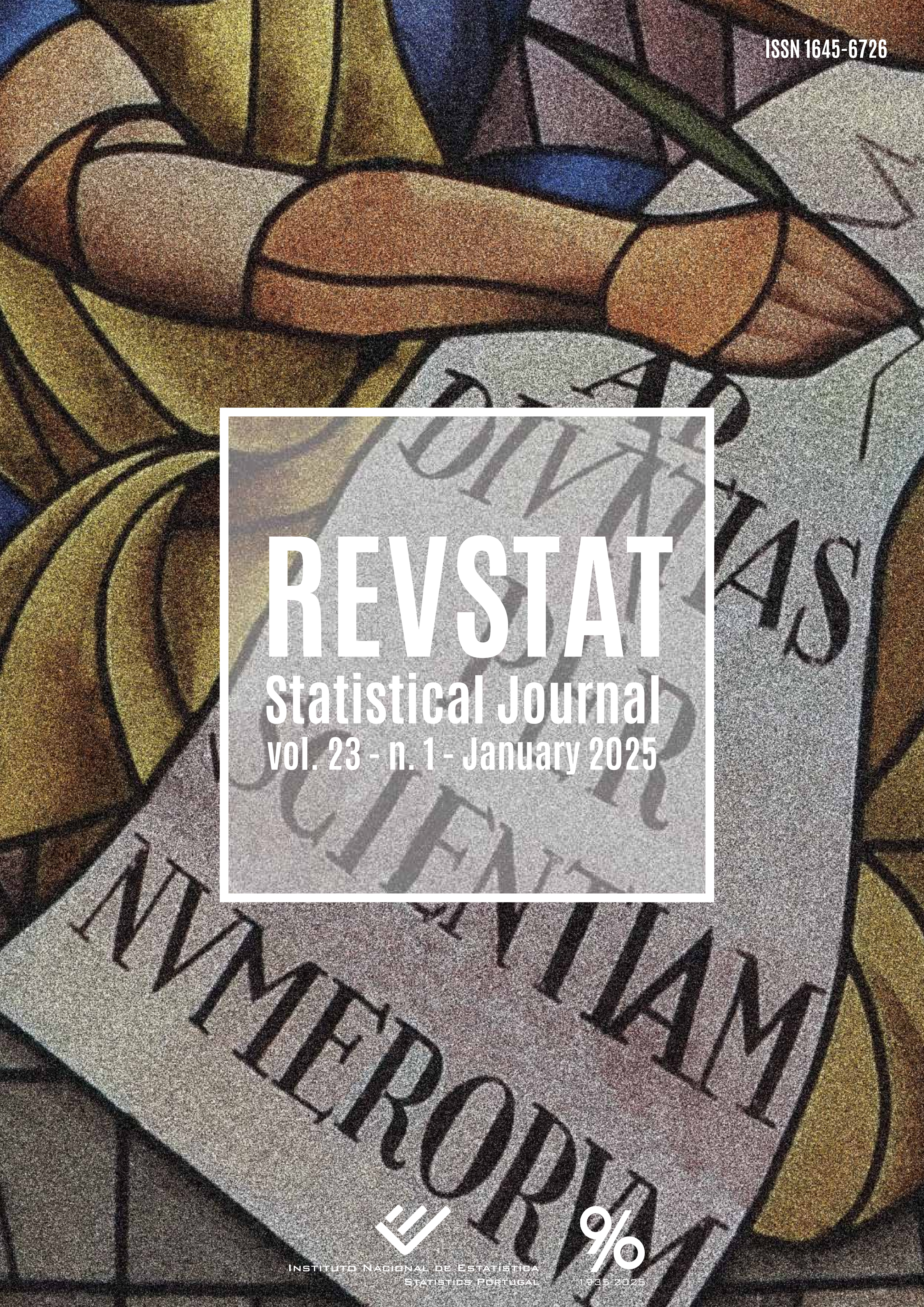


ISSN 1645-6726

REVSTAT
Statistical Journal

vol. 23 - n. 1 - January 2025

Instituto Nacional de Estatística
Statistics Portugal
1935-2025

REVSTAT

Statistical Journal

vol. 23 - n. 1 - January 2025

REVSTAT — Statistical Journal, vol.23, n. 1 (January 2025)

vol.1, 2003-   . - Lisbon : Statistics Portugal, 2003-   .

Continues: Revista de Estatística = ISSN 0873-4275.

ISSN 1645-6726 ; e-ISSN 2183-0371

Publisher – Statistics Portugal

Layout-Graphic Design – *Carlos Perpétuo* | **Cover Design*** – *Helena Nogueira*

Edition - 130 copies | **Legal Deposit Registration** - 191915/03 | **Price** [VAT included] - € 9,00

**image*: stain glass window by Abel Manta (1888-1982)

INDEX

Abstracted/indexed in: *Journal Citation Reports™ – JCR (Clarivate); DOAJ – Directory of Open Access Journals; Current Index to Statistics; Google Scholar; Mathematical Reviews® (MathSciNet®); Zentralblatt für Mathematic; SCImago Journal & Country Rank; Scopus®.*

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 1–17
https://doi.org/10.57805/revstat.v23i1.492

Control Monitoring Schemes for Percentiles of Generalized Exponential Distribution with Hybrid Censoring

Authors: Shovan Chowdhury
– Quantitative Methods and Operations Management Area,
Indian Institute of Management,
Kozhikode, India
shovanc@iimk.ac.in

Amarjit Kundu ✉
– Department of Mathematics, Raiganj University,
West Bengal, India
bapai_k@yahoo.com

Bidhan Modok
– Department of Mathematics, Raiganj University,
West Bengal, India
bidhanmodok95@gmail.com

Received: June 2021 Revised: December 2022 Accepted: December 2022

Abstract:

- In this article, a parametric bootstrap "control monitoring scheme" equivalently known as "control chart", is proposed for process monitoring of percentiles of the generalized exponential distribution for type-I hybrid censored data assuming in-control parameters to be unknown. Monte Carlo simulations are carried out to evaluate the in-control and out-of-control performance of the proposed scheme in terms of average run lengths. Conventional Shewhart-type scheme is also proposed under the same set-up asymptotically and compared with bootstrap scheme using a skewed data set. Finally, an application of the proposed scheme is shown from clinical practice.

Keywords:

- *average run length; control monitoring scheme; false alarm rate; generalized exponential distribution; hybrid censoring; parametric bootstrap.*

AMS Subject Classification:

- 62P30 (Primary), 62F40 (Secondary), 62N01 (Secondary).

✉ Corresponding author

1. INTRODUCTION

Researchers have often shown interest in developing control charts, also known as control monitoring scheme, for monitoring percentiles of an underlying distribution in reliability studies. Padgett and Spurrier (1990) and Nichols and Padgett (2005) argued in favor of monitoring lower percentile of strength distribution over average to confirm the quality of carbon fiber strength to be in control. Chowdhury *et al.* (2021) emphasized on monitoring the upper (lower) percentile of the proportion of non-conforming (conforming) units, as an upward (downward) shift in the upper (lower) percentile indicates a deterioration in quality. All the available control schemes for monitoring distribution quantiles (see for example, Padgett and Spurrier, 1990; Nichols and Padgett, 2005; Erto and Pallotta, 2007; Lio and Park, 2008, 2010; Lio *et al.* (2014); Erto *et al.* (2015); Chiang *et al.*, 2017, 2018; Leiva *et al.*, 2022; and Ma *et al.*, 2022) used parametric bootstrap method with different distributional assumptions under classical or/and Bayesian set-up. Additionally, Chiang *et al.* (2018) used model selection approaches to choose between competing underlying distributions. Due to the non-availability of closed form expressions of the sampling distribution of the percentiles, computational methods such as parametric bootstrap is used to obtain the control limits. For more discussion on the bootstrap technique and its advantages, one can refer to Efron and Tibshirani (1993), Liu and Tang (1996), Jones and Woodall (1998) and Seppala *et al.* (1996).

The results obtained from the aforementioned papers are useful and valuable, and can be applied to complete data setting only. In practice, reliability data are skewed and censored. Recently Vining *et al.* (2016) emphasized on using censored data in reliability studies as customers expected products and processes to perform with high quality over the entire expected lifetime of the product/process. Most of the available schemes for censored data monitor mean of a process. Few papers are found in the literature for monitoring percentiles of a process using censored data. Haghighi *et al.* (2015) proposed control charts for the quantiles of the Weibull distribution for type-II censored data, based on the distribution of a pivotal quantity conditioned on ancillary statistics. Wang *et al.* (2018) proposed EWMA and CUSUM charts for monitoring the lower Weibull percentiles under complete data and type-II censoring using the same approach as used in Haghighi *et al.* (2015). Encouraged by these findings, in this paper, a control monitoring scheme is proposed based on bootstrap method using hybrid censoring which generalizes control monitoring schemes under type-I and type-II censoring.

In type-I censoring scheme, the experiment is aborted after a pre-decided time $T = x_0$; whereas in type-II censoring, the termination is subject to failure of a pre-fixed number of items r. The hybrid censoring scheme which is popularly known in the literature as type-I hybrid censoring scheme was initially introduced by Epstein (1954) and can be considered as a mixture of type-I and type-II censoring schemes. It can be described briefly as follows: Suppose n identical units are put on an experiment. Now if $X_{1:n}, ..., X_{n:n}$ are the ordered lifetimes of the units, then the experiment is aborted either when a pre-chosen number $r(< n)$ out of n items has failed or when a pre-determined time x_0 has elapsed. Hence the life test can be terminated at a random time $X^* = \min\{X_{r:n}, T\}$. One of the following two types of observations can be witnessed under type-I hybrid censoring scheme.

Case I: $\{X_{1:n} < ... < X_{r:n}\}$ if $X_{r:n} < x_0$.

Case II: $\{X_{1:n} < ... < X_{d:n} < T\}$ if $d+1 \leq r < n$ and $x_0 \leq X_{r:n}$.

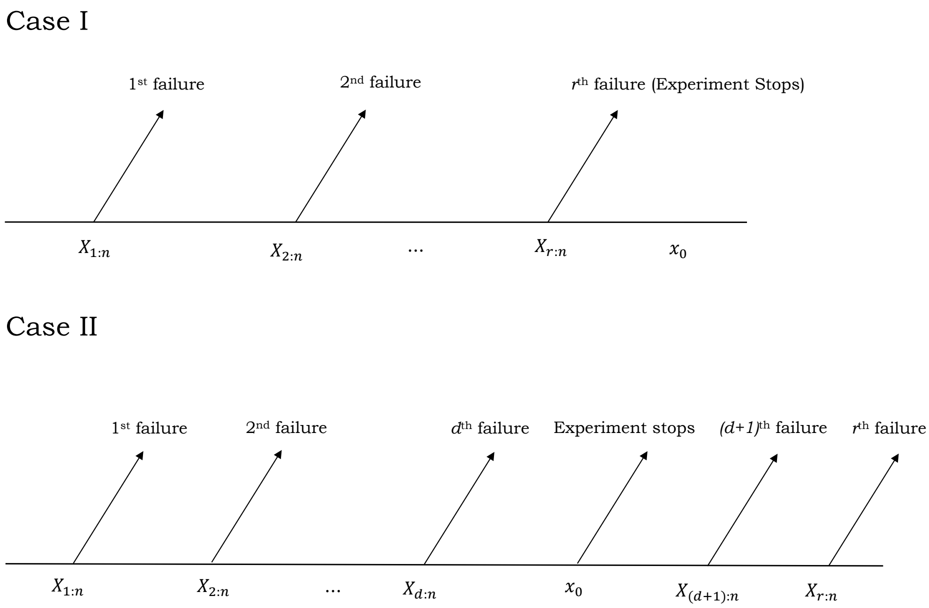


Figure 1: Schematic illustration of type-I hybrid censoring scheme.

In reliability studies, two-parameter Weibull is the most popular distribution to the practitioners. Gupta and Kundu (1999) proposed the two parameter generalized exponential (GE) distribution as an alternative to the Weibull and studied its properties extensively. The scale and shape parameters of the GE distribution bring quite a bit of flexibility in the distribution to analyze any positive real data. Both the Weibull and GE distributions have increasing or decreasing failure rates depending on the shape parameter. Many authors pointed out (see for example, Bain, 1976) that since the hazard function of a GE distribution is bounded above or bounded below as opposed to Weibull which is unbounded, the GE may be more appropriate as a population model when the items in the population are in a regular maintenance environment. The hazard rate may increase initially, but after some times the system reaches a stable condition because of maintenance. Therefore, if it is known that the data are from a regular maintenance environment, it may make more sense to fit the GE distribution over the Weibull. As opposed to Weibull distribution, GE represents a parallel system of independent and identically distributed exponential components. GE has likelihood ratio ordering on the shape parameter indicating the possibility of constructing a uniformly most powerful test for testing a one-sided hypothesis on the shape parameter keeping the scale parameter known. The Weibull distribution doesn't enjoy any such ordering properties and hence no such uniformly most powerful test exists for Weibull. One of the disadvantages of Weibull can be pointed out that the asymptotic convergence to normality for the distribution of the maximum likelihood estimators is very slow (Bain, 1976). Therefore most of the asymptotic inferences may not be very accurate unless the sample size is very large. For a detailed comparison between Weibull and GE, one can refer to Gupta and Kundu (2001). Motivated by these findings, GE is chosen as the underlying distribution to develop a bootstrap control monitoring scheme for hybrid censored data.

The rest of the paper is organized as follows. Section 2 provides the statistical background of the paper. The proposed bootstrap and Shewhart-type control monitoring schemes for GE percentiles with hybrid censored data are introduced in Section 3. Section 4 is devoted to the practical implementation of the schemes including tabulation of the control limits and average run length (ARL). Simulation results of both in-control (IC) and out-of-control (OOC) performance of the bootstrap scheme are presented in Section 4. The effectiveness of the proposed scheme is evaluated in Section 5 using a skewed data set. Bootstrap control monitoring scheme for type-I and type-II censored data are also obtained in Section 5 as a special case and are compared with bootstrap chart under hybrid censoring scheme. An application of the proposed scheme is shown from clinical practice in Section 6. Section 7 concludes the paper.

2. STATISTICAL FRAMEWORK

2.1. Maximum likelihood estimators

Let X be a random variable following two parameter GE distribution with the shape parameter $\theta > 0$ and scale parameter $\lambda > 0$. Then probability density function (pdf) and cumulative distribution function (cdf) of X are given by

$$f(x|\theta,\lambda) = \theta\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\theta-1} \tag{2.1}$$

and

$$F(x|\theta,\lambda) = \Big(1-e^{-\lambda x}\Big)^{\theta}. \tag{2.2}$$

Let ξ_p be the $100p^{\text{th}}$ percentile of the GE distribution and is obtained as

$$\xi_p = -\frac{1}{\lambda}\ln\Big(1-p^{\frac{1}{\theta}}\Big). \tag{2.3}$$

Now, let $x_{i_1}, x_{i_2}, ..., x_{i_n}$ be i^{th} in-control (IC) random subgroup of size n $(i = 1, 2, ..., k)$ drawn from phase I process following GE distribution as in (2.1). On the basis of the observed data and ignoring the additive constant, the log-likelihood function under hybrid censoring (for Case I and II as introduced in Section 1) is given by

$$\begin{aligned} L(\theta,\lambda|data) &= d\ln\theta + d\ln\lambda - \lambda\sum_{i=1}^{d} x_{i:n} + (\theta-1)\sum_{i=1}^{d}\ln(1-e^{-\lambda x_{i:n}}) \\ &+ (n-d)\ln\Big(1-(1-e^{-\lambda c})^{\theta}\Big). \end{aligned} \tag{2.4}$$

Note that for Case I, $d = r$ and $c = x_{r:n}$, and for Case II, $0 \le d \le r-1$ and $c = x_0$. Also it can be shown that for $\lambda \to 0$, and for any fixed θ, maximum likelihood estimators (MLE) of θ and λ do not exist when $d = 0$. Assuming d to be positive, the MLEs $\hat{\theta}$ and $\hat{\lambda}$ are obtained by maximizing the log-likelihood function (2.4), and subsequently solving the non-linear equations

$$\frac{\partial L}{\partial\theta} = 0, \ \frac{\partial L}{\partial\lambda} = 0.$$

As closed-form solutions of these two equations are not available, EM algorithm is used to obtain the MLEs. Let the observed data and the censored data be denoted by $\mathbf{X} = (x_{1:n}, ..., x_{d:n})$ and $\mathbf{Y} = (y_1, ..., y_{n-d})$ respectively. Here for given d, $\mathbf{Y}$ is not observable and hence can be thought of as missing data. The combination of $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ forms the complete data set. Ignoring the additive constant, the log-likelihood function of the uncensored data set, denoted by $L_c(\theta, \lambda|Z)$ is given by

$$\begin{aligned} L_c(\theta, \lambda|data) &= n\ln\theta + n\ln\lambda - \lambda\left(\sum_{i=1}^{d} x_{i:n} + \sum_{i=1}^{n-d} y_i\right) \\ &+ (\theta-1)\left(\sum_{i=1}^{d}\ln\left(1-e^{-\lambda x_{i:n}}\right) + \sum_{i=1}^{n-d}\ln\left(1-e^{-\lambda y_i}\right)\right). \end{aligned} \tag{2.5}$$

Now for 'E'-step of the EM algorithm, one needs to compute the pseudo log-likelihood function as $L_s(\theta, \lambda|data) = E(L_c(Z; \theta, \lambda)|X)$, obtained as

$$\begin{aligned} L_s(\theta, \lambda|data) &= n\ln\theta + n\ln\lambda - \lambda\sum_{i=1}^{d} x_{i:n} + (\theta-1)\sum_{i=1}^{d}\ln\left(1-e^{-\lambda x_{i:n}}\right) \\ &- \lambda\sum_{i=1}^{n-d} E[Y_i|Y_i > c] + (\theta-1)\sum_{i=1}^{n-d} E\left[\ln(1-e^{-\lambda Y_i})|Y_i > c\right]. \end{aligned} \tag{2.6}$$

Now the 'M'-step involves maximization of the pseudo log-likelihood function given in (2.6). Therefore, if at the k^{th} stage the estimate of (θ, λ) is (θ^k, λ^k), then $(\theta^{(k+1)}, \lambda^{(k+1)})$ can be obtained by maximizing

$$\begin{aligned} g(\theta, \lambda) &= n\ln\theta + n\ln\lambda - \lambda\sum_{i=1}^{d} x_{i:n} + (\theta-1)\sum_{i=1}^{d}\ln\left(1-e^{-\lambda x_{i:n}}\right) \\ &- \lambda(n-d)A\left(c, \theta^{(k)}, \lambda^{(k)}\right) + (\theta-1)(n-d)B\left(c, \theta^{(k)}, \lambda^{(k)}\right), \end{aligned} \tag{2.7}$$

where

$$\begin{aligned} A(c, \theta, \lambda) &= -\frac{\theta}{\lambda(1-F(c, \theta, \lambda))}u(\lambda c, \theta), \\ B(c, \theta, \lambda) &= \frac{1}{\theta(1-F(c, \theta, \lambda))}\left[\left(1-e^{-c\lambda}\right)^{\theta}\left(1-\theta\ln\left(1-e^{-c\lambda}\right)\right) - 1\right]. \end{aligned}$$

The maximization of (2.7) can be performed by using similar technique as of Gupta and Kundu (2001). First, $\lambda^{(k+1)}$ can be obtained by solving a fixed point type equation $h(\lambda) = \lambda$, where the function $h(\lambda)$ is defined as

$$h(\lambda) = \left[\frac{1}{n}\sum_{i=1}^{d} x_{i:n} + \frac{n-d}{n}A - \frac{1}{n}\left(\hat{\theta}(\lambda) - 1\right)\sum_{i=1}^{d}\frac{x_{i:n}e^{-\lambda x_{i:n}}}{1-e^{-\lambda x_{i:n}}}\right]^{-1},$$

with $A = A\left(c, \theta^{(k)}, \lambda^{(k)}\right)$, $B = B\left(c, \theta^{(k)}, \lambda^{(k)}\right)$ and $\hat{\theta}(\lambda) = -\frac{n}{\sum_{i=1}^{d}\ln\left(1-e^{-\lambda x_{i:n}}\right) + (n-d)B}$. Once $\lambda^{(k+1)}$ is obtained, $\theta^{(k+1)}$ is obtained by solving the equation $\theta^{(k+1)} = \hat{\theta}\left(\lambda^{(k+1)}\right)$. For more detail on the estimation of GE parameters under hybrid censoring (see Kundu and Pradhan, 2009).

The MLE of the $100p^{\text{th}}$ percentiles, denoted by $\hat{\xi_p}$, is also obtained as

$$\hat{\xi_p} = -\frac{1}{\hat{\lambda}}\ln\left(1 - p^{\frac{1}{\hat{\theta}}}\right). \tag{2.8}$$

2.2. Asymptotic properties

An outline of the Fisher information matrix and asymptotic properties of the estimators are discussed here. For more detail, one may refer to Gupta and Kundu (2001) and Kundu and Pradhan (2009). Using the missing value principle of Louis (1982), it can be written that

$$\text{Observed information} = \text{Complete information} - \text{Missing information}, \tag{2.9}$$

and can be expressed as

$$I_X(\Theta) = I_W(\Theta) - I_{W|X}(\Theta), \tag{2.10}$$

with $\Theta = (\theta, \lambda)$; X = the observed vector; W = the complete data; $I_W(\Theta)$ = the complete information; $I_{W|X}(\Theta)$ = the missing information. The complete information $I_W(\Theta)$ is given by

$$I_W(\Theta) = -E\left[\frac{\partial^2 L_c(W;\Theta)}{\partial\Theta^2}\right],$$

with the Fisher information matrix of the censored observations being written as

$$I_{W|X}(\Theta) = -(n-d)E_{Z|X}\left[\frac{\partial^2 \ln f_Z(z|X,\Theta)}{\partial\Theta^2}\right].$$

The asymptotic variance covariance matrix of $\hat{\Theta}$, the MLE of Θ, can be obtained by inverting $I_X\left(\hat{\Theta}\right)$. The elements of the matrix $I_X(\Theta)$ for the complete data set can be obtained in Kundu and Pradhan (2009).

Let $\hat{\xi}_p\left(\hat{\Theta}_n\right)$ be the value of ξ_p at $\Theta = \hat{\Theta}_n$, obtained from (2.3) and calculated on the basis of n observations. Then as in Chiang *et al.* (2017), it can be shown that $\hat{\xi}_p\left(\hat{\Theta}_n\right)$ follows asymptotic normal distribution with mean $\xi_p(\Theta)$ and variance $\frac{1}{n}\nabla\xi_p^T(\mathbf{\Theta})\mathbf{I_Y}^{-1}(\mathbf{\Theta})\nabla\xi_p(\mathbf{\Theta})$, where $\nabla\xi_p(\Theta)$ is the gradient of $\xi_p(\Theta)$ with respect to Θ. In practice, $\mathbf{I_Y}(\mathbf{\Theta})$ is replaced by the observed Fisher Information matrix $\hat{\mathbf{I}}_Y\left(\hat{\mathbf{\Theta}}_n\right)$, obtained by substituting the unknown parameters θ and λ by their respective MLEs.

3. CONSTRUCTION OF PROPOSED CONTROL MONITORING SCHEMES

3.1. Charting procedure for the bootstrap hybrid-censored control (BHCC) monitoring scheme

Here, the bootstrap hybrid-censored control (BHCC) monitoring scheme for GE percentiles is developed using the following charting procedure.

Step-1: Collect and establish k reference samples $X_m = (x_{i1}, x_{i2}, ..., x_{im})$ of size m each from an IC process (Phase I process) following GE cdf $F(x|\theta,\lambda)$ as in (2.2).

Step-2: Obtain the MLEs of θ and λ from Step-1 under hybrid censoring following the procedure described in Section 2 and estimate the cdf as $F(x|\hat{\theta}, \hat{\lambda})$.

Step-3: Generate a bootstrap sample of size m, $x_1^*, x_2^*, ..., x_m^*$, from $F(x|\hat{\theta}, \hat{\lambda})$ as obtained in Step-2.

Step-4: Obtain the MLEs of θ and λ under hybrid censoring using the bootstrap sample obtained in Step-3, and denote these as θ^* and λ^*.

Step-5: Using (2.3) and (2.8), compute the bootstrap estimate of the $100p^{\text{th}}$ percentile as

$$\hat{\xi}_p^* = -\frac{1}{\lambda^*} \ln\Big(1 - p^{\frac{1}{\theta^*}}\Big). \tag{3.1}$$

Step-6: Repeat Steps 3-5 large number of times (B) to obtain bootstrap estimates of $\hat{\xi}_p^*$, denoted by $\hat{\xi}_{1p}^*, \hat{\xi}_{2p}^*, ..., \hat{\xi}_{Bp}^*$.

Step-7: Using B bootstrap estimates as obtained in Step 6, find the $\frac{\nu}{2}^{\text{th}}$ and $(1-\frac{\nu}{2})^{\text{th}}$ empirical percentiles as the lower control limit (LCL) and upper control limit (UCL) respectively to construct a two-sided BHCC chart, where ν is the false alarm rate (FAR) defined as the probability that an observation is considered as out of control (OOC) when the process is actually IC. Here, empirical sample percentiles are obtained following a method proposed by Hyndman and Fan (1996).

Step-8: Sequentially observe the j^{th} phase II (test) sample $Y_{j:m} = (Y_{j1}, Y_{j2}, ..., Y_{jm})$ of size $m, j = 1, 2, ...$

Step-9: Sequentially obtain $\hat{\xi}_{jp}$ using (3.1) after obtaining MLEs of the parameters under hybrid censoring scheme using the j^{th} test sample as described in Step-5.

Step-10: Plot $\hat{\xi}_{jp}$ against LCL and UCL as obtained in Step-7 of the Phase I process.

Step-11: If $\hat{\xi}_{jp}$ falls in between the LCL and UCL, then the process is assumed to be in-control, otherwise, an OOC signal is activated.

3.2. Charting procedure for the Shewhart-type hybrid-censored control (SHCC) monitoring scheme

Shewhart-type control monitoring scheme for the percentiles of GE distribution, named as SHCC scheme is derived in this section following the asymptotic properties of the MLEs obtained in Section 2.2. The steps for designing the SHCC scheme for $100p^{\text{th}}$ percentile of proportion, $\xi_p(\Theta)$, are described as follows.

In phase I, samples are drawn from in-control process following GE distribution in k independent random subgroups of size m each with $n = m \times k$ being the total sample size.

Step-1: As described in Section 2.1, the MLEs $\hat{\Theta}_n = \left(\hat{\theta}_n, \hat{\lambda}_n\right)$ are computed on the basis of n in-control sample values of Phase I. Then the asymptotic standard error of $\hat{\xi}_{p,m}\left(\hat{\Theta}_m\right)$ is computed as

$$SE_{\xi_{p,m}} = \sqrt{\frac{1}{m}\nabla\xi_p^T\left(\hat{\boldsymbol{\Theta}}_n\right)\mathbf{I}_Y^{-1}\left(\hat{\boldsymbol{\Theta}}_n\right)\nabla\xi_p\left(\hat{\boldsymbol{\Theta}}_n\right)}, \tag{3.2}$$

where $\nabla\xi_p\left(\hat{\boldsymbol{\Theta}}_n\right)$ is the gradient of $\xi_p(\boldsymbol{\Theta})$ at $\boldsymbol{\Theta} = \hat{\boldsymbol{\Theta}}_n$. $\mathbf{I}_Y^{-1}\left(\hat{\boldsymbol{\Theta}}_n\right)$ is calculated following the procedure as described in Section 2.2.

Step-2: The MLEs $\hat{\boldsymbol{\Theta}}_m^j$ of Θ and $\xi_p^j\left(\hat{\boldsymbol{\Theta}}_m^j\right)$ are calculated based on j^{th} $(j = 1, 2, ..., k)$ IC samples of size m each. The sample mean of $\xi_p^j\left(\hat{\boldsymbol{\Theta}}_m^j\right)$s is calculated as

$$\bar{\xi}_p\left(\hat{\boldsymbol{\Theta}}_m\right) = \frac{1}{k}\sum_{i=1}^{k}\xi_p^j\left(\hat{\boldsymbol{\Theta}}_m^j\right).$$

Step-3: The Shewhart-type control monitoring scheme has the center line $CL_{SH} = \bar{\xi}_p\left(\hat{\boldsymbol{\Theta}}_m\right)$. If ν is the false alarm rate (FAR), then for $0 < \nu < 1$, the upper and lower control limits of the SHCC scheme are found to be

$$UCL_{SH} = \bar{\xi}_p\left(\hat{\boldsymbol{\Theta}}_m\right) + z_{(1-\nu/2)}SE_{\xi_{p,m}},$$

and

$$LCL_{SH} = \bar{\xi}_p\left(\hat{\boldsymbol{\Theta}}_m\right) - z_{(1-\nu/2)}SE_{\xi_{p,m}},$$

respectively, where $z_{(1-\nu/2)}$ is the $(1-\nu/2)^{\text{th}}$ quantile of standard normal distribution.

4. SIMULATION STUDY

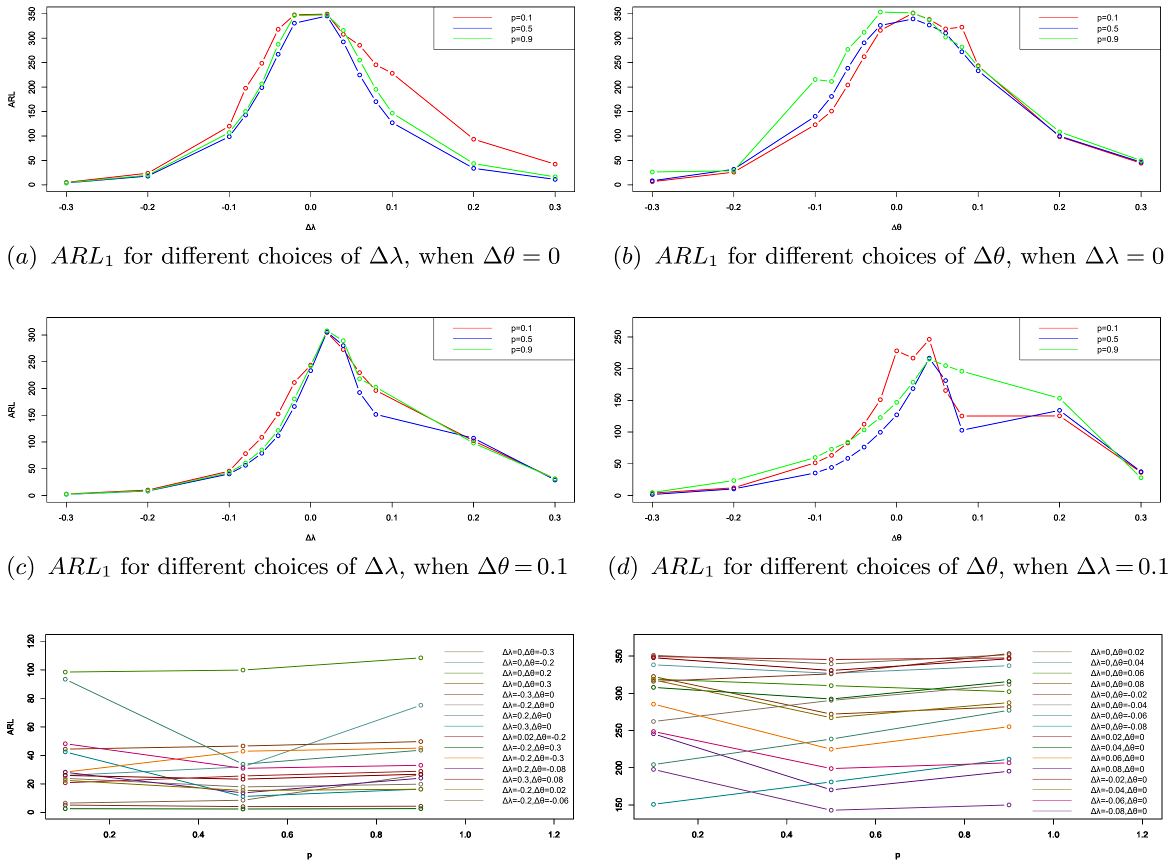
In this section, the IC and OOC performances of the proposed BHCC monitoring scheme are evaluated through a comprehensive simulation study. Numerical computations in R (version 4.0.2) based on Monte-Carlo simulations are used to determine the average UCL and LCL. The MLEs of the parameters θ and λ are obtained for the pair ($\theta = 5.5$, $\lambda = 0.05$). The control limits are obtained based on $B = 5,000$ bootstrap samples. The simulations are carried out with different bootstrap sample sizes m with $k = 20$ subgroups, different percentiles ($p = 0.1, 0.5, 0.9$), different levels of FAR ($\nu = 0.1, 0.005, 0.0027, 0.002, 0.001$) and the following censoring schemes: Scheme $1:$ $m = 25$, $r = 15$, $x_0 = 55$; Scheme $2:$ $m = 25$, $r = 20$, $x_0 = 55$; Scheme $3:$ $m = 40$, $r = 30$, $x_0 = 55$; Scheme $4:$ $m = 40$, $r = 35$, $x_0 = 55$; Scheme $5:$ $m = 25$, $r = 15$, $x_0 = 70$; Scheme $6:$ $m = 25$, $r = 20$, $x_0 = 70$; Scheme $7:$ $m = 40$, $r = 30$, $x_0 = 70$; and Scheme $8:$ $m = 40$, $r = 35$, $x_0 = 70$. The performance of the scheme is assessed by run length, defined as the number of cases required to observe the first OOC signal. For each simulation, the run length is obtained, followed by obtaining the average run length (ARL) and the standard deviation of run length (SDRL) by using $5,000$ simulation runs.

4.1. IC monitoring scheme performance

The estimated IC control limits of the BHCC scheme are displayed in Table 1 of the supplementary article, along with the respective ARL and SDRL as the scheme performance measures, denoted by ARL_0 and $SDRL_0$ respectively. It is easy to show that the reciprocal of FAR is same as the nominal (theoretical) ARL, viz. for $\nu = 0.1, 0.005, 0.0027, 0.002$ and 0.001, the nominal ARL should be equal to $10, 200, 370$, 500 and 1000 respectively. In general, the smaller ARLs indicate narrower control limits, while ARLs larger than 370 specifies wider limits that the bootstrap control schemes give fewer false signals. The simulated values of ARL_0 in Table 1 are found to be closer to the theoretical results implying that the BHCC monitoring scheme for percentiles perform well with skewed data. As the bootstrap sample size (m) increases, the estimated control limits get closer together. Moreover, for fixed m, the control limits become farther apart as the percentile (p) increases. Also, $SDRL_0$ is found to be closer to the ARL_0, satisfying the theoretical result of the geometric distribution used as the run length model.

4.2. OOC monitoring scheme performance

The OOC performance of the BHCC monitoring scheme is investigated by measuring impact of changes in the IC parameter estimates on ARL. In other words, the phase II sample is considered taken from $GE(\theta + \Delta\theta, \lambda + \Delta\lambda)$, while the IC sample comes from $GE(\theta, \lambda)$. The effects of shifts ($\Delta\theta$ and/or $\Delta\lambda$) in the parameters of the GE distribution on ARL of the percentile scheme is examined and exhibited in Table 2 of the supplementary article. In general, the simulation results reveal that for fixed m, r, and x_0, the OOC ARL values (denoted by ARL_1) for the percentiles decrease sharply with both downward and upward small, medium and large shifts in the parameters indicating the effectiveness and usefulness of the scheme. However, the speed of detection varies depending on the type of shifts, the parameters, and the percentile being considered. Except for minor sampling fluctuations, in general, the monitoring scheme detects OOC signal in percentiles faster for downward shifts than the upward shifts (refer Table 2 and Figure 2). In particular, when θ is IC, the ARLs around 50^{th} percentile are smaller than the other percentiles for both upward and downward shifts in λ as is evident from Table 2 and Figure 2(a). For example, for a 4% decrease (increase) in λ when θ is IC ($\Delta\theta = 0$), there is about 27.8% (21%) reduction in the ARL of the 50^{th} percentile. On the other hand, when λ is IC, the ARLs for the lower percentiles (around 10^{th} percentile) is found to be smaller than the other percentiles for downward (upward) shift in θ (refer Table 2 and Figure 2(b)). For example, there is about 44.8% (13.8%) reduction in the ARL of the 10^{th} percentile for a 6% decrease (increase) in θ when λ is IC. From Table 2 and Figure 2(c) it is also clear that for 10% deviation in θ the ARLs around 50^{th} percentile are smaller than the other percentiles for both upward and downward shifts in λ. Again, from Figure 2(d) it can also be observed that, for 10% deviation in λ the ARLs around 50^{th} percentile are smaller than the other percentiles for both upward and downward shifts in θ.



(a) ARL_1 for different choices of $\Delta\lambda$, when $\Delta\theta = 0$

(b) ARL_1 for different choices of $\Delta\theta$, when $\Delta\lambda = 0$

(c) ARL_1 for different choices of $\Delta\lambda$, when $\Delta\theta = 0.1$

(d) ARL_1 for different choices of $\Delta\theta$, when $\Delta\lambda = 0.1$

(e) ARL_1 for different choices of p

(f) ARL_1 for different choices of p

Figure 2: Graphs of ARL_1 for different choices of $\Delta\lambda$, $\Delta\theta$ and p.

5. ILLUSTRATIVE EXAMPLE WITH COMPARISONS

In this section, the BHCC and SHCC monitoring schemes are illustrated by a numerical example which records the waiting times (in minutes) of 100 customers before getting their services (see Ghitany *et al.*, 2008). The BHCC scheme is then compared with bootstrap scheme with Type I and Type II censoring. Various summary measures of the data set can be found below:

Min	5%	10%	25%	50%	75%	90%	95%	Max
0.800	1.895	2.880	4.675	8.100	13.000	19.090	21.955	38.500

First, the Weibull and GE distributions are compared for fitting the data set. For Weibull model, the MLEs of the shape and scale parameters are found to be 1.458 and 10.954 respectively with Kolmogorov-Smirnov test (K-S) statistic value $D = 0.0577$, and p-value, $p = 0.8927$. For GE model, the MLEs are obtained as $\widehat{\theta} = 2.183$ and $\widehat{\lambda} = 0.159$ with $D = 0.0402$ and $p = 0.9970$. The histogram of the data and two fitted densities are provided in Figure 3. The fit results confirm that the GE distribution provides a better fit than Weibull in this case. Moreover, logarithm of the ratio of maximized likelihood (RML), defined as $T = \log L = l_{GE}(\hat{\theta}, \hat{\lambda}) - l_{WE}(\hat{shape}, \hat{scale}) = -317.0884 - (-318.7261) = 1.6377 > 0$ indicates to choose GE distribution over Weibull.

