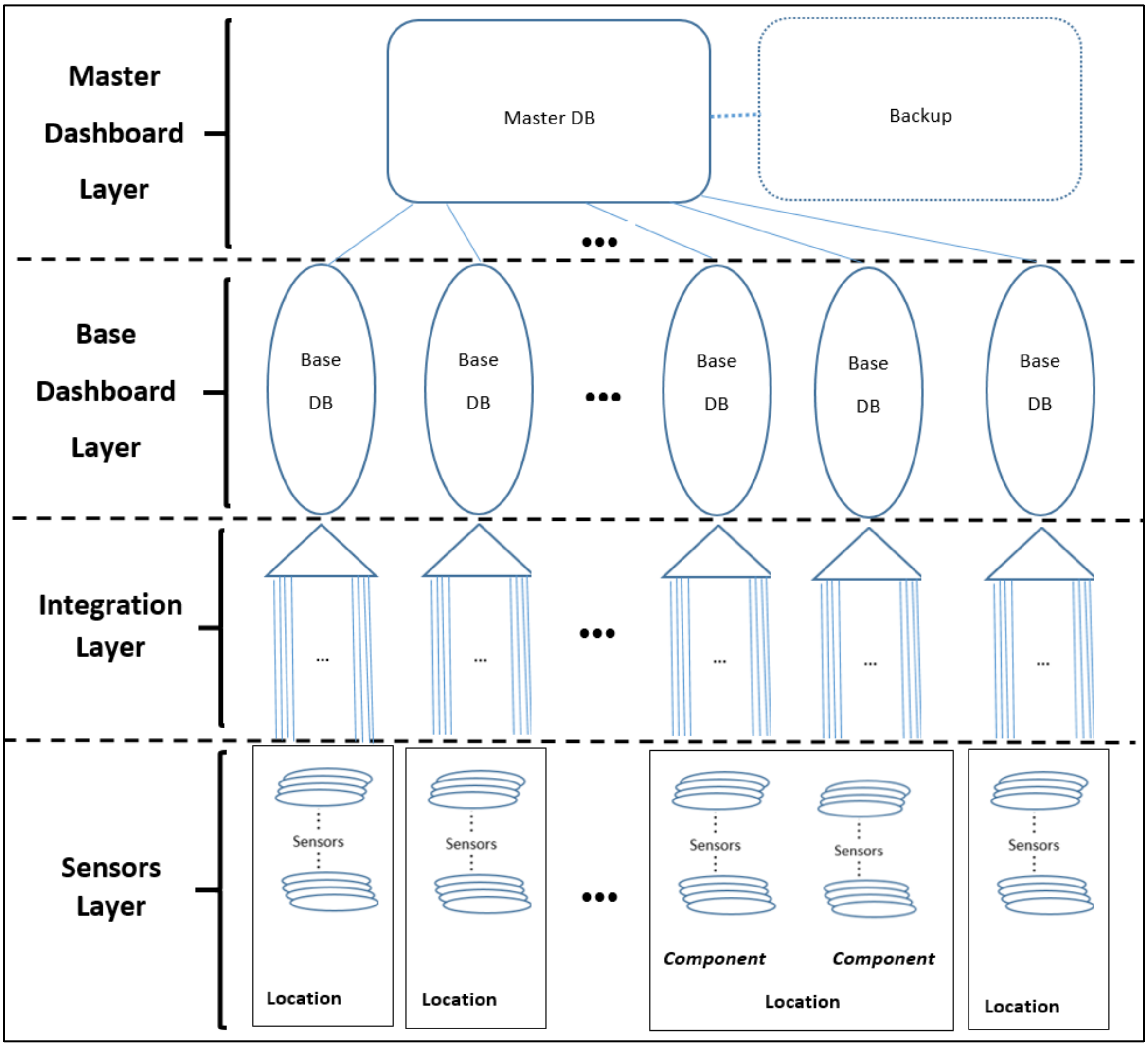


Figure 1: *Notional DNMS structure*

III. Modeling

The modeling approaches presented here are simplistic, favoring analytical tractability and ease of exposition. (Section IV offers additional discussions.) For any real world application, one could of course incorporate the specifics that characterize the subject network.