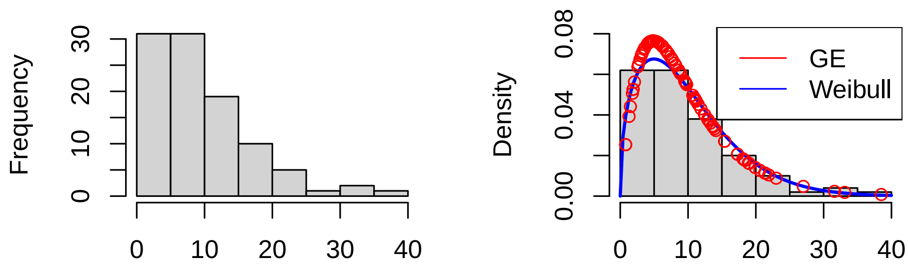


Figure 3: Histogram and density plot of waiting times of 100 patients.

In order to achieve service excellence, the bank may find extreme percentiles of the waiting times worth investigating over the average waiting time. An upward shift in the upper percentile of the waiting times indicates deterioration in the service quality and requires monitoring. In view of this objective, the BHCC chart is constructed for monitoring 90^{th} percentile of the waiting times. The complete data is censored either at the waiting time of the first 60% of the total number of customers ($r = 60$) or at the waiting time of 10 minutes ($x_0 = 10$), whichever occurs earlier. The censoring time $x_0 = 10$ is chosen near to 60^{th} percentile. The complete set of 100 observations is considered as four ($k = 4$) reference samples of size $m = 25$ each. The MLEs of θ and λ under the stated hybrid censoring scheme are obtained as $\widehat{\theta} = 1.760$ and $\widehat{\lambda} = 0.127$ respectively. Using these MLEs, $B = 5,000$ bootstrap samples of size $m = 25$ each are drawn with $r = 15$ (60% of the subgroup size) and $x_0 = 10$. Following the steps $4 - 7$ in subsection 3.1, and using $\nu = 0.0027$ as FAR, the control limits of the BHCC scheme for the 90^{th} percentile are obtained as $UCL = 78.597$, $CL = 26.255$ and $LCL = 12.702$, while the same for the SHCC scheme are computed as $UCL_{SH} = 23.711$, $LCL_{SH} = 19.966$ and $CL_{SH} = 21.839$. It is observed that both schemes provide asymmetric control limits from the respective CL, while the SHCC scheme has narrower interval than the BHCC scheme. Twenty subgroups of size $m = 25$ each are generated from the OOC process under similar hybrid censoring plan having shape parameters $\theta = 2.024$ and $\lambda = 0.108$ (15% increase in θ and 15% decrease in λ).

The OOC performance of the BHCC and SHCC schemes for the 90^{th} percentile are presented in Figure 4 and Figure 5 respectively. The BHCC scheme is able to produce OOC signals quite efficiently with five 90^{th} percentile points falling above the UCL with the first OOC signal being obtained at test sample 2 indicating effectiveness of the scheme in terms of quick detection as well. On the other hand, nine OOC signals are produced by the SHCC scheme with test sample 2 producing the first OOC signal. It is to be noted here that the SHCC scheme grossly underestimates the IC ARL (computation of IC ARL for SHCC scheme is not shown for brevity) due to the narrow band of the control limits which may eventually produce false OOC signals.

Next, bootstrap monitoring scheme is used for type-I (denoted as BT^{I}CC) and type-II (denoted as BT^{II}CC) censored data coming from the GE distribution and their performance is compared with the BHCC monitoring scheme with the same data set and the procedure as used before. The control monitoring schemes for type-I and type-II censored data can be derived as a special case of hybrid censored data for $r = n$ and $T = x_{n:n}$ respectively. The MLEs of θ and λ under type-I censoring with $x_0 = 10$ are obtained as $\widehat{\theta} = 1.803$ and $\widehat{\lambda} = 0.131$ respectively, while the same under type-II censoring with $r = 60$ are found to be

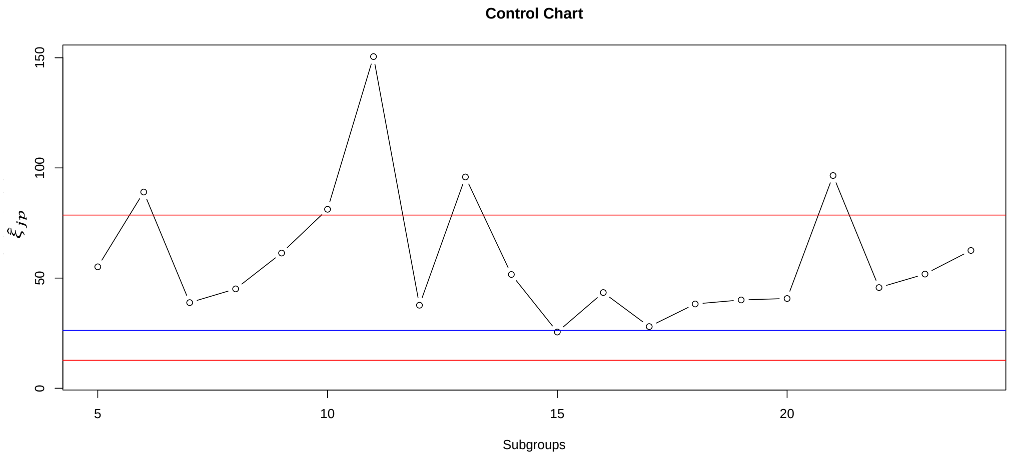


Figure 4: BHCC monitoring scheme for 90^{th} percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 78.597$, $CL = 26.255$, $LCL = 12.702$.

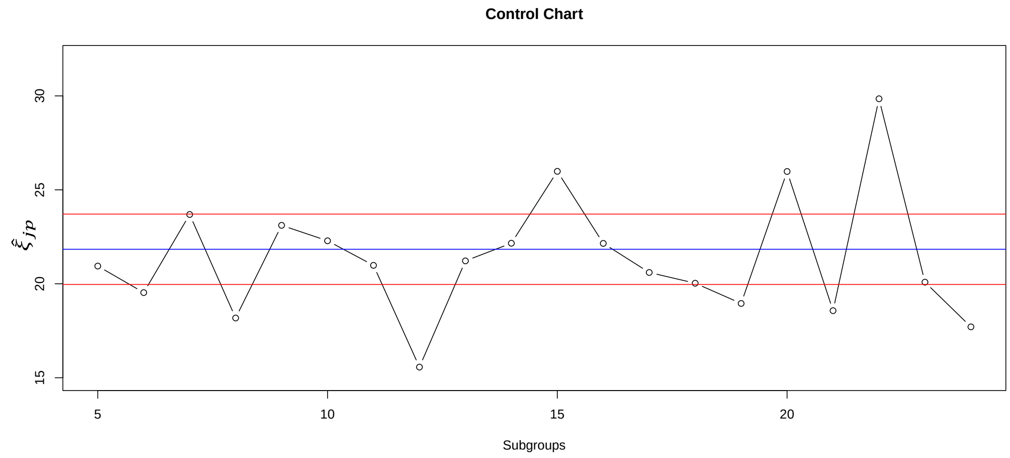


Figure 5: SHCC monitoring scheme for 90^{th} percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL_{SH} = 23.711$, $CL_{SH} = 21.839$, $LCL_{SH} = 19.966$; Example 1.

$\widehat{\theta} = 1.766$ and $\widehat{\lambda} = 0.127$ respectively. While the control limits of BT^ICC scheme for the 90^{th} percentile are obtained as $UCL = 76.812, CL = 15.833$ and $LCL = 12.066$, the same for the BT^{II}CC scheme are calculated as $UCL = 44.144, CL = 24.770$, and $LCL = 12.599$. Both the schemes provide asymmetric control limits with the BT^{II}CC scheme having narrower interval than the BT^ICC scheme. After the first four IC subgroups, twenty subgroups of size $m = 25$ each are generated from the OOC process having $\theta = 2.073$ and $\lambda = 0.111$ (15% increase in θ and 15% decrease in λ). Figure 6 and Figure 7 provide the OOC performance of the control monitoring schemes for the 90^{th} percentile. Figure 6 shows that the type-I censored scheme is able to generate three OOC points falling above the UCL with the first being produced at test sample 12. The type-II censored scheme as is shown in Figure 7 produces two OOC signals just above the UCL with test sample 9 providing the first signal. It is evident from the data analysis that the hybrid censored control monitoring scheme performs better than type-I and type-II censored control monitoring schemes in terms of both frequency and speed of detection of OOC signals.

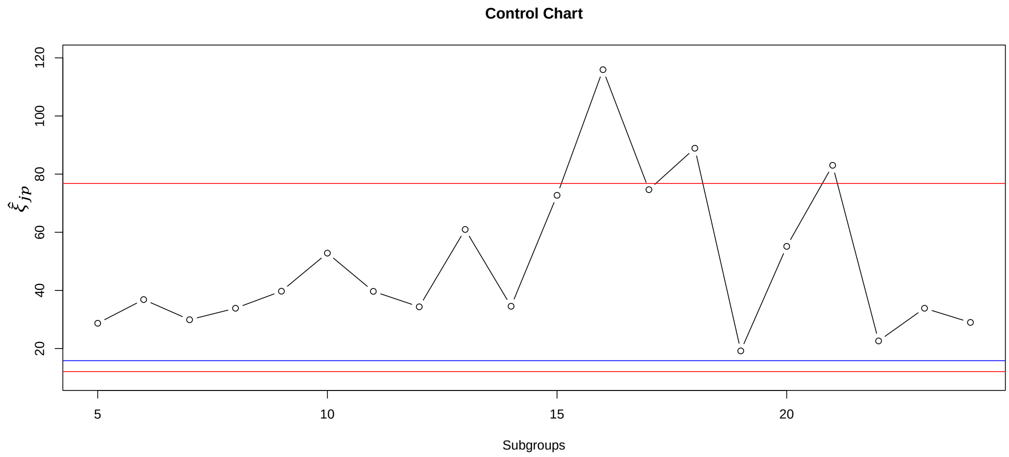


Figure 6: BT^{I}CC monitoring scheme for 90$^{\text{th}}$ percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 76.812$, $CL = 15.833$, $LCL = 12.066$.

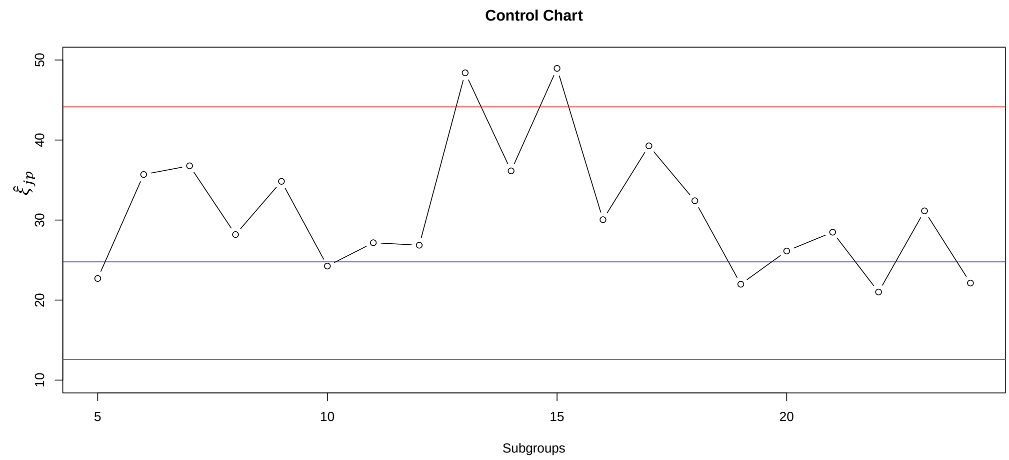


Figure 7: BT^{II}CC monitoring scheme for 90$^{\text{th}}$ percentile of the waiting time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 44.144$, $CL = 24.770$, $LCL = 12.599$.

6. APPLICATION

This section provides an application of the BHCC monitoring scheme in clinical practice. The scheme is used to monitor the top percentile of the survival times of 120 patients (see Hamedani, 2013) with breast cancer obtained from a large hospital in a period from 1929 to 1938. The histogram of the data set as shown in Figure 8 and the summary measures below suggest the skewed nature of the data set.

Min	5%	10%	25%	50%	75%	90%	95%	Max
0.3	6.585	10.110	17.800	40.000	60.000	105.400	125.050	154.0

The MLEs of θ and λ for the complete data set coming from the GE distribution are found to be $\widehat{\theta} = 1.649$ and $\widehat{\lambda} = 0.029$ respectively. The fitted density is provided in Figure 8. The one sample K-S statistic and corresponding p-value are found to be 0.0717 and 0.5681 respectively.

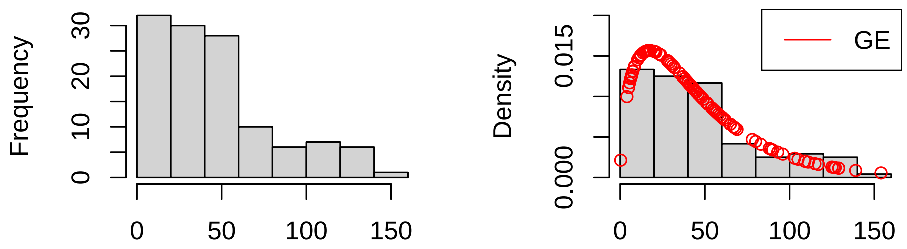


Figure 8: Histogram and density plot of survival times of 120 patients with breast cancer.

The fit results recommend GE distribution to model the survival time data and subsequent development of BHCC scheme. The complete sample data is split into six subgroups of size 20 each. Under hybrid censoring with $r = 72$ and $x_0 = 60$, the estimates of the GE parameters are obtained as $\widehat{\theta} = 1.415$ and $\widehat{\lambda} = 0.024$. Using 5,000 bootstrap samples of size $m = 20$ each with $r = 12$ and $x_0 = 60$, the control limits of the BHCC monitoring scheme for the 90$^{\text{th}}$ percentile are evaluated as $UCL = 322.248, CL = 115.577, LCL = 48.912$. Next, twenty phase II samples of size $m = 20$ each are generated from the process under similar hybrid censoring plan with $\Delta\theta = 0.15$ and $\Delta\lambda = -0.15$ to develop the BHCC monitoring scheme for the 90$^{\text{th}}$ percentile as presented in Figure 9. The scheme has been able to detect OOC signals at 2$^{\text{nd}}$, 5$^{\text{th}}$ and 11$^{\text{th}}$ samples. For the same data set, the control limits for the BT^{I}CC monitoring scheme for 90$^{\text{th}}$ percentile with $x_0 = 60$ are found to be $UCL = 291.965, CL = 106.950, LCL = 51.203$. Figure 10 shows that this scheme has been able to detect only one OOC signal at the 20$^{\text{th}}$ sample. On the other hand, BT^{II}CC monitoring scheme for 90$^{\text{th}}$ percentile with $r = 72$ is presented in Figure 11 with $UCL = 232.296, CL = 115.565, LCL = 48.846$. Figure 11 shows that this scheme also detects only one OOC signal at the 9$^{\text{th}}$ sample. The frequency and speed of detection of OOC signals further justify the use of BHCC monitoring scheme over BT^{I}CC and BT^{II}CC monitoring schemes for the percentiles of survival time in a healthcare set-up.

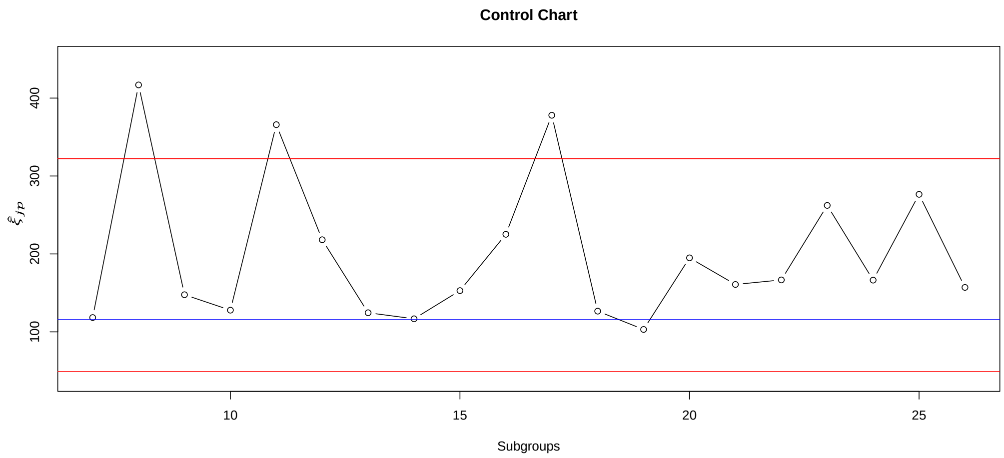


Figure 9: BHCC monitoring scheme for 90$^{\text{th}}$ percentile of the survival time data with $\Delta\theta = 0.15$,$\Delta\lambda = -0.15$, $UCL = 322.248$, $CL = 115.577$, $LCL = 48.912$.

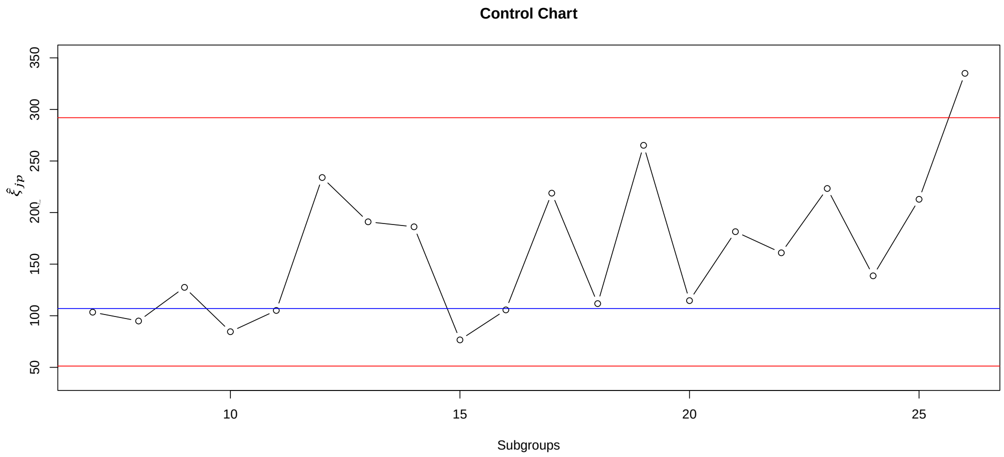


Figure 10: BT^{I}CC monitoring scheme for 90$^{\text{th}}$ percentile of the survival time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 291.965$, $CL = 106.950$, $LCL = 51.203$.

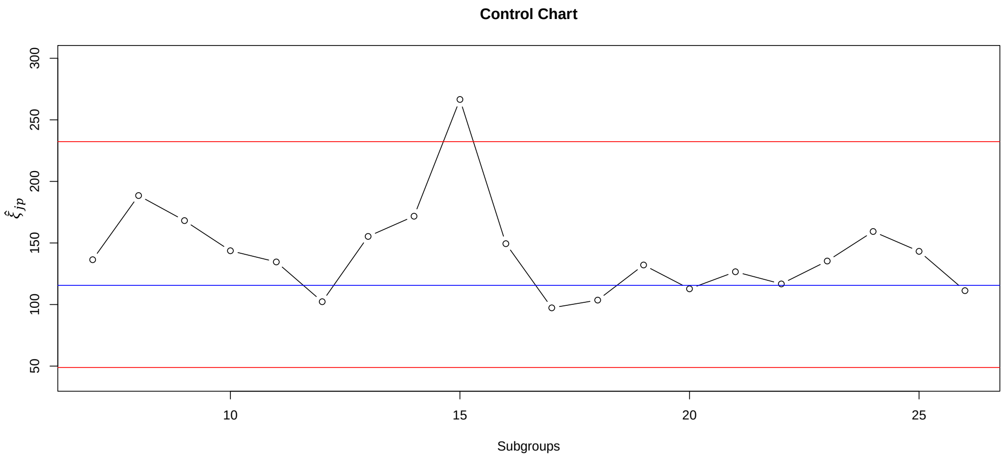


Figure 11: BT^{II}CC monitoring scheme for 90$^{\text{th}}$ percentile of the survival time data with $\Delta\theta = 0.15$, $\Delta\lambda = -0.15$, $UCL = 232.296$, $CL = 115.565$, $LCL = 48.846$.

7. CONCLUDING REMARKS

In this work, hybrid censoring is employed to develop control monitoring schemes for percentiles of GE distribution using bootstrap and asymptotic methods. Bootstrap monitoring schemes for type-I and type-II censored data are also developed under similar set-up as a special case of hybrid censoring plan. An extensive simulation study is conducted to evaluate the IC and OOC performance of the schemes. The hybrid censored schemes are found to be effective in the detection of OOC signals in terms of both magnitude and speed as demonstrated by a real data set. One application from healthcare is also provided to establish the effectiveness of the schemes. In this sense, the present work is the first attempt to apply a new censoring scheme in the process control and generalizes available control monitoring schemes for the GE data. As a scope for future research, hybrid censored schemes may be proposed under Bayesian set-up measuring uncertainty in the parameter(s). One can also think of using progressive censoring scheme to construct such control mechanism. For highly reliable products, accelerated life testing scheme may be employed under various censoring plans for the same purpose.

DATA AVAILABILITY STATEMENT

The data sets used in this manuscript are available in Ghitany *et al.* (2008) and Hamedani (2013).

ACKNOWLEDGMENTS

The authors are thankful to the reviewer, the associate editor and the editor-in-chief for their constructive comments and suggestions which have improved the earlier version of this manuscript.

REFERENCES

Bain, L. J. (1976). *Statistical Analysis of Reliability and Life Testing Model*. Marceland Dekker Inc., New York.

Chiang, J. Y., Jiang, N., Brown, T. N., Tsai, T. R., and Lio, Y. L. (2017). Control charts for generalized exponential distribution percentiles. *Communications in Statistics – Simulation and Computation*, 46(10):7827–7843.

Chiang, J. Y., Lio, Y. L., Ng, H. K. T., Tsai, T. R., and Li, T. (2018). Robust bootstrap control charts for percentiles based on model selection approaches. *Computers & Industrial Engineering*, 123:119–133.

Chowdhury, S., Kundu, A., and Modok, B. (2021). Bootstrap beta control chart for monitoring proportion data. *International Journal of Quality & Reliability Management*, `https://www.emerald.com/insight/content/doi/10.1108/IJQRM-09-2020-0287/full/html`.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.

Epstein, B. (1954). Truncated life tests in the exponential case. *The Annals of Mathematical Statistics*, 25(3):555–564.

Erto, P. and Pallotta, G. (2007). A new control chart for Weibull technological processes. *Quality Technology & Quantitative Management*, 4(4):553–567.

Erto, P., Pallota, G., and Mastrangelo, C. M. (2015). A semi-empirical Bayesian chart to monitor Weibull percentiles. *Scandinavian Journal of Statistics*, 42(3):701–712.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4):493–506.

Gupta, R. D. and Kundu, D. (1999). Generalized exponential distribution. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family; an alternative to gamma and Weibull. *Biometrical Journal*, 43:117–130.

Haghighi, F., Pascual, F., and Castagliola, P. (2015). Conditional control charts for Weibull quantiles under type-II censoring. *Quality and Reliability Engineering International*, 31(8):1649–1664.

Hamedani, G. (2013). The Zografos-Balakrishnan log-logistic distribution: Properties and applications. *Journal of Statistical Theory and Applications*, 12(3):225–244.

Hyndman, R. J. and Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician*, 50:361–365.

Jones, L. A. and Woodall, W. Y. (1998). The performance of bootstrap control charts. *Journal of Quality Technology*, 30:362–375.

Kundu, D. and Pradhan, B. (2009). Estimating the parameters of the generalized exponential distribution in presence of hybrid censoring. *Communications in Statistics – Theory and Methods*, 38(12):2030–2041.

Leiva, V., Santos, R. A. D., Saulo, H., Marchant, C., and Lio, Y. (2022). Bootstrap control charts for quantiles based on log?symmetric distributions with applications to the monitoring of reliability data. *Quality and Reliability Engineering International*, `https://doi.org/10.1002/qre.3072`.

Lio, Y. L. and Park, C. (2008). A bootstrap control chart for Birnbaum-Saunders percentiles. *Quality and Reliability Engineering International*, 24:585–600.

Lio, Y. L. and Park, C. (2010). A bootstrap control chart for inverse Gaussian percentiles. *Journal of Statistical Computation and Simulation*, 80:287–299.

Lio, Y. L., Tsai, T. R., Aslam, M., and Jiang, N. (2014). Control charts for monitoring Burr type-X percentiles. *Communications in Statistics – Simulation and Computation*, 43:761–776.

Liu, R. Y. and Tang, J. (1996). Control charts for dependent and independent measurements based on the bootstrap. *Journal of the American Statistical Association*, 91:1694–1700.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, B*, 44(2):226–233.

Ma, Z., Park, C., and Wang, M. (2022). A robust bootstrap control chart for the log-logistic percentiles. *Journal of Statistical Theory and Practice*, 16(1):1–23.

Nichols, M. D. and Padgett, W. J. (2005). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, 22:141–151.

Padgett, W. J. and Spurrier, J. D. (1990). Shewhart-type charts for percentiles of strength distributions. *Journal of Quality Technology*, 22:283–288.

Seppala, T., Moskowitz, H., Plante, R., and Tang, J. (1996). Statistical process control via the subgroup bootstrap. *Journal of Quality Technology*, 27:139–153.

Vining, G., Kuahci, M., and Pedersen, S. (2016). Recent advances and future directions for quality engineering. *Quality and Reliability Engineering International*, 32(3):863–875.

Wang, F. K., Bizuneh, B., and Cheng, X. B. (2018). New control charts for monitoring the Weibull percentiles under complete data and Type-II censoring. *Quality and Reliability Engineering International*, 34(3):403–416.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 19–45
https://doi.org/10.57805/revstat.v23i1.453

Statistical Inferences to the Parameter and Reliability Characteristics of Gamma-mixed Rayleigh Distribution under Progressively Censored Data with Application

Authors: Kousik Maiti ✉
– Department of Computer Science and Engineering,
Indian Institute of Technology Roorkee,
Roorkee-247667, India
kousikulu@gmail.com

Suchandan Kayal
– Department of Mathematics, National Institute of Technology Rourkela,
Rourkela-769008, India
suchandan.kayal@gmail.com , kayals@nitrkl.ac.in

Received: April 2022 Revised: December 2022 Accepted: December 2022

Abstract:

- We consider estimation of the model parameters and the reliability characteristics of a gamma-mixed Rayleigh distribution based on progressively type-II censored sample (PT-IICS). The sufficient condition for existence and uniqueness of the maximum likelihood estimates (MLE) is obtained. We compute MLEs using the expectation maximization (EM) algorithm. Asymptotic confidence intervals are constructed. Confidence intervals using the bootstrap-p and bootstrap-t methods are constructed. Bayes estimates are derived. Highest posterior density (HPD) credible intervals are derived using the importance sampling method. Prediction estimates and associated prediction equal-tail intervals under one-sample and two-sample frameworks are obtained. A simulation study is conducted. Finally, a real dataset is considered and analyzed.

Keywords:

- *EM algorithm; observed Fisher information matrix; Bayes estimates; Bayesian prediction estimates; HPD credible interval.*

AMS Subject Classification:

- 62F10, 62F15, 62F40, 62N01.

✉ Corresponding author

1. INTRODUCTION

In recent years, life testing experiments are less preferred because of being time consuming and expensive. In many situations, use of complete sample is neither possible nor desirable. In such cases, the sample needs to be censored. Censoring is a condition in which the value of observation is partially known and incomplete. Among different types of censoring schemes, the two basic censoring schemes are type-I and type-II. In the type-I censoring scheme, the life testing experiment terminates at a pre-specified time T, whereas, the type-II censoring scheme terminates when one has m number of failures. For applications and importance of these schemes, we refer to Lawless (2011) and Cohen (2016). The main drawback of these censoring schemes is that they do not allow removal of the items in between other than the termination point. To overcome such drawback, a more general censoring scheme, known as the progressive censoring was introduced in the literature. It can be classified into progressive type-I and progressive type-II censoring schemes. In the progressive type-I censoring scheme, let the number of items used in a life testing experiment be n. In this scheme, $R_1, R_2, ..., R_m$ items are randomly withdrawn at pre-specified time points $T_1, T_2, ..., T_m$, respectively. The test will be terminated at prefixed time point T_m in this scheme. Now, we describe the PT-IICS. Consider n number of total units at initial time on an experiment. We remove randomly R_1 number of survival units when first failure time $X_{1:m:n}$ is observed. This process continues till the m-th failure occurs. We assume that the m-th failure takes place at time $X_{m:m:n}$ and the remaining number of surviving units is $R_m = n - (m + \sum_{i=1}^{m-1} R_i)$. Henceforth, we denote $R = (R_1, R_2, ..., R_m)$ and $X = (X_{1:m:n}, X_{2:m:n}, ..., X_{m:m:n})$ for the censoring scheme and the PT-IICS, respectively. Due to several applications, various inferential procedures based on PT-IICS have been established for many lifetime distributions. For instance, see Muhammed and Almetwally (2020), Nik *et al.* (2021), Albalawi *et al.* (2022) and the references contained therein.

A random variable X is said to follow a gamma-mixed Rayleigh distribution if its probability density and cumulative distribution functions are respectively given by $(\alpha, \beta > 0)$

$$f_X(x; \alpha, \beta) = \frac{\alpha\beta^\alpha x}{(x^2+\beta^2)^{(\alpha/2)+1}} \quad \text{and} \quad F_X(x;\alpha,\beta) = 1 - \frac{\beta^\alpha}{(x^2+\beta^2)^{\alpha/2}}, \tag{1.1}$$

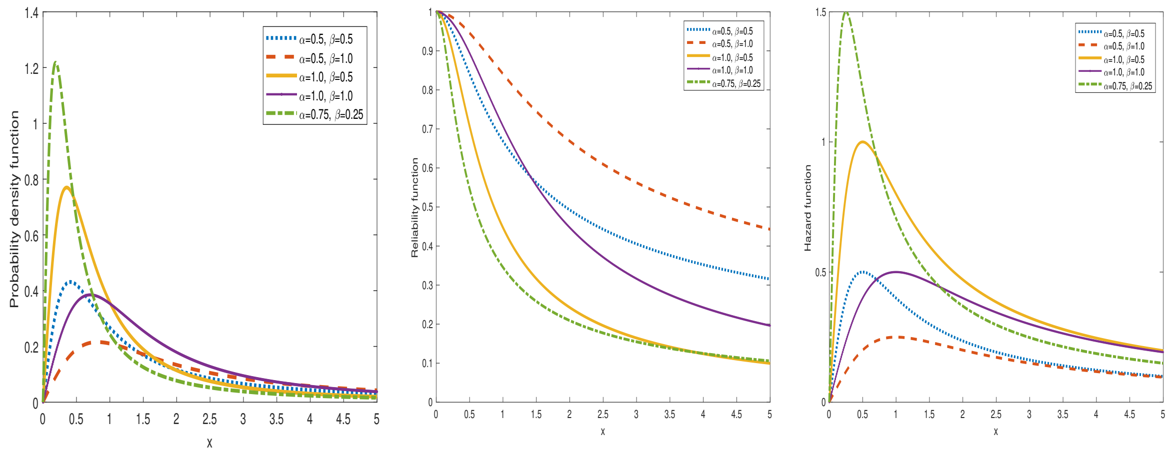
where $x > 0$. Here, α is known as the shape parameter and β is known as the scale parameter. The reliability function and the hazard function of this distribution are respectively obtained as

$$r(x;\alpha,\beta) = \frac{\beta^\alpha}{(x^2+\beta^2)^{\alpha/2}} \quad \text{and} \quad h(x;\alpha,\beta) = \frac{x\alpha}{x^2+\beta^2}, \tag{1.2}$$

where $x > 0$ and $\alpha, \beta > 0$. Various shapes of the probability density, reliability and hazard functions of the gamma-mixed Rayleigh distribution are depicted in Figures 1(a), 1(b) and 1(c), respectively. Differentiating $h(x;\alpha,\beta)$ with respect to x, we obtain

$$\frac{dh(x;\alpha,\beta)}{dx} = \frac{\alpha(\beta+x)(\beta-x)}{(x^2+\beta^2)^2} = \begin{cases} >0, & \text{for } x < \beta \\ <0, & \text{for } x > \beta \\ =0, & \text{for } x = \beta. \end{cases} \tag{1.3}$$

Thus, the hazard function of the gamma-mixed Rayleigh distribution is increasing for $x < \beta$ and decreasing for $x > \beta$, for any value of $\alpha > 0$. Figure 1(c) shows that the hazard of the gamma-mixed Rayleigh distribution is hump-shaped, that is, the hazard is increasing early and eventually begins declining. One may refer to Sarhan *et al.* (2013) for similar study on the exponentiated generalized linear exponential distribution. This type of hazard is often used in modeling data related to survival after successful surgery, where there is an initial increase in risk due to infection or other complications just after the procedure, followed by a steady decline in risk as the patient recovers (see Klein and Moeschberger, 1997).



(a) **(b)** **(c)**

Figure 1: The plots of the (a) density (b) reliability and (c) hazard functions based on different values of the parameters.

The Bayesian prediction of the unknown observation is an important problem. Various authors have studied prediction problems based on the PT-IICS. Kayal *et al.* (2017) obtained the prediction intervals and estimates for future observations in one-sample and two-sample problems for the Chen distribution. Similar problem was studied by Arabi *et al.* (2019) for the Poisson-exponential distribution when PT-IICS is available. For flexible Weibull distribution, Bdair *et al.* (2019) considered Bayesian prediction problem based on the progressive type-II censored data. Very recently, Maiti and Kayal (2019) obtained prediction estimates and intervals for future observations in one-sample and two-sample problems for the generalized Fréchet distribution from Bayesian point of view. To the best of our knowledge, nobody has considered the gamma-mixed Rayleigh distribution with distribution function given by (1.1) for the purpose of statistical inference and Bayesian prediction based on the PT-IICS. In this paper, we address the problem of inference and prediction when the PT-IICS is available from gamma-mixed Rayleigh distribution.

The rest of the paper is organized as follows. In the next section, we obtain MLEs for the unknown parameters, reliability and hazard functions. The existence and uniqueness of the MLEs have been studied. The EM algorithm is described to compute the proposed MLEs. Section 3 deals with the construction of various interval estimates. In Section 4, we derive Bayes estimates with respect to three loss functions. Two approaches are adopted to compute approximate Bayes estimates. Importance sampling method is used to compute HPD credible intervals. Further, in Section 5, we derive Bayesian prediction and interval estimates. In Section 6, we carry out a simulation study to compare the performance of the proposed estimates. A real life dataset is considered for the illustration purpose. Finally, Section 7 concludes the paper.

2. MAXIMUM LIKELIHOOD ESTIMATION

In this section, we derive MLEs of α and β of the gamma-mixed Rayleigh distribution based on the PT-IICS. Using invariance property of the MLE, the MLEs of $r(x)$ and $h(x)$ can be obtained. The likelihood function of α and β is given by

$$L(\alpha,\beta \mid x) = K\prod_{i=1}^{m}\big(1-F_X(x_{i:m:n};\alpha,\beta)\big)^{R_i} f_X(x_{i:m:n};\alpha,\beta), \tag{2.1}$$

where the constant $K = n(n-(R_1+1))(n-(\sum_{j=1}^{2}R_j+2))\cdots(n-\sum_{j=1}^{m-1}(R_j+1))$ and $x=(x_{1:m:n},x_{2:m:n},...,x_{m:m:n})$. The log-likelihood function of α and β is obtained as

$$\begin{aligned}\ell \;=\; \ell(\alpha,\beta \mid x) &\propto m\ln\alpha + m\alpha\ln\beta + \sum_{i=1}^{m}\ln x_{i:m:n} + \alpha\ln\beta\sum_{i=1}^{m}R_i \\ &\quad - \sum_{i=1}^{m}\Big(\frac{\alpha}{2}(1+R_i)+1\Big)\ln(x_{i:m:n}^2+\beta^2).\end{aligned} \tag{2.2}$$

The likelihood equations of α and β are

$$m\Big(\frac{1}{\alpha}+\ln\beta\Big) + \ln\beta\sum_{i=1}^{m}R_i - \frac{1}{2}\sum_{i=1}^{m}(R_i+1)\ln\big(\beta^2+x_{i:m:n}^2\big) = 0 \tag{2.3}$$

and

$$\alpha\Big(\sum_{i=1}^{m}R_i+m\Big) - 2\beta^2\sum_{i=1}^{m}\frac{\big(\frac{\alpha}{2}(R_i+1)+1\big)}{\beta^2+x_{i:m:n}^2} = 0, \tag{2.4}$$

respectively. The MLEs of α and β can be obtained after solving (2.3) and (2.4) simultaneously. These are difficult to obtain in explicit form. The above system of nonlinear equations can be solved by solving a two-dimensional optimization problem. In this case, one may use the Newton-Raphson algorithm. However, the standard Newton-Raphson method does not converge in some cases. We use EM algorithm to compute the MLEs of α and β, which is described below. Note that the EM algorithm was introduced by Dempster *et al.* (1977). Prior to the computation, we discuss the condition under which the MLEs exist and are unique.

Theorem 2.1. *The MLEs of α and β for $(\alpha,\beta)\in(0,\infty)\times(0,\infty)$ exist and are unique under the PT-IICS, provided $x_{i:m:n}>\beta$ holds, for $i=1,...,m$.*

Proof: We show that the maximum value of the log-likelihood function $\ell(\alpha,\beta \mid x)$ exists and also unique for $(\alpha,\beta)\in(0,\infty)\times(0,\infty)$. One may refer to the papers by Cancho *et al.* (2011) and Khan and Mitra (2019) for similar study in other estimation problems. The second order partial derivatives of the log-likelihood function ℓ with respect to α and β are given by

$$\frac{\partial^2\ell}{\partial\alpha^2} \;=\; -\frac{m}{\alpha^2} < 0, \tag{2.5}$$

$$\frac{\partial^2\ell}{\partial\beta^2} \;=\; -\frac{\alpha(\sum_{i=1}^{m}R_i+m)}{\beta^2} - \sum_{i=1}^{m}(\alpha(R_i+1)+2)\frac{(x_{i:m:n}^2-\beta^2)}{(x_{i:m:n}^2+\beta^2)^2} < 0, \tag{2.6}$$

if $x_{i:m:n}>\beta$. Therefore, for fixed $\alpha(\beta)$, ℓ is a strictly concave function with respect to $\beta(\alpha)$.

For fixed β, we get

$$\lim_{\alpha\to 0} \ell(\alpha,\beta \mid x) = -\infty \quad \text{and} \quad \lim_{\alpha\to\infty} \ell(\alpha,\beta \mid x) = -\infty.$$

Similarly, for fixed α, we have $\lim_{\beta\to 0} \ell(\alpha,\beta \mid x) = -\infty$ and $\lim_{\beta\to\infty} \ell(\alpha,\beta \mid x) = -\infty$. So, for fixed $\alpha(\beta)$, ℓ is a unimodal function with respect to $\beta(\alpha)$. Again,

$$\lim_{\alpha\to 0,\ \beta\to 0} \ell(\alpha,\beta \mid x) = -\infty, \quad \lim_{\alpha\to\infty,\ \beta\to 0} \ell(\alpha,\beta \mid x) = -\infty,$$
$$\lim_{\alpha\to 0,\ \beta\to\infty} \ell(\alpha,\beta \mid x) = -\infty, \quad \lim_{\alpha\to\infty,\ \beta\to\infty} \ell(\alpha,\beta \mid x) = -\infty.$$

Let $(\alpha_0,\beta_0) \in (0,\infty)\times(0,\infty)$ and $\ell(\alpha_0,\beta_0 \mid x) = \rho$. Further, set

$$D = \Big\{(\alpha,\beta) : (\alpha,\beta) \in (0,\infty)\times(0,\infty),\ \ell(\alpha,\beta \mid x) \geq \rho\Big\}.$$

So, D is a closed and bounded set, hence D is compact set. Note that the function ℓ is continuous with respect to (α,β). Thus, ℓ has a maximum value for some $(\alpha,\beta) \in D$. Suppose that at $(\alpha_1,\beta_1) \in (0,\infty)\times(0,\infty)$, the function ℓ has maximum. Now, we have to show that (α_1,β_1) is unique. We observe that

$$\ell(\alpha_1,\beta_1 \mid x) > \ell(\alpha_1,\beta \mid x) > \ell(\alpha,\beta \mid x),$$

for $(\alpha,\beta) \in (0,\infty)\times(0,\infty)$, which ensures the desired uniqueness. □

2.1. EM algorithm

The EM algorithm is mainly used to compute the MLEs of the unknown parameters in cases where the likelihood equations cannot be solved explicitly. EM algorithm has two steps: the expectation (E) step and the maximization (M) step. The E-step involves computation of the pseudo log-likelihood function. The M-step involves maximization of the pseudo log-likelihood function. Let the observed sample and censored data be denoted by $X = (X_{1:m:n}, X_{2:m:n}, ..., X_{m:m:n})$ and $Z = (Z_1, Z_2, ..., Z_m)$, respectively, where Z_j is a $1\times R_j$ vector $(Z_{j1}, Z_{j2}, ..., Z_{jR_j})$, for $j = 1,2,...,m$. Note that the complete sample is a combination of the observed sample and the censored data. Denote the complete sample by $W = (X,Z)$. The likelihood function of the complete sample (see Ng *et al.*, 2002) is given by

$$L_C(W;\alpha,\beta) = \prod_{j=1}^{m}\left[f_X(x_{j:m:n};\alpha,\beta)\prod_{k=1}^{R_j} f_Z(z_{jk};\alpha,\beta)\right]. \tag{2.7}$$

Then, the log-likelihood function for the complete sample is

$$\begin{aligned}\ell_C(W;\alpha,\beta) \;=\; & n\ln(\alpha\beta^{\alpha}) + \sum_{j=1}^{m}\left[\ln x_{j:m:n} + \sum_{k=1}^{R_j}\ln z_{jk}\right.\\ & \left. - \left(\frac{\alpha}{2}+1\right)\left(\sum_{k=1}^{R_j}\ln(z_{jk}^2+\beta^2) + \ln(x_{j:m:n}^2+\beta^2)\right)\right].\end{aligned} \tag{2.8}$$

In the E-step, the conditional expectation of the log-likelihood function $\ell_C(W;\alpha,\beta)$ is obtained. This is known as the pseudo log-likelihood function. This can be obtained from $\ell_C(W;\alpha,\beta)$ by replacing any function of z_{jk} say $\psi(z_{jk})$ with $E[\psi(Z_{jk})|Z_{jk} > x_{j:m:n}]$. Thus, the pseudo log-likelihood function is obtained as

$$\begin{aligned}(2.9)\qquad \ell_s(\alpha,\beta) &= n(\ln\alpha\beta^{\alpha}) + \sum_{j=1}^{m}\ln x_{j:m:n} + \sum_{j=1}^{m} R_j A(x_{j:m:n};\alpha,\beta)\\ &\quad - \left(\frac{\alpha}{2}+1\right)\sum_{j=1}^{m}\Big(R_j B(x_{j:m:n};\alpha,\beta) + \ln(x_{j:m:n}^2+\beta^2)\Big),\end{aligned}$$

where

$$\begin{aligned}(2.10)\qquad A(x_{j:m:n};\alpha,\beta) &= E[\ln Z_{jk}|Z_{jk} > x_{j:m:n}]\\ &= \alpha(x_j^2+\beta^2)^{\alpha/2}\int_{x_{j:m:n}}^{\infty}\frac{t\ln t}{(t^2+\beta^2)^{(\alpha/2)+1}}\,dt\end{aligned}$$

and

$$\begin{aligned}(2.11)\qquad B(x_{j:m:n};\alpha,\beta) &= E\big[\ln(Z_{jk}^2+\beta^2)|Z_{jk} > x_{j:m:n}\big]\\ &= \ln(x_{j:m:n}^2+\beta^2) + \frac{2}{\alpha}.\end{aligned}$$

In the M-step, we maximize the pseudo log-likelihood function given by (2.9) obtained in E-step after substituting the values of (2.10) and (2.11) in (2.9). Let $(\alpha^{(p)},\beta^{(p)})$ be an estimate of (α,β) at p-th stage. The corresponding updated estimate $(\alpha^{(p+1)},\beta^{(p+1)})$ can be obtained by maximizing

$$\begin{aligned}(2.12)\qquad \ell_s^*(\alpha,\beta) &= n(\ln\alpha\beta^{\alpha}) + \sum_{j=1}^{m}\ln x_{j:m:n} + \sum_{j=1}^{m} R_j A(x_{j:m:n};\alpha^{(p)},\beta^{(p)})\\ &\quad - \left(\frac{\alpha}{2}+1\right)\sum_{j=1}^{m}\Big(R_j B(x_{j:m:n};\alpha^{(p)},\beta^{(p)}) + \ln(x_{j:m:n}^2+\beta^2)\Big)\end{aligned}$$

with respect to α and β. Now, we compute $\beta^{(p+1)}$ using fixed point iteration method (see Kundu and Pradhan, 2009). The corresponding estimate is obtained by solving the equation

$$(2.13)\qquad \exp\left\{\frac{1}{2n}\sum_{j=1}^{m}(BR_j + \ln(\beta^2 + x_{j:m:n}^2)) - \frac{1}{\hat\alpha(\beta)}\right\} = \beta,$$

where

$$(2.14)\qquad \hat\alpha(\beta) = \left(n - \sum_{j=1}^{m}\frac{\beta^2}{\beta^2 + x_{j:m:n}^2}\right)^{-1}\sum_{j=1}^{m}\frac{2\beta^2}{\beta^2 + x_{j:m:n}^2}$$

with $B = B(x_{j:m:n};\alpha^{(p)},\beta^{(p)})$. We estimate $\beta^{(p+1)}$. The updated estimate $\alpha^{(p+1)}$ can be obtained from $\alpha^{(p+1)} = \hat\alpha(\beta^{(p+1)})$ using (2.14). The algorithm is provided below.

Step-1: Set $p = 0$. Based on the starting value $(\alpha^{(0)}, \beta^{(0)})$, we estimate the parameters α and β.

Step-2: Calculate $B = B(x_{j:m:n}; \alpha^{(p)}, \beta^{(p)})$ from the observed sample $X = x$ and the parameters $\alpha^{(p)}, \beta^{(p)}$.

Step-3: Update (α, β) as $(\alpha^{(p+1)}, \beta^{(p+1)})$.

Step-4: If $|(\alpha^{(p+1)}, \beta^{(p+1)}) - (\alpha^{(p)}, \beta^{(p)})| \le \epsilon$ ($\epsilon > 0$ very small tolerance), then we get the MLEs of the parameters α and β.

Step-5: If $|(\alpha^{(p+1)}, \beta^{(p+1)}) - (\alpha^{(p)}, \beta^{(p)})| > \epsilon$, then set $p = p + 1$ and go to the step 1.

Denote the MLEs of α and β by $\hat{\alpha}$ and $\hat{\beta}$. Replacing α and β with $\hat{\alpha}$ and $\hat{\beta}$, the MLEs of the reliability and hazard functions are respectively obtained as $(x > 0)$

$$\hat{r}(x) = \frac{\beta^{\alpha}}{(x^2 + \beta^2)^{\alpha/2}}\Big|_{(\alpha,\beta)=(\hat{\alpha},\hat{\beta})} \quad \text{and} \quad \hat{h}(x) = \frac{x\alpha}{x^2 + \beta^2}\Big|_{(\alpha,\beta)=(\hat{\alpha},\hat{\beta})}. \tag{2.15}$$

Remark 2.1. The main advantage of the EM algorithm is that computations are straightforward and does not require second and higher order derivatives.

3. INTERVAL ESTIMATES

In this section, we obtain $100(1 - \varphi)\%$ confidence intervals for the parameters, reliability and hazard functions based on PT-IICS. Two techniques are used. First, we discuss the construction of asymptotic confidence intervals. It is noted that to apply this procedure, we need the concept of observed Fisher information matrix. Louis (1982) first derived the observed Fisher information matrix using missing information based on the EM algorithm. The observed Fisher information matrix is used to construct the asymptotic confidence intervals. According to Louis, the observed information equals to the complete information minus the missing information. That is, $I_X(\theta) = I_W(\theta) - I_{W|X}(\theta)$, where $I_X(\theta)$, $I_W(\theta)$ and $I_{W|X}(\theta)$ are the observed information, complete information and missing information, respectively. Denote $\theta = (\alpha, \beta)$. The complete information matrix $I_W(\theta)$ is given as

$$I_W(\theta) = -E\left[\frac{\partial^2 \ell_C(W;\theta)}{\partial \theta^2}\right] = \begin{pmatrix} \frac{n}{\alpha^2} & -\frac{2n}{\beta(\alpha+2)} \\ -\frac{2n}{\beta(\alpha+2)} & \frac{4n\alpha}{\beta^2(\alpha+4)} \end{pmatrix}. \tag{3.1}$$

Again, the missing information $I_{W|X}(\theta)$ at j-th failure time $x_{j:m:n}$ is obtained as

$$I_{W|X}^{j:m:n}(\theta) = \begin{pmatrix} -b_{20}(x_{j:m:n};\alpha,\beta) & -b_{11}(x_{j:m:n};\alpha,\beta) \\ -b_{11}(x_{j:m:n};\alpha,\beta) & -b_{02}(x_{j:m:n};\alpha,\beta) \end{pmatrix},$$

where

$$b_{20}(x_{j:m:n};\alpha,\beta) = -\frac{1}{\alpha^2}, \qquad b_{11}(x_{j:m:n};\alpha,\beta) = \frac{2\beta}{(\alpha+2)(x_{j:m:n}^2 + \beta^2)},$$

$$b_{02}(x_{j:m:n};\alpha,\beta) = \frac{\alpha}{(x_{j:m:n}^2 + \beta^2)}\left[\frac{(x_{j:m:n}^2 - \beta^2)}{(x_{j:m:n}^2 + \beta^2)} + \frac{2(\alpha+2)\beta^2}{(\alpha+4)(x_{j:m:n}^2 + \beta^2)} - 1\right].$$

Thus, the total missing information $I_{W|X}(\theta)$ is given as

$$I_{W|X}(\theta) = \sum_{j=1}^{m} R_j I_{W|X}^{j:m:n}(\theta). \tag{3.2}$$

From the 2×2 order matrices given by (3.1) and (3.2), we compute the observed Fisher information matrix of α and β as

$$I_X(\theta) = \begin{pmatrix} d_{20} & d_{11} \\ d_{11} & d_{02} \end{pmatrix}, \tag{3.3}$$

where

$$\begin{aligned} d_{20} &= \frac{1}{\alpha^2}\left(n - \sum_{j=1}^{m} R_j\right), \qquad d_{11} = -\frac{2}{\beta(\alpha+2)}\left[n - \frac{\beta^2 \sum_{j=1}^{m} R_j}{(x_{j:m:n}^2 + \beta^2)}\right] \quad \text{and} \\ d_{02} &= \frac{4n\alpha}{\beta^2(\alpha+4)} + \frac{\alpha \sum_{j=1}^{m} R_j}{(x_{j:m:n}^2 + \beta^2)}\left[\frac{(x_{j:m:n}^2 - \beta^2)}{(x_{j:m:n}^2 + \beta^2)} + \frac{2(\alpha+2)\beta^2}{(\alpha+4)(x_{j:m:n}^2 + \beta^2)} - 1\right]. \end{aligned}$$

In this part, we obtain asymptotic confidence intervals using (i) normal approximation (NA) of the MLE and (ii) the log-transformed (NL) MLE methods. We omit the details of this method to maintain brevity. For the formulas for the NA and NL approaches, see Lee and Cho (2017) and Maiti and Kayal (2020, 2021).

3.1. Bootstrap confidence intervals

It is seen in the previous subsection that to obtain the approximate confidence intervals of the unknown model parameters, it is required to derive second order derivatives which is cumbersome. So, we consider bootstrap technique, which is simpler than NA and NL methods. In particular, we adopt percentile bootstrap (Boot-p) and bootstrap-t (Boot-t) techniques. Here, we describe the procedure how to obtain confidence intervals using Boot-p method. First, we obtain the MLEs of $\eta = (\alpha, \beta, r(x), h(x))$. Denote the MLEs of η by $\hat{\eta} = (\hat{\alpha}, \hat{\beta}, \hat{r}(x), \hat{h}(x))$. Now, based on $\hat{\alpha}$ and $\hat{\beta}$, the bootstrap sample $x^* = (x_1^*, x_2^*, ..., x_n^*)$ has to be generated. We compute $\hat{\eta}^* = (\hat{\alpha}^*, \hat{\beta}^*, \hat{r}^*(x), \hat{h}^*(x))$ based on x^*. Repeat this procedure for 1000 times to get $\hat{\eta}_1^*, \hat{\eta}_2^*, ..., \hat{\eta}_{1000}^*$, where $\hat{\eta}_i^* = (\hat{\alpha}_i^*, \hat{\beta}_i^*, \hat{r}_i^*(x), \hat{h}_i^*(x)),\ i = 1, 2, ..., 1000$. Next, we arrange $\hat{\eta}_i^*$'s in ascending order and denote $\hat{\eta}_{(1)}^* \le \hat{\eta}_{(2)}^* \le \cdots \le \hat{\eta}_{(1000)}^*$. Thus, the $100(1-\varphi)\%$ approximate bootstrap-p confidence interval for η is obtained as (L, U), where $L = \hat{\eta}_{(\frac{i\varphi}{2})}^*$ and $U = \hat{\eta}_{(i(1-\frac{\varphi}{2}))}^*$. The percentile bootstrap confidence interval of η at 95% level of confidence is $(\hat{\eta}_{(25)}^*, \hat{\eta}_{(975)}^*)$. For small sample size, the Boot-p method does not perform well. In this subsection, we discuss Boot-t method, which is simple to apply compared to Boot-p method. We obtain $\hat{\eta}^* = (\hat{\alpha}^*, \hat{\beta}^*, \hat{r}^*(x), \hat{h}^*(x))$ similar to the procedure as mentioned in Boot-p method. Then, based on the bootstrap sample $x^* = (x_1^*, x_2^*, ..., x_n^*)$, we compute the variance-covariance matrix $I_X^{*-1}(\hat{\alpha}^*, \hat{\beta}^*)$. For $i = 1, 2, ..., 1000$, calculate the value of the statistic $T_{\eta_i}^* = (\hat{\eta}_i^* - \hat{\eta}_i)/\sqrt{\widehat{var}(\hat{\eta}_i^*)}$. Then, we arrange in the ascending order and get $T_{\eta_{(1)}}^* \le T_{\eta_{(2)}}^* \le \cdots \le T_{\eta_{(1000)}}^*$. Now, the $100(1-\varphi)\%$ approximate bootstrap-t confidence interval for η is given by (L, U), where $L = T_{\eta_{(\frac{i\varphi}{2})}}^*$ and $U = T_{\eta_{(i(1-\frac{\varphi}{2}))}}^*$. The approximate Boot-$t$ confidence interval of η at 95% level of confidence is $(T_{\eta_{(25)}}^*, T_{\eta_{(975)}}^*)$.

4. BAYESIAN ESTIMATION

In this section, we obtain Bayes estimates of the unknown parameters α, β and the reliability characteristics $r(t), h(t)$ of the gamma-mixed Rayleigh distribution based on PT-IICS. Three loss functions have been considered: (i) squared error loss (SEL) function, (ii) LINEX loss function and (iii) entropy loss function. The SEL function is a balance type loss function. That is, when this loss function is used, the overestimation as well as underestimation do not have any effect on the estimation problem. However, there are situations, where the squared error loss function is not suitable. For example, when we estimate reliability of a rocket, the underestimation is dangerous than the overestimation. Further, the overestimation is severe than the underestimation when estimating the water level of bank of river in a flood-prone area. We also consider two asymmetric loss functions (LINEX and entropy) which are useful to deal with this type of situations. Let δ be an estimator of the unknown parameter ϕ. Then, Table 1 represents Bayes estimates of ϕ under the squared error, LINEX and entropy loss functions. In Table 1, ω and κ are both non-zero real numbers. For $\kappa = -1$, the Bayes estimate with respect to the entropy loss function reduces to that under the squared error function. To obtain the Bayes estimates, one needs to consider prior distributions for the unknown model parameters. It is well known that the joint conjugate prior is not available when both the parameters are not known. Further, there is no clear methodology to choose an appropriate prior (see Arnold and Press, 1983) for a Bayesian estimation problem.

Table 1: Loss functions and the corresponding form of the Bayes estimates.

Name of the loss functions	Form of the loss functions	Form of the Bayes estimates
SEL	$l_s(\phi,\delta) = (\delta-\phi)^2$	$E_\phi(\phi\vert\, x)$
LINEX	$l_\ell(\phi,\delta) = \exp\{\omega(\delta-\phi)\} - \omega(\delta-\phi) - 1$	$-\frac{1}{\omega}\ln(E_\phi(\exp\{-\omega\phi\}\vert\, x))$
Entropy	$l_e(\phi,\delta) = (\delta/\phi)^\kappa - \kappa\ln(\delta/\phi) - 1$	$\left[E_\phi(\phi^{-\kappa}\vert\, x)\right]^{-\frac{1}{\kappa}}$

Note that the gamma distribution is versatile for adjusting different shapes of the density function. It has a log-concave density function in the interval $(0,\infty)$. Jeffery's prior can be obtained as a special case of the gamma prior. Due to these facts, various authors have considered independent gamma distributions as the priors for different Bayesian estimation problems. See, for instance, Kundu (2008), Huang and Wu (2012) and Maiti and Kayal (2020). Here, we assume independent gamma priors for α and β. Let $\alpha \sim \text{Gamma}(a_1, a_2)$ and $\beta \sim \text{Gamma}(a_3, a_4)$, when $\text{Gamma}(a_1, a_2)$ and $\text{Gamma}(a_3, a_4)$ represent gamma distributions with scale and shape parameters $1/a_2, a_1$ and $1/a_4, a_3$, respectively. The probability density functions of $\text{Gamma}(a_1, a_2)$ and $\text{Gamma}(a_3, a_4)$ are given by

$$g_1(\alpha; a_1, a_2) \propto \alpha^{a_1-1}\exp\{-\alpha a_2\} \quad \text{and} \quad g_2(\beta; a_3, a_4) \propto \beta^{a_3-1}\exp\{-\beta a_4\},$$

respectively, where $\alpha, \beta > 0$ and $a_1, a_2, a_3, a_4 > 0$. The hyper-parameters in the prior distributions are assumed to be known. After some simplification, the posterior distribution of

α, β given $X = x$ is obtained as

$$\Pi(\alpha, \beta|x) \propto \frac{\Pi_1(\alpha, \beta, x)}{\int_0^\infty \int_0^\infty \Pi_1(\alpha, \beta, x) d\alpha d\beta}, \tag{4.1}$$

where the joint distribution of α, β and X is given by

$$\Pi_1(\alpha, \beta, x) \propto \alpha^{m+a_1-1}\beta^{m\alpha+a_3-1}\exp\{-(\alpha a_2 + \beta a_4)\}\prod_{i=1}^{m}\frac{x_{i:m:n}\beta^{\alpha R_i}}{(x_{i:m:n}^2 + \beta^2)^{\frac{\alpha}{2}(1+R_i)+1}}. \tag{4.2}$$

Thus, for any arbitrary estimand $g(\alpha, \beta)$, the Bayes estimates with respect to the LINEX and entropy loss functions are respectively obtained as

$$\hat{g}_{bl} = -\frac{1}{\omega}\ln\left[\frac{\int_0^\infty \int_0^\infty \exp\{-\omega g(\alpha, \beta)\}\Pi_1(\alpha, \beta, x) d\alpha d\beta}{\int_0^\infty \int_0^\infty \Pi_1(\alpha, \beta, x) d\alpha d\beta}\right] \quad \text{and} \tag{4.3}$$

$$\hat{g}_{be} = \left[\frac{\int_0^\infty \int_0^\infty g^{-\kappa}(\alpha, \beta)\Pi_1(\alpha, \beta, x) d\alpha d\beta}{\int_0^\infty \int_0^\infty \Pi_1(\alpha, \beta, x) d\alpha d\beta}\right]^{-\frac{1}{\kappa}}. \tag{4.4}$$

As mentioned before, the Bayes estimate with respect to the SEL function can be obtained from (4.4) when $\kappa = -1$. Note that the required Bayes estimates of $\alpha, \beta, r(x)$ and $h(x)$ with respect to the LINEX and entropy loss functions can be computed after substituting $\alpha, \beta, r(x)$ and $h(x)$ in the place of $g(\alpha, \beta)$ in (4.3) and (4.4), respectively. Choosing values of the hyper-parameters is always an important task from Bayesian point of view. Below, we propose a method in this purpose.

Remark 4.1. We generate m samples from a gamma-mixed Rayleigh distribution with distribution function given by (1.1). For each of this m samples, we obtain the MLEs of the model parameters, which are denoted by $\hat{\alpha}^j$ and $\hat{\beta}^j$, $j = 1, 2, ..., m$. The mean and variance of the gamma prior distribution with density function $g_1(\alpha; a_1, a_2)$ are $\frac{a_1}{a_2}$ and $\frac{a_1}{a_2^2}$, respectively. Further, the mean and variance of the MLEs of α for m samples are $\frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j$ and $\frac{1}{m-1}\sum_{j=1}^m(\hat{\alpha}^j - \frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j)^2$, respectively. Therefore, the mean and variance of the MLEs are equal to $\frac{a_1}{a_2}$ and $\frac{a_1}{a_2^2}$, respectively. That is,

$$\frac{a_1}{a_2} = \frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j \quad \text{and} \quad \frac{a_1}{a_2^2} = \frac{1}{m-1}\sum_{j=1}^m\left(\hat{\alpha}^j - \frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j\right)^2.$$

Solving these equations, we get

$$a_1 = \frac{\left(\frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j\right)^2}{\frac{1}{m-1}\sum_{j=1}^m\left(\hat{\alpha}^j - \frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j\right)^2} \quad \text{and} \quad a_2 = \frac{\frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j}{\frac{1}{m-1}\sum_{j=1}^m\left(\hat{\alpha}^j - \frac{1}{m}\sum_{j=1}^m \hat{\alpha}^j\right)^2}.$$

In a similar manner, the hyper-parameters a_3 and a_4 can be obtained from the above equations by replacing $\hat{\alpha}^j$ with $\hat{\beta}^j$.

4.1. Computational methods

In the above section, we see that the proposed Bayes estimates are in the form of the ratio of two integrals. These integrals can not be evaluated in terms of some closed-form expressions. So, we use two approaches in order to get approximate values of the Bayes estimates. One of these is proposed by Lindley (1980). Other is due to Chen and Shao (1999).

4.1.1. Lindley's approximation method

In this subsection, we discuss the Bayes estimates of $\alpha, \beta, r(x)$ and $h(x)$ using Lindley's approximation technique. The detailed derivations are omitted to maintain brevity. We refer to Lee and Cho (2017) and Maiti and Kayal (2021) for detailed derivation of the Bayes estimates using this method. First, we consider LINEX loss function. With respect to this loss function, the Bayes estimate of α is given by

$$\hat{\alpha}_{bl} = -\frac{1}{\omega}\ln\Big[\exp\{-\omega\alpha\} + (1/2)\omega\exp\{-\omega\alpha\}\left[\omega\tau_{11} - A(\alpha,\beta)\right]\Big]\Big|_{(\alpha,\beta)=(\hat{\alpha},\hat{\beta})}, \tag{4.5}$$

where $A(\alpha,\beta) = \{l_{30}\tau_{11}^2 + l_{03}\tau_{21}\tau_{22} + 3l_{21}\tau_{11}\tau_{12} + l_{12}(\tau_{11}\tau_{22} + 2\tau_{21}^2) + 2p_1\tau_{11} + 2p_2\tau_{12}\},\ l_{ij} = \frac{\partial^{i+j}l}{\partial\alpha^i\partial\beta^j};\ i,j = 0,1,2,3;\ i+j = 3,\ p_1 = \frac{\partial p}{\partial\alpha},\ p_2 = \frac{\partial p}{\partial\beta}$ and p is equal to the logarithm of joint prior distribution of α and β. The Bayes estimate of α with respect to the entropy loss function is

$$\hat{\alpha}_{be} = \Big[\alpha^{-\kappa} + (1/2)\kappa\alpha^{-(\kappa+1)}\big[(\kappa+1)\alpha^{-1}\tau_{11} - A(\alpha,\beta)\big]\Big]^{-\frac{1}{\kappa}}\Big|_{(\alpha,\beta)=(\hat{\alpha},\hat{\beta})}. \tag{4.6}$$

The Bayes estimate of α with respect to the squared error loss function can be obtained from (4.6) substituting $\kappa = -1$. Further, the Bayes estimates of $\beta, r(x)$ and $h(x)$ with respect to the squared error, LINEX and entropy loss functions can be derived similarly.

4.1.2. Importance sampling method

In the previous subsection, we obtain the Bayes estimates using Lindley's approximation method. One disadvantage of this method is that it requires higher order partial derivatives of the log-likelihood function. Further, the Lindley's approximation can not be used to construct highest posterior density (HPD) credible intervals. In this subsection, we describe importance sampling method which is free from the higher order partial derivatives. It is also used to compute HPD credible intervals. To apply importance sampling method, we need to rewrite the joint posterior distribution of α, β given $X = x$ in (4.1) as

$$\Pi(\alpha,\beta|X=x) \propto Gamma_{\alpha}(m+a_1,a_2)Gamma_{\beta|\alpha}(m\alpha+a_3,a_4)h(\alpha,\beta) \tag{4.7}$$

where

$$h(\alpha,\beta) = a_4^{-(m\alpha+a_3)}\prod_{i=1}^{m}\beta^{\alpha R_i}x_{i:m:n}(x_{i:m:n}^2+\beta^2)^{-(\frac{\alpha}{2}(1+R_i)+1)}.$$

At first, we generate α from gamma distribution $Gamma_{\alpha}(m+a_1,a_2)$. Next, β is generated from the $Gamma_{\beta|\alpha}(m\alpha+a_3,a_4)$ distribution. We repeat this procedure 1000 times to obtain $(\alpha_1,\beta_1),(\alpha_2,\beta_2),...,(\alpha_{1000},\beta_{1000})$. Thus, the Bayes estimates of a parametric function $g(\alpha,\beta)$ under LINEX and entropy loss functions are respectively given by

$$\hat{g}_{bl} = -\frac{1}{\omega}\ln\left[\frac{\sum_{i=1}^{1000}\exp\{-\omega g(\alpha_i,\beta_i)\}h(\alpha_i,\beta_i)}{\sum_{i=1}^{1000}h(\alpha_i,\beta_i)}\right] \tag{4.8}$$

and

$$\hat{g}_{be} = \left[\frac{\sum_{i=1}^{1000} g(\alpha_i,\beta_i)^{-\kappa} h(\alpha_i,\beta_i)}{\sum_{i=1}^{1000} h(\alpha_i,\beta_i)}\right]^{-\frac{1}{\kappa}}. \tag{4.9}$$

We compute the Bayes estimates of $\alpha, \beta, r(x)$ and $h(x)$ substituting $\alpha, \beta, r(x)$ and $h(x)$ in place of $g(\alpha, \beta)$, respectively in (4.8) and (4.9) under LINEX and entropy loss functions. Using the concept of importance sampling method, one can derive HPD credible intervals for the unknown parameters α, β and reliability characteristics $r(x), h(x)$. The derivation of the credible intervals have been skipped from this paper due to sake of conciseness. One may refer to Kundu and Raqab (2015) and Rastogi and Tripathi (2014) for elaborate discussion on the derivation of the HPD credible interval for some lifetime distributions.

5. BAYESIAN PREDICTION AND INTERVAL ESTIMATION

In the previous section, we study the Bayesian estimation for the unknown parameters, reliability and the hazard functions. Here, we discuss Bayesian prediction for the future observations based on the PT-IICS taken from the gamma-mixed Rayleigh distribution. We compute the corresponding prediction intervals. There have been a lot of efforts from various authors in prediction problems. For some recent references, please refer to Dey *et al.* (2018) and Bdair *et al.* (2019). This section is divided into two subsections. The following subsection deals with one-sample prediction problem.

5.1. One-sample prediction and Bayesian prediction interval (BPI)

Suppose n number of total independent life testing units are subjected to an experiment. Let $x = (x_{1:m:n}, x_{2:m:n}, ..., x_{m:m:n})$ be the observed progressively type-II censored sample. The censoring scheme is taken as $R = (R_1, R_2, ..., R_m)$. Let $y_i = (y_{i1}, y_{i2}, ..., y_{iR_i})$ represent the ordered lifetimes of the units which are censored at the i-th failure $x_{i:m:n}$. The future observations to be predicted based on x are $y = (y_{ip}; i = 1, 2, ..., m; p = 1, 2, ..., R_i)$. The conditional density y under the given information can be obtained as

$$\begin{aligned} f_1(y|x,\alpha,\beta) &= p\binom{R_i}{p}\sum_{k=0}^{p-1}(-1)^{p-k-1}\binom{p-1}{k} f(y)(1-F(y))^{R_i-k-1} \\ &\quad \times(1-F(x_i))^{k-R_i}, \qquad y > x_{i:m:n}. \end{aligned} \tag{5.1}$$

The distribution function is

$$F_1(y|x,\alpha,\beta) = p\binom{R_i}{p}\sum_{k=0}^{p-1}\frac{(-1)^{p-k-1}}{R_i-k}\binom{p-1}{k}\left[1-(1-F(x_i))^{k-R_i}(1-F(y))^{R_i-k}\right]. \tag{5.2}$$

Notice that the posterior predictive density and distribution functions are respectively given by

$$f_1^*(y|x) = \int_0^\infty \int_0^\infty f_1(y|x,\alpha,\beta)\Pi(\alpha,\beta|x)d\alpha d\beta \tag{5.3}$$

and

$$F_1^*(y|x) = \int_0^\infty \int_0^\infty F_1(y|x,\alpha,\beta)\Pi(\alpha,\beta|x)d\alpha d\beta. \tag{5.4}$$

The Bayesian predictive estimate of y under LINEX and entropy loss functions are respectively given by

$$\widehat{y}_l = -\frac{1}{\omega}\ln\left[\int_{x_i}^\infty \exp\{-\omega y\} f_1^*(y|x)dy\right] = -\frac{1}{\omega}\ln[E(P_1(\alpha,\beta)|x)] \tag{5.5}$$

and

$$\widehat{y}_e = \left[\int_{x_i}^\infty y^{-\kappa} f_1^*(y|x)dz\right]^{-\frac{1}{\kappa}} = [E(P_2(\alpha,\beta)|x)]^{-\frac{1}{\kappa}}, \tag{5.6}$$

where

$$P_1(\alpha,\beta) = \int_{x_i}^\infty \exp\{-\omega y\} f_1(y|x,\alpha,\beta)dy \quad \text{and} \quad P_2(\alpha,\beta) = \int_{x_i}^\infty y^{-\kappa} f_1(y|x,\alpha,\beta)dy.$$

Note that above integrals can not be computed analytically. Thus, one needs to use numerical technique in order to compute the predictive estimates. In this purpose, we use importance sampling methods as mentioned in Subsection 4.1.2. Equations (5.5) and (5.6) can be evaluated using importance sampling method as

$$\begin{aligned} \widehat{y}_l &= -\left(\frac{1}{\omega}\right)\ln\left[\frac{\sum_{i=1}^{1000} P_1(\alpha_i,\beta_i)h(\alpha_i,\beta_i)}{\sum_{i=1}^{1000} h(\alpha_i,\beta_i)}\right] \quad \text{and} \\ \widehat{y}_e &= \left[\frac{\sum_{i=1}^{1000} P_2(\alpha_i,\beta_i)h(\alpha_i,\beta_i)}{\sum_{i=1}^{1000} h(\alpha_i,\beta_i)}\right]^{-1/\kappa}, \end{aligned} \tag{5.7}$$

respectively. Next, Bayesian prediction interval is obtained. The prior predictive survival function $S_1(t|x,\alpha,\beta)$ is obtained as

$$S_1(t|x,\alpha,\beta) = \frac{P(y > t|x,\alpha,\beta)}{P(y > x_{i:m:n}|x,\alpha,\beta)} = \frac{\int_t^\infty f_1(u|x,\alpha,\beta)du}{\int_{x_{i:m:n}}^\infty f_1(u|x,\alpha,\beta)du}.$$

The posterior survival function is given by

$$S_1^*(t|x) = \int_0^\infty \int_0^\infty S_1(t|x,\alpha,\beta)\Pi(\alpha,\beta|x)d\alpha d\beta. \tag{5.8}$$

Equation (5.8) can be evaluated using importance sampling method under SEL function as

$$S_1^*(t|x) = \frac{\sum_{i=1}^{1000} S_1(t|x,\alpha_i,\beta_i)h(\alpha_i,\beta_i)}{\sum_{i=1}^{1000} h(\alpha_i,\beta_i)}. \tag{5.9}$$

We obtain two sided $100(1-\varphi)\%$ equal-tail symmetric predictive interval (L,U) by solving the following non-linear equations

$$S_1^*(L|x) = 1 - \frac{\varphi}{2} \quad \text{and} \quad S_1^*(U|x) = \frac{\varphi}{2}. \tag{5.10}$$

The algorithm to obtain the lower bound L and the upper bound U from $S_1^*(t|x) = \eta$, where t is L or U and $\eta = (1 - \frac{\varphi}{2})$ or $\frac{\varphi}{2}$ is described below.

Step-1: Set initial value $t = t_0$.

Step-2: Calculate $S_1^*(t|x) = \frac{\sum_{i=1}^{1000} S_1(t|x,\alpha_i,\beta_i)h(\alpha_i,\beta_i)}{\sum_{i=1}^{1000} h(\alpha_i,\beta_i)}$.

Step-3: If $S_1^*(t|x) < \eta$, then increase t value otherwise decrease the value of t.

Step-4: Repeat steps 2 and 3 until $S_1^*(t|x) \simeq \eta$.