The operational mission of a DNMS is to systematically monitor and regularly report on the cybersecurity health of the organization's network. This entails regularly scanning the network, updating detailed data on detected defects, and summarizing results in dashboard displays. All of these activities are to be accomplished every d days within the backdrop of ongoing organization business activities. The organization-level requirement is that with some prescribed high probability, P, the DNMS will successfully complete each of its fundamental mission functions within any d-day operational window. For large networks, nominal values of d might be as long as 5-7 days or as short as 2-3 days. To support DNMS design decompositions and prospects for reliability inference based on disparate information sources, the mission success probability P is parsed into subsystem and component element performance requirements. Begin by writing $P = P_1 P_2 P_3 P_4$, where each distinct P_i, $i = 1,2,3,4$, is the probability that the i-th level of the DNMS (as defined in Section II and portrayed

In Figure 1) will successfully execute each of its mission essential functions within an operational performance window. Imposing equal apportionment translates to setting $P_1 = P_2 = P_3 = P_4$ and obtaining $P_i = P^{1/4}$, $i = 1,2,3,4$. This simplification enjoys the practical advantage of framing subsequent calculations in terms of a single parameter to be provided by the organization's management. The subsequent derivations examine each DNMS level separately, consider relevant operational and maintenance profiles, and obtain associated performance probability and element reliability requirements. These are translated to specifications of mean time between operational mission failures (MTBOMF) for individual DNMS element types, i.e., reliability requirements, a format that is more amenable for classes of subsystem design and testing analyses.

III.i. Sensors

A single sensor is tasked to scan a designated portion of the entire network sometime within each d-day performance window, record the collected data, and disseminate to the integration layer. These primary functions must be accomplished in time to allowing adequate opportunity for the integration layer and dashboards to complete their related data processing within that same d-day span. Accordingly, it is plausible to assume that the sensor undertakes an initial execution attempt relatively early within the operational performance window, the probability of it being successful in its initial foray is p_1, and that success entails operating without fault for t_1 consecutive hours (small compared to $24d$ hours). If the initial attempt is successful, no further sensor operation is required until the next d-day performance window arrives. If unsuccessful, a failure event, or lack of a success event, is automatically registered and diagnostic steps are initiated to determine the failure cause and restore or replace the sensor to an as good as new state. This conceptually includes the possibility of temporarily reassigning another sensor to complete the subject sensor's original obligations. The subsequent attempt likewise may be successful or fail, with the same probabilistic characteristics. It is assumed that the value of t_1 and the capabilities of the DNMS-specific logistical support processes, including consideration of availability and restoration times, could enable up to a_1 attempts for the sensor to complete a suitable execution within the desired timeframe. Nominal values of a_1 may be in the vicinity of $d/2$.

The probability that the sensor successfully completes its operational mission takes the form

$$1 - (1 - p_1)^{a_1} = 1 - \left(1 - e^{-t_1/\theta_1}\right)^{a_1}, \tag{1}$$

upon imposing a standard exponential time to failure distribution and set the associated MTBOMF value equal to θ_1. This result holds for a single sensor. The DNMS, however, comprises different sensor types and varying counts for each. Say the total number of sensors is n_1 (which could be hundreds for a large network). Treating their behaviors as being identical and independent, the probability that the entire sensor layer successfully executes a d-day operational mission is

$$\left[1 - \left(1 - e^{-t_1/\theta_1}\right)^{a_1}\right]^{n_1}. \tag{2}$$

Equating this to the prescribed DNMS mission performance requirement of $P_1 = P^{1/4}$, one deduces the associated MTBOMF value, i.e., the reliability requirement for a DNMS sensor element:

$$\theta_1 = -t_1/ln\left[1 - \left(1 - P^{1/4n_1}\right)^{1/a_1}\right]. \tag{3}$$

(Some adhere to the common definition that reliability is the probability that an item will perform its intended function for a specified time interval under stated conditions [4] – which, in this paper's setting, aligns more closely to the formulation $p_1 = e^{-t_1/\theta_1}$.) For the representative set of values t_1 = 4 hours, P = 0.999, n_1 = 500, and a_1 = 2, (3) yields θ_1 = 5654 hours.

III.ii. Integration Layer

Three different concepts for how a DNMS integration layer may be structured are considered here. Each description is accompanied by its own derivation of the associated MTBOMF requirement ascribed to a single resident element.