5.2. Two-sample prediction and BPI

In this section, we derive Bayesian two-sample prediction estimate for future observation based on the PT-IICS. It is noted that the two-sample plan is applied in which the observed sample is the PT-IICS and $Z_1 < Z_2 < \cdots < Z_T$ be the unobserved future observations from the same sample, yet to be observed. The predictive density function of Z_j can be written as

$$f(z_j|\alpha, \beta) = j\binom{T}{j}\sum_{p=0}^{j-1}(-1)^{j-1-p}\binom{j-1}{p}[1 - F(z_j)]^{T-1-p}f(z_j). \tag{5.11}$$

Again, the posterior prediction density function is obtained as

$$f^*(z_j|x) = \int_0^\infty \int_0^\infty f(z_j|\alpha, \beta)\Pi(\alpha, \beta|x)d\alpha d\beta.$$

Further, the Bayesian predictive estimate of Z_j under LINEX and entropy loss functions are respectively obtained as

$$\widehat{z}_{j_l} = -\left(\frac{1}{\omega}\right)\ln\left[\frac{\sum_{i=1}^{1000} T_1(\alpha_i, \beta_i)h(\alpha_i, \beta_i)}{\sum_{i=1}^{1000} h(\alpha_i, \beta_i)}\right]$$

and

$$\widehat{z}_{j_e} = \left[\frac{\sum_{i=1}^{1000} T_2(\alpha_i, \beta_i)h(\alpha_i, \beta_i)}{\sum_{i=1}^{1000} h(\alpha_i, \beta_i)}\right]^{-1/\kappa},$$

where

$$T_1(\alpha, \beta) = \int_0^\infty \exp\{-\omega z_j\}f(z_j|\alpha, \beta)dz_j \quad \text{and} \quad T_2(\alpha, \beta) = \int_0^\infty z_j^{-\kappa}f(z_j|\alpha, \beta)dz_j.$$

Next, Bayesian prediction interval is obtained. The predictive posterior survival function is given by

$$S_1^*(z_j|x) = \int_0^\infty \int_0^\infty S_1(z_j|x, \alpha, \beta)\Pi(\alpha, \beta|x)d\alpha d\beta,$$

where

$$S_1(z_j|x, \alpha, \beta) = \frac{\int_{z_j}^\infty f_1(u|x, \alpha, \beta)du}{\int_{x_{i:m:n}}^\infty f_1(u|x, \alpha, \beta)du}.$$

The above integration can be approximated using importance sampling method. Further, to obtain the two-sided $100(1-\varphi)\%$ equal-tail symmetric prediction interval (L, U) for Z_j, we have to solve the non-linear equations given by

$$S_1^*(L|x) = 1 - \frac{\varphi}{2} \quad \text{and} \quad S_1^*(U|x) = \frac{\varphi}{2}. \tag{5.12}$$

6. SIMULATION RESULTS AND REAL DATA ANALYSIS

In this section, we first carry out simulation study to observe the performance of the proposed estimates. Next, we consider a real dataset for illustrative purpose.

6.1. Simulation results

This subsection is devoted to the comparative study of the proposed estimates. For this purpose, we generate 1000 progressive type-II censored samples from gamma-mixed Rayleigh distribution. We consider various combinations of (n, m) as $(35, 20), (35, 35), (50, 45)$ and $(50, 50)$. The actual values of α and β are taken as 0.5 and 0.25, respectively. The actual values of $r(x)$ and $h(x)$ are 0.405461 and 0.324324, respectively for $x = 1.5$. There is no reason of taking the value of x as 1.5. One may consider other values of x too. The simulation study has been carried out for other values of x, but not presented here for brevity. For other values of x, similar behaviour of the proposed methods have been observed. The simulation is carried out using the statistical software R (Vienna, Austria; `https://www.r-project.org/`), version 4.1.0. In Table 2, the estimated values of the hyper-parameters are presented for different values of m. For the purpose of the Bayesian estimates, we take $\omega = -0.25, 0.001$ and $\kappa = -0.5, 0.5$ for LINEX and entropy loss functions, respectively. Further, for each n, three different censoring schemes such as progressive type-II, type-II and complete sample have been used for simulation study. These schemes are presented in Table 3. It is known that the type-II censoring scheme is a special case of the progressive type-II censoring scheme.

Table 2: Values of the hyper-parameters for different m.

(α, β)	m	a_1	a_2	a_3	a_4
(0.5, 0.25)	20	0.29516	0.10370	0.29308	0.19850
	35	0.79587	0.50618	0.78481	0.98676
	45	1.41808	1.18262	1.38071	2.30699
	50	1.88165	1.75665	1.84799	3.42774

Table 3: Different censoring schemes (CS).

Scheme	Category	m	$(R_1, R_2, ..., R_m)$
Progressive type-II censoring	Pr-IIc	Odd	$\left(R_{\frac{m+1}{2}} = n - m,\ R_i = 0;\ \ i \neq \frac{m+1}{2}\right)$
		Even	$\left(R_{m/2} = n - m,\ R_i = 0;\ i \neq \frac{m}{2}\right)$
Type-II censoring	Ty-IIc		$(R_m = n - m,\ R_i = 0;\ i \neq m)$
Complete case	Cc		$(R_i = 0;\ i = 1 \sim m)$

Tables 4 and 5 present the average and mean squared error (MSE) values of the MLEs and the Bayes estimates for (α, β) and $(r(x), h(x))$, respectively. The 1st column is for (n, m), the 2nd column is for various censoring schemes (CS), 3rd column is for the estimands. Here, estimands are the unknown parameters α, β and the reliability characteristics $r(x), h(x)$.

Table 4: Average and MSE values of estimates for the parameters α and β.

(n,m)	CS	Parameter	EM	Method	SEL	LINEX		EL	
						$\omega=-0.25$	$\omega=0.001$	$\kappa=-0.5$	$\kappa=0.5$
			Avg (MSE)		Avg (MSE)	Avg (MSE)	Avg (MSE)	Avg (MSE)	Avg (MSE)
(35,20)	Pr-IIc	α	0.577322 (0.045870)	Lin	0.604350 (0.010889)	0.609516 (0.011994)	0.604329 (0.010884)	0.586034 (0.007402)	0.550292 (0.002529)
				Imp	0.615036 (0.012560)	0.617727 (0.012989)	0.615021 (0.012558)	0.608165 (0.010962)	0.601033 (0.008704)
		β	0.294180 (0.018021)	Lin	0.305454 (0.003075)	0.306914 (0.003239)	0.305448 (0.003074)	0.295391 (0.002060)	0.276136 (0.000683)
				Imp	0.320065 (0.003605)	0.323106 (0.004121)	0.320040 (0.003603)	0.305264 (0.003116)	0.303128 (0.001605)
	Ty-IIc	α	0.650948 (0.166828)	Lin	0.739859 (0.057532)	0.748100 (0.061553)	0.739825 (0.057516)	0.712266 (0.045057)	0.652946 (0.023393)
				Imp	0.742216 (0.064523)	0.748506 (0.067051)	0.742177 (0.064507)	0.0.73152 (0.060087)	0.730816 (0.053506)
		β	0.322548 (0.041284)	Lin	0.356579 (0.011359)	0.358578 (0.011789)	0.356571 (0.011357)	0.343506 (0.008743)	0.316460 (0.004417)
				Imp	0.358136 (0.018642)	0.360861 (0.021394)	0.358132 (0.018637)	0.351134 (0.016752)	0.342618 (0.011306)
(,35)	Cc	α	0.517253 (0.015562)	Lin	0.537146 (0.001380)	0.538660 (0.001495)	0.537140 (0.001379)	0.531160 (0.000971)	0.518915 (0.000358)
				Imp	0.541207 (0.002163)	0.546251 (0.002237)	0.541206 (0.002162)	0.537645 (0.002088)	0.524005 (0.001463)
		β	0.259399 (0.006465)	Lin	0.277225 (0.000741)	0.278008 (0.000784)	0.277222 (0.000741)	0.270984 (0.000440)	0.258125 (0.000066)
				Imp	0.284130 (0.001053)	0.287056 (0.001134)	0.284087 (0.001051)	0.282670 (0.001041)	0.278009 (0.000915)
(50,45)	Pr-IIc	α	0.527374 (0.014947)	Lin	0.561924 (0.003834)	0.563149 (0.003988)	0.561918 (0.003834)	0.557074 (0.003257)	0.546668 (0.002178)
				Imp	0.564010 (0.003952)	0.567732 (0.004139)	0.564008 (0.003948)	0.558507 (0.003760)	0.556072 (0.003427)
		β	0.276102 (0.008370)	Lin	0.309227 (0.003508)	0.309901 (0.003588)	0.309224 (0.003507)	0.303973 (0.002913)	0.292121 (0.001774)
				Imp	0.312564 (0.003567)	0.318007 (0.003644)	0.312561 (0.003565)	0.307715 (0.003340)	0.296405 (0.003197)
	Ty-IIc	α	0.528481 (0.015516)	Lin	0.564288 (0.004133)	0.565528 (0.004294)	0.564284 (0.004132)	0.559387 (0.003527)	0.548837 (0.002385)
				Imp	0.570806 (0.004215)	0.577130 (0.004362)	0.570806 (0.004210)	0.561010 (0.004100)	0.560377 (0.003761)
		β	0.276786 (0.008606)	Lin	0.310822 (0.003699)	0.311502 (0.003782)	0.310819 (0.003699)	0.305526 (0.003083)	0.293530 (0.001895)
				Imp	0.315542 (0.003720)	0.320566 (0.003935)	0.315537 (0.003720)	0.312147 (0.003565)	0.307081 (0.003416)
(,50)	Cc	α	0.527521 (0.011298)	Lin	0.573868 (0.005456)	0.574784 (0.005592)	0.573864 (0.005456)	0.570163 (0.004923)	0.561857 (0.003826)
				Imp	0.579013 (0.005521)	0.581451 (0.005640)	0.579011 (0.005520)	0.560891 (0.005314)	0.560071 (0.005281)
		β	0.279152 (0.008391)	Lin	0.339271 (0.007969)	0.339828 (0.008069)	0.339268 (0.007969)	0.334464 (0.007134)	0.3216673 (0.005136)
				Imp	0.341553 (0.007974)	0.346086 (0.008213)	0.341550 (0.007971)	0.338880 (0.007718)	0.320799 (0.007428)

The average values and the MSEs of the MLEs are presented in 4th column. Note that the MLEs are computed based on EM algorithm. We present two methods Lindley's approximation (Lin) and importance sampling (Imp) in fifth column. In 6–10th columns, the average and MSE values of the Bayes estimates with respect to the squared error, LINEX and entropy loss functions are presented. The MSE values of each estimate are placed inside the parenthesis.

Table 5: Average and MSE values of the estimates for $r(x)$ and $h(x)$.

(n,m)	CS	Parameter	EM Avg (MSE)	Method	SEL Avg (MSE)	LINEX $\omega=-0.25$ Avg (MSE)	LINEX $\omega=0.001$ Avg (MSE)	EL $\kappa=-0.5$ Avg (MSE)	EL $\kappa=0.5$ Avg (MSE)
(35,20)	Pr-IIc	$r(x)$	0.398127 (0.006980)	Lin	0.392919 (0.001572)	0.393820 (0.001364)	0.392916 (0.001571)	0.388221 (0.002977)	0.378922 (0.007040)
				Imp	0.421631 (0.002423)	0.425102 (0.002608)	0.421630 (0.002416)	0.405181 (0.002335)	0.402219 (0.002286)
		$h(x)$	0.365147 (0.014572)	Lin	0.382504 (0.003385)	0.384415 (0.003611)	0.382496 (0.003384)	0.372067 (0.002279)	0.351927 (0.000762)
				Imp	0.386712 (0.003461)	0.405312 (0.003516)	0.386710 (0.003460)	0.381207 (0.003070)	0.364315 (0.001004)
	Ty-IIc	$r(x)$	0.390138 (0.008294)	Lin	0.359881 (0.002078)	0.360923 (0.001984)	0.359876 (0.002078)	0.354171 (0.002631)	0.343389 (0.001853)
				Imp	0.361864 (0.003105)	0.367701 (0.002281)	0.361861 (0.003104)	0.346105 (0.002845)	0.331611 (0.002506)
		$h(x)$	0.396838 (0.034886)	Lin	0.459716 (0.018331)	0.462730 (0.019156)	0.459704 (0.018328)	0.444301 (0.014394)	0.411965 (0.007681)
				Imp	0.466071 (0.02160)	0.463102 (0.02377)	0.466070 (0.02159)	0.446265 (0.016492)	0.413509 (0.015423)
(,35)	Cc	$r(x)$	0.405129 (0.003941)	Lin	0.403232 (0.000497)	0.403732 (0.000299)	0.40323 (0.000495)	0.400726 (0.000224)	0.395755 (0.000142)
				Imp	0.414506 (0.000534)	0.418187 (0.000436)	0.414501 (0.000530)	0.409461 (0.000303)	0.407196 (0.000287)
		$h(x)$	0.332928 (0.005467)	Lin	0.344692 (0.000415)	0.345272 (0.000439)	0.344689 (0.000415)	0.341173 (0.000284)	0.334047 (0.000145)
				Imp	0.347236 (0.000521)	0.356043 (0.000540)	0.347230 (0.000518)	0.347991 (0.000477)	0.340564 (0.000420)
(50,45)	Pr-IIc	$r(x)$	0.409594 (0.003151)	Lin	0.409331 (0.000150)	0.409721 (0.000181)	0.409331 (0.000150)	0.407411 (0.000138)	0.403577 (0.000113)
				Imp	0.416460 (0.000213)	0.418805 (0.000227)	0.416456 (0.000212)	0.415643 (0.000186)	0.408037 (0.000172)
		$h(x)$	0.337868 (0.005060)	Lin	0.357965 (0.001132)	0.358433 (0.001163)	0.357965 (0.001132)	0.355143 (0.000950)	0.349185 (0.000618)
				Imp	0.371653 (0.001276)	0.366203 (0.001315)	0.371652 (0.001276)	0.364966 (0.001219)	0.361500 (0.001207)
	Ty-IIc	$r(x)$	0.409480 (0.003157)	Lin	0.408961 (0.000224)	0.409349 (0.000251)	0.408960 (0.000222)	0.407044 (0.000202)	0.403218 (0.000150)
				Imp	0.415008 (0.000415)	0.417648 (0.000430)	0.415000 (0.000415)	0.410651 (0.000407)	0.409472 (0.000389)
		$h(x)$	0.338443 (0.005220)	Lin	0.359309 (0.001224)	0.359779 (0.001257)	0.359307 (0.001224)	0.356457 (0.001033)	0.350424 (0.000681)
				Imp	0.376511 (0.001672)	0.385205 (0.001842)	0.376510 (0.001670)	0.364006 (0.001258)	0.362380 (0.001029)
(,50)	Cc	$r(x)$	0.409277 (0.003093)	Lin	0.421297 (0.000251)	0.421678 (0.000263)	0.421296 (0.000251)	0.419397 (0.000194)	0.415464 (0.000100)
				Imp	0.425031 (0.000271)	0.428014 (0.000284)	0.425030 (0.000269)	0.423330 (0.000253)	0.421643 (0.000234)
		$h(x)$	0.337988 (0.003660)	Lin	0.362957 (0.001492)	0.363316 (0.001520)	0.362956 (0.001492)	0.360747 (0.001327)	0.355944 (0.000910)
				Imp	0.366172 (0.001781)	0.368008 (0.001845)	0.366171 (0.001780)	0.364031 (0.001542)	0.358813 (0.001325)

Table 6 represents average lengths of 95% confidence intervals and the HPD credible intervals for $\alpha, \beta, r(x)$ and $h(x)$. We have tabulated one-sample and two-sample Bayesian prediction estimates and 95% prediction intervals for future observation in Tables 7 and 8, respectively.

Table 6: Average lengths of 95% interval estimates of α, β, $r(x)$ and $h(x)$.

(n,m)	CS		Asymptotic		Bootstrap		HPD
			NA	NL	Boot-p	Boot-t	
(35,20)	Pr-IIc	α	0.648062	0.682627	0.576212	0.543509	0.267133
		β	0.329015	0.346433	0.311642	0.300701	0.186809
		$r(x)$	0.322288	0.331721	0.277547	0.246013	0.167990
		$h(x)$	0.418470	0.441055	0.378052	0.320005	0.210808
	Ty-IIc	α	0.858970	0.922660	0.761805	0.721656	0.531964
		β	0.400653	0.426912	0.346010	0.300891	0.222437
		$r(x)$	0.338097	0.350497	0.294644	0.260079	0.190059
		$h(x)$	0.580588	0.629159	0.510660	0.493112	0.301243
(,35)	Cc	α	0.420716	0.432409	0.375604	0.310064	0.176117
		β	0.301196	0.318404	0.265203	0.228561	0.133991
		$r(x)$	0.252868	0.257088	0.210705	0.189620	0.117867
		$h(x)$	0.276372	0.284284	0.236081	0.194051	0.140660
(50,45)	Pr-IIc	α	0.427858	0.439689	0.346033	0.310446	0.188656
		β	0.360034	0.386089	0.306770	0.264126	0.160533
		$r(x)$	0.220637	0.223362	0.176136	0.143088	0.100362
		$h(x)$	0.246997	0.252462	0.184461	0.136504	0.086420
	Ty-IIc	α	0.433796	0.446078	0.379100	0.341444	0.213064
		β	0.364740	0.391709	0.310788	0.306171	0.175757
		$r(x)$	0.220169	0.222880	0.150991	0.123404	0.068944
		$h(x)$	0.250175	0.255833	0.197005	0.158817	0.100888
(,50)	Cc	α	0.374461	0.382373	0.334671	0.280062	0.133785
		β	0.292170	0.305690	0.245508	0.224999	0.145871
		$r(x)$	0.211673	0.214052	0.131106	0.108841	0.043649
		$h(x)$	0.230374	0.234809	0.169507	0.120889	0.056643

In the both sample prediction problems, the values of the parameters, hyper-parameters, ω and κ are taken same. Here, we consider $p = 1, 2, 3$ for 1st and 7th failure stages in one-sample prediction, and $j = 1, 2, 3$ for $T = m$ in two-sample prediction. From the numerical values, the following discussions can be drawn.

1. From Table 2, it is observed that with increasing values of m, values of the hyper-parameters a_1, a_2, a_3 and a_4 increase.

2. From the tabulated values in Table 4, we notice that the Bayes estimates perform better than the MLE in terms of the MSE. Further, the Bayes estimates for positive values of ω and κ are better than that for negative values of ω and κ in terms of the average values and MSEs. The simulated average values of the estimates approach towards the true value when (n, m) increases. Further, MSE decreases when (n, m) increases. Similar observation is noticed for the case of complete sample. As expected, the behavior of the Bayes estimates under SEL function and the LINEX loss function is approximately same for small values of ω (here $\omega = 0.001$). It is seen that in general, the progressive type-II censoring schemes produces better result than type-II censoring scheme in terms of the average (Avg) and MSE values. Similar behavior of the estimates of $r(x)$ and $h(x)$ (presented in Table 5) can be pointed out. The abbreviation EL is used for entropy loss function.

3. In Table 6, it is observed that for asymptotic confidence intervals, the NA method provides better estimates than NL method. For the case of bootstrap confidence intervals, Boot-t method performs better than Boot-p method. However, among the computed five intervals, HPD credible intervals give the best performance. Also, it is noticed that the average length decreases when effective sample size increases. When comparing progressive type-II censoring and type-II censoring plans, the progressive type-II plan provides better result.

4. From Table 7, we observe that the values of the predictive estimates based on progressive type-II censoring scheme are larger than that for type-II censoring scheme. The values of the predictive estimates and prediction lengths increase as i, p increase. When the effective sample size (m) increases, the predictive estimate values and predictive interval lengths decrease. Similar observation can be noticed from Table 8 for two-sample prediction problem.

Table 7: One-sample prediction values and 95% prediction intervals for future observations.

(n,m)	CS	i	p	SEL	LINEX		EL		L	U	Width
					$\omega=-0.25$	$\omega=0.001$	$\kappa=-0.5$	$\kappa=0.5$			
(35,20)	Pr-IIc	1	1	0.108452	0.117826	0.108417	0.104236	0.102013	0.033642	0.30036	0.266718
			2	0.124011	0.152041	0.124010	0.110628	0.109972	0.080750	0.421644	0.340894
			3	0.177314	0.182622	0.177306	0.154082	0.147880	0.019462	0.430405	0.410943
		7	1	0.240586	0.285261	0.240585	0.231178	0.200864	0.108235	0.671046	0.562811
			2	0.293324	0.374406	0.293319	0.274152	0.261997	0.152083	0.782603	0.630520
			3	0.315852	0.418077	0.315847	0.300919	0.281485	0.072634	0.766852	0.694218
	Ty-IIc	1	1	0.075164	0.106172	0.075158	0.052800	0.046281	0.008285	0.394668	0.386383
			2	0.108273	0.134867	0.108266	0.095076	0.052997	0.052347	0.460809	0.408462
			3	0.152972	0.160723	0.152897	0.123699	0.120758	0.089046	0.501897	0.412851
		7	1	0.192640	0.248671	0.192578	0.164284	0.128670	0.097825	0.840869	0.743044
			2	0.228068	0.278526	0.227972	0.196099	0.180907	0.130884	0.956252	0.825368
			3	0.276291	0.286070	0.276188	0.241046	0.211220	0.172691	1.068736	0.896045
(,35)	Cc	1	1	0.059483	0.080963	0.059476	0.024317	0.018605	0.016180	0.156838	0.140658
			2	0.089255	0.118047	0.089250	0.070965	0.059672	0.019941	0.192223	0.172282
			3	0.136974	0.145052	0.136970	0.114441	0.118064	0.056427	0.279560	0.223133
		7	1	0.149753	0.215093	0.149748	0.137570	0.119329	0.082594	0.275598	0.193004
			2	0.191426	0.255305	0.191422	0.158973	0.152834	0.113426	0.339898	0.226472
			3	0.248593	0.276009	0.248497	0.216852	0.208440	0.064285	0.324402	0.260117
(50,45)	Pr-IIc	1	1	0.087745	0.102351	0.087742	0.061882	0.026954	0.007920	0.183156	0.175236
			2	0.116562	0.139095	0.116558	0.098726	0.072609	0.025834	0.235675	0.209841
			3	0.150768	0.164964	0.150760	0.130999	0.109556	0.072440	0.366447	0.294007
		7	1	0.174109	0.245108	0.173981	0.167168	0.150699	0.100826	0.583458	0.482632
			2	0.228347	0.264052	0.228337	0.184623	0.159082	0.119950	0.650769	0.530819
			3	0.298067	0.345223	0.298060	0.265214	0.221704	0.075301	0.685883	0.610582
	Ty-IIc	1	1	0.067653	0.899425	0.067647	0.050715	0.041532	0.006715	0.249486	0.242771
			2	0.075989	0.125008	0.075988	0.071324	0.068227	0.028600	0.423756	0.395156
			3	0.126706	0.140764	0.123803	0.093587	0.074553	0.053428	0.464197	0.410769
		7	1	0.176572	0.207206	0.176568	0.120975	0.096408	0.031407	0.584353	0.552946
			2	0.205034	0.235607	0.205027	0.185209	0.172136	0.125855	0.758676	0.632821
			3	0.246174	0.264084	0.246172	0.208461	0.174507	0.131252	0.797154	0.665902
(,50)	Cc	1	1	0.038965	0.064027	0.038957	0.034208	0.028497	0.006783	0.116741	0.109958
			2	0.061794	0.097659	0.061788	0.058993	0.037806	0.009397	0.192057	0.182660
			3	0.106455	0.128709	0.106449	0.097808	0.070845	0.053129	0.296908	0.243779
		7	1	0.130846	0.198432	0.130840	0.100975	0.074588	0.034628	0.16496	0.130332
			2	0.174050	0.226741	0.174043	0.164317	0.130894	0.093459	0.234368	0.140909
			3	0.215686	0.231606	0.215679	0.196309	0.164320	0.116237	0.32676	0.210523

Table 8: Two-sample prediction values and 95% prediction intervals for future observations.

(n, m)	CS	j	SEL	LINEX		EL		L	U	Width
				$\omega=-0.25$	$\omega=0.001$	$\kappa=-0.5$	$\kappa=0.5$			
(35,20)	Pr-IIc	1	0.035652	0.084621	0.035649	0.029411	0.016425	0.004351	0.562487	0.558136
		2	0.118036	0.140993	0.118030	0.084403	0.042692	0.009422	0.700663	0.691241
		3	0.156008	0.162308	0.155792	0.110699	0.097183	0.013140	0.733492	0.720352
	Ty-IIc	1	0.010546	0.043119	0.010537	0.007234	0.004977	0.000782	0.712836	0.712054
		2	0.061582	0.075408	0.061561	0.038947	0.029425	0.005279	0.84835	0.843071
		3	0.109776	0.158223	0.109770	0.102432	0.092564	0.006233	0.889221	0.882988
(,35)	Cc	1	0.007642	0.009751	0.007636	0.004318	0.003707	0.000824	0.388925	0.388101
		2	0.021274	0.048030	0.021259	0.016799	0.012973	0.005348	0.441291	0.435943
		3	0.057293	0.132947	0.057288	0.033741	0.030912	0.009425	0.497781	0.488356
(50,45)	Pr-IIc	1	0.026423	0.053190	0.026418	0.024083	0.014654	0.000693	0.36197	0.361277
		2	0.080145	0.117522	0.080140	0.055920	0.046728	0.008432	0.448753	0.440321
		3	0.129506	0.127001	0.129489	0.107418	0.086947	0.026488	0.540278	0.513790
	Ty-IIc	1	0.008824	0.015283	0.008819	0.005271	0.004725	0.000707	0.484861	0.484154
		2	0.041672	0.049382	0.041672	0.028526	0.021310	0.008262	0.559655	0.551393
		3	0.010243	0.129743	0.010240	0.091253	0.058291	0.003714	0.616722	0.613008
(,50)	Cc	1	0.005281	0.006994	0.005274	0.002867	0.002173	0.000848	0.302409	0.301561
		2	0.017126	0.024907	0.017121	0.004892	0.003282	0.000437	0.361289	0.360852
		3	0.022809	0.037615	0.022800	0.019264	0.014066	0.002640	0.424486	0.421846

6.2. Real data analysis

In this subsection, we consider real life dataset representing the times to breakdown of an insulating fluid between electrodes recorded at the voltage of 34 kV (minutes) in a life test. The dataset is introduced by Nelson (2016) and used by Soliman (2005). The dataset is presented below.

0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67
4.85	6.50	7.3	8.01	8.27	12.06	31.75	32.52
33.91	36.71	72.89					

For the purpose of goodness of fit test, we consider various methods such as log-likelihood criterion, Kolmogorov-Smirnov (KS) statistic, Akaike's-information criterion (AIC), the associated second-order information criterion (AICc) and Bayesian information criterion (BIC). The values of the MLEs and the five goodness of fit test statistics are presented in Table 9.

Table 9: The MLEs, KS, log-likelihood, AIC, AICc and BIC values for the real dataset.

Distribution	MLEs		KS	$\ln L$	BIC	AICc	AIC
	Shape	Scale					
G-MR(α, β)	$\hat{\alpha} = 0.795210$	$\hat{\beta} = 2.392015$	0.135509	-70.34277	146.5744	145.4355	144.6855
HL(λ)		$\hat{\lambda} = 0.088745$	0.332880	-71.97299	146.8904	146.1813	145.946
IExpHL(α, θ)	$\hat{\alpha} = 0.426676$	$\hat{\theta} = 0.801178$	0.264552	-74.03980	153.9685	152.8296	152.0796
IW(α, λ)	$\hat{\alpha} = 2.038295$	$\lambda = 1.119888$	0.329144	-75.25765	156.4042	155.2653	154.5153
GF $(\alpha, \lambda, \sigma)$	$\hat{\alpha} = 7.465586$, $\hat{\lambda} = 0.354321$	$\hat{\sigma} = 7.260802$	0.667178	-95.1172	199.0677	197.8344	196.2344

The numerical values in Table 9 suggest that the gamma-mixed Rayleigh (G-MR) distribution fits the data well compared to the half-logistic (HL), inverted exponentiated half-logistic (IExpHL), inverse Weibull (IW) and generalized Fréchet (GF) distributions. Now, we compute the proposed estimates for the unknown parameters, reliability and hazard functions. In Table 10, we consider progressive type-II censored sample with total sample size $n = 19$, the failure sample size $m = 14$. We adopt various schemes for the purpose of computation. Here, we consider three schemes say Pr-IIc, Ty-IIc and Cc, that is $(R_1, R_2, ..., R_m)$ =(0*6,5,0*7), (0*13,5) and (0*19), respectively. Note that (0*3,2) denotes the censoring scheme $(0, 0, 0, 2)$.

Table 10: Progressive type-II censored data for the real dataset.

i	1	2	3	4	5	6	7
$x_{i:m:n}$	0.19	0.78	1.31	2.78	4.15	4.67	4.85
i	8	9	10	11	12	13	14
$x_{i:m:n}$	8.01	8.27	12.06	31.75	33.91	36.71	72.89

We take all the hyperparameter values as zero. We present the average values of the proposed estimates of α, β in Table 11 and $r(x)$ and $h(x)$ in Table 12. Table 13 represents 95% interval estimates of α, β, $r(x)$ and $h(x)$. Further, we have tabulated one-sample and two-sample predicted values and 95% prediction intervals in Tables 14 and 15, respectively. Here, we obtain one-sample prediction estimates of the lifetime of first three units at i-th failure and two-sample prediction estimates of the lifetime of first three units and size of sample $T = 10$. The plots of the probability density functions of five different models and histogram for real dataset are presented in Figure 2. In Figures 3 and 4, the plots of the density, distribution, reliability and hazard functions of gamma-mixed Rayleigh distribution under Pr-IIc, Ty-IIc and Cc schemes are depicted.

Table 11: Estimates of α and β for the real dataset.

(n,m)	CS		EM	Method	SEL	LINEX		EL	
						$\omega = -0.25$	$\omega = 0.001$	$\kappa = -0.5$	$\kappa = 0.5$
			Avg		Avg	Avg	Avg	Avg	Avg
(19,14)	Pr-IIc	α	0.776441	Lin	0.399239	0.399847	0.399239	0.411029	0.437988
				Imp	0.367051	0.376001	0.367050	0.381976	0.387007
		β	3.356965	Lin	1.362572	1.455717	1.362637	1.463404	1.652837
				Imp	1.335429	1.389600	1.335427	1.412632	1.486753
	Ty-IIc	α	0.294786	Lin	0.206869	0.207328	0.206867	0.205371	0.205485
				Imp	0.176532	0.193725	0.176530	0.162305	0.167035
		β	1.12691	Lin	0.502474	0.537084	0.502364	0.495457	0.531081
				Imp	0.464582	0.499007	0.464578	0.446396	0.468757
(,19)	Cc	α	0.795210	Lin	0.570614	0.575673	0.570595	0.563006	0.558460
				Imp	0.523781	0.568766	0.523780	0.512525	0.504817
		β	2.392015	Lin	1.465001	1.477988	1.464669	1.429069	1.429344
				Imp	1.400864	1.459764	1.400860	1.387562	1.385258

Table 12: Estimates of $r(x)$ and $h(x)$ for the real dataset.

(n, m)	CS		EM	Method	SEL	LINEX		EL	
						$\omega = -0.25$	$\omega = 0.001$	$\kappa = -0.5$	$\kappa = 0.5$
			Avg		Avg	Avg	Avg	Avg	Avg
(19,14)	Pr-IIc	$r(x)$	0.931769	Lin	0.867124	0.868877	0.867125	0.867672	0.868731
				Imp	0.846209	0.847669	0.846207	0.847008	0.847460
		$h(x)$	0.086149	Lin	0.152817	0.152628	0.152818	0.151136	0.148734
				Imp	0.186422	0.187994	0.186421	0.177537	0.153480
	Ty-IIc	$r(x)$	0.860480	Lin	0.773932	0.773973	0.773932	0.773969	0.774245
				Imp	0.748209	0.748867	0.748208	0.748452	0.748666
		$h(x)$	0.125622	Lin	0.132865	0.133130	0.132864	0.128547	0.120052
				Imp	0.164263	0.168117	0.164261	0.123786	0.119007
(,19)	Cc	$r(x)$	0.876466	Lin	0.801888	0.801855	0.801888	0.802035	0.802425
				Imp	0.774826	0.780074	0.774824	0.775314	0.776174
		$h(x)$	0.149630	Lin	0.211578	0.211800	0.211577	0.206734	0.189875
				Imp	0.230761	0.236482	0.230760	0.214776	0.206782

Table 13: 95% interval estimates of α, β, $r(x)$ and $h(x)$ based on the real dataset.

(n, m)	CS		Asymptotic		Bootstrap		HPD
			NA	NL	Boot-p	Boot-t	
(19,14)	Pr-IIc	α	(0.03697, 1.5159)	(0.29957, 2.01244)	(0.11275, 1.32586)	(0.08563, 1.16478)	(0.00794, 0.86115)
		β	(0.00000, 7.20239)	(1.06772, 10.5544)	(2.86452, 9.98328)	(3.05617, 9.06852)	(0.09908, 4.43365)
		$r(x)$	(0.83953, 1.02401)	(0.84394, 1.02873)	(0.80581, 0.95486)	(0.82068, 0.96246)	(0.84105, 0.94857)
		$h(x)$	(0.00000, 0.19222)	(0.02515, 0.29512)	(0.05946, 0.24653)	(0.07563, 0.23213)	(0.07884, 0.18947)
	Ty-IIc	α	(0.08674, 0.50284)	(0.14554, 0.59706)	(0.10466, 0.47148)	(0.09462, 0.41656)	(0.12035, 0.2984)
		β	(0.00000, 2.65557)	(0.29025, 4.37535)	(0.12630, 2.24429)	(0.21364, 2.06421)	(0.41286, 1.51018)
		$r(x)$	(0.68785, 1.03311)	(0.70406, 1.05165)	(0.62482, 0.90287)	(0.66946, 0.88035)	(0.72451, 0.87172)
		$h(x)$	(0.03414, 0.21710)	(0.06065, 0.26020)	(0.07567, 0.22233)	(0.08745, 0.19816)	(0.10526, 0.1824)
(,19)	Cc	α	(0.21443, 1.37599)	(0.38309, 1.65070)	(0.33456, 1.43563)	(0.27664, 1.35645)	(0.42784, 1.31559)
		β	(0.00000, 4.81147)	(0.86994, 6.57719)	(1.01503, 4.92624)	(1.21536, 4.02611)	(1.36485, 2.53242)
		$r(x)$	(0.73417, 1.01876)	(0.74512, 1.03096)	(0.68356, 0.90118)	(0.70599, 0.88412)	(0.75086, 0.88837)
		$h(x)$	(0.00223, 0.29703)	(0.05587, 0.40071)	(0.04312, 0.25221)	(0.10537, 0.26529)	(0.13458, 0.25346)

Table 14: One-sample prediction values and 95% prediction intervals for future observations.

(n,m)	CS	i	p	SEL	LINEX		EL		BPI
					$\omega=-0.25$	$\omega=0.001$	$\kappa=-0.5$	$\kappa=0.5$	
(19,14)	Pr-IIc	1	1	4.864253	5.283172	4.864247	4.064281	3.642513	(3.410321, 5.572681)
			2	5.123467	6.017240	5.123460	4.618267	4.042580	(3.843775, 6.333201)
			3	5.507412	6.724315	5.507403	4.970824	4.386421	(4.210341, 7.262675)
		5	1	7.129901	9.058243	7.129882	6.527802	6.201358	(5.462305, 10.305467)
			2	8.105347	11.25680	8.105339	7.201634	7.053412	(6.76250, 11.738928)
			3	8.753105	12.88402	8.753087	7.836405	7.643187	(7.335213, 12.599454)
	Ty-IIc	1	1	2.560428	3.526411	2.557964	2.134526	2.074091	(1.761855, 3.837296)
			2	4.068825	4.672553	4.068822	2.760048	2.496253	(2.142392, 5.058906)
			3	4.748263	4.958240	4.748257	3.209992	2.897385	(2.437218, 5.321908)
		5	1	3.580742	5.272538	3.580739	3.336142	2.582773	(2.161007, 5.509298)
			2	5.23189	6.856842	5.231817	4.219138	4.000876	(3.496125, 7.196189)
			3	7.100923	7.594625	7.100916	6.735364	5.436582	(4.430564, 8.493101)
(,19)	Cc	1	1	2.182731	2.382467	2.182726	1.766528	1.623407	(0.942521, 2.676136)
			2	3.067269	3.631854	3.067261	2.578532	2.247716	(1.432610, 4.259329)
			3	3.854727	5.582428	3.854718	2.891759	2.374685	(1.796821, 6.011103)
		5	1	2.978550	4.297582	2.978544	2.432080	2.178441	(1.800672, 4.810808)
			2	3.352725	5.317162	3.352716	2.70553	2.484336	(2.134255, 5.565777)
			3	5.034900	5.924856	5.033875	3.885664	3.100858	(2.704073, 7.589788)

Table 15: Two-sample prediction values and 95% prediction intervals for future observations.

(n,m)	CS	j	SEL	LINEX		EL		BPI
				$\omega=-0.25$	$\omega=0.001$	$\kappa=-0.5$	$\kappa=0.5$	
(19,14)	Pr-IIc	1	1.724826	2.074824	1.724813	1.376145	1.265064	(1.077346, 2.385841)
		2	2.064354	2.467026	2.064348	1.750517	1.305119	(1.213075, 2.826527)
		3	4.846177	3.794121	4.846169	3.526443	3.344056	(2.19428, 4.92259)
	Ty-IIc	1	1.255674	1.699764	1.255670	1.096628	0.946491	(0.631672, 1.886354)
		2	1.462812	2.152316	1.462805	1.224056	1.064583	(0.816055, 2.28461)
		3	2.803286	3.45821	2.803179	2.437182	2.175564	(1.430592, 3.802253)
(,19)	Cc	1	0.860765	1.180676	0.860761	0.681231	0.620711	(0.462854, 1.257712)
		2	1.113446	1.445251	1.113437	0.846616	0.726489	(0.371066, 1.623857)
		3	2.045084	2.152647	2.045076	1.615233	1.086961	(0.794823, 2.343845)

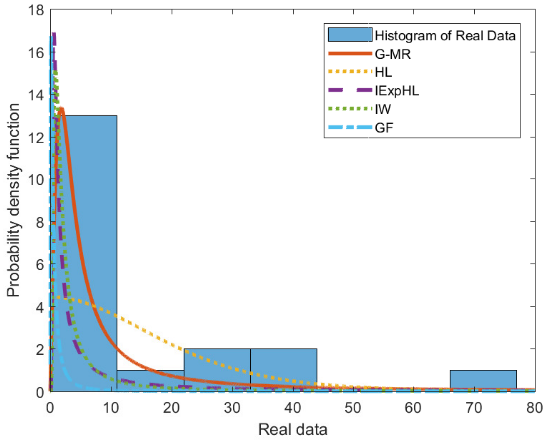


Figure 2: The histogram of the real dataset and the plots of the probability density functions of the fitted G-MR, HL, IExpHL, IW, GF models.

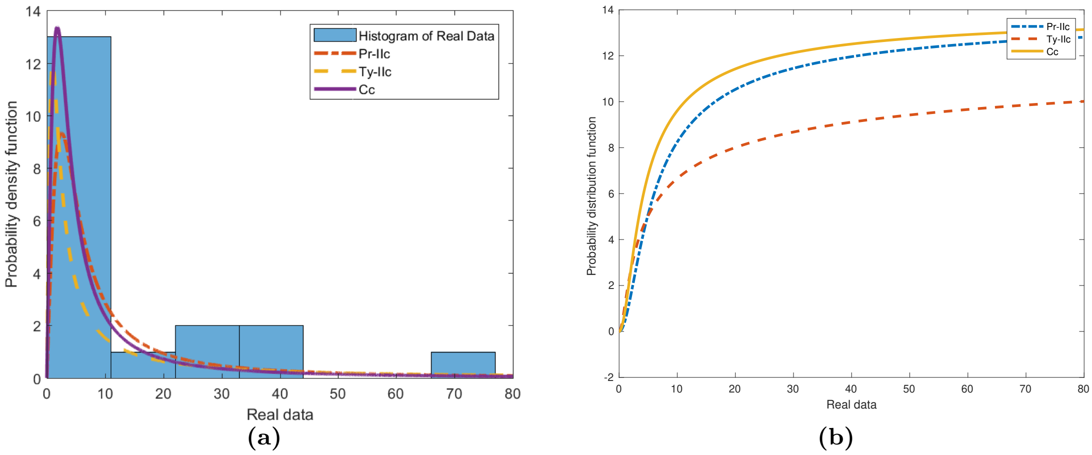


Figure 3: The plots of the (a) density and (b) distribution functions of the gamma-mixed Rayleigh distribution based on different censoring schemes.

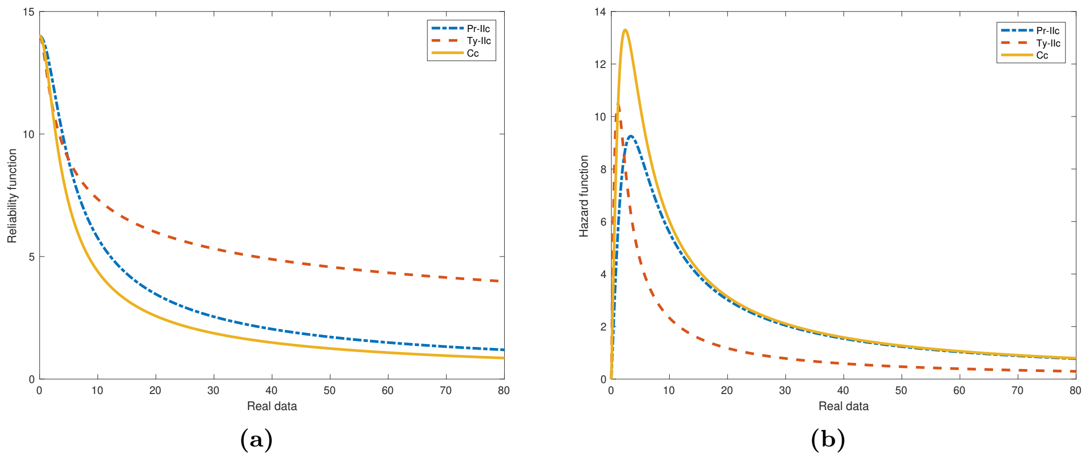


Figure 4: The plots of the (a) reliability and (b) hazard functions based on different censoring schemes.

7. CONCLUDING REMARKS

In industrial life tests, reliability analysis and clinical trials, the type-II progressive censoring methodology has become quite popular for analyzing lifetime data. It allows for random removals of the remaining survival units at each failure time. In this article, we considered inference and prediction problems for the gamma-mixed Rayleigh distribution when progressive type-II censored sample is available. We obtained conditions under which the MLEs exist and are unique, then derived the MLEs using EM algorithm. The Bayes estimates have been computed with respect to three loss functions, such as squared error, LINEX and entropy loss functions. Two approximations say Lindley approximation and importance sampling method have been used for the computation of the Bayes estimates. We also derived confidence and credible intervals using various methods. Specifically, we have obtained asymptotic, bootstrap-p and bootstrap-t confidence intervals and highest posterior density credible interval. Further, we discussed Bayesian prediction problems. One-sample and two-sample prediction problems have been considered. An elaborate simulation study was conducted for the comparison of the proposed estimates. From the simulation study, it has been observed that the Bayes estimates perform better than the MLEs in terms of the MSE values. Further, the Bayes estimates for positive values of ω and κ are better than that for negative values of ω and κ in terms of the average values and MSEs. For the present problem, we recommend the Bayes estimates to use for the case of point estimation. It has been observed that for asymptotic confidence intervals, the NA method provides better estimates than NL method in the sense of the average lengths. For the case of bootstrap confidence intervals, Boot-t method performs better than Boot-p method. However, among the computed five intervals, HPD credible intervals give the best performance. Among all the interval estimates, we recommend HPD credible interval estimate. In addition to these, we have also computed predictive estimates. It has been noticed that when the effective sample size increases, the predictive estimates and predictive interval lengths decrease. Finally, we considered a real life dataset representing the times to breakdown of an insulating fluid between electrodes recorded at the voltage of 34 kV (minutes) in a life test for illustrative purposes.

ACKNOWLEDGMENTS

The authors would like to thank the Editor-in-chief, an associate editor and an anonymous referee for careful reading and valuable comments, which have led to this improved version. The authors also wish to thank Dr. Sugata Gangopadhyay for his valuable suggestions while preparing the draft of this manuscript. One of the authors Kousik Maiti thanks the Ministry of Electronics & Information Technology (MeitY), Government of India for the financial assistantship received to carry out this research work.

REFERENCES

Albalawi, O., Kabdwal, N. C., Azhad, Q. J., Hora, R., and Alsaedi, B. S. (2022). Estimation of the generalized logarithmic transformation exponential distribution under progressively Type-II censored data with application to the COVID-19 mortality rates. *Mathematics*, 10(7):1015.

Arabi, B. R., Noori, A. M., Gurunlu, A. O., Singh, S., and Vasfi, M. (2019). Estimation and prediction for the Poisson-exponential distribution based on Type-II censored data. *American Journal of Mathematical and Management Sciences*, 38(1):96–115.

Arnold, B. C. and Press, S. J. (1983). Bayesian inference for Pareto populations. *Journal of Econometrics*, 21(3):287–306.

Bdair, O., Awwad, R. A., Abufoudeh, G., and Naser, M. (2019). Estimation and prediction for flexible Weibull distribution based on progressive Type II censored data. *Communications in Mathematics and Statistics*, `https://doi.org/10.1007/s40304-018-00173-0`, 1–23.

Cancho, V. G., Louzada, N. F., and Barriga, G. D. (2011). The Poisson-exponential lifetime distribution. *Journal of Econometrics*, 55(1):677–686.

Chen, M. H. and Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics*, 8(1):69–92.

Cohen, A. C. (2016). *Truncated and Censored Samples: Theory and Applications*, CRC press.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–38.

Dey, S., Nassar, M., Maurya, R. K., and Tripathi, Y. M. (2018). Estimation and prediction of Marshall–Olkin extended exponential distribution under progressively type-II censored data. *Journal of Statistical Computation and Simulation*, 88(12):2287–2308.

Huang, S. R. and Wu, S. J. (2012). Bayesian estimation and prediction for Weibull model with progressive censoring. *Journal of Statistical Computation and Simulation*, 82(11):1607–1620.

Kayal, T., Tripathi, Y. M., Singh, D. P., and Rastogi, M. K. (2017). Estimation and prediction for Chen distribution with bathtub shape under progressive censoring. *Journal of Statistical Computation and Simulation*, 87(2):348–366.

Khan, R. A. and Mitra, M. (2019). Estimation issues in the Exponential–Logarithmic model under hybrid censoring. *Statistical Papers*, 1–32.

Klein, J. P. and Moeschberger, M. L. (1997). Statistics for biology and health. *Stat. Biol. Health*, New York, 27238.

Kundu, D. (2008). Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics*, 50(2):144–154.

Kundu, D. and Pradhan, B. (2009). Estimating the parameters of the generalized exponential distribution in presence of hybrid censoring. *Communications in Statistics – Theory and Methods*, 38(12):2030–2041.

Kundu, D. and Raqab, M. Z. (2015). Estimation of $R = P[Y < X]$ for three-parameter generalized Rayleigh distribution. *Journal of Statistical Computation and Simulation*, 85(4):725–739.

Lawless, J. F. (2011). *Statistical Models and Methods for Lifetime Data*. Vol. 362, John Wiley & Sons.

Lee, K. and Cho, Y. (2017). Bayesian and maximum likelihood estimations of the inverted exponentiated half logistic distribution under progressive Type II censoring. *Journal of Applied Statistics*, 44(5):811–832.

Lindley, D. V. (1980). Approximate Bayesian methods. *Trabajos de Estadistica Y de Investigacion Operativa*, 31(1):223–245.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):226–233.

Maiti, K. and Kayal, S. (2019). Estimation for the generalized Fréchet distribution under progressive censoring scheme. *International Journal of System Assurance Engineering and Management*, 10(5):1276–1301.

Maiti, K. and Kayal, S. (2020). Estimating reliability characteristics of the log-logistic distribution under progressive censoring with two applications. *Annals of Data Science*, `https://doi.org/10.1007/s40745-020-00292-y`, 1–40.

Maiti, K. and Kayal, S. (2021). Estimation of parameters and reliability characteristics for a generalized Rayleigh distribution under progressive type-II censored sample. *Communications in Statistics – Simulation and Computation*, 50(11):3669–3698.

Muhammed, H. Z. and Almetwally, E. M. (2020). Bayesian and non-Bayesian estimation for the bivariate inverse Weibull distribution under progressive type-II censoring. *Annals of Data Science*, `https://doi.org/10.1007/s40745-020-00316-7`, 1–32.

Nelson, W. B. (2016). *Applied Life Data Analysis*, Vol. 577, John Wiley & Sons.

Ng, H., Chan, P., and Balakrishnan, N. (2002). Estimation of parameters from progressively censored data using EM algorithm. *Computational Statistics & Data Analysis*, 39(4):371–386.

Nik, A. S., Asgharzadeh, A., and Raqab, M. Z. (2021). Estimation and prediction for a new Pareto-type distribution under progressive type-II censoring. *Mathematics and Computers in Simulation*, 190:508–530.

Rastogi, M. K. and Tripathi, Y. M. (2014). Parameter and reliability estimation for an exponentiated half-logistic distribution under progressive type II censoring. *Journal of Statistical Computation and Simulation*, 84(8):1711–1727.

Sarhan, A. M., Abd, E. A. A., and Alasbahi, I. A. (2013). Exponentiated generalized linear exponential distribution. *Applied Mathematical Modelling*, 37(5):2838–2849.

Soliman, A. A. (2005). Estimation of parameters of life from progressively censored data using Burr-XII model. *IEEE Transactions on Reliability*, 54(1):34–42.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 47–52
https://doi.org/10.57805/revstat.v23i1.530

On a Characterization of Exponential and Double Exponential Distributions

Authors: Reza Rastegar
– Occidental Petroleum Corporation,
Houston, TX 77046, USA
reza_rastegar2@oxy.com

Alexander Roitershtein ✉
– Department of Mathematics, Texas A&M University,
College Station, TX 77843, USA
roiterst@tamu.edu

Received: September 2022 Revised: December 2022 Accepted: January 2023

Abstract:

- Recently, G. Yanev (2020) obtained a characterization of the exponential family of distributions in terms of a functional equation for certain mixture densities. The purpose of this note is twofold: we extend Yanev's theorem by relaxing a restriction on the sign of mixture coefficients and, in addition, obtain a similar characterization for the Laplace family of distributions.

Keywords:

- *hypoexponential distribution; Laplace distribution; characterization of distributions; sums of independent random variables; characteristic functions; functional equations.*

AMS Subject Classification:

- 62E10, 60G50, 60E10.

✉ Corresponding author

1. INTRODUCTION

Our aim is to prove certain characterizations of the exponential and double exponential families of distributions. We will use the notation $X \sim \mathcal{E}(\lambda)$ to indicate that X is an exponential random variable with parameter $\lambda > 0$, that is $P(X > x) = e^{-\lambda x}$ for all $x > 0$. We will write $X \in \mathcal{E}$ if $X \sim \mathcal{E}(\lambda)$ for some $\lambda > 0$. Similarly, will write $X \in \mathcal{L}$ if X has a Laplace (double exponential) distribution (Kotz *et al.*, 2001), that is, for some $\lambda > 0$ and $Y \sim \mathcal{E}(\lambda)$,

$$P(X > x) = \frac{1}{2}\Big(P(Y > x) + P(-Y > x)\Big) = \frac{\lambda}{2}\int_x^\infty e^{-\lambda|y|}\,dy, \qquad \forall\, x \in \mathbb{R}.$$

For the exponential random variable we have:

Theorem 1.1. *Let X be a random variable and $\mu_1, ..., \mu_n$ be distinct non-zero real numbers. Let $\varphi(t) = E\big(e^{itX}\big)$, $t \in \mathbb{R}$, be the characteristic function of X, and suppose that φ is infinitely differentiable at zero and, furthermore,*

$$\prod_{k=1}^n \varphi(\mu_k t) = \sum_{k=1}^n \theta_k \varphi(\mu_k t), \qquad t \in \mathbb{R}, \tag{1.1}$$

where

$$\theta_k = \prod_{j=1, j\neq k}^n \frac{\mu_k}{\mu_k - \mu_j}, \qquad k = 1, ..., n. \tag{1.2}$$

If, in addition,

$$\sum_{(k_1,...,k_n)\in W_{n,m}} \prod_{j=1}^n \mu_j^{k_j} \neq \sum_{k=1}^n \mu_k^m \quad \text{for any integer } m \geq 2, \tag{1.3}$$

where

$$W_{n,m} := \left\{(k_1, ..., k_n) \in \mathbb{Z}^n : k_j \geq 0 \text{ and } \sum_{j=1}^n k_j = m\right\}, \tag{1.4}$$

then, either $P(X = 0) = 1$ or $E(X) \neq 0$ and $X \cdot \operatorname{sign}\big(E(X)\big) \sim \mathcal{E}(\lambda)$ with $\lambda = 1/E(X)$.

The proof of the theorem is given in Section 2. Theorem 1.1 is an extension of a similar result of G. Yanev (2020) obtained under the additional assumption that the coefficients μ_k are positive. In that case, the key technical condition (1.3) is trivial as the left-hand sides contains the μ_k^m terms and hence is always larger than the right-hand side.

To ensure the existence of the derivatives of φ at zero one can impose Cramér's condition, namely assume that there is $t_0 > 0$ such that $E\big(e^{tX}\big) < \infty$ for all $t \in (-t_0, t_0)$. Note also that the equality in (1.3) for any fixed $m \in \mathbb{N}$ describes a low-dimensional manifold in $\mathbb{R}^n$, and hence Theorem 1.1 is true for almost every vector $(\mu_1, ..., \mu_n)$ chosen at random from a continuous distribution on $\mathbb{R}^n$.

The identity in (1.1) with θ_k introduced in (1.2) holds for any $X \in \mathcal{E}$, and Theorem 1.1 can be seen as a converse to this result. Let $X_1, ..., X_n$, $n \geq 2$, be independent copies of a random variable X. Equations (1.1) and (1.2) give an expression of the characteristic function of the sum

$$S = \mu_1 X_1 + \cdots + \mu_n X_n \tag{1.5}$$

as a linear combination of $\varphi(\mu_k t)$'s. If $X \sim \mathcal{E}(\lambda)$, then $\varphi(t) = \frac{\lambda}{\lambda - it}$, and thus (1.1) is the partial fraction decomposition of the complex-valued rational function $\psi(t) := E\big(e^{itS}\big)$. In the particular case when $X \in \mathcal{E}$ and $\mu_k = \frac{1}{L-k+1}$ for some integer $L > n$, the random variable S/λ is distributed as the n-th order statistic of a sample of L independent copies of X (this is the Rényi representation of order statistics; see, for instance, David and Nagaraja, 2004, p. 18). For further background and earlier versions (particular cases) of Yanev's characterization theorem see Arnold and Villasenor (2013), Milošević and Obradović (2016) and Yanev (2020).

It was pointed out in Yanev (2020) that an extension of their result to a more general class of coefficients μ_k would be of interest from the viewpoint of both theory and applications[1]. When all the coefficients μ_k are positive and X is an exponential random variable, the random variable $S = \sum_{k=1}^{n} \mu_k X_k$ has a hypoexponential distribution. When some of the coefficients are negative, S is a difference of two hypoexponential random variables. Some applications of such differences are discussed, for instance, in Li and Li (2019). An insightful theoretical exploration of the densities of hypoexponential distributions can be found in Belton *et al.* (2022).

We remark that the theorem is not true if the particular form of the coefficients θ_k in (1.2) is not enforced. For instance, for the Laplace distribution we have:

Theorem 1.2. *Let X be a random variable and $\mu_1, ..., \mu_n$ be distinct positive numbers. Let $\varphi(t) = E\big(e^{itX}\big)$, $t \in \mathbb{R}$, be the characteristic function of X, and suppose that φ is infinitely differentiable at zero and, furthermore, (1.1) holds with*

$$\theta_k = \prod_{j=1, j \neq k}^{n} \frac{\mu_k^2}{\mu_k^2 - \mu_j^2}, \qquad k = 1, ..., n. \tag{1.6}$$

Then, either $P(X = 0) = 1$ or X has a Laplace distribution.

The result is closely related to the one stated in Theorem 1.1 because $X \in \mathcal{L}$ implies that for a suitable $Y \in \mathcal{E}$,

$$E\big(e^{itX}\big) = \frac{1}{2}\Big(E\big(e^{itY}\big) + E\big(e^{-itY}\big)\Big).$$

The proof of the theorem is similar to that of Theorem 1.1, and therefore is omitted. The key technical ingredient of the proof, namely an analogue of Lemma 2.1 for Laplace distributions, follows immediately from Lemma 2-(iii) in Yanev (2020), and the rest of the proof of Theorem 1.1 can be carried over verbatim to the double exponential setup of Theorem 1.2.

We conclude the introduction with a brief discussion of condition (1.3). The equality with $n = 2$ and some $m \geq 2$ reads $\sum_{j=0}^{m} \mu_1^j \mu_2^{m-j} = \mu_1^m + \mu_2^m$, which is equivalent to

[1] See also a recent preprint (Yanev, 2022), where a special case of exponential convolutions with repeated coefficients is considered.

$\frac{\mu_1^{m+1}-\mu_2^{m+1}}{\mu_1-\mu_2} = \mu_1^m + \mu_2^m$. The latter implies that $\mu_2^{m-1} = \mu_1^{m-1}$, and hence m is odd and $\mu_2 = -\mu_1$. In that case, (1.1) becomes

$$\varphi(t)\varphi(-t) = \frac{1}{2}\big(\varphi(t) + \varphi(-t)\big), \qquad t \in \mathbb{R}. \tag{1.7}$$

The equation is satisfied when X is a Bernoulli random variable with $P(X=0) = P(X=a) = \frac{1}{2}$ for some constant $a > 0$, in which case $\varphi(t) = \frac{1}{2}\big(1 + e^{iat}\big)$. More generally, (1.7) holds if and only if $\varphi(t) = \frac{1}{2}\big(1 + e^{i\rho(t)}\big)$, where $\rho : \mathbb{R} \to \mathbb{R}$ is an odd function. Unfortunately, we are not aware of any example where φ in this form would be a characteristic function of a random variable beyond the linear case $\rho(t) = at$ and linear fractional $\rho(t) = \frac{\lambda - ti}{\lambda + ti}$, $\rho(t) = \frac{\lambda + ti}{\lambda - ti}$ which correspond to, respectively, $X \in \mathcal{E}(\lambda)$ and $-X \in \mathcal{E}(\lambda)$.

Our proof technique differs significantly from the one used in Yanev (2020). However, interestingly enough, both rely on the validity of (1.3). Nevertheless, we believe that the following might be true:

Conjecture. For $n \geq 3$, (1.3) is an artifact of the proof and is not necessary.

2. PROOF OF THEOREM 1.1

The following is a suitable version of Lemma 2-(iii) in Yanev (2020).

Lemma 2.1. *Assume (1.3). Then, for any integer $m \geq 2$,*

$$\sum_{k=1}^{n} \theta_k \mu_k^m \neq \sum_{k=1}^{n} \mu_k^m.$$

Proof of Lemma 2.1: Let $Y_1, ..., Y_n$ be i. i. d. random variables such that $Y_k \in \mathcal{E}(1)$ for each $k = 1, ..., n$. Similarly to (1.5), define

$$\widetilde{S} = \mu_1 Y_1 + \cdots + \mu_n Y_n.$$

Let $\widetilde{\varphi}(t) = E\big(e^{itY_1}\big) = \frac{1}{1-it}$. Thus, for $E\big(e^{i\widetilde{S}t}\big) = \prod_{k=1}^{n} \frac{1}{1-it\mu_k}$ we have a partial fraction decomposition similar to (1.1), namely

$$E\big(e^{i\widetilde{S}t}\big) = \sum_{k=1}^{n} \frac{\theta_k}{1 - it\mu_k} = \sum_{k=1}^{n} \theta_k \widetilde{\varphi}(\mu_k t), \qquad t \in \mathbb{R},$$

where the coefficients θ_k are introduced in (1.2). Differentiating m times and taking in account that $E(Y_1^m) = m!$, we obtain

$$(-i)^m \frac{d^m}{dt^m} E\big(e^{i\widetilde{S}t}\big)\Big|_{t=0} = E\big(\widetilde{S}^m\big) = \sum_{k=1}^{n} \theta_k \mu_k^m E\big(Y_1^m\big) = m! \sum_{k=1}^{n} \theta_k \mu_k^m.$$

Recall $W_{n,m}$ from (1.4). Using the multinomial expansion

$$\widetilde{S}^m = \big(\mu_1 Y_1 + \cdots + \mu_n Y_n\big)^m = \sum_{(k_1,...,k_n) \in W_{n,m}} \frac{m!}{k_1! \cdots k_n!} \prod_{j=1}^{n} (\mu_j Y_j)^{k_j},$$

and the fact that $E(Y_1^k) = k!$ for any $k \in \mathbb{N}$, we obtain:

$$\begin{aligned} \sum_{k=1}^{n} \theta_k \mu_k^m &= \frac{1}{m!} E\big(\widetilde{S}^m\big) = \sum_{(k_1,\dots,k_n)\in W_{n,m}} \frac{1}{k_1!\cdots k_n!} \prod_{j=1}^{n} \mu_j^{k_j} E\big(Y_j^{k_j}\big) \\ &= \sum_{(k_1,\dots,k_n)\in W_{n,m}} \prod_{j=1}^{n} \mu_j^{k_j}, \end{aligned} \tag{2.1}$$

which yields the result in view of (1.3). □

Differentiating both sides of (1.1) m times we obtain the identity

$$\frac{d^m}{dt^m} \prod_{k=1}^{n} \varphi(\mu_k t)\bigg|_{t=0} = \sum_{k=1}^{n} \theta_k \mu_k^m \varphi^{(m)}(0), \qquad m \geq 2. \tag{2.2}$$

In view of Lemma 2.1 and the fact that $\varphi(0) = 1$, these identities can be used to determine all the derivatives of φ at zero in terms of $\varphi'(0)$, first $\varphi''(0)$ in terms of the parameter $\varphi'(0)$, then $\varphi'''(0)$ in terms of $\varphi'(0)$ and $\varphi''(0)$, and hence in terms of $\varphi'(0)$ only, and so on. For instance, (2.2) with $m = 2$ yields

$$\varphi''(0) \sum_{k=1}^{n} \mu_k^2(\theta_k - 1) = \big(\varphi'(0)\big)^2 \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mu_k \mu_j.$$

Note that $\sum_{k=1}^{n} \mu_k^2(\theta_k - 1) \neq 0$ by Lemma 2.1 with $m = 2$.

In general, for an arbitrary $m \in \mathbb{N}$, (2.2) can be written as

$$\varphi^{(m)}(0) \sum_{k=1}^{n} \mu_k^m(\theta_k - 1) = F_m\big(\mu_1, \dots, \mu_n, \varphi'(0), \dots, \varphi^{(m-1)}(0)\big),$$

where the multivariate functional $F_m(\cdot)$ in the right-hand side is independent of $\varphi^{(m)}(0)$. An explicit form of F_m is given by the general Leibnitz rule (an extension of the product differentiation rule to higher derivatives):

$$F_m = \sum_{(k_1,\dots,k_n)\in V_{n,m}} \frac{m!}{k_1!\cdots k_n!} \prod_{j=1}^{n} \mu_j^{k_j} \varphi^{(k_j)}(0),$$

where (cf. (1.4))

$$V_{n,m} := \left\{ (k_1, \dots, k_n) \in \mathbb{Z}^n : \ 0 \leq k_j < m \ \text{ and } \sum_{j=1}^{n} k_j = m \right\}.$$

In view of Lemma 2.1, $\sum_{k=1}^{n} \mu_k^m(\theta_k - 1) \neq 0$, and thus we can define the derivative of φ at zero inductively, using the formula

$$\varphi^{(m)}(0) = \frac{F_m\big(\mu_1, \dots, \mu_n, \varphi'(0), \dots, \varphi^{(m-1)}(0)\big)}{\sum_{k=1}^{n} \mu_k^m(\theta_k - 1)}.$$

Let now $Z \in \mathcal{E}$ and $\psi(t) = E\big(e^{itZ}\big)$. The derivatives $\psi''(0), \psi'''(0), \dots$ as functions of the parameter $\psi'(0)$ can be in principle derived using the same inductive algorithm. Therefore, $\varphi'(0) = 0$ implies $P(X = 0) = 1$ while $\varphi'(0) = \psi'(0) = \lambda^{-1}$ for some $\lambda > 0$ implies that $\varphi^{(m)}(0) = \psi^{(m)}(0)$ for all $m \in \mathbb{N}$, and hence (since φ is analytic under the conditions of the theorem) $\varphi(t) = \psi(t) = \frac{\lambda}{\lambda - it}$ as desired. Finally, the case $\varphi(0) = -\lambda^{-1} < 0$ can be reduced to the previous one by switching from X to $-X$ in the above argument. □

REFERENCES

Arnold, B. C. and Villasenor, J. A. (2013). Exponential characterizations motivated by the structure of order statistics in samples of size two. *Statist. Probab. Lett.*, 83:596–601.

Belton, A., Guillot, D., Khare, A., and Putinar, M. (2022). Hirschman-Widder densities. *Appl. Comput. Harmon. Anal.*, 60:396–425.

David, H. A. and Nagaraja, H. N. (2004). *Order Statistics*. Wiley.

Kotz, S., Kozubowski, T. J., and Podgórski, K. (2001). *The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser.

Li, K. H. and Li, C. T. (2019). Linear combination of independent exponential random variables. *Methodol. Comput. Appl. Probab.*, 21:253–277.

Milošević, B. and Obradović, M. (2016). Some characterizations of the exponential distribution based on order statistics. *Appl. Anal. Discrete Math.*, 10:394–407.

Yanev, G. P. (2020). Exponential and hypoexponential distributions: some characterizations. *Mathematics*, 8:2207.

Yanev, G. P. (2022). On characterization of the exponential distribution via hypoexponential distributions. `arXiv: 2204.00867`.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 53–77
https://doi.org/10.57805/revstat.v23i1.501

Optimal Imputation Methods under Stratified Ranked Set Sampling

Authors: Shashi Bhushan
– Department of Statistics, University of Lucknow, Lucknow, India, 226007
shashi.py@gmail.com

Anoop Kumar ✉
– Department of Statistics, Central University of Haryana, Mahendergarh, India, 123031
anoop.asy@gmail.com

Received: September 2022 Revised: January 2023 Accepted: January 2023

Abstract:

- It is long familiar that the stratified ranked set sampling ($SRSS$) is more efficient than ranked set sampling (RSS) and stratified random sampling ($StRS$). The existence of missing values may alter the final inference of any study. This paper is a fundamental effort to suggest some combined and separate imputation methods in the presence of missing data under $SRSS$. The proposed imputation methods become superior than the mean imputation method, ratio imputation method, Diana and Perri (2010) type imputation method, and Sohail *et al.* (2018) type imputation methods. A simulation study is administered over two hypothetically drawn asymmetric populations.

Keywords:

- *missing values; imputation; stratified ranked set sampling.*

AMS Subject Classification:

- 62D05.

✉ Corresponding author

1. INTRODUCTION

The dilemma of missing value is very usual in a sample survey and its presence can spoil the traditional results. Therefore, it becomes essential to resolve the problem of missing values in a data set. The well-known imputation technique is used to replace the missing values. Three basic concepts on missing values were suggested by Rubin (1976), such as missing at random (MAR), observed at random (OAR), and parameter distribution (PD). Several renowned authors like Lee *et al.* (1994), Singh and Horn (2000), Singh and Deo (2003), Singh (2009), and Singh and Valdes (2009) introduced various imputation methods in the presence of missing values. Heitjan and Basu (1996) exhibited a difference between missing at random and missing completely at random ($MCAR$) approach. Thereafter, Ahmed *et al.* (2006), Kadilar and Cingi (2008), Diana and Perri (2010) and Bhushan and Pandey (2016, 2018), Mohamed *et al.* (2016), Prasad (2017), Bouza *et al.* (2020), Bouza-Herrera and Viada (2021), and Bhushan *et al.* (2018, 2022) utilized $MCAR$ strategy in their study for the imputation of missing values.

In real life, situations may emerge where it is either difficult to measure the study variable or indeed expensive but can be ranked either visually or by any cost free method. In such circumstances, McIntyre (1952) proposed the idea of ranked set sampling (RSS), which is superior to simple random sampling but did not furnish any mathematical support. Takahasi and Wakimoto (1968) extended the idea of McIntyre (1952) and provided the obligatory mathematical foundation to the theory of RSS. Samawi (1996) envisaged the idea of $SRSS$ superior to $StRS$. Samawi and Siam (2003) introduced combined and separate ratio estimators under $SRSS$. Mandowara and Mehta (2014) considered modified ratio estimators under $SRSS$. Linder *et al.* (2015) investigated the regression estimator under $SRSS$. Khan and Shabbir (2016) suggested Hartley-Ross type unbiased estimators under RSS and $SRSS$. Recently, Saini and Kumar (2018) suggested the ratio estimator using quartile as an auxiliary information under $SRSS$.

In sample surveys, when each group contains very small observations, then each observation becomes essential to draw conclusions. Further use of such kind of data set consisting of missing values may vitiate the final conclusion and decrease the efficiency of the estimator as well. In order to tackle with such kind of problems, Bouza-Herrera and Al-Omari (2011) suggested mean imputation and ratio methods for the median estimator under RSS. Al-Omari and Bouza (2014) introduced ratio estimators of the population mean with missing values under RSS. Sohail *et al.* (2018) suggested ratio type imputation methods under RSS.

In this paper, we suggest some imputation methods in the presence of missing data under $SRSS$. The rest paper is arranged in subsequent sections. In the next section, we discuss the sampling methodology along with the notations used throughout the manuscript. In Section 3, the combined and separate imputation methods are reviewed. In Section 4, we have suggested combined and separate classes of imputation methods. The theoretical comparisons of combined and separate imputation methods are given in Section 5, whereas Section 6 deals with the simulation study conducted in favour of theoretical findings. Lastly, the conclusion is given in Section 7.

2. METHODOLOGY AND NOTATIONS

The procedure of ranked set sampling consists of drawing m simple random samples of size m from the population. These m units are now ranked within each set with respect to the variable of interest, say x. The first smallest unit is quantified from the first set for the measurement of the auxiliary variable along with the associated study variables. The unit with the second smallest rank is quantified from the second ranked set for the measurement of the auxiliary variables along with the associated study variable and the process is carried on as far as the m-th smallest unit is quantified from the last set. The above process is known as a cycle. The repetition of this whole procedure up to k times furnishes $n = mk$ ranked set samples.

The stratified ranked set sampling is a sampling procedure analogous to stratified random sampling, which is based on splitting a population into L mutually exclusive and exhaustive strata and a ranked set sample of $n_h = m_h k$ units are measured within each stratum such that $h = 1, 2, ..., L$. The sampling is accomplished independently across the strata. Thus, $SRSS$ scheme can be supposed to a collection of L separate ranked set samples.

Consider a finite population U comprised of N measurable units with values y_i, $i \in U$. Let a stratified ranked set sample of size $n = m_h k$ be chosen from U to estimate the population mean of the study variable y. Let r be the number of responding elements out of n sampled elements. Let P be the probability that i-th respondent associated with a responding class A and $(1 - P)$ be the probability that i-th respondent associated with the non-responding class $\bar{A}$. Moreover, note that $s = A \cup \bar{A}$ and let the values y_i, $i \in A$ be observable for each characteristic, but for the characteristic $i \in \bar{A}$ the values are missing and require imputation in order to establish the complete frame of data to draw a reasonable inference. The auxiliary variable x will be used to execute the imputation of missing values and let the ranking be performed over the auxiliary variables as well.

The succeeding notations would be used from the beginning to end in the case of combined estimators.

Let $\bar{y}_{r,srss} = \bar{Y}(1 + \epsilon_0)$, $\bar{x}_{r,srss} = \bar{X}(1 + \epsilon_1)$, $\bar{x}_{n,srss} = \bar{X}(1 + \epsilon_2)$ such that $E(\epsilon_0) = E(\epsilon_1) = E(\epsilon_2) = 0$ and

$$E(\epsilon_0^2) = \sum_{h=1}^{L} W_h^2 \left(\frac{C_{y_h}^2}{m_h k P} - \frac{1}{m_h^2 k P} \sum_{i=1}^{m_h} \frac{\tau_{y_h}^2}{\bar{Y}^2} \right) = \sum_{h=1}^{L} W_h^2 \left(\gamma^* C_{y_h}^2 - D_{y_h}^{2^*} \right) = I_0^*,$$

$$E(\epsilon_1^2) = \sum_{h=1}^{L} W_h^2 \left(\frac{C_{x_h}^2}{m_h k P} - \frac{1}{m_h^2 k P} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}^2} \right) = \sum_{h=1}^{L} W_h^2 \left(\gamma^* C_{x_h}^2 - D_{x_h}^{2^*} \right) = I_1^*,$$

$$E(\epsilon_2^2) = \sum_{h=1}^{L} W_h^2 \left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}^2} \right) = \sum_{h=1}^{L} W_h^2 \left(\gamma C_{x_h}^2 - D_{x_h}^2 \right) = I_1,$$

$$E(\epsilon_0, \epsilon_1) = \sum_{h=1}^{L} W_h^2 \left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h k P} - \frac{1}{m_h^2 k P} \sum_{i=1}^{m_h} \frac{\tau_{x_h y_h}}{\bar{X}\bar{Y}} \right)$$

$$= \sum_{h=1}^{L} W_h^2 \left(\gamma^* \rho_{x_h y_h} C_{x_h} C_{y_h} - D_{x_h y_h}^* \right) = I_{01}^*,$$

$$
\begin{aligned}
E(\epsilon_0,\epsilon_2) &= \sum_{h=1}^{L} W_h^2 \left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h y_h}}{\bar{X}\bar{Y}} \right) \\
&= \sum_{h=1}^{L} W_h^2 (\gamma \rho_{x_h y_h} C_{x_h} C_{y_h} - D_{x_h y_h}) = I_{01}, \\
E(\epsilon_1,\epsilon_2) &= \sum_{h=1}^{L} W_h^2 \left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}^2} \right) = \sum_{h=1}^{L} W_h^2 \left(\gamma C_{x_h}^2 - D_{x_h}^2 \right) = I_1,
\end{aligned}
$$

where $\gamma^* = 1/m_h kP$, $\gamma = 1/m_h k$, $\tau_{y_h} = (\mu_{y_h} - \bar{Y}_h)$, $\tau_{x_h} = (\mu_{x_h} - \bar{X}_h)$ and $\tau_{x_h y_h} = (\mu_{x_h} - \bar{X}_h) \cdot \cdot (\mu_{y_h} - \bar{Y}_h)$. Also, $C_{x_h} = S_{x_h}/\bar{X}$ and $C_{y_h} = S_{y_h}/\bar{Y}$ are the coefficients of variation of auxiliary variable x and study variable y, respectively.

In the case of separate estimators, the following notations will be used throughout the paper.

Let $\bar{y}_{r,h[rss]} = \bar{Y}_h(1 + e_{0_h})$, $\bar{x}_{r,h(rss)} = \bar{X}_h(1 + e_{1_h})$, $\bar{x}_{n,h(rss)} = \bar{X}_h(1 + e_{2_h})$ such that $E(e_{0_h}) = E(e_{1_h}) = E(e_{2_h}) = 0$ and

$$
\begin{aligned}
E(e_{0_h}^2) &= \left(\frac{C_{y_h}^2}{m_h kP} - \frac{1}{m_h^2 kP} \sum_{i=1}^{m_h} \frac{\tau_{y_h}^2}{\bar{Y}_h^2} \right) = \left(\gamma^* C_{y_h}^2 - M_{y_h}^{2^*} \right) = J_0^*, \\
E(e_{1_h}^2) &= \left(\frac{C_{x_h}^2}{m_h kP} - \frac{1}{m_h^2 kP} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}_h^2} \right) = \left(\gamma^* C_{x_h}^2 - M_{x_h}^{2^*} \right) = J_1^*, \\
E(e_{2_h}^2) &= \left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}_h^2} \right) = \left(\gamma C_{x_h}^2 - M_{x_h}^2 \right) = J_1, \\
E(e_{0_h}, e_{1_h}) &= \left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h kP} - \frac{1}{m_h^2 kP} \sum_{i=1}^{m_h} \frac{\tau_{x_h y_h}}{\bar{X}_h \bar{Y}_h} \right) = \left(\gamma^* \rho_{x_h y_h} C_{x_h} C_{y_h} - M_{x_h y_h}^* \right) = J_{01}^*, \\
E(e_{0_h}, e_{2_h}) &= \left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h y_h}}{\bar{X}_h \bar{Y}_h} \right) = (\gamma \rho_{x_h y_h} C_{x_h} C_{y_h} - M_{x_h y_h}) = J_{01}, \\
E(\epsilon_{1_h}, \epsilon_{2_h}) &= \left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}_h^2} \right) = \left(\gamma C_{x_h}^2 - M_{x_h}^2 \right) = J_1,
\end{aligned}
$$

where $\tau_{y_h} = (\mu_{y_h} - \bar{Y}_h)$, $\tau_{x_h} = (\mu_{x_h} - \bar{X}_h)$ and $\tau_{x_h y_h} = (\mu_{x_h} - \bar{X}_h)(\mu_{y_h} - \bar{Y}_h)$, $C_{x_h} = S_{x_h}/\bar{X}_h$ and $C_{y_h} = S_{y_h}/\bar{Y}_h$.