One conceivable structure of an integration layer, Design I, connects each sensor via a directed pathway to its assigned base dashboard. There are $n_2 = n_1$ such conduits, each determining an individual integration layer element, and the parameter definitions and logic underlying the development of (3) transfer straightforwardly to obtain

$$\theta_{2,\mathrm{I}} = -t_2/ln\left[1-\left(1-P^{(1/4n_2)}\right)^{1/a_2}\right]. \quad (4)$$

Considering the values $t_2 = 4$ hours, $P = 0.999$, $n_2 = 500$, and $a_2 = 3$, which vary from their sensor level counterparts only in that the number of attempts has been increased from 2 to 3, (4) leads to a reduced MTBOMF requirement of $\theta_{2,\mathrm{I}} = 502$ hours.

A variation of the preceding construct incorporates a set of additional elements, data aggregation devices, one interfacing with each unique base dashboard. The data flow corresponding to Design II is sensor ⇨ pathway ⇨ data aggregation device ⇨ base dashboard. The equal apportionment principle allocates a mission success probability of $P_2^{1/2}$ to each class of elements in Design II. For an individual conduit element, the required MTBOMF threshold thus can be read directly from (4):

$$\theta_{2,\mathrm{II(c)}} = -t_2/ln\left[1-\left(1-P^{(1/8n_2)}\right)^{1/a_2}\right]. \quad (5)$$

Retaining the input specifications from the immediately preceding numerical example, insertion into (5) leads to the higher MTBOMF requirement of $\theta_{2,\mathrm{II(c)}} = 633$ hours (consistent with the notion that the P in (4) effectively is increased to $P^{1/2}$ in (5)). For the data aggregation devices, the appropriate count of elements is n_3, the number of base dashboards (which is substantially smaller than n_2). Thus the associated MTBOMF requirement for a single data aggregator is simply

$$\theta_{2,\mathrm{II(a)}} = -t_2/ln\left[1-\left(1-P^{(1/8n_3)}\right)^{1/a_2}\right]. \quad (6)$$

For illustration purposes consider the same set of input parameters as in the two preceding examples, but substantially reduce the number of elements down by two orders of magnitude to $n_3 = 5$. The resultant MTBOMF requirement declines considerably to $\theta_{2,\mathrm{II(a)}} = 135$ hours. The pairing (5) and (6) assume that their values for the functional operational times and numbers of attempts available within the operational performance window are identical to their respective counterparts in (4). If need be, these can be adjusted appropriately.

Design III retains the presence of the data aggregation devices and accompanying assumptions, but excludes the antecedent pathways. Accordingly, the form of (4) holds and revising the relevant count of elements yields

$$\theta_{2,\mathrm{III}} = -t_2/ln\left[1-\left(1-P^{(1/4n_3)}\right)^{1/a_2}\right]. \quad (7)$$

Relative to (6), the power of P has been increased by a factor of two and the value of $\theta_{2,\mathrm{III}}$ will decrease commensurably. For the identical parameterization, the application of (7) gives $\theta_{2,\mathrm{III}} =$ 107 hours.

III.iii. Base Dashboards

For modeling purposes, the layout of base dashboards parallels that of Design III for the integration layer – as there is a one-to-one correspondence between data aggregation devices and base dashboards. Rewriting (7) to allow for possible changes in operations times and allowable number of tries to complete those operations, it follows that

$$\theta_3 = -t_3/ln\left[1-\left(1-P^{(1/4n_3)}\right)^{1/a_3}\right], \quad (8)$$

which differs from (7) merely by the multiplicative factor t_3/t_2. Here the value of t_3 includes the time needed to ingest the data from the integration layer as well as system on-time for displaying data summaries and supporting user needs. Setting t_3 = 10 hours (2.5 times t_2), P = 0.999, n_3 = 5, and a_3 = 3, (8) yields a MTBOMF requirement of θ_3 = 267 hours – an increase of 150 percent compared to the comparable value given for (7).

III.iv. Master Dashboard

The operational profile for the master dashboard includes ingesting summary level data from each of the base dashboards, updating the backup master dashboard with that content, enabling bilateral information flows with the subordinate dashboards, and supporting continuous monitoring of the state of cybersecurity across the entire organization. The associated number of operating hours is t_4 hours per business day, totaling t_4d hours over a d-day performance window. A nominal value for t_4 is 10 hours. If the master dashboard loses some essential functionality, the backup master dashboard will be fully activated to serve as a substitute and maintain operations. Since the backup is running in a warm standby mode, the timing and nature of the manifested failure will determine whether the up-to-date summary data already has been mirrored in the backup, can be transferred from the "failed" master dashboard to the backup, or needs to be ingested anew by the backup.