3. RECAP OF IMPUTATION METHODS

In this section, we consider a concise recap of existing prominent combined and separate imputation methods under $SRSS$.

3.1. Combined imputation methods

The mean method of imputation under $SRSS$ is given by

$$y^c_{.i_m} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

The sequent estimator is given by

$$T^c_m = \bar{y}_{r,srss}$$

where $\bar{y}_{r,srss} = \sum_{h=1}^{L} W_h \bar{y}_{h[rss]}$ is the stratified ranked set sample mean of study variable y. Also, $W_h = N_h/N$ is the weight of stratum h and N_h and N are the size of stratum h and total population size, respectively.

The imputation methods are categorized into three situations under the availability of auxiliary informations:

Situation I: When $\bar{X}$ is known and $\bar{x}_{n,srss}$ is utilized.

Situation II: When $\bar{X}$ is known and $\bar{x}_{r,srss}$ is utilized.

Situation III: When $\bar{X}$ is unknown and $\bar{x}_{n,srss}$, $\bar{x}_{r,srss}$ are utilized.

The classical combined ratio type imputation methods are defined under $SRSS$ as:

Situation I

$$y^c_{.i_{R_1}} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\bar{y}_{r,rss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Situation II

$$y^c_{.i_{R_2}} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\bar{y}_{r,rss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Situation III

$$y^c_{.i_{R_3}} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\bar{y}_{r,rss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

The sequent estimators are

$$T^c_{R_1} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right),$$

$$T^c_{R_2} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right),$$

$$T^c_{R_3} = \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right),$$

where $\bar{x}_{n,srss} = \sum_{h=1}^{L} W_h \bar{x}_{h(rss)}$ is the stratified ranked set sample mean of auxiliary variable x.

Following Diana and Perri (2010), we define the regression imputation methods to impute the missing value under $SRSS$ as:

Situation I

$$y^c_{.i_{DP_1}} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_r + \frac{n}{n-r} b_1(\bar{X} - \bar{x}_{n,srss}) y_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

Situation II

$$y^c_{.i_{DP_2}} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_r + \frac{n}{n-r} b_2(\bar{X} - \bar{x}_{r,srss}) y_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

Situation III

$$y^c_{.i_{DP_3}} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_r + \frac{n}{n-r} b_3(\bar{x}_{n,srss} - \bar{x}_{r,srss}) y_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

The sequent combined estimators under the above situations are given by

$$\begin{aligned} T^c_{DP_1} &= \bar{y}_{r,srss} + b_1(\bar{X} - \bar{x}_{n,srss}), \\ T^c_{DP_2} &= \bar{y}_{r,srss} + b_2(\bar{X} - \bar{x}_{r,srss}), \\ T^c_{DP_3} &= \bar{y}_{r,srss} + b_3(\bar{x}_{n,srss} - \bar{x}_{r,srss}). \end{aligned}$$

Following Sohail *et al.* (2018), one may envisage a combined class of ratio type imputation methods under $SRSS$ for the imputation of missing values as:

Situation I

$$y^c_{.iS_1} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\beta_1} - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iS_4} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_4\bar{x}_{n,srss}+(1-\beta_4)\bar{X}}\right) - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Situation II

$$y^c_{.iS_2} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\beta_2} - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iS_5} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_5\bar{x}_{r,srss}+(1-\beta_5)\bar{X}}\right) - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Situation III

$$y^c_{.iS_3} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\beta_3} - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iS_6} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_6\bar{x}_{r,srss}+(1-\beta_6)\bar{x}_{n,srss}}\right) - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

The sequent estimators are given by

$$T^c_{S_1} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\beta_1},$$
$$T^c_{S_2} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\beta_2},$$
$$T^c_{S_3} = \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\beta_3},$$
$$T^c_{S_4} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_4\bar{x}_{n,srss} + (1-\beta_4)\bar{X}}\right),$$
$$T^c_{S_5} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_5\bar{x}_{r,srss} + (1-\beta_5)\bar{X}}\right),$$
$$T^c_{S_6} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_6\bar{x}_{r,srss} + (1-\beta_6)\bar{x}_{n,srss}}\right),$$

where $\beta_i;\ \ i = 1, 2, ..., 6$ are suitably chosen optimizing scalars.

Appendix A of supplementary file contains the minimum mean square error (MSE) of the sequent estimators consisting of different imputation methods.

3.2. Separate imputation methods

The separate mean method of imputation under $SRSS$ is given by

$$y^s_{.i_m} = \begin{cases} y_i & \text{for } i \in A_h, \\ \bar{y}_{r,h[rss]} & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimator is given by

$$T^s_m = \sum_{h=1}^{L} W_h \bar{y}_{r,h[rss]},$$

where $\bar{y}_{r,h[rss]} = \frac{1}{m_h k}\sum_{i=1}^{m_h}\sum_{j=1}^{k} y_{h[i]j}$ is the ranked set sample mean of study variable in stratum h.

The separate imputation methods are categorized into three situations under the availability of auxiliary informations:

Situation I: When $\bar{X}$ is known and $\bar{x}_{n,h(rss)}$ is utilized.

Situation II: When $\bar{X}$ is known and $\bar{x}_{r,h(rss)}$ is utilized.

Situation III: When $\bar{X}$ is unknown and $\bar{x}_{n,h(rss)}$, $\bar{x}_{r,h(rss)}$ are utilized.

The classical separate ratio type imputation method is described under $SRSS$ as:

Situation I

$$y^s_{.i_{R_1}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation II

$$y^s_{.i_{R_2}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation III

$$y^s_{.i_{R_3}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimators are given by

$$T^s_{R_1} = \sum_{h=1}^{L} W_h\left[\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)\right],$$

$$T^s_{R_2} = \sum_{h=1}^{L} W_h\left[\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)\right],$$

$$T^s_{R_3} = \sum_{h=1}^{L} W_h\left[\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)\right].$$

On the lines of Diana and Perri (2010), we define a separate regression imputation method under $SRSS$ as:

Situation I

$$y^s_{.i_{DP_1}} = \begin{cases} y_i & \text{for } i \in A_h, \\ y_{r,h[rss]} + \frac{n}{n-r}b_1(\bar{X} - \bar{x}_{n,h(rss)}) & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation II

$$y^s_{.i_{DP_2}} = \begin{cases} y_i & \text{for } i \in A_h, \\ y_{r,h[rss]} + \frac{n}{n-r}b_2(\bar{X} - \bar{x}_{r,h(rss)}) & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation III

$$y^s_{.i_{DP_3}} = \begin{cases} y_i & \text{for } i \in A_h, \\ y_{r,h[rss]} + \frac{n}{n-r}b_3(\bar{x}_{n,h(rss)} - \bar{x}_{r,h(rss)}) & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent separate estimators under the above situations are given by

$$T^s_{DP_1} = \sum_{h=1}^{L} W_h[\bar{y}_{r,h[rss]} + b_{1_h}(\bar{X}_h - \bar{x}_{n,h(rss)})],$$

$$T^s_{DP_2} = \sum_{h=1}^{L} W_h[\bar{y}_{r,h[rss]} + b_{2_h}(\bar{X}_h - \bar{x}_{r,h(rss)})],$$

$$T^s_{DP_3} = \sum_{h=1}^{L} W_h[\bar{y}_{r,h[rss]} + b_{3_h}(\bar{x}_{n,h(rss)} - \bar{x}_{r,h(rss)})].$$

Motivated by Sohail *et al.* (2018), we define a separate class of ratio type imputation methods under $SRSS$ as:

Situation I

$$y^s_{.is_1} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\beta_{1_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

$$y^s_{.is_4} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{4_h}\bar{x}_{n,h(rss)}+(1-\beta_{4_h})\bar{X}_h}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation II

$$y^s_{.is_2} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\beta_{2_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.is_5} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{5_h}\bar{x}_{r,h(rss)}+(1-\beta_{5_h})\bar{X}_h}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation III

$$y^s_{.is_3} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\beta_{3_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.is_6} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{6_h}\bar{x}_{r,h(rss)}+(1-\beta_{6_h})\bar{x}_{n,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimators are given by

$$T^s_{S_1} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\beta_{1_h}},$$

$$T^s_{S_2} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\beta_{2_h}},$$

$$T^s_{S_3} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\beta_{3_h}},$$

$$T^s_{S_4} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{4_h}\bar{x}_{n,h(rss)} + (1-\beta_{4_h})\bar{X}_h}\right),$$

$$T^s_{S_5} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{5_h}\bar{x}_{r,h(rss)} + (1-\beta_{5_h})\bar{X}_h}\right),$$

$$T^s_{S_6} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{6_h}\bar{x}_{r,h(rss)} + (1-\beta_{6_h})\bar{x}_{n,h(rss)}}\right),$$

where $\beta_{i_h};\ i = 1, 2, ..., 6$ are suitably opted scalars.

Appendix B of supplementary file contains the minimum mean square error (MSE) of the sequent estimators consisting of different imputation methods.

4. PROPOSED IMPUTATION METHODS

The crux of this paper is binary:

1. To propose some efficient combined and separate imputation methods for the estimation of population mean $\bar{Y}$.
2. To determine the effect of the correlation coefficient, coefficient of skewness, and coefficient of kurtosis over the efficiency of the imputation procedures.

4.1. Combined imputation methods

Following Bhushan and Pandey (2016, 2018), we envisage nine new imputation methods under the three situations specified in the former section as:

Situation I

$$y^c_{.iSA_1} = \begin{cases} \alpha_1 y_i & \text{for } i \in A, \\ \alpha_1 \bar{y}_{r,srss} + \frac{n\theta_1}{n-r}(\bar{x}_{n,srss} - \bar{X}) & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_4} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_4 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\theta_4} - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_7} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_7 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{X}+\theta_7(\bar{x}_{n,srss}-\bar{X})}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Situation II

$$y^c_{.iSA_2} = \begin{cases} \alpha_2 y_i & \text{for } i \in A, \\ \alpha_2 \bar{y}_{r,srss} + \frac{n\theta_2}{n-r}(\bar{x}_{r,srss} - \bar{X}) & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_5} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_5 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\theta_5} - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_8} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_8 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{X}+\theta_8(\bar{x}_{r,srss}-\bar{X})}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Situation III

$$y^c_{.iSA_3} = \begin{cases} \alpha_3 y_i & \text{for } i \in A, \\ \alpha_3 \bar{y}_{r,srss} + \frac{n\theta_3}{n-r}(\bar{x}_{r,srss} - \bar{x}_{n,srss}) & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_6} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_6 \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\theta_6} - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_9} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_9 \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{n,srss}+\theta_9(\bar{x}_{n,srss}-\bar{x}_{r,srss})}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Under the above situations, the sequent estimators are given by

$$\begin{aligned}
T^c_{SA_1} &= \alpha_1 \bar{y}_{r,srss} + \theta_1(\bar{x}_{n,srss} - \bar{X}),\\
T^c_{SA_2} &= \alpha_2 \bar{y}_{r,srss} + \theta_2(\bar{x}_{r,srss} - \bar{X}),\\
T^c_{SA_3} &= \alpha_3 \bar{y}_{r,srss} + \theta_3(\bar{x}_{r,srss} - \bar{x}_{n,srss}),\\
T^c_{SA_4} &= \alpha_4 \bar{y}_{r,srss} \left(\frac{\bar{X}}{\bar{x}_{n,srss}} \right)^{\theta_4},\\
T^c_{SA_5} &= \alpha_5 \bar{y}_{r,srss} \left(\frac{\bar{X}}{\bar{x}_{r,srss}} \right)^{\theta_5},\\
T^c_{SA_6} &= \alpha_6 \bar{y}_{r,srss} \left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}} \right)^{\theta_6},\\
T^c_{SA_7} &= \alpha_7 \bar{y}_{r,srss} \left[\frac{\bar{X}}{\bar{X} + \theta_7(\bar{x}_{n,srss} - \bar{X})} \right],\\
T^c_{SA_8} &= \alpha_8 \bar{y}_{r,srss} \left[\frac{\bar{X}}{\bar{X} + \theta_8(\bar{x}_{r,srss} - \bar{X})} \right],\\
T^c_{SA_9} &= \alpha_9 \bar{y}_{r,srss} \left[\frac{\bar{x}_{n,srss}}{\bar{x}_{n,srss} + \theta_9(\bar{x}_{r,srss} - \bar{x}_{n,srss})} \right],
\end{aligned}$$

where $\alpha_1, \alpha_2, ..., \alpha_9$ and $\theta_1, \theta_2, ..., \theta_9$ are the suitably chosen scalars.

Theorem 4.1. *The MSE of the sequent estimators consisting of the proposed imputation methods is given by*

$$\begin{aligned}
MSE(T^c_{SA_1}) &= (\alpha_1 - 1)^2 \bar{Y}^2 + \alpha_1^2 \bar{Y}^2 I_0^* + \theta_1^2 \bar{X}^2 I_1 + 2\alpha_1\theta_1 \bar{X}\bar{Y} I_{01},\\
MSE(T^c_{SA_2}) &= (\alpha_2 - 1)^2 \bar{Y}^2 + \alpha_2^2 \bar{Y}^2 I_0^* + \theta_2^2 \bar{X}^2 I_1^* + 2\alpha_2\theta_2 \bar{X}\bar{Y} I_{01}^*,\\
MSE(T^c_{SA_3}) &= \begin{bmatrix} (\alpha_3 - 1)^2 \bar{Y}^2 + \alpha_3^2 \bar{Y}^2 I_0^* + \theta_3^2 \bar{X}^2 \{I_1^* - I_1\} \\ +2\alpha_3\theta_3 \bar{X}\bar{Y}\{I_{01}^* - I_{01}\} \end{bmatrix},\\
MSE(T^c_{SA_4}) &= \bar{Y}^2 \begin{bmatrix} 1 + \alpha_4^2 \{ 1 + I_0^* + \theta_4(2\theta_4 + 1) I_1 - 4\theta_4 I_{01} \} \\ -2\alpha_4 \left\{ 1 - \theta_4 I_{01} + \frac{\theta_4(\theta_4+1)}{2} I_1 \right\} \end{bmatrix},\\
MSE(T^c_{SA_5}) &= \bar{Y}^2 \begin{bmatrix} 1 + \alpha_5^2 \{ 1 + I_0^* + \theta_5(2\theta_5 + 1) I_1^* - 4\theta_5 I_{01}^* \} \\ -2\alpha_5 \left\{ 1 - \theta_5 I_{01}^* + \frac{\theta_5(\theta_5+1)}{2} I_1^* \right\} \end{bmatrix},\\
MSE(T^c_{SA_6}) &= \bar{Y}^2 \begin{bmatrix} 1 + \alpha_6^2 \{ 1 + I_0^* + \theta_6(2\theta_6 + 1)(I_1^* - I_1) - 4\theta_6(I_{01}^* - I_{01}) \} \\ -2\alpha_6 \left\{ 1 - \theta_6(I_{01}^* - I_{01}) + \frac{\theta_6(\theta_6+1)}{2}(I_1^* - I_1) \right\} \end{bmatrix},\\
MSE(T^c_{SA_7}) &= \bar{Y}^2 \begin{bmatrix} 1 + \alpha_7^2 \{ 1 + I_0^* + 3\theta_7^2 I_1 - 4\theta_7 I_{01} \} \\ -2\alpha_7 \{ 1 + \theta_7^2 I_1 - \theta_7 I_{01} \} \end{bmatrix},\\
MSE(T^c_{SA_8}) &= \bar{Y}^2 \begin{bmatrix} 1 + \alpha_8^2 \{ 1 + I_0^* + 3\theta_8^2 I_1^* - 4\theta_8 I_{01}^* \} \\ -2\alpha_8 \{ 1 + \theta_8^2 I_1^* - \theta_8 I_{01}^* \} \end{bmatrix},\\
MSE(T^c_{SA_9}) &= \bar{Y}^2 \begin{bmatrix} 1 + \alpha_9^2 \{ 1 + I_0^* + 3\theta_9^2 (I_1^* - I_1) - 4\theta_9(I_{01}^* - I_{01}) \} \\ -2\alpha_9 \{ 1 + \theta_9^2 (I_1^* - I_1) - \theta_9(I_{01}^* - I_{01}) \} \end{bmatrix}.
\end{aligned}$$

Proof: Appendix C of supplementary file contains a summary of the derivations. The derivations can easily be done using Taylor series expansion. □

Theorem 4.2. *The minimum MSE of the sequent estimators consisting of the proposed imputation methods are*

$$minMSE(T^c_{SA_i}) = \bar{Y}^2(1-\alpha_{i(opt)}) = \bar{Y}^2\left(1-\frac{A_i^2}{B_i}\right);\ i=1,2,3, \tag{4.1}$$

$$minMSE(T^c_{SA_j}) = \bar{Y}^2\left(1-\frac{A_j^2}{B_j}\right);\ j=4,5,6, \tag{4.2}$$

$$minMSE(T^c_{SA_k}) = \bar{Y}^2\left(1-\frac{A_k^2}{B_k}\right);\ k=7,8,9. \tag{4.3}$$

Proof: Appendix C of supplementary file contains a summary of the derivations. □

Corollary 4.1. *The proposed sequent estimators $T^c_{SA_i},\ i=1,2,3$ dominate the proposed sequent estimators $T^c_{SA_j},\ j=4,5,6$, iff*

$$\alpha_{i(opt)} > \frac{A_j^2}{B_j} \tag{4.4}$$

and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.4).

Proof: By comparing the minimum MSEs of the proposed estimators from (4.1) and (4.2), we get (4.4). □

Corollary 4.2. *The proposed sequent estimators $T^c_{SA_i},\ i=1,2,3$ dominate the proposed sequent estimators $T^c_{SA_k},\ k=7,8,9$, iff*

$$\alpha_{i(opt)} > \frac{A_k^2}{B_k} \tag{4.5}$$

and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.5).

Proof: On comparing the minimum MSEs of the proposed estimators from (4.1) and (4.3), we get (4.5). □

Corollary 4.3. *The proposed sequent estimators $T^c_{SA_j},\ i=4,5,6$ dominate the proposed sequent estimators $T^c_{SA_k},\ k=7,8,9$, iff*

$$\frac{A_j^2}{B_j} > \frac{A_k^2}{B_k} \tag{4.6}$$

and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.6).

Proof: On comparing the minimum MSEs of the proposed estimators from (4.2) and (4.3), we get (4.6). □

The only way to determine if (4.4), (4.5), and (4.6) are true in practise is through the computational analysis done in Section 6.

4.2. Separate imputation methods

On the lines of Bhushan and Pandey (2016, 2018), we suggest nine new separate imputation methods under the three situations discussed in the earlier section as:

Situation I

$$y^s_{.i_{SA_1}} = \begin{cases} \alpha_{1_h} y_i & \text{for } i \in A_h, \\ \alpha_{1_h} \bar{y}_{r,h[rss]} + \frac{n\theta_{1_h}}{n-r}(\bar{x}_{n,h(rss)} - \bar{X}_h) & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_4}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{4_h} \bar{y}_{r,h[rss]} \left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\theta_{4_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_7}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{7_h} \bar{y}_{r,h[rss]} \left(\frac{\bar{X}_h}{\bar{X}_h + \theta_{7_h}(\bar{x}_{n,h(rss)} - \bar{X}_h)}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation II

$$y^s_{.i_{SA_2}} = \begin{cases} \alpha_{2_h} y_i & \text{for } i \in A_h, \\ \alpha_{2_h} \bar{y}_{r,h[rss]} + \frac{n\theta_{2_h}}{n-r}(\bar{x}_{r,h(rss)} - \bar{X}_h) & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_5}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{5_h} \bar{y}_{r,h[rss]} \left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\theta_{5_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_8}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{8_h} \bar{y}_{r,h[rss]} \left(\frac{\bar{X}_h}{\bar{X} + \theta_{8_h}(\bar{x}_{r,h(rss)} - \bar{X}_h)}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

Situation III

$$y^s_{.i_{SA_3}} = \begin{cases} \alpha_{3_h} y_i & \text{for } i \in A_h, \\ \alpha_{3_h} \bar{y}_{r,h[rss]} + \frac{n\theta_{3_h}}{n-r}(\bar{x}_{r,h(rss)} - \bar{x}_{n,h(rss)}) & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_6}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{6_h} \bar{y}_{r,h[rss]} \left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\theta_{6_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_9}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{9_h} \bar{y}_{r,h[rss]} \left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{n,h(rss)} + \theta_{9_h}(\bar{x}_{n,h(rss)} - \bar{x}_{r,h(rss)})}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimators consisting of the above imputation methods are

$$T^s_{SA_1} = \sum_{h=1}^{L} W_h[\alpha_{1_h} \bar{y}_{r,h[rss]} + \theta_{1_h}(\bar{x}_{n,h(rss)} - \bar{X}_h)],$$

$$T^s_{SA_2} = \sum_{h=1}^{L} W_h[\alpha_{2_h} \bar{y}_{r,h[rss]} + \theta_{2_h}(\bar{x}_{r,h(rss)} - \bar{X}_h)],$$

$$T^s_{SA_3} = \sum_{h=1}^{L} W_h[\alpha_{3_h} \bar{y}_{r,h[rss]} + \theta_{3_h}(\bar{x}_{r,h(rss)} - \bar{x}_{n,h(rss)})],$$

$$T^s_{SA_4}=\sum_{h=1}^{L}W_h\alpha_{4_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\theta_{4_h}},$$
$$T^s_{SA_5}=\sum_{h=1}^{L}W_h\alpha_{5_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\theta_{5_h}},$$
$$T^s_{SA_6}=\sum_{h=1}^{L}W_h\alpha_{6_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\theta_{6_h}},$$
$$T^s_{SA_7}=\sum_{h=1}^{L}W_h\alpha_{7_h}\bar{y}_{r,h[rss]}\left[\frac{\bar{X}_h}{\bar{X}_h+\theta_{7_h}(\bar{x}_{n,h(rss)}-\bar{X}_h)}\right],$$
$$T^s_{SA_8}=\sum_{h=1}^{L}W_h\alpha_{8_h}\bar{y}_{r,h[rss]}\left[\frac{\bar{X}_h}{\bar{X}_h+\theta_{8_h}(\bar{x}_{r,h(rss)}-\bar{X}_h)}\right],$$
$$T^s_{SA_9}=\sum_{h=1}^{L}W_h\alpha_{9_h}\bar{y}_{r,h[rss]}\left[\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{n,h(rss)}+\theta_{9_h}(\bar{x}_{r,h(rss)}-\bar{x}_{n,h(rss)})}\right],$$

where $\alpha_{1_h},\alpha_{2_h},...,\alpha_{9_h}$ and $\theta_{1_h},\theta_{2_h},...,\theta_{9_h}$ are suitably chosen scalars.

Theorem 4.3. *The MSE of the sequent estimators consisting of the proposed imputation methods is given by*

$$MSE(T^s_{SA_1})=\sum_{h=1}^{L}W_h^2\left[\,(\alpha_{1_h}-1)^2\bar{Y}_h^2+\alpha_{1_h}^2\bar{Y}_h^2J_0^*+\theta_{1_h}^2\bar{X}_h^2J_1+2\alpha_{1_h}\theta_{1_h}\bar{X}_h\bar{Y}_hJ_{01}\,\right],$$
$$MSE(T^s_{SA_2})=\sum_{h=1}^{L}W_h^2\left[\,(\alpha_{2_h}-1)^2\bar{Y}_h^2+\alpha_{2_h}^2\bar{Y}_h^2J_0^*+\theta_{2_h}^2\bar{X}_h^2J_1^*+2\alpha_{2_h}\theta_{2_h}\bar{X}_h\bar{Y}_hJ_{01}^*\,\right],$$
$$MSE(T^s_{SA_3})=\sum_{h=1}^{L}W_h^2\left[\begin{array}{l}(\alpha_{3_h}-1)^2\bar{Y}_h^2+\alpha_{3_h}^2\bar{Y}_h^2J_0^*+\theta_{3_h}^2\bar{X}_h^2\{J_1^*-J_1\}\\+2\alpha_{3_h}\theta_{3_h}\bar{X}_h\bar{Y}_h\{J_{01}^*-J_{01}\}\end{array}\right],$$
$$MSE(T^s_{SA_4})=\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\begin{array}{l}1+\alpha_{4_h}^2\{\,1+J_0^*+\theta_{4_h}(2\theta_{4_h}+1)J_1-4\theta_{4_h}J_{01}\,\}\\-2\alpha_{4_h}\left\{1-\theta_{4_h}J_{01}+\frac{\theta_{4_h}(\theta_{4_h}+1)}{2}J_1\right\}\end{array}\right],$$
$$MSE(T^s_{SA_5})=\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\begin{array}{l}1+\alpha_{5_h}^2\{\,1+J_0^*+\theta_{5_h}(2\theta_{5_h}+1)J_1^*-4\theta_{5_h}J_{01}^*\,\}\\-2\alpha_{5_h}\left\{1-\theta_{5_h}J_{01}^*+\frac{\theta_{5_h}(\theta_{5_h}+1)}{2}J_1^*\right\}\end{array}\right],$$
$$MSE(T^s_{SA_6})=\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\begin{array}{l}1+\alpha_{6_h}^2\{\,1+J_0^*+\theta_{6_h}(2\theta_{6_h}+1)(J_1^*-J_1)-4\theta_{6_h}(J_{01}^*-J_{01})\,\}\\-2\alpha_{6_h}\left\{\,1-\theta_{6_h}(J_{01}^*-J_{01})+\frac{\theta_{6_h}(\theta_{6_h}+1)}{2}(J_1^*-J_1)\,\right\}\end{array}\right],$$
$$MSE(T^s_{SA_7})=\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\begin{array}{l}1+\alpha_{7_h}^2\{\,1+J_0^*+3\theta_{7_h}^2J_1-4\theta_{7_h}J_{01}\,\}\\-2\alpha_{7_h}\{1+\theta_{7_h}^2J_1-\theta_{7_h}J_{01}\}\end{array}\right],$$
$$MSE(T^s_{SA_8})=\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\begin{array}{l}1+\alpha_{8_h}^2\{\,1+J_0^*+3\theta_{8_h}^2J_1^*-4\theta_{8_h}J_{01}^*\,\}\\-2\alpha_{8_h}\{1+\theta_{8_h}^2J_1^*-\theta_{8_h}J_{01}^*\}\end{array}\right],$$
$$MSE(T^s_{SA_9})=\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\begin{array}{l}1+\alpha_{9_h}^2\{\,1+J_0^*+3\theta_{9_h}^2(\,J_1^*-J_1\,)-4\theta_{9_h}(J_{01}^*-J_{01})\,\}\\-2\alpha_{9_h}\{\,1+\theta_{9_h}^2(J_1^*-J_1)-\theta_{9_h}(J_{01}^*-J_{01})\,\}\end{array}\right].$$

Proof: Appendix C of supplementary file contains a summary of the derivations. The derivations can easily be done using Taylor series expansion. □

Theorem 4.4. *The minimum MSE of the sequent estimators consisting of the proposed imputation methods is given by*

$$minMSE(T^s_{SA_i}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 (1 - \alpha_{i_h(opt)}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A_{i_h}^2}{B_{i_h}}\right);\ i = 1, 2, 3, \tag{4.7}$$

$$minMSE(T^s_{SA_j}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A_{j_h}^2}{B_{j_h}}\right);\ j = 4, 5, 6, \tag{4.8}$$

$$minMSE(T^s_{SA_k}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A_{k_h}^2}{B_{k_h}}\right);\ k = 7, 8, 9. \tag{4.9}$$

Proof: Appendix C of supplementary file contains a summary of the derivations. □

Corollary 4.4. *The proposed sequent estimators $T^s_{SA_i},\ i = 1, 2, 3$ dominate the proposed sequent estimators $T^s_{SA_j},\ j = 4, 5, 6$, iff*

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \alpha_{i_h(opt)} > \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A_{i_h}^2}{B_{i_h}}\right) \tag{4.10}$$

and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.10).

Proof: On comparing the minimum MSEs of the proposed estimators from (4.7) and (4.8), we get (4.10). □

Corollary 4.5. *The proposed sequent estimators $T^s_{SA_i},\ i = 1, 2, 3$ dominate the proposed sequent estimators $T^s_{SA_k},\ k = 4, 5, 6$, iff*

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \alpha_{i_h(opt)} > \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A_{k_h}^2}{B_{k_h}}\right) \tag{4.11}$$

and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.11).

Proof: By comparing the minimum MSEs of the proposed estimators from (4.7) and (4.9), we get (4.11). □

Corollary 4.6. *The proposed sequent estimators $T^s_{SA_j},\ j = 4, 5, 6$ dominate the proposed sequent estimators $T^s_{SA_k},\ k = 7, 8, 9$, iff*

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A_{j_h}^2}{B_{j_h}}\right) > \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A_{k_h}^2}{B_{k_h}}\right) \tag{4.12}$$

and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.12).

Proof: By comparing the minimum MSEs of the proposed estimators from (4.8) and (4.9), we get (4.12). □

The only way to determine if (4.10), (4.11), and (4.12) are true in practise is through the computational analysis done in Section 6.

5. OPTIMALITY CONDITIONS

In this section, we obtain the optimality conditions under two heads, namely, optimality conditions for combined imputation methods and the optimality conditions for separate imputation methods.

5.1. Optimality conditions for the combined imputation methods

By comparing the minimum MSE of the suggested combined imputation methods $y^{c}_{.i_{SA_i}}$, $i = 1, 2, ..., 9$ from (4.1) and (4.2) with the minimum MSE of the other existing combined imputation methods from $(A.1)$, $(A.2)$, $(A.3)$, $(A.4)$, $(A.8)$, $(A.9)$, $(A.10)$, $(A.14)$, $(A.15)$, and $(A.16)$, respectively, given in Appendix A of supplementary file, we get the following optimality conditions:

$$MSE(T^c_m) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^*,$$

$$MSE(T^c_{R_1}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* - I_1 + 2I_{01},$$

$$MSE(T^c_{R_2}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* - I_1^* + 2I_{01}^*,$$

$$MSE(T^c_{R_3}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* - I_1 - I_1^* + 2I_{01}^*,$$

$$MSE(T^c_{DP_1}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{I_{01}^2}{I_1},$$

$$MSE(T^c_{DP_2}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{{I_{01}^*}^2}{I_1^*},$$

$$MSE(T^c_{DP_3}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{(I_{01}^* - I_{01})^2}{(I_1^* - I_1)},$$

$$MSE(T^c_{S_1}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{I_{01}^2}{I_1},$$

$$MSE(T^c_{S_2}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{{I_{01}^*}^2}{I_1^*},$$

$$MSE(T^c_{S_3}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{(I_{01}^* - I_{01})^2}{(I_1^* - I_1)}.$$

The optimality of the suggested combined imputation methods can be justified under the above conditions.

5.2. Optimality conditions for the separate imputation methods

By comparing the minimum MSE of the proposed imputation methods $y^{s}_{.iSA_i}$, $i = 1, 2, ..., 9$ given in (4.7) and (4.8) with the minimum MSE of the other existing imputation methods given in $(B.17)$, $(B.18)$, $(B.19)$, $(B.20)$, $(B.24)$, $(B.25)$, $(B.26)$, $(B.30)$, $(B.31)$, and $(B.32)$, respectively, given in Appendix B of supplementary file, we get the following optimality conditions:

$$MSE(T^s_m) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 J_1^*,$$

$$MSE(T^s_{R_1}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 [J_0^* + J_1 - 2J_{01}],$$

$$MSE(T^s_{R_2}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 [J_0^* + J_1^* - 2J_{01}^*],$$

$$MSE(T^s_{R_3}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 [J_0^* + J_1^* + J_1 - 2J_{01}^*],$$

$$MSE(T^s_{DP_1}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[J_0^* - \frac{J_{01}^2}{J_1}\right],$$

$$MSE(T^s_{DP_2}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[J_0^* - \frac{{J_{01}^*}^2}{J_1^*}\right],$$

$$MSE(T^s_{DP_3}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[J_0^* - \frac{\left(J_{01}^* - J_{01}\right)^2}{\left(J_1^* - J_1\right)}\right],$$

$$MSE(T^s_{S_1}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[J_0^* - \frac{J_{01}^2}{J_1}\right],$$

$$MSE(T^s_{S_2}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[J_0^* - \frac{{J_{01}^*}^2}{J_1^*}\right],$$

$$MSE(T^s_{S_3}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[J_0^* - \frac{\left(J_{01}^* - J_{01}\right)^2}{\left(J_1^* - J_1\right)}\right].$$

Under the above conditions, the optimality of the proposed separate imputation methods can be ascertained.

5.3. Comparison of proposed combined and separate imputation methods

By comparing the minimum MSE of the proposed combined and separate classes of imputation methods from (4.1), (4.2) and (4.7), (4.8), we get

$$(5.1)\quad minMSE(T^c_{SA_i}) - minMSE(T^s_{SA_i}) = \sum_{h=1}^{L} \left[(\bar{Y}^2 - W_h^2 \bar{Y}_h^2) - \left(\bar{Y}^2 \frac{A_i^2}{B_i} - W_h^2 \bar{Y}_h^2 \frac{A^2_{i_h}}{B_{i_h}} \right) \right].$$

If the sequent estimators are conclusive and the relationship between auxiliary and study variables within each stratum is a straight line passing through the origine, then the last term of (5.1) is miniscule and it vanished.

In addition, except R_h becomes invariant from stratum to stratum, the separate estimators perform better in each stratum provided the sample in each stratum is to be sufficiently large so that the approximate formula for $MSE(T^s_{SA_i}),\ i = 1, 2, ..., 9$ is valid and the cumulative bias that can affect the proposed estimators is negligible, whereas the proposed combined estimators are to be highly advocated with only a small sample in each stratum (see, Cochran, 1977).

6. SIMULATION STUDY

To highlight the properties and to access the performance of the proposed imputation methods, motivated by Singh and Horn (1998), simulations were carried out over two artificially generated asymmetric populations such as gamma and exponential of size $N = 2100$ units each with variables X and Y whose values are given by

$$y_i = 8.2 + \sqrt{(1-\rho^2_{xy})}\ y^*_i + \rho_{xy}\left(\frac{S_y}{S_x}\right)x^*_i,$$
$$x_i = 4.2 + x^*_i.$$

where x^*_i and y^*_i are independent variates of proportional distribution. Each population is divided into three equal strata and a stratified ranked set sample of size 9 with set size 3 and number of cycles 3 is drawn from each stratum with the help of the methodology described in Section 2. With 10000 iterations, the percent relative efficiency (PRE) of the sequent estimators with respect to the conventional mean estimator was obtained as

$$PRE = \frac{\frac{1}{10000}\sum_{i=1}^{10000}(T_m - \bar{Y})^2}{\frac{1}{10000}\sum_{i=1}^{10000}(T^* - \bar{Y})^2} \times 100,$$

where T^* is the existing and proposed combined and separate class of estimators.

The findings of the simulation are disclosed from Table 1 to Table 4 through PRE for reasonably chosen values of correlation coefficient $\rho_{xy} = 0.6, 0.7, 0.8, 0.9$ and fair choice of response probability $P = 0.4, 0.6$.

From Table 1 to Table 4, consisting of the simulation results of two asymmetric populations, namely, gamma and exponential, we have seen that the proposed combined and separate imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ dominate the other existing imputation methods for reasonably chosen values of the correlation coefficient. We have also seen that the proposed combined and separate imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 4, 5, 6$ perform better among the proposed class of imputation methods under situations I, II and III. Moreover, it is also seen that the PRE of the proposed imputation methods under situations I, II and III in both populations decreases with the increase in asymmetry.

Table 1: PRE of proposed combined estimators at $P = 0.4$.

ρ_{xy}	0.6	0.7	0.8	0.9
T_m^c	100	100	100	100
$x^* \sim \Gamma(0.5, 1.5)$				
$y^* \sim \Gamma(1.5, 2)$				
Skewness of y	1.3292	1.4081	1.6083	1.9768
Kurtosis of y	5.4000	5.6181	6.5418	8.4693
Situation I				
$T_{SA_i}^c,\ i = 1, 7$	106.0998	105.5777	105.4108	104.7131
$\mathbf{T_{SA_4}^c}$	**106.1067**	**105.5846**	**105.4176**	**104.7198**
$T_{R_1}^c$	100.2267	99.3709	98.3220	95.5744
$T_{DP_1}^c / T_{S_i}^c,\ i = 1, 4$	105.983	105.46	105.3005	104.6141
Situation II				
$T_{SA_i}^c,\ i = 2, 8$	115.0805	113.7576	113.4881	111.7679
$\mathbf{T_{SA_5}^c}$	**115.1003**	**113.7771**	**113.5075**	**111.7865**
$T_{R_1}^c$	97.5580	95.7573	93.4999	87.5916
$T_{DP_2}^c / T_{S_i}^c,\ i = 2, 5$	114.9637	113.6399	113.3778	111.6689
Situation III				
$T_{SA_i}^c,\ i = 3, 9$	108.1050	107.4685	107.3868	106.5397
$\mathbf{T_{SA_6}^c}$	**108.1161**	**107.4795**	**107.3977**	**106.5502**
$T_{R_1}^c$	97.3431	96.3414	95.0160	91.2945
$T_{DP_3}^c / T_{S_i}^c,\ i = 3, 6$	107.9882	107.3508	107.2765	106.4406
$x^* \sim Exp(3.0)$				
$y^* \sim Exp(2.0)$				
Skewness of y	1.4612	1.3814	1.3734	1.4769
Kurtosis of y	5.9268	5.4885	5.4119	5.8395
Situation I				
$T_{SA_i}^c,\ i = 1, 7$	106.2035	106.5419	106.7350	105.8748
$\mathbf{T_{SA_4}^c}$	**106.2096**	**106.5481**	**106.7411**	**105.8808**
$T_{R_1}^c$	99.7239	100.2227	100.2960	98.5947
$T_{DP_1}^c / T_{S_i}^c,\ i = 1, 4$	106.1075	106.4467	106.6419	105.7829
Situation II				
$T_{SA_i}^c,\ i = 2, 8$	112.9619	112.5542	112.8376	111.6621
$\mathbf{T_{SA_5}^c}$	**112.9805**	**112.5726**	**112.856**	**111.6799**
$T_{R_2}^c$	89.3047	88.4097	88.1260	85.9139
$T_{DP_2}^c / T_{S_i}^c,\ i = 2, 5$	112.8659	112.4590	112.7445	111.5702
Situation III				
$T_{SA_i}^c,\ i = 3, 9$	105.3578	105.7988	105.8817	105.2694
$\mathbf{T_{SA_6}^c}$	**105.3684**	**105.8092**	**105.8920**	**105.2796**
$T_{R_3}^c$	87.5260	88.2363	87.8974	86.9790
$T_{DP_3}^c / T_{S_i}^c,\ i = 3, 6$	105.2617	105.7036	105.7886	105.1775

Table 2: PRE of proposed combined estimators at $P = 0.6$.

ρ_{xy}	0.6	0.7	0.8	0.9
T_m^c	100	100	100	100
$x^* \sim \Gamma(0.5, 1.5)$				
$y^* \sim \Gamma(1.5, 2)$				
Skewness of y	1.3292	1.4081	1.6083	1.9768
Kurtosis of y	5.4000	5.6181	6.5418	8.4693
Situation I				
$T_{SA_i}^c,\ i = 1, 7$	109.3365	108.5032	108.2434	107.1516
$\mathbf{T_{SA_4}^c}$	**109.3437**	**108.5104**	**108.2505**	**107.1584**
$T_{R_1}^c$	100.3407	99.0588	97.5032	93.5045
$T_{DP_1}^c / T_{S_i}^c,\ i = 1, 4$	109.2587	108.4248	108.1698	107.0855
Situation II				
$T_{SA_i}^c,\ i = 2, 8$	115.3909	114.0088	113.6905	111.9040
$\mathbf{T_{SA_5}^c}$	**115.4041**	**114.0218**	**113.7034**	**111.9163**
$T_{R_2}^c$	98.2571	96.3813	94.0605	88.0659
$T_{DP_2}^c / T_{S_i}^c,\ i = 2, 5$	115.3131	113.9304	113.6170	111.8379
Situation III				
$T_{SA_i}^c,\ i = 3, 9$	105.1641	104.7748	104.7360	104.2151
$\mathbf{T_{SA_6}^c}$	**105.1690**	**104.7796**	**104.7407**	**104.2197**
$T_{R_3}^c$	97.9303	97.2720	96.3820	93.8046
$T_{DP_3}^c / T_{S_i}^c,\ i = 3, 6$	105.0863	104.6964	104.6625	104.1491
$x^* \sim Exp(3.0)$				
$y^* \sim Exp(2.0)$				
Skewness of y	1.4612	1.3814	1.3734	1.4769
Kurtosis of y	5.9268	5.4885	5.4119	5.8395
Situation I				
$T_{SA_i}^c,\ i = 1, 7$	109.5194	110.0755	110.3901	109.0041
$\mathbf{T_{SA_4}^c}$	**109.5258**	**110.0819**	**110.3965**	**109.0104**
$T_{R_1}^c$	99.5862	100.3351	100.4455	97.9046
$T_{DP_1}^c / T_{S_i}^c,\ i = 1, 4$	109.4554	110.0121	110.3281	108.9429
Situation II				
$T_{SA_i}^c,\ i = 2, 8$	113.7547	113.6404	113.9884	112.5204
$\mathbf{T_{SA_5}^c}$	**113.7669**	**113.6525**	**114.0006**	**112.5322**
$T_{R_2}^c$	91.4958	91.0133	90.8178	88.2189
$T_{DP_2}^c / T_{S_i}^c,\ i = 2, 5$	113.6907	113.577	113.9264	112.4592
Situation III				
$T_{SA_3}^c$	103.4189	103.5222	103.5633	103.3902
$\mathbf{T_{SA_6}^c}$	**103.4436**	**103.5267**	**103.5677**	**103.4246**
$T_{R_3}^c$	90.6449	90.7375	90.4534	89.9166
$T_{DP_3}^c / T_{S_i}^c,\ i = 3, 6$	103.3749	103.4588	103.5013	103.3590

Table 3: PRE of proposed separate estimators at $P = 0.4$.

ρ_{xy}	0.6	0.7	0.8	0.9
T_m^s	100	100	100	100
$x^* \sim \Gamma(0.5, 1.5)$				
$y^* \sim \Gamma(1.5, 2)$				
Skewness of y	1.3292	1.4081	1.6083	1.9768
Kurtosis of y	5.4000	5.6181	6.5418	8.4693
Situation I				
$T_{SA_i}^s,\ i = 1, 7$	106.0962	105.5742	105.4076	104.7104
$\mathbf{T_{SA_4}^s}$	**106.1027**	**105.5807**	**105.4141**	**104.7167**
$T_{R_1}^s$	100.6188	99.7795	98.7804	96.1227
$T_{DP_1}^s/T_{S_i}^s,\ i = 1, 4$	105.983	105.4600	105.3005	104.6141
Situation II				
$T_{SA_i}^s,\ i = 2, 8$	115.0769	113.7541	113.4850	111.7652
$\mathbf{T_{SA_5}^s}$	**115.0954**	**113.7724**	**113.5031**	**111.7826**
$T_{R_2}^s$	98.5741	96.7889	94.6180	88.8216
$T_{DP_2}^s/T_{S_i}^s,\ i = 2, 5$	114.9637	113.6399	113.3778	111.6689
Situation III				
$T_{SA_i}^s,\ i = 3, 9$	108.1014	107.4650	107.3836	106.5370
$\mathbf{T_{SA_6}^s}$	**108.1119**	**107.4753**	**107.3939**	**106.5469**
$T_{R_3}^s$	81.4466	80.3429	77.9883	72.6504
$T_{DP_3}^s/T_{S_i}^s,\ i = 3, 6$	107.9882	107.3508	107.2765	106.4406
$x^* \sim Exp(3.0)$				
$y^* \sim Exp(2.0)$				
Skewness of y	1.4612	1.3814	1.3734	1.4769
Kurtosis of y	5.9268	5.4885	5.4119	5.8395
Situation I				
$T_{SA_i}^s,\ i = 1, 7$	106.2010	106.5393	106.7324	105.8724
$\mathbf{T_{SA_4}^s}$	**106.2067**	**106.5451**	**106.7382**	**105.8781**
$T_{R_1}^s$	100.0305	100.5150	100.5927	98.9217
$T_{DP_1}^s/T_{S_i}^s,\ i = 1, 4$	106.1075	106.4467	106.6419	105.7829
Situation II				
$T_{SA_i}^s,\ i = 2, 8$	112.4594	112.5516	112.8350	111.6596
$\mathbf{T_{SA_5}^s}$	**112.4770**	**112.5691**	**112.8525**	**111.6766**
$T_{R_2}^s$	90.1295	89.1992	88.9243	86.7481
$T_{DP_2}^s/T_{S_i}^s,\ i = 2, 5$	112.8659	112.4590	112.7445	111.5702
Situation III				
$T_{SA_i}^s,\ i = 3, 9$	105.3552	105.7962	105.8791	105.2670
$\mathbf{T_{SA_6}^s}$	**105.3654**	**105.8061**	**105.889**	**105.2767**
$T_{R_3}^s$	74.6746	73.9110	73.4349	71.8010
$T_{DP_3}^s/T_{S_i}^s,\ i = 3, 6$	105.2617	105.7036	105.7886	105.1775

Table 4: PRE of proposed separate estimators at $P = 0.6$.

ρ_{xy}	0.6	0.7	0.8	0.9
T_m^s	100	100	100	100
$x^* \sim \Gamma(0.5, 1.5)$				
$y^* \sim \Gamma(1.5, 2)$				
Skewness of y	1.3292	1.4081	1.6083	1.9768
Kurtosis of y	5.4000	5.6181	6.5418	8.4693
Situation I				
$T_{SA_i}^s,\ i = 1, 7$	109.3341	108.5009	108.2413	107.1498
$\mathbf{T_{SA_4}^s}$	**109.3409**	**108.5076**	**108.2479**	**107.1562**
$T_{R_1}^s$	100.9318	99.6695	98.1811	94.2940
$T_{DP_1}^s / T_{S_i}^s,\ i = 1, 4$	109.2587	108.4248	108.1698	107.0855
Situation II				
$T_{SA_i}^s,\ i = 2, 8$	115.3885	114.0065	113.6884	111.9022
$\mathbf{T_{SA_5}^s}$	**115.4008**	**114.0187**	**113.7005**	**111.9138**
$T_{R_2}^s$	99.2688	97.4083	95.1744	89.2927
$T_{DP_2}^s / T_{S_i}^s,\ i = 2, 5$	115.3131	113.9304	113.6170	111.8379
Situation III				
$T_{SA_i}^s,\ i = 3, 9$	105.1618	104.7725	104.7339	104.2133
$\mathbf{T_{SA_6}^s}$	**105.1663**	**104.7770**	**104.7383**	**104.2176**
$T_{R_3}^s$	75.3238	74.4041	72.0020	66.8509
$T_{DP_3}^s / T_{S_i}^s,\ i = 3, 6$	105.0863	104.6964	104.6625	104.1491
$x^* \sim Exp(3.0)$				
$y^* \sim Exp(2.0)$				
Skewness of y	1.4612	1.3814	1.3734	1.4769
Kurtosis of y	5.9268	5.4885	5.4119	5.8395
Situation I				
$T_{SA_i}^s,\ i = 1, 7$	109.5177	110.0737	110.3884	109.0025
$\mathbf{T_{SA_4}^s}$	**109.5238**	**110.0798**	**110.3944**	**109.0084**
$T_{R_1}^s$	100.0457	100.7759	100.8935	98.3895
$T_{DP_1}^s / T_{S_i}^s,\ i = 1, 4$	109.4554	110.0121	110.3281	108.9429
Situation II				
$T_{SA_i}^s,\ i = 2, 8$	113.7531	113.8386	113.9867	112.5188
$\mathbf{T_{SA_5}^s}$	**113.7646**	**113.8502**	**113.9983**	**112.5300**
$T_{R_2}^s$	92.3101	91.7963	91.6107	89.0465
$T_{DP_2}^s / T_{S_i}^s,\ i = 2, 5$	113.6907	113.7770	113.9264	112.4592
Situation III				
$T_{SA_i}^s,\ i = 3, 9$	103.4372	103.5205	103.5615	103.4185
$\mathbf{T_{SA_6}^s}$	**103.4417**	**103.5247**	**103.5658**	**103.4228**
$T_{R_3}^s$	70.0312	69.5489	69.0562	67.4148
$T_{DP_3}^s / T_{S_i}^s,\ i = 3, 6$	103.3749	103.4588	103.5013	103.3590

7. CONCLUSION

This paper is the outset to suggest some combined and separate classes of imputation methods along with their properties for the estimation of population mean in the presence of missing data using $SRSS$. The theoretical conditions are derived under which the proposed combined and separate classes of imputation methods are justified. In order to enhance the theoretical findings and to determine the effect of skewness and kurtosis over PRE, a simulation study is accomplished on two asymmetric populations viz. gamma and exponential with reasonable choice of correlation coefficient ρ_{xy} and probability of non responding units P. It is noticed from the perusal of theoretical and simulation results that:

1. The proposed combined and separate of imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ always perform better than the combined and separate mean imputation method $y^c_{.i_m}$ and $y^s_{.i_m}$, ratio imputation methods $y^c_{.i_{R_j}}$ and $y^s_{.i_{R_j}}$, $j = 1, 2, 3$ and their own conventional counterparts for different values of correlation coefficient ρ_{xy}, coefficient of skewness β_1 and coefficient of kurtosis β_2.
2. The proposed combined and separate imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 4, 5, 6$ are best among the proposed classes of imputation methods under situations I, II and III.
3. The PRE of the proposed combined and separate classes of imputation methods $y^c_{.i_{SA_j}}$, $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ and their conventional counterparts under situations I, II and III are contrary to the asymmetry which is similar to the results of McIntyre (1952), Dell and Clutter (1972) and Bhushan and Kumar (2022) where they chose a wide range of skewed distributions and concluded that the asymmetry shows adverse effect over the efficiency of the estimators.
4. The suggested combined classes of imputation methods $y^c_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ are superior than the suggested separate classes of imputation methods $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ in situations I, II and III.

Therefore, due to the dominance of the proposed imputation methods over the existing imputation methods, we recommend them to survey persons for their real life problems.

APPENDIXES A-B-C. Supplementary file

Supplementary data to this article can be found online.

ACKNOWLEDGMENTS

The authors are extremely grateful to the learned referee for his valuable comments and Editor-in-Chief.

REFERENCES

Ahmed, M. S., Al-Titi, O., Al-Rawi, Z., and Abu-Dayyeh, W. (2006). Estimation of a population mean using different imputation methods. *Statistics in Transition*, 7(6):1247–1264.

Al-Omari, A. and Bouza, C. (2014). Ratio estimators of the population mean with missing values using ranked set sampling. *Environmetrics*, 26(2):67–76.

Bhushan, S. and Kumar, A. (2022). Novel log type class of estimators under ranked set sampling. *Sankhya B*, 84:421–447.

Bhushan, S. and Pandey, A. P. (2016). Optimal imputation of missing data for estimation of population mean. *Journal of Statistics and Management System*, 19(6):755–769.

Bhushan, S. and Pandey, A. P. (2018). Optimality of ratio type estimation methods for population mean in presence of missing data. *Communication in Statistics – Theory and Methods*, 47(11):2576–2589.

Bhushan, S., Pandey, A. P., and Pandey, A. (2018). On optimality of imputation methods for estimation of population mean using higher order moments of an auxiliary variable. *Communication in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2018.1500595

Bhushan, S., Kumar, A., Pandey, A. P., and Singh, S. (2022). Estimation of population mean in presence of missing data under simple random sampling. *Communications in Statistics – Simulation and Computation*, https://doi.org/10.1080/03610918.2021.2006713.

Bouza-Herrera, C. N. and Al-Omari, A. I. (2011). Ranked set estimation with imputation of the missing observations: The median estimator. *Revista Investigation Operacional*, 32(1):30–37.

Bouza, C. N., Viada, C., and Vishwakarma, G. K. (2020). Studying the total under missingness by guessing the value of a superpopulation model for imputation. *Revista Investigation Operacional*, 41(7):979–989.

Bouza-Herrera, C. N. and Viada, C. E. (2021). Imputation of individual values of a variable using product predictors. *Revista Investigation Operacional*, 42(3):321–325.

Cochran, W. G. (1977). *Sampling Techniques*. John Wiley and Sons, New York.

Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, 28:545–555.

Diana, G. and Perri, P. F. (2010). Improved estimators of the population mean for missing data. *Communication in Statistics – Theory and Methods*, 39:3245–3251.

Heitjan, D. F. and Basu, S. (1996). Distinguishing 'Missing at Random' and 'Missing Completely at Random'. *The American Statistician*, 50:207–213.

Kadilar, C. and Cingi, H. (2008). Estimators for the population mean in the case of missing data. *Communication in Statistics – Theory and Methods*, 37:2226–2236.

Khan, L. and Shabbir, J. (2016). Hartley-Ross type unbiased estimators using ranked set sampling and stratified ranked set sampling. *North Carolina Journal of Mathematics and Statistics*, 2:10–22.

Lee, H., Rancourt, E., and Sarndal, C. E. (1994). Experiments with variance estimation from survey data with imputed values. *Journal of Official Statistics*, 10:231–243.

Linder, D. F., Samawi, H., Yu, L., Chatterjee, A., Huang, Y., and Vogel, R. (2015). On stratified ranked set sampling for regression estimators. *Journal of Applied Statistics*, 42(12):2571–2583.

Mandowara, V. L. and Mehta, N. (2014). Modified ratio estimators using stratified ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, 43(3):461–471.

McIntyre, G. A. (1952). A method of unbiased selective sampling using ranked set. *Australian Journal Agricultural Research*, 3:385–390.

Mohamed, C., Sedory, S. A., and Singh, S. (2016). Imputation using higher order moments of an auxiliary variable. *Communication in Statistics – Simulation and Computation*, 46:6588–6617.

Prasad, S. (2017). Ratio exponential type estimators with imputation for missing data in sample surveys. *Model Assisted Statistics and Applications*, 12(2):95–106.

Rubin, R. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

Saini, M. and Kumar, A. (2018). Ratio estimators using stratified random sampling and stratified ranked set sampling. *Life Cycle Reliability and Safety Engineering*, 84(5):931–945.

Samawi, H. M. (1996). Stratified ranked set sampling. *Pakistan Journal of Statistics*, 12(1):9–16.

Samawi, H. M. and Siam, M. I. (2003). Ratio estimation using stratified ranked set sampling. *Metron – International Journal of Statistics*, LXI(1):75–90.

Singh, S. and Deo, B. (2003). Imputation by power transformation. *Statistical Papers*, 44:555–579.

Singh, S. (2009). A new method of imputation in survey sampling. *Statistics: A Journal of Theoretical and Applied Statistics*, 43(5):499–511.

Singh, S. and Horn, S. (1998). An alternative survey in multi-character survey. *Metrika*, 48:99–107.

Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. *Metrika*, 51:267–276.

Singh, S. and Valdes, S. R. (2009). Optimal method of imputation in survey sampling. *Applied Mathematical Sciences*, 3(35):1727–1737.

Sohail, M. U., Shabbir, J., and Ahmed, S. (2018). A class of ratio type estimators for imputing the missing values under rank set sampling. *Journal of Statistical Theory and Practice*, 1–31.

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 20:1–31.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 79–100
https://doi.org/10.57805/revstat.v23i1.438

Data Analytics and Distribution Function Estimation via Mean Absolute Deviation: Nonparametric Approach

Author: Elsayed A. H. Elamir
– Department of Management and Marketing, University of Bahrain, Kingdom of Bahrain
shabib@uob.edu.bh

Received: March 2022 Revised: January 2023 Accepted: January 2023

Abstract:

- Mean absolute deviation function is used to explore the pattern and the distribution of the data graphically to enable analysts gaining greater understanding of raw data and to foster a quick and a deep understanding of the data as an important basis for successful data analytics. Furthermore, new nonparametric approaches for estimating the cumulative distribution function based on the mean absolute deviation function are proposed. These new approaches are meant to be a general nonparametric class that includes the empirical distribution function as a special case. Simulation study reveals that the Richardson extrapolation approach has a major improvement in terms of average squared errors over the classical empirical estimators and has comparable results with smooth approaches such as cubic spline and constrained linear spline for practically small samples. The properties of the proposed estimators are studied. Moreover, the Richardson approach has been applied to real data analysis and has been used to estimate the hazardous concentration five percent.

Keywords:

- *empirical distribution function; nonparametric estimation; numerical differentiation; Richardson extrapolation; skewness; uniform consistency.*

AMS Subject Classification:

- 62G30, 62G32.

✉ Corresponding author

1. INTRODUCTION

Data analytic techniques are very important to explore the structure and the distribution of the data to enable analysts gaining greater understanding of the raw data. Data is often collected in large, unstructured volumes from various sources and analysts must first understand and develop a comprehensive view of the data before using it in further analysis; see, Healy (2019). Nowadays, data exploration has established as a mandatory phase in every data science project. Typical plots include scatter plot, box plot, quantile quantile plot and many more have been used as a graphical approach to learn about distributions, correlations, outliers, trends, and other data characteristics; see, Tukey (1987), Gandomi and Haider (2015) and Cumming and Finch (2005). One main advantage of data exploration graphically is to learn about characteristics and potential problems of a data set without the need to formulate assumptions about the data beforehand and to foster a quick and a deep understanding of the data as an important basis for successful and efficient data science projects; see, Matt and Joshua (2019), Runkler (2020), James *et al.* (2013), Healy (2019), and Larson-Hall (2017). The estimation of a distribution function is not only a fascinating problem by itself, but it also emerges naturally in real-world problems in a variety of scientific domains including commerce, hydrology, and environmental sciences. As a result, a variety of nonparametric approaches for tackling this problem have risen in different disciplines; see, Efromovich (2001), Cheng and Peng (2002), and Charles *et al.* (2010). The risk term, or natural hazard, appears to be closely tied to the distribution function in many circumstances. Scientists want to know the likelihood of a large earthquake, the likelihood of high wind speed or hurrican, and the hazard of low levels; see, Baszczynska (2016), Xue and Wang (2010), Babu *et al.* (2002), Erdogan *et al.* (2019) and Mombeni *et al.* (2021).

The population mean absolute deviation (MAD) about any value v can be written as

$$\Delta_X(v) = E|X - v|, v \in R. \tag{1.1}$$

This function is usually used as a direct measure of the scale for any distribution about chosen v such as mean absolute deviation about population mean (μ)

$$\Delta_X(\mu) = E|X - \mu|$$

and mean absolute deviation about population median (M)

$$\Delta_X(M) = E|X - M|.$$

These measures offer a direct indication of the dispersion of a random variable about its mean and median, respectively, and have many applications in different fields; see, Dodge (2002), Pham-Gia and Hung (2001), Gorard (2005), Elamir (2012) and Habib (2012).

There are two main aims for this article. Since the mean absolute deviation function characterizes the distribution function and gives a dispersive ordering of probability distributions, the first aim is to use the mean absolute deviation function to explore the pattern in the data graphically. The second aim is to use the first derivative of the mean absolute function to estimate the population distribution function where a new method based on Richardson extrapolation approach is proposed.

This article is organized as follows. MAD function representation is explained in Section 2. MAD plot is proposed in Section 3. Some uses of MAD function are introduced in Section 4. Distribution function in terms of MAD is presented in Section 5. Several nonparametric estimation approaches for distribution functions are derived in Section 6. Simulation study is conducted to study the properties of proposed estimators in terms of average mean square in Section 7. Ricardson extrapolation approximation is applied to acute toxicity values in Section 8. Section 9 is devoted to conclusion.

2. MAD FUNCTION AND ITS REPRESENTATION

Let $X_1, ..., X_n$ be an independent and identically a random sample from a continuous distribution function $F_X(.)(0 < F < 1)$, density $f_X(.)(f \geq 0)$, mean $\mu = E(X)$, median $M = Med(X)$, standard deviation $\sigma = \sqrt{E(x-\mu)^2}$, indicator function $I_{i \leq k}$ be 1 if $i \leq k$, 0 else, and $X_{(1)}, ..., X_{(n)}$ be the corresponding order statistics. Another representation of MAD in terms of distribution function is given by Munoz-Perez and Sanchez-Gomez (1990) as

(2.1) $$\Delta_X(v) = v[2F_X(v) - 1] + E(X) - 2E[XI_{X \leq v}]$$

and its first derivative

(2.2) $$\acute{\Delta}_X(v) = 2F_X(v) - 1.$$

For more details, see, Habib (2012).

Theorem 2.1. *The mean absolute deviation about $v(v \in R)$ is minimized when v is the median and it is a convex function.*

Proof: Since $F_X(M) = 0.5$, the first derivative of MAD function at median (M) is zero $\acute{\Delta}_X(M) = 0$ with positive second derivative $\acute{\acute{\Delta}}_X(M) = 2f_X(M) > 0$. Therefore, $\Delta_X(v) = E|X - v|$ has a minimum value at $v = M$. Where $\acute{\Delta}_X(v) = 2F_X(v) - 1$ and $\acute{\acute{\Delta}}_X(v) = 2f_X(v) \geq 0$ for all $v \in R$, then $\Delta_X(v)$ is a convex function. □

Munoz-Perez and Sanchez-Gomez (1990) prove that $\Delta_X(v)$ characterizes the distribution function and give a dispersive ordering of probability distributions as it satisfies the following conditions: (1) there is only a finite number of discontinuity points in the derivative, (2) it is a convex function on real line (R), (3) $\lim_{x \to \infty} \acute{\Delta}_X(x) = 1$ and $\lim_{x \to -\infty} \acute{\Delta}_X(x) = -1$ and (4) $\lim_{x \to \infty}[\Delta_X(x) - x] = -E(X)$, and $\lim_{x \to -\infty}[\Delta_X(x) + x] = E(X)$. Since $\Delta_X(v)$ satisfies the above conditions, there subsists a unique distribution function which has $\Delta_X(v)$ as its dispersion function.

3. MEAN ABSOLUTE DEVIATION PLOT (MAD PLOT)

The MAD plot can be introduced as

(3.1) $$X_{axis} = v_i, \text{ versus } Y_{axis} = \Delta_X(v_i), \text{ for each } v_i = x_i \text{ and } i = 1, ..., n$$

with two straight lines at

$$\mu - v_i \text{ and } v_i - \mu.$$

This plot represents data on x-axis and the mean absolute deviation at each $v = x$ on the y-axis that includes mean absolute deviation about mean and median as special cases. In other words, it is a simple curve plot between the actual data and its mean absolute deviation at each point. Figure 1 displays the MAD plot for the standard normal distribution using the quantile function for standard normal from R-software (2021) $v_i = qnorm(p = (i - 0.5)/n,$ $\mu = 0, \sigma = 1)$, and $i = 1, ..., 100$ with two straight lines $\mu - v_i$ and $v_i - \mu$ that show the degree of approximation with $\Delta_X(v_i)$.

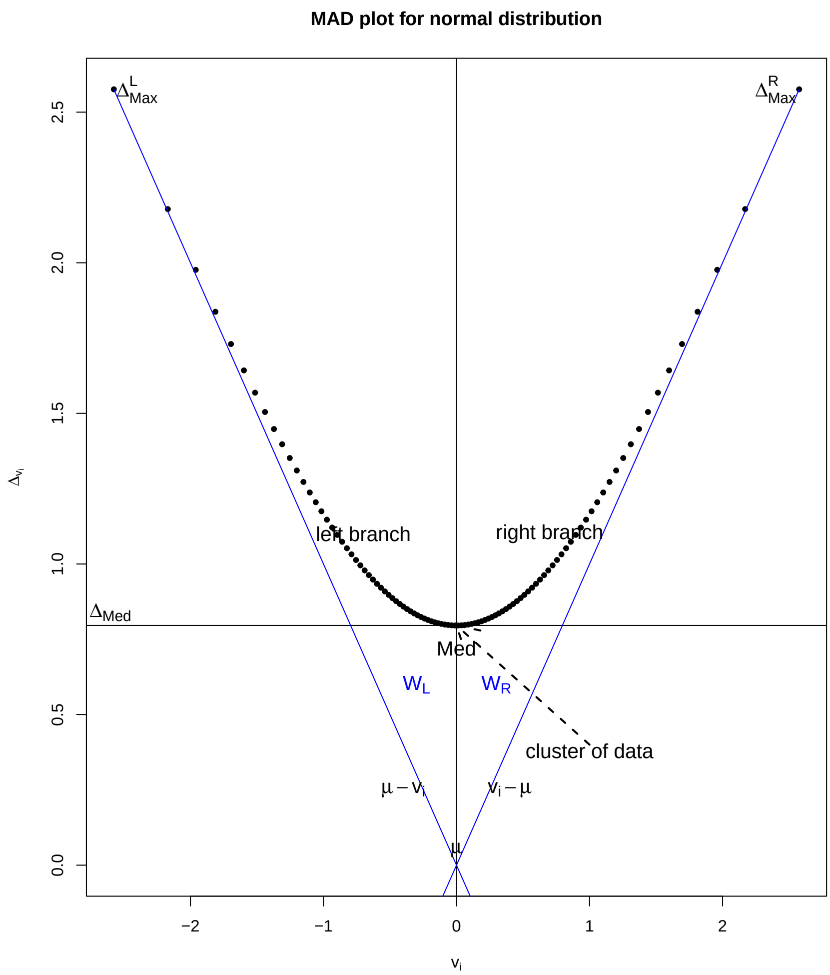


Figure 1: MAD function plot for standard normal distribution using $qnorm((i - 0.5)/n)$, and $i = 1, ..., 100$.

For the standard normal distribution, the MAD function formulates a parabola curve, or a quadratic function that has a minimum at median $\Delta_X(M) = \Delta_{Med}$ (mean absolute deviation about median) and reflects a lot of information that includes:

- the location measure median (M) on x-axis is at min $\Delta_X(v)$ and the mean (μ) is at the intersection of straight lines,
- the scale measure $\Delta_X(M)$ (mean absolute deviation about median) is at the minimum of the MAD function,

- the right MAD branch $\Delta_X(v)I_{v>M}$ and its maximum Δ^R_{Max} give an indication of spread out of the data and tail length in the right side of median,
- the left MAD branch $\Delta_X(v)I_{v<M}$ and its maximum Δ^L_{Max} give an indication of spread out of the data and tail length in the left side of median,
- two straight lines $\mu - v_i$ and $v_i - \mu$ give the degree of approximation with $\Delta_X(v_i)$,
- the wideness between MAD function, median, straight lines $\mu - v_i$ and $v_i - \mu$ (right wideness W_R, and left widenes W_L) give an indication of the direction, degree of skewness and peakedness,
- the cluster of data may give an indication about modality.

Furthermore, the MAD function could be divided to the right MAD function ($\Delta^+_X(v)$) and the left MAD function ($\Delta^-_X(v)$) as

$$\Delta_X(v) = E|X - v| = E(X - v)^+ + E(X - v)^- = \Delta^+_X(v) + \Delta^-_X(v)$$

and

$$\Delta^+_X(v) = E[(X - v)I_{X>v}] \text{ and } \Delta^-_X(v) = E[(v - X)I_{X\leq v}].$$

The relationship between the straight lines and the MAD functions can be written as

$$E(X - v) = \Delta^+_X(v) - \Delta^-_X(v).$$

Therefore, when $v = \mu$, we have $\Delta^+_X(v) = \Delta^-_X(v)$, also if $v = M$, we have $\mu - M = \Delta^+_X(M) - \Delta^-_X(M)$ that could be considered as a measure of skewness; see, Munoz-Perez and Sanchez-Gomez (1990) and Habib (2012).

4. USES OF MAD FUNCTIONS

4.1. Wideness and skewness

The area between the right straight line, median and right MAD branch can be defined in standard form as

$$W_R = \Delta^+_X(M)/\sigma$$

and

$$\Delta^+_X(M) = E[(X - M)I_{X>M}].$$

We may consider W_R as the right wideness measure which reflect how much the right MAD branch is away from the right straight line ($v_i - \mu$) and median. Since $\Delta \leq \sigma$ by Jensen's inequality, it is straightforward to prove that $0 \leq W_R \leq 1$. In terms of data, when the value of W_R is near 1 it indicates big wideness or stretch out from the median, in other words, the data will be spread out far away from median in the right side. If the value of W_R is near 0 it indicates small wideness from the median, in other words, the data will be concentrated near the median in the right side.

Also, the area between the left straight line, median and the left MAD branch (left wideness) can be defined in standard form as

$$W_L = \Delta_X^-(M)/\sigma$$

and

$$\Delta_X^-(M) = E[(M - X)I_{X \leq M}].$$

We may consider W_L as the left wideness measure which reflect how much the left MAD branch is away from left straight line $(\mu - v_i)$ and median. It is straightforward to prove that $0 \leq W_L \leq 1$. In terms of data, when the value of W_L is near 1 it indicates big wideness or stretches out from the median. In other words, the data will be spread out far away from median in the left side. If the value of W_L is near 0 it indicates small wideness from the median, in other words, the data will be closed to the median in the left side.

The general measure of wideness for a distribution in terms of right and left wideness may be defined as

$$W = W_L + W_R = (\Delta_X^-(M) + \Delta_X^+(M))/\sigma = \Delta_X(M)/\sigma.$$

We may consider W as a measure of total wideness between MAD function and the two straight lines. This measure will be very useful for symmetric distributions where it may be related to what is called platykurtic or flatness that had been used as a test for normal distribution; see, Geary (1935) and Elamir (2012). The interpretation of this measure especially for symmetric distributions in terms of data can be as follows. If W is near 1, the distribution of the data is "strong curved inwards", near zero strong "curved outwards", and 0.798 for normal distribution. The tightness between MAD function and the two straight lines may be defined as a complement of wideness as

$$L = 1 - W = (\sigma - \Delta_X(M))/\sigma.$$

The standardized distance between the standard deviation of the population and the mean absolute deviation about median. This measure is very useful for symmetric distributions where it may be related to what is called leptokurtic (peakedness). As σ getting far away from $\Delta_X(M)$, the more leptokurtic (more data concentration about median). Since $\Delta_X(M) \leq \sigma$, then $0 \leq L \leq 1$. The values of W_R, W_L, W and L are presented in Table 1 for some selected symmetric distributions. The distributions that have big flatness as Beta (1, 1) and Beta (0.1, 0.1) have W near 1 while the distributions with strong peakedness such as $t(df = 3)$ have W with a low value.

Table 1: Wideness and leptokurtic for some symmetric distributions.

Distribution	W_R	W_L	W	L
Beta(0.1,0.1)	0.483	0.483	0.966	0.034
Beta(1,1)	0.433	0.433	0.866	0.134
Normal	0.399	0.399	0.798	0.202
Logistic	0.382	0.382	0.765	0.235
Laplace	0.355	0.355	0.71	0.29
$t(df = 3)$	0.326	0.326	0.652	0.348

With respect to the tails of the distribution, different measures may be proposed from Figure 1 as

$$T_R = \Delta^R_{Max}/\sigma \text{ and } T_L = \Delta^L_{Max}/\sigma.$$

Alternatively,

$$T_{R1} = \sigma/\Delta^R_{Max} \text{ and } T_{L1} = \sigma/\Delta^L_{Max}.$$

where

$$\Delta^R_{Max} = Max.[(v - X)I_{X\leq v}] \text{ and } \Delta^L_{Max} = Max.[(X - v)I_{X>v}].$$

All values of these measures are more than or equal 0 with no upper value. These measures give an indication about tail length. For the first two measures, the small values around 1 indicate short tails, around 3 medium tails while larger values indicate long tail. For the other two measures T_{R1} and T_{L1}, the values around zero indicate very long tails, values around 0.30 indicate medium tails while values around one indicate short tails. Note that the values of the above measures will depend on the sample size. The first measure of skewness in terms of wideness can be defined as

$$SK_1 = W_R - W_L = (\mu - M)/\sigma.$$

This measure is equivalent to Groeneveld and Meeden (1984) measure of skewness. The second measure of skewness in terms of tailedness can be proposed as

$$SK_2 = T_R - T_L = (\Delta^R_{Max} - \Delta^L_{Max})/\sigma.$$

Alternatively,

$$SK_{21} = T_{L1} - T_{R1} = \sigma/\Delta^L_{Max} - \sigma/\Delta^R_{Max}.$$

Figure 2 displays a MAD plot for Beta (0.1,0.1), normal, Laplace and exponential distributions using their quantile function for each distribution in R-software with $n = 300$. It may conclude that:

- the location measures median and mean values are located at minimum of MAD curve and intersection between two straight lines on x axis, respectively, while the dispersion measure $\Delta_X(M)$ is located on y axis at minimum of MAD curve,
- tails measures (T_R and T_L) near 1 may give an indication of short tail such as beta distribution while around 3 may give an indication of medium tail such as normal distribution,
- equal wideness (W_R, W_L) may give an indication of symmetric distributions such as beta, normal and Laplace, while not equal measures are indication of skewed distributions such as exponential,
- for symmetric distributions, a value of wideness as 0.966 may give an indication of strong "curved inwards" such as beta (0.1,0.1) and value as 0.71 may give an indication of data close to the median.

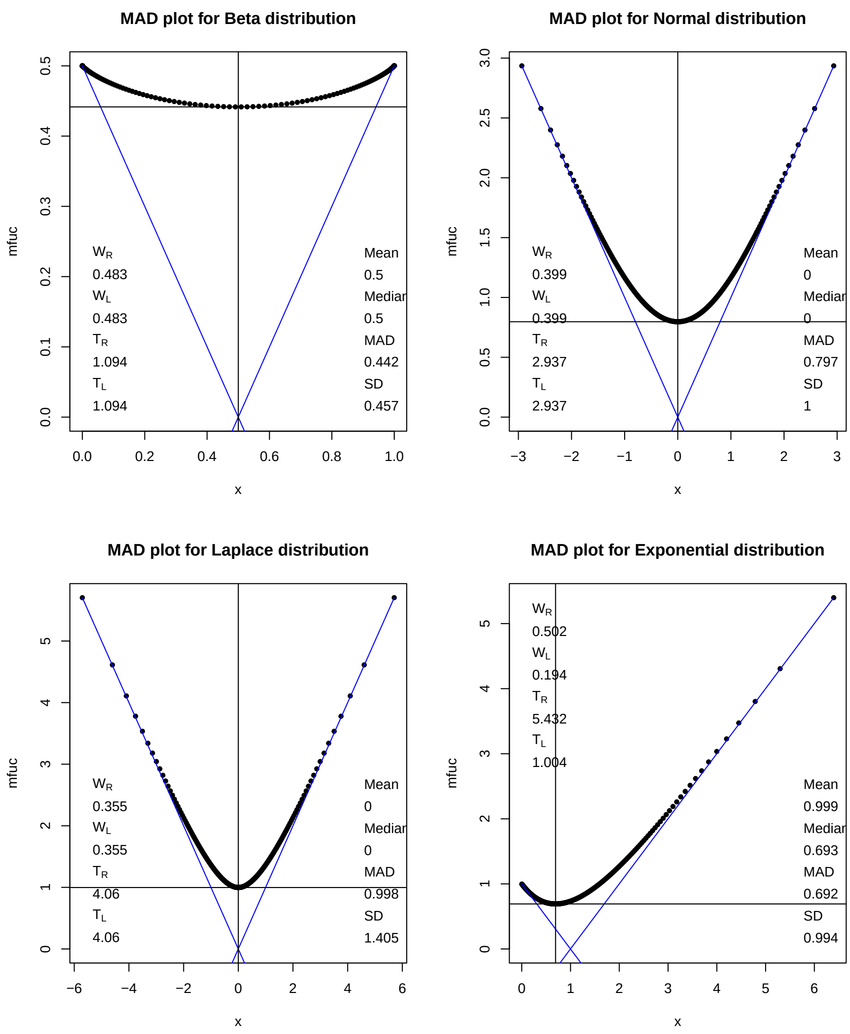


Figure 2: MAD plot for Beta (0.1,0.1), normal, Laplace and exponential distributions using quantile function for each distribution in R-software with $p = (i - 0.5)/n$, and $i = 1, ..., 300$.