A pragmatic perspective, consistent with the explicit design choice of a warm standby backup vice a hot standby, would not consider a one to two hour period for users of the master dashboard being deprived of wholly updated summary data as constituting an operational mission failure (OMF). The corresponding likelihood of mission success is the Poisson probability of no more than one failure occurring over the prescribed mission time, and the associated MTBOMF requirement value is the unique solution to the equation

$$P^{1/4} = e^{-(t_4d/\theta_4)}[1+(t_4d/\theta_4)]. \quad (9)$$

The right-hand-side of (9) is the standard formula for hot standby reliability [5], and is appropriate here under the relaxed interpretation of a master dashboard OMF. For the values P = 0.999, t_4 = 10 hours, and d = 5 days, (9) yields θ_1 = 2219 hours.

To support the development of a model representation that accommodates broader definitions of OMFs, the parameter α is introduced to denote the probability that when operationalized the backup master dashboard need not ingest updated summary from the base dashboards. Additionally, a harsher definition of master dashboard success is imposed, demanding no break in the currency of data summary presentations. Under this construct, a simple generalization of (9) follows:

$$P^{1/4} = e^{-(t_4d/\theta_4)}[1+\alpha(t_4d/\theta_4)]. \quad (10)$$

The limiting value α = 1 recovers (9), while the other extreme α = 0 gives no credit whatsoever for redundancy. Setting α = 0.5 and retaining the immediately preceding example inputs, (10) determines a substantially higher MTBOMF requirement of θ_1 = 99,963 hours. Even for α = 0.9, the calculated MTBOMF is 20191 hours, more than nine times the corresponding threshold presented earlier for (9). Clearly the stricter interpretation of a master dashboard OMF establishes considerably

higher reliability requirements and could motivate transitioning to a hot standby design.

III.v. Combined Dashboards Perspective

Under some circumstances, it may be reasonable in reliability calculations to treat the base and master dashboards as being identical. Relevant considerations include commonality of software platforms, software modules, and hardware components, and similarity of failure mode histories. When plausible, paired equations from III.3 and III.4 can be consolidated into a single equation representative of dashboards as a whole – after reapportionment of the DNMS mission success probability. For example, combining (9) with the appropriate transformation of (8) leads to the formulation

$$P^{1/2} = \left[1 - \left(1 - e^{-t_3/\theta_{3,4}}\right)^{a_3}\right]^{n_3} e^{-(t_4 d/\theta_{3,4})}\left[1 + \left(t_4 d/\theta_{3,4}\right)\right], \tag{11}$$

where the new notation $\theta_{3,4}$ denotes the common MTBOMF value ascribed to all of the dashboards. As the right-hand-side is a monotonically increasing function of $\theta_{3,4}$, (11) possesses a unique solution. From the collection of example input values presented earlier, $P = 0.999$, $t_3 = 10$, $a_3 = 3$, $n_3 = 5$, $t_4 = 10$, and $d = 5$, it follows that $\theta_{3,4} = 1567$ hours. This determination of the MTBOMF requirement lies between the two separate MTBOMF requirements calculated previously for (8) and (9), but is considerably closer to the latter, i.e., the influence of the master dashboard dominates. This would be even more so the case if the role of (9) in this example were to be replaced by the more demanding (10).

IV. Discussion

This paper develops tractable models for the reliability of a DNMS architecture comprised of four distinct levels with varying operational and maintenance profiles. They offer informative insights to contractors responsible for proposing, designing, deploying, and supporting the DNMS, seeking to balance design and operational implementation investments against formally prescribed system performance demands or possibly even subject to potential monetary penalties were the deployed DNMS to incur operational performance shortfalls. The straightforward model representations also can be utilized by DNMS host organizations to establish formal reliability demonstration requirements and to guide the development of operational and logistical support processes.

Additionally, both integrators and customers can utilize the framework to assess emerging reliability data from a deployed DNMS and to contemplate specific types of potential design refinements, both architectural and procedural. When conducting dedicated reliability demonstrations or scoring emerging results, care must be taken in defining what constitutes an OMF. For example, minor technical glitches that are nearly immediately remedied via automated or manually induced system reboots may be practically inconsequential. Also, follow-on DNMS integration activities naturally will occur over the deployed lifetime of the DNMS (e.g., coincident with changes to the organization's landscape or the application of routine software upgrades for individual classes of DNMS elements), and these can be expected to engender some initial sets of misconfiguration problems and start-up failures. Whether these should count as OMFs against the DNMS or as a separate category of system failures may depend on the purpose of the immediate reliability analyses and the specifics of any relevant formal requirements contractually imposed on DNMS integrator teams.