5. DISTRIBUTION FUNCTION IN TERMS OF MAD

According to Munoz-Perez and Sanchez-Gomez (1990), the MAD function can be rewritten in terms of indicator function as

$$\Delta_X(v) = E[(X - v)I_{X>v}] + E[(v - X)I_{X \leq v}].$$

The first derivative of $\Delta_X(v)$ with respect to v can be obtained as

$$\acute{\Delta}_X(v) = -[1 - F_X(v)] + F_X(v).$$

Therefore, the distribution function can be re-expressed in three MAD functions as follows. In terms of the derivative of MAD function ($\acute{\Delta}_X(v)$) as

$$F_X(v) = 0.5 + 0.5\acute{\Delta}_X(v). \qquad (5.1)$$

In terms of the derivative of right MAD function ($\acute{\Delta}_X^+(v)$) as

$$F_X(v) = 1 + \acute{\Delta}_X^+(v). \tag{5.2}$$

Finally, in terms of the derivative of left MAD function as

$$F_X(v) = \acute{\Delta}_X^-(v). \tag{5.3}$$

The right and left MAD can characterize the distribution function because they are primeval function of F_x; see, Munoz-Perez and Sanchez-Gomez (1990). Figure 3 displays the MAD functions ($\Delta_X(v), \acute{\Delta}_X^+(v)$ and $\acute{\Delta}_X^-(v)$ plot of standardized data from beta, exponential, Laplace and normal distributions. It may conclude that:

- $\acute{\Delta}_X^-(v)$ is monotone increasing function with minimum 0 and intersection with $\acute{\Delta}_X^+(v)$ at mean,
- $\acute{\Delta}_X^+(v)$ is monotone decreasing function with minimum 0 and intersection with $\acute{\Delta}_X^-(v)$ at mean,
- $\acute{\Delta}_X^-(v)$ and $\acute{\Delta}_X^+(v)$ have joint points with $\Delta_X(v)$ at extreme ends.

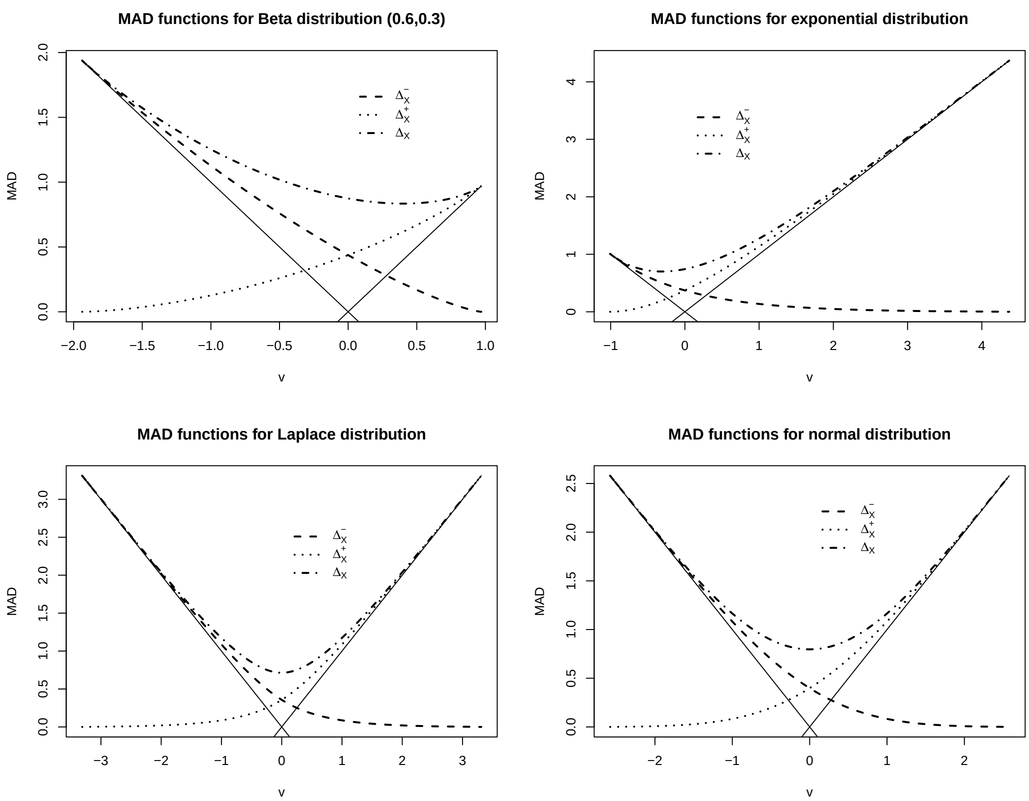


Figure 3: MAD functions ($\Delta_X(v), \Delta_X^+(v)$ and $\Delta_X^-(v)$) for standardized data from beta (0.6, 0.3), exponential, Laplace and normal distributions based on quantile functions with $q = (i - 0.5)/n$ and $n = 100$.

The most common non-parametric estimator for the underlying distribution function F is specified by the empirical cumulative distribution function (ecdf). The ecdf is defined by

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \le x}.$$

$\hat{F}_n(x)$ has good statistical properties such as: (1) it is a nondecreasing function with jumps of size $1/n$ at each order statistic, (2) it is bounded between zero and one, (3) it is first order efficient based on minimax criteria, for every x, $h = n\hat{F}_n(x)$ has a binomial distribution $(n, p = F_X(x))$, (4) and for large n, $\sqrt{n}(\hat{F}_n(x) - F(x)) \sim N(0, F_X(x)(1 - F_X(x))$; see Dvoretzky *et al.* (1956), Lehmann and Casella (1998), and Csaki (1984). Furthermore, the empirical distribution function is the nonparametric maximum likelihood estimator of F and has an important role in nonparametric bootstrap and simulation; see, Haddou and Perron (2006) and Efron and Tibshirani (1993).

6. ESTIMATION OF DISTRIBUTION FUNCTION USING MAD

When it is difficult to analytically obtain the derivative of the function $\Delta_X^-(v)$

$$F_X(v) = \lim_{h \to 0} \frac{\Delta_X^-(v+h) - \Delta_X^-(v)}{h}. \tag{6.1}$$

The numerical derivative can be used to obtain a good approximation to the true function $F_X(v)$. In the following we assume that this limit exists, i.e., $\Delta_X^-(v)$ is differentiable at $x = v$. By using the numerical differentiation, it could consider nonparametric estimators of the population distribution function $F_X(v)$ using a random sample $X_1, ..., X_n$ of size n. Consider the pairs of data

$$(x_i = x_{(i)}, y_i, i = 1, ..., n),$$

where $x_{(i)}$ is the observed order data and y_i is the estimated left MAD function that can be obtained from data as

$$y_i = g(x_{(i)}) = \hat{\Delta}_X^-(v) = \hat{E}[(v - x)I_{x \le v}] = \frac{1}{n}\sum_{j=1}^{n}(v_i - x_j)I_{x_j \le v_i} \;\text{ for }\; v_i = x_{(i)},\;\; i = 1, ..., n.$$

The nonparametric estimates of $F_X(v)$ can be derived numerically using several approaches as follows.

6.1. Forward difference approach

By using Taylor series,

$$g(x+h) = g(x) + \acute{g}(x)h + \frac{\acute{\acute{g}}(x)}{2}h^2 + \cdots + \frac{g^{k-1}(x)}{(k-1)!}h^{k-1}.$$

See, Burden and Faires (2011) and Levy (2012).

The first derivative in terms of first two terms,

$$\acute{g}(x) \approx \frac{g(x+h)-g(x)}{h} - \frac{h^2}{2}\acute{g}(\zeta), \quad \zeta \in (x, x+h).$$

This is the first order approximation $O(h)$. Therefore, a forward difference approach is

$$\acute{g}(x) \approx \frac{g_{i+1}-g_i}{h} + O(h).$$

By considering $h = x_{i+1} - x_i$ and with one-sided 1 at the endpoints of the data set, an estimation of $F_X(v)$ can be approximated by two terms Taylor series expansion as

$$\hat{F}_O(v) = \begin{cases} \dfrac{y_{i+1}-y_i}{x_{i+1}-x_i}, & i = 1, ..., n-1, \\ 1, & i = n. \end{cases} \tag{6.2}$$

Theorem 6.1. *With one-sided 1 at the endpoint of the data set, the forward difference approach is*

$$\hat{F}_O(v) = \begin{cases} \dfrac{i}{n}, & i = 1, ..., n-1, \\ 1, & i = n. \end{cases} \tag{6.3}$$

Proof: By considering $y_i = \frac{1}{n}\sum_{j=1}^{n}(v_i - x_j)I_{x_j \le v_i}$ for each $v_i = x_i$, $i = 1, ..., n$. For $v_1 = x_1$ then $y_1 = 0$, for $v_2 = x_2$ then $y_2 = \frac{1}{n}(x_2 - x_1)$, for $v_3 = x_3$ then $y_3 = \frac{1}{n}[(x_3 - x_1) + (x_3 - x_2)] = \frac{1}{n}[2x_3 + ny_2 - x_2 - x_2] = \frac{1}{n}[2(x_3 - x_2) + ny_2]$, ... for $v_i = x_i$ then $y_i = \frac{1}{n}[(x_i - x_1) + \cdots + (x_i - x_{i-1})] = \frac{1}{n}[(i-1)(x_i - x_{i-1}) + ny_{i-1}]$, therefore,

$$y_i - y_{i-1} = \frac{(i-1)}{n}(x_i - x_{i-1}),\ i = 2, ..., n$$

and

$$y_{i+1} - y_i = \frac{i}{n}(x_{i+1} - x_i),\ i = 1, ..., n-1.$$

This shows that the forward difference approach for the left MAD function is just the empirical distribution function i/n, $i = 1, ..., n$ and has equal jumping value $1/n$. □

6.2. Backward difference approach

Similarly, a backward differencing estimation of $F_X(v)$ can be approximated by two terms Taylor series expansion as

$$\hat{F}_B(v) = \begin{cases} \dfrac{y_i - y_{i-1}}{x_i - x_{i-1}}, & i = 2, ..., n, \\ 0, & i = 1. \end{cases} \tag{6.4}$$

For more details, see, Burden and Faires (2011) and Levy (2012).

Theorem 6.2. *With one-sided 0 at the endpoint of the data set, the backward difference approach is*

$$\hat{F}_B(v) = \begin{cases} \frac{i-1}{n}, & i = 2, ..., n, \\ 0, & i = 1. \end{cases} \tag{6.5}$$

Proof: By noting that $y_i - y_{i-1} = \frac{(i-1)}{n}(x_i - x_{i-1}),\ i = 2, ..., n.$ □

This shows that the backward difference approach for the left MAD function is just the empirical distribution function $\frac{i-1}{n}$, $i = 1, ..., n$ and has equal jumping value $1/n$.

6.3. Centre difference approach

A more accurate scheme can be derived using Taylor series

$$g(x+h) = g(x) + \acute{g}(x)h + \frac{\acute{\acute{g}}(x)}{2}h^2 + \frac{g^{(3)}(\zeta_1)}{6}h^3$$

and

$$g(x-h) = g(x) - \acute{g}(x)h + \frac{\acute{\acute{g}}(x)}{2}h^2 - \frac{g^{(3)}(\zeta_2)}{6}h^3.$$

By subtracting, the second order approximation ($O(h^2)$) of the first derivative is

$$\acute{g}(x) = \frac{g(x+h) - g(x-h)}{2h} - \frac{h^2}{6}g^{(3)}(\zeta) = \frac{g(x+h) - g(x-h)}{2h} + O(h^2), \quad \zeta \in (x-h, x+h).$$

With two-sided $1/n$ and $(n-1)/n$ at the endpoints of the data set, an estimate of $F_X(v)$ can be approximated by Taylor series expansion as

$$\hat{F}_C(v) = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1}, & i = 1, \\ \frac{y_{i+1} - y_{i-1}}{x_{i+1} - x_{i-1}}, & i = 2, ..., n-1, \\ \frac{y_n - y_{n-1}}{x_n - x_{n-1}}, & i = n. \end{cases} \tag{6.6}$$

Theorem 6.3. *With two-sided $1/n$ and $(n-1)/n$ at the endpoints of the data set, the centre difference approach is*

$$\hat{F}_C(v) = \begin{cases} \frac{1}{n}, & i = 1, \\ \frac{i}{n} - \frac{x_i - x_{i-1}}{n(x_{i+1-x_{i-1}})}, & i = 2, ..., n-1, \\ \frac{n-1}{n}, & i = n. \end{cases} \tag{6.7}$$

Proof: By noting that $y_i - y_{i-1} = \frac{(i-1)}{n}(x_i - x_{i-1}),\ i = 2, ..., n.$ □

This shows that the centre difference approach does not have an equal jumping function, but it is jumping by unequal quantity that depends on the ratio $(x_i - x_{i-1})/n(x_{i+1} - x_{i-1})$. Note that $(x_i - x_{i-1})/n(x_{i+1} - x_{i-1})$ is less than 1 and tends to 0 for $n \to \infty$. Also, noting that for $n = 2, ..., n-1$, we have

$$\left|\hat{F}_C(v) - \hat{F}_n(v)\right| = \left|\frac{i}{n} - \frac{(x_i - x_{i-1})}{n(x_{i+1} - x_{i-1})} - \frac{i}{n}\right| = \frac{(x_i - x_{i-1})}{n(x_{i+1} - x_{i-1})} < \frac{1}{n}.$$

$\hat{F}_C(v)$ is strongly uniformly consistent as $n \to \infty$; see, Serfling (1980).

6.4. FC-Hermite approach

With one-sided difference at the endpoints of the data set, an accurate estimate of $F_X(v)$ can be proposed by what is known as FC-Hermite approach or Hermite spline interpolation from Fritsch and Carlson (1980) and "splinefun" given in stats-package R-software (2021) as

(6.8) $$\hat{F}_{FCH}(v) = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1}, & i = 1, \\ 0.5\left(\dfrac{y_{i+1} - y_{i-1}}{x_{i+1} - x_{i-1}} + \dfrac{y_i - y_{i-1}}{x_i - x_{i-1}}\right), & i = 2, ..., n-1, \\ \dfrac{y_n - y_{n-1}}{x_n - x_{n-1}}, & i = n. \end{cases}$$

It can be noted that this approach combines the forward and backward approaches by using three data points at $i+1, i$ and $i-1$. The FC-Hermite approach can be rewritten in a very simple form as

$$\hat{F}_{FCH}(v) = \begin{cases} \dfrac{1}{n}, & i = 1, \\ \dfrac{2i-1}{2n}, & i = 2, ..., n-1, \\ \dfrac{n-1}{n}, & i = n. \end{cases}$$

This approach has equal jumping value $1/n$ except for first and last values and is related to Hazen (1914) plotting position. Also, note that for $n = 2, ..., n-1$, we have

$$\left|\hat{F}_{FCH}(v) - \hat{F}_n(v)\right| = \left|\frac{i}{n} - \frac{0.5}{n} - \frac{i}{n}\right| = \frac{1}{2n} < \frac{1}{n}.$$

$\hat{F}_{FCH}(v)$ is strongly uniformly consistent as $n \to \infty$; see, Serfling (1980).

6.5. Forward-backward-center approach

It might be very useful to use the combination of different approaches to increase accuracy of distribution function estimation. Another proposed estimate for $F_X(v)$ can be obtained by combining forward, backward and centre approaches as

$$\hat{F}_{OBC}(v) = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1}, & i = 1, \\ \dfrac{1}{3}\left(\dfrac{y_i - y_{i-1}}{x_i - x_{i-1}} + \dfrac{y_{i+1} - y_{i-1}}{x_{i+1} - x_{i-1}} + \dfrac{y_{i+1} - y_i}{x_{i+1} - x_i}\right), & i = 2, ..., n-1, \\ \dfrac{y_n - y_{n-1}}{x_n - x_{n-1}}, & i = n. \end{cases} \tag{6.9}$$

This can be rewritten as

$$\hat{F}_{OBC}(v) = \begin{cases} \dfrac{1}{n}, & i = 1, \\ \dfrac{3i-1}{3n} - \dfrac{(x_i - x_{i-1})}{3n(x_{i+1} - x_{i-1})}, & i = 2, ..., n-1, \\ \dfrac{n-1}{n}, & i = n. \end{cases}$$

This approach uses three data points at $i-1, i$, and i+1 and has an advantage of having non equal jumping values. Also, note that for $n = 2, ..., n-1$, we have

$$\left|\hat{F}_{OBC}(v) - \hat{F}_n(v)\right| = \left|\frac{i}{n} - \frac{(x_i - x_{i-1})}{3n(x_{i+1} - x_{i-1})} - \frac{i}{n}\right| = \frac{1}{3n}\left|1 - \frac{(x_i - x_{i-1})}{(x_{i+1} - x_{i-1})}\right| < \frac{1}{n}.$$

$\hat{F}_{OBC}(v)$ is strongly uniformly consistent as $n \to \infty$; see, Serfling (1980).

6.6. Richardson extrapolation approach

When applying low order formulas, Richardson's extrapolation is employed to achieve high accuracy results. As pointed out by Burden and Faires (2011), Richardson extrapolation is significantly more effective with even power than when all powers of h are used because the averaging process creates results with errors $O(h^2)$, $O(h^4)$ and $O(h^6)$, ..., with essentially no increase in computation, over the results with errors, $O(h), O(h^2), O(h^3), ...$; see, Burden and Faires (2011).

Theorem 6.4. *An improved approximation for distribution function estimation based on Richardson extrapolation is*

$$\hat{F}_R(v) = \acute{g}(x) = \frac{2^{2p_1}G(h) - G(2h)}{2^{2p_1} - 1} + O(h^{2p_2}). \tag{6.10}$$

Proof: Assume that $G(h)$ be a difference formula with step-size h, approximating $\acute{g}(x)$ as

$$G(h) = \acute{g}(x) + a_1 h^{2p_1} + a_2 h^{2p_2} + a_3 h^{2p_3} + \cdots$$

Note $p_1 < p_2 < p_3, ...$ and a_i are constants. Therefore,

$$\acute{g}(x) = G(h) - a_1 h^{2p_1} - a_2 h^{2p_2} - a_3 h^{2p_3} - \cdots = G(h) + O(h^{2p_1}).$$

Hence, if we consider $G(h)$ as an approximation to $\acute{g}(x)$, the error is $O(h^{p_1})$. Sure, if h tends to zero, $G(h) \to \acute{g}(x)$. If $G(h)$ is computed for step-size $2h$, then

$$G(2h) = \acute{g}(x) + a_1 2^{2p_1} h^{2p_1} + a_2 2^{2p_2} h^{2p_2} + a_3 2^{2p_3} h^{2p_3} + \cdots$$

Multiplying $G(h)$ by 2^{2p_1} and subtract from $G(2h)$ we obtain

$$\acute{g}(x) = \frac{2^{2p_1} G(h) - G(2h)}{2^{2p_1} - 1} + \left[a_2 2^{2p_1} - a_2 2^{2p_2}\right] h^{2p_1} + \cdots = \frac{2^{2p_1} G(h) - G(2h)}{2^{2p_1} - 1} + O(h^{2p_2}).$$

Hence,

$$\acute{g}(x) \approx \frac{2^{2p_1} G(h) - G(2h)}{2^{2p_1 - 1}}. \tag{6.11}$$

□

This new approximation is of order $O(h^{2p_2})$.

The estimate $\hat{F}_R(v)$ can be simply obtained from R software (package "pracma"; see, Borchers, 2021) using function "numdiff (f=function, x)". Because $\hat{F}_R(v)$ is bounded by 0 and 1 and estimated numerically, it may not become in some cases nondecreasing. In these cases, $\hat{F}_R(v)$ needs to be adjusted to become monotonic nondecreasing. One of the good methods that can be used is bounded isotonic regression introduced by Barlow *et al.* (1972) and Balabdaoui *et al.* (2009). The PAVA algorithm has been used to find this solution and has been implemented in the R package OrdMonReg (Balabdaoui *et al.*, 2009) under the function BoundedIsoMean; see, Balabdaoui *et al.* (1980). This function can produce an estimate that is bounded by $0, 1$ and monotone nondecreasing. The adjusted $\hat{F}_{Ra}(v)$ is estimated via function

$$\hat{F}_{Ra} = \text{BoundedIsoMean}(y = \hat{F}_R(v), w = 1/n, a = 0, b = 1)$$

based on OrdMonReg package in R software.

Theorem 6.5. *The Richardson extrapolation estimator $\hat{F}_{Ra}(v)$ is strongly uniformly consistent*

$$\sup_v \left| \hat{F}_{Ra}(v) - F_X(x) \right| \to 0 \; w.p.1.$$

Proof: Let $F_n(v)$ be the empirical distribution function. It may write

$$\left| \hat{F}_{Ra}(v) - F_X(x) \right| \leq \left| \hat{F}_{Ra}(v) - F_n(v) \right| + |F_n(v) - F_X(v)|.$$

It is well known from Serfling (1980) that

$$\sup_v \left| \hat{F}_n(v) - F_X(x) \right| \to 0 \; w.p.1,$$

and

$$\sup_v \left| \hat{F}_{Ra}(v) - F_n(v) \right| \leq \frac{1}{n}$$

tends to 0 when $n \to \infty$. Therefore,

$$\sup_v \left| \hat{F}_{Ra}(v) - F_X(x) \right| \to 0 \; w.p.1.$$

□

Theorem 6.6. *The Richardson extrapolation estimator $\hat{F}_{Ra}(v)$ has an asymptotic normal distribution as*

$$\sqrt{n}\Big(\hat{F}_{Ra}(v) - F_X(v)\Big) \xrightarrow{d} N(0, F_X(v)(1-F_X(v))).$$

Proof: Let $F_n(v)$ be the empirical distribution function and $F_X(v)$ be the true function. It is well known that; see, Serfling (1980),

$$\sqrt{n}\Big(\hat{F}_n(v) - F_X(v)\Big) \xrightarrow{d} N(0, F_X(v)(1-F_X(v))).$$

Then

$$\sqrt{n}\Big(\hat{F}_{Ra}(v) - F_n(v)\Big) \le \frac{1}{\sqrt{n}}.$$

As $n \to \infty$

$$\sqrt{n}\Big(\hat{F}_{Ra}(v) - F_X(v)\Big) \xrightarrow{d} N(0, F_x(v)(1-F_X(v))). \qquad \square$$

7. SIMULATION

Simulation study is conducted to evaluate the performance of the proposed approaches. Five mixture normal distributions that used in Xue and Wang (2010) are implemented to compare the proposed approaches results with their results. These distributions are Gaussian distribution (G), distribution no. 3 (strongly skewed distribution (SS)), distribution no. 5 (outlier (OU)), distribution no. 7 (separated bimodal distribution (SB)), distribution no. 14 (smooth comb (SC)). These distributions cover a wide range of shapes and they are given in Table 2, for more details; see, Marron and Wand (1992). From every distribution, 1000 simulated samples of sizes 20, 50 and 200 are generated, respectively. The estimators $\hat{F^*}_n$ (empirical), $\hat{F}_O$ (forward), $\hat{F}_B$ (backward), $\hat{F}_C$ (centre), $\hat{F}_{FCH}$ (FC-Hermite), $\hat{F}_{OBC}$ (forward-backward-centre) and $\hat{F}_{Ra}$ (adjusted Richardson extrapolation) are computed.

Table 2: Distribution functions used in the simulation study.

Name	Distribution
Standard normal distribution (G)	$N(0,1)$
Strongly skewed distribution #3 (SS)	$\sum_{l=0}^{7} \frac{1}{8} N\Big(3\Big(\big(\frac{2}{3}\big)^l - 1\Big), \big(\frac{2}{3}\big)^{2l}\Big)$
Outlier distribution #5 (OU)	$\frac{1}{10}N(0,1) + \frac{1}{10}N\Big(0, \big(\frac{1}{10}\big)^2\Big)$
Separated bimodal distribution #7 (SB)	$\frac{1}{2}N\Big(-\frac{3}{2}, \big(\frac{1}{2}\big)^2\Big) + \frac{1}{2}N\Big(\frac{3}{2}, \big(\frac{1}{2}\big)^2\Big)$
Smooth comb #14 (SC)	$\sum_{l=0}^{5} \Big(\frac{2^{5-l}}{63}\Big) N\Big(\frac{65-96/2^l}{21}, \Big(\frac{32/63}{2^l}\Big)^2\Big)$

It should be noted that these estimators are not smoothed, and they will be compared with empirical $\hat{F^*}_n$ and smoothed estimators (linear spline (PS1), cubic spline (PS3), constrained linear spline (CPS1), and constrained cubic spline (CPS3)) given in Xue and Wang (2010) in terms of averaged squared errors (ASE) that is defined as

$$ASE_{\hat{F}} = \frac{1}{n}\sum_{i=1}^{n}\Big[\hat{F}_X(v_i) - F_X(v_i)\Big]^2.$$

Table 3: Averaged squares errors (ASE) of all estimators ($\times 10^3$).

	n	$PS1^*$	$PS3^*$	$CPS1^*$	$CPS3^*$	$\hat{F^*}_n$	$\hat{F}_O$	$\hat{F}_B$	$\hat{F}_C$	$\hat{F}_{OBC}$	$\hat{F}_{FCH}$	$\hat{F}_{Ra}$
G	20	6.97	6.86	8.08	7.17	8.53	8.48	8.56	8.70	8.09	7.84	7.42
	50	2.87	2.76	3.13	2.97	3.29	3.31	3.30	3.35	3.24	3.19	2.99
	200	0.82	0.75	0.83	0.78	0.84	0.84	0.83	0.83	0.82	0.81	0.77
SS	20	8.07	7.36	8.70	7.89	9.02	8.58	8.66	8.89	8.21	7.90	7.24
	50	3.38	2.98	3.33	3.25	3.51	3.42	3.32	3.42	3.30	3.26	3.17
	200	0.84	0.79	0.83	0.82	0.84	0.84	0.84	0.83	0.83	0.83	0.83
OU	20	8.25	8.10	8.47	9.20	8.71	8.52	8.47	8.80	8.16	7.92	7.92
	50	3.38	3.33	3.41	3.38	3.46	3.32	3.32	3.38	3.27	3.22	3.22
	200	0.78	0.77	0.79	0.82	0.81	0.78	0.78	0.79	0.78	0.77	0.77
SB	20	8.07	7.86	8.39	8.04	8.56	8.93	7.82	8.60	8.00	7.73	7.01
	50	3.19	3.12	3.20	3.14	3.32	3.37	3.15	3.31	3.20	3.16	2.92
	200	0.79	0.77	0.79	0.78	0.81	0.84	0.83	0.84	0.83	0.82	0.79
SC	20	8.20	7.80	8.28	7.98	8.56	8.62	8.48	8.82	8.14	7.85	7.18
	50	3.23	3.19	3.27	3.25	3.36	3.34	3.39	3.43	3.31	3.26	3.25
	200	0.81	0.79	0.81	0.83	0.82	0.82	0.84	0.84	0.83	0.80	0.82

$*$ indicate that the results in these columns are from Xue and Wang (2010), G: standard normal, SS: strongly skewed, OU: outliers, SB: separated bimodal, SC, smooth com (PS1), cubic spline (PS3), constrained linear spline (CPS1), and constrained cubic spline (CPS3).

The results of the simulation study are given in Table 3 that illustrates that:

- The ASE decreases for all estimators with increasing n,
- The estimators $\hat{F}_O$, $\hat{F}_B$, and $\hat{F}_C$ have almost the same ASE as $\hat{F^*}_n$ and this is expected where all of them are some types of general class of empirical functions; see Cunnane (1978) and Hosking and Wallis (1995),
- The estimator $\hat{F}_{OBC}$ has improved ASE over classical empirical estimators, for example, if the distribution is normal and sample size is 20, there is improvement about 5% in ASE over $\hat{F^*}_n, \hat{F}_O, \hat{F}_B$,
- The estimator $\hat{F}_{FCH}$ is surprised as it is very simple and has a very good improvement in terms of ASE. In all cases, there is an improvement about 10% in ASE over $\hat{F^*}_n, \hat{F}_O, \hat{F}_B$, about 4% over $\hat{F}_C$, and less improvement about 2% than $\hat{F}_{Ra}$,
- The estimator $\hat{F}_{Ra}$ has a major improvement about 12% over $\hat{F^*}_n, \hat{F}_O, \hat{F}_B$, medium improvement about 6% over $\hat{F}_C$, and small improvement about 2% over $\hat{F}_{FCH}$. The $\hat{F}_{Ra}$ is very comparable to two spline smooth unconstrained estimators PS1 and PS3 in terms of ASE,
- With respect to two monotone nondecreasing constrained splines (CPS1 and CPS3), $\hat{F}_{Ra}$ has a very competitive ASE with all studied distributions and $\hat{F}_{FCH}$ has ASE almost as same as CPS1 and CPS3,
- For a large n such as 200, the performance of all estimators is comparable in terms of ASE.

8. APPLICATION

In ecotoxicology, lognormal and loglogistic distributions are applied to fit a data. A low percentile 5% is of great interest where the hazardous concentration 5% (HC5) is explained as the value of pollutant concentration protecting 95% of the species; see Posthuma *et al.* (2010). There is a data set 'endosulfant' in R software package "fitdistplus"; see, Delignette-Muller and Dutang (2021). This data includes acute toxicity values (ATV) for the organochlorine pesticide 'endosulfan' (geometric mean of LC50 and EC50 values in $ug.L^{(-1)}$, tested on Australian and non-Australian laboratory-species; see, Hose and Van den Brink (2004). Figure 4 displays the MAD plot and Cullen and Frey graph; see, Cullen and Frey (1999). MAD plot shows a very weak right wideness and a very weak wideness in the left side. The distribution is very strong right skewed ($k = 0.26$) and has a very long right tail $\hat{T}_R = 6.707$, while very short left tail $\hat{T}_L = 0.266$. The skewness based on tails is 6.441 (very strong). Cullen and Frey graph is an indicative graph where it shows the relationship between Pearson skewness squared and kurtosis. For given data, the skewness is 5.076 and kurtosis is 30.728 that may suggest a lognormal distribution as a good candidate to fit the data. Moreover, Muller and Dutang (2014, 2021) used lognormal, loglogistic, Pareto and Burr III distributions to fit a suitable distribution for ATV data.

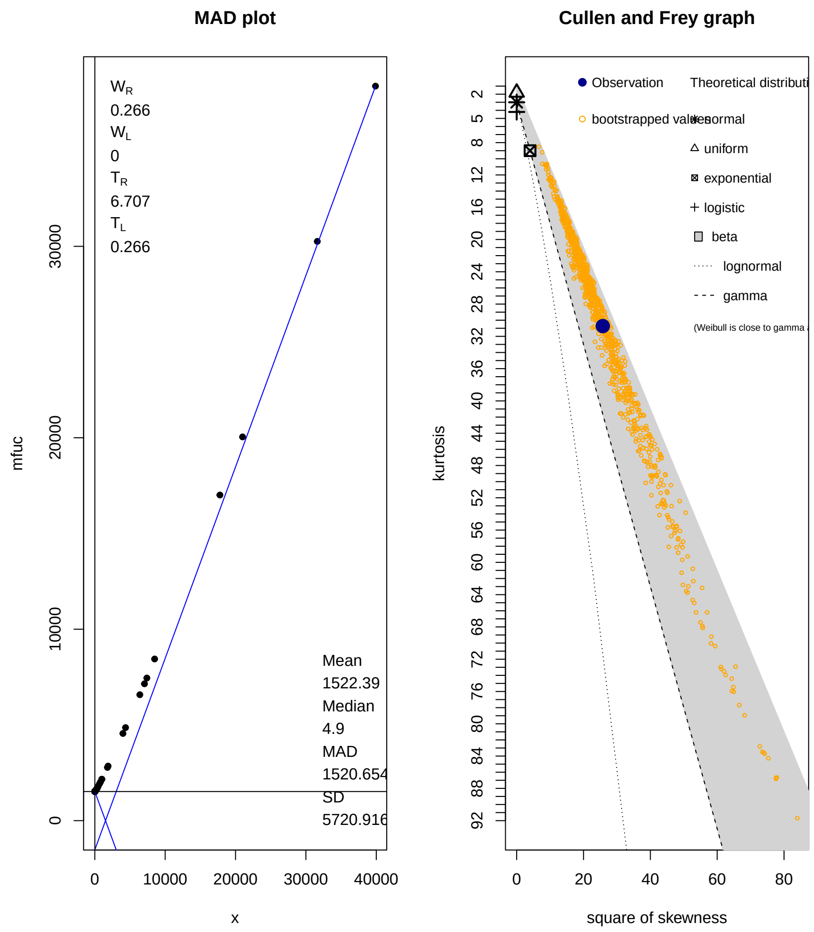


Figure 4: MAD plot and Cullen and Frey graph for acute toxicity values data (ATV).

The proposed adjusted Richardson approximation ($\hat{F}_{Ra}$) is used to estimate nonparametric distribution function of ATV data. Also, 99% pointwise confidence intervals based on normal approximation for ($\hat{F}_{Ra}$) are obtained. In Figure 5 the estimated distribution function and 99% confidence interval are plotted along with the estimated parametric distribution functions; for more details about this estimation; see, Delignette-Muller and Dutang (2021). For the lognormal distribution it has a good fit in the right tail while a bad fit in the left tail due to a high probability at left tail. The loglogistic does not fit from both tails. The two-parameter Pareto and three-parameter distributions have a good fitting in left tail while a worse fitting in the right tail. As concluded by Delignette-Muller and Dutang (2021), none of the four distributions correctly described the right tail observed in ATV data, but the left tail seems to be better described by Burr III distribution; see also, Hose and Van den Brink (2004). They estimated the HC5 value using Burr III distribution as 0.294 while the HC5 from the data is 0.20. The Richardson approximation for HC5 is computed using interpolation as 0.242 while the estimation of HC5 using empirical distribution function is 0.161 (forward approach).

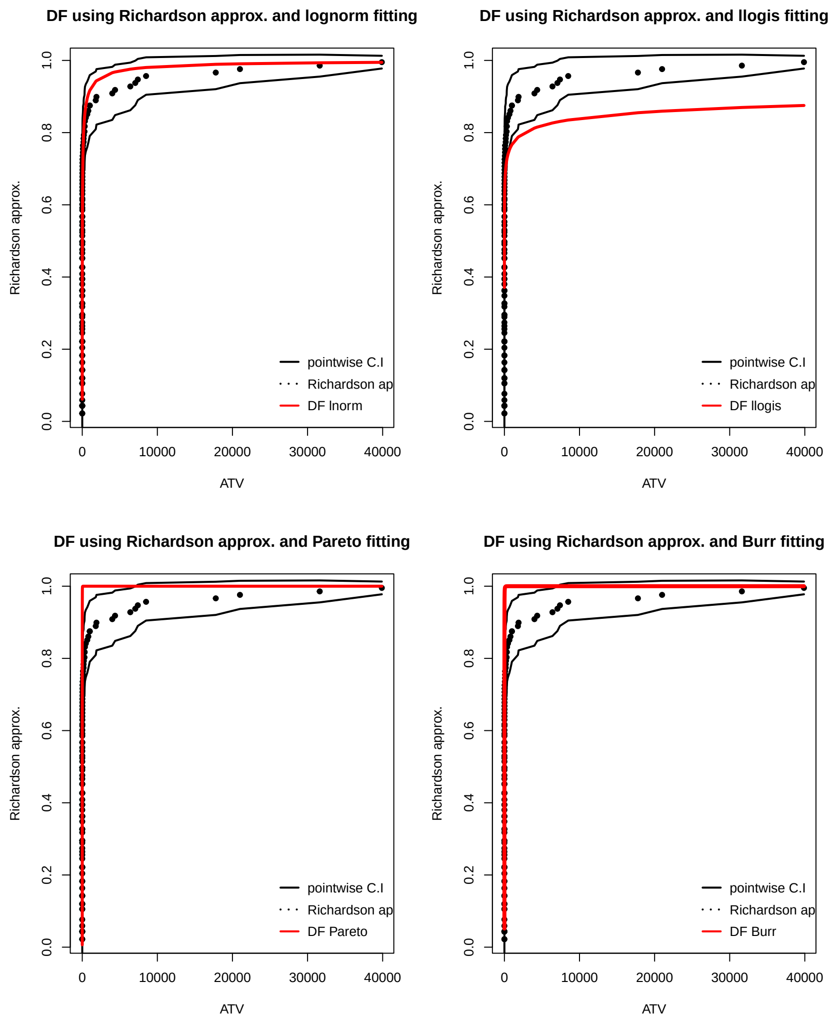


Figure 5: Plots of the estimated distribution function using Richardson extrapolation approximation with its 99% pointwise confidence intervals, and the estimated distribution functions from lognormal, loglogistic, Pareto and Burr III models for ATV data.

9. CONCLUSION

The usefulness of the mean absolute deviation function is introduced in two directions. Firstly, it was used to explore the pattern and the structure in the data graphically through the wideness and tailedness concepts. The wideness reflected information about how much the mean absolute function is away from the straight lines $(v_i - \mu)$ and $(\mu - v_i)$. This created right, left and overall wideness measures. These measures reflected skewness in the data and used to reflect the flatness and peakedness in symmetric distributions. The tailedness reflected information about how long the right and left tails in the distribution of data via the maximum of right and left mean absolute deviation functions. Secondly, a new method based on Richardson extrapolation approach was proposed to estimate the population distribution function. In general, six approaches developed that included forward, backward, central, mix and FC-Hermite interpolation and Richardson extrapolation approaches. Simulation study was implemented using different distributions that represented different shapes such as bell-shaped, separated-bimodal, strong-skewed, smooth-comb and outliers. Three estimators showed improvement in terms of averaged squared errors