An alternative to relying on simple models of the type that constructed in this paper would be to pursue detailed simulation modeling of end-to-end DNMS performance steps over the course of a subject d-day performance window. In addition to incurring requisite time and resource costs, such an approach would be confronted by several analytical challenges. First, definitive operating

profiles cannot be readily discerned. Recall that DNMS is an addition to an organization's existing functionalities and primary operational missions. From that perspective, the daily implementation of DNMS is of secondary importance and there are numerous options, depending on the organization's current operational priorities, of when and how DNMS will be activated and utilized during a particular d-day cycle. Likewise, maintenance events integral to DNMS diagnostic, replace, and restore processes cannot be precisely characterized.

The modeling constructs in this paper account for DNMS employment uncertainties via simplified but plausible representations that embrace DNMS implementation realities. The emphasis is on total operational time for each DNMS element type, vice detailed event-to-event sequencing. Further, the derivations focus on the number of attempts available to an element for completing its assigned operational mission, instead of modeling the detailed specifics of how logistical support processes enable multiple tries to be realized. This is compatible with conventional expressions of operational maintenance requirements (e.g., resolve help desk tickets by the end of the next business day) and can embrace formulations of a spectrum of support responsiveness. To pursue analytical objectives beyond those explicitly considered in this paper, the current model forms could be embedded, as appropriate, into simple simulations tied to coarsely defined events (e.g., operational days or manifested OMFs). For instance, the effects of dynamically evolving support processes readily could be played out over extended operational periods. Other analytical issues that could be addressed by similar methods are discussed below.

One simplifying assumption made consistently herein is that times to failures are governed by memoryless one-parameter exponential distributions. This is a common pragmatic approach for setting reliability requirements [2]. Alternative time-to-failure distributions could be postulated, in which case consideration would need to be given to the impact of repair/restore/replace maintenance events and the interpretation of reliability for planning and assessment purposes. In particular, different classes of recurring events may convert the 'fixed' DNMS element to 'good-as-new' or 'bad-as-old' states [5]. If the former holds universally, then the choice of the distribution is irrelevant as far as the probability of mission accomplishment calculations are concerned. Specific choices for distributions and innate parameters may, however, be of interest for tracking the demonstrated capabilities of deployed systems and projecting future performance.

Throughout the derivations, each d-day performance window implicitly is treated as probabilistically independent and identical. When $d = 5$, corresponding to a standard work week, the weekend days can be expected to offer ample time to recover before the onset of the next window. For values of $d < 5$, the lack of an early mission success in a given operational period may precipitate additional renewal efforts to prepare adequately for the advent of a follow-on performance window. If the assumption of independence cannot be defended plausibly, Markov chain methods [6] may be appropriate.

This paper's modeling framework conceptually could be expanded to incorporate explicit considerations of cost criteria encompassing design, operation, and supportability expenses, especially within the context of financial incentives associated with demonstrated operational performance of the DNMS over time. For example, contractor models relating design costs to DNMS element reliabilities can be utilized to trade off investments against possible performance-based penalties or bonuses. Similarly, the developer may determine that directly funding additional logistical support capabilities may be cost-effective in the long run.

References

[1] National Research Council, Reliability Issues for DOD Systems: Report of a Workshop, The National Academies Press, Washington, DC, 2002. (https://doi.org/10.17226/10561)

[2] National Research Council, *Reliability Growth: Enhancing Defense System Reliability*, The National Academies Press, Washington, DC, 2002. (https://doi.org/10.17226/18987)

[3] National Research Council, *Innovations in Software Engineering for Defense Systems*, The National Academies Press, Washington, DC, 2003. (https://doi.org/10.17226/10561)

[4] U.S. Department of Defense, Test and Evaluation of System Reliability, Availability, and Maintainability: A Primer, Report No. DoD 3235.1-H, Washington, DC, 1982.

[5] Beichelt, F. and Tittman, P., Reliability and Maintenance: Networks and Systems, CRC Press, Boca Raton, 2012.

[6] Privault, N., Understanding Markov Chains: Examples and Applications, Springer, Singapore, 2018.

ISSN 1932-2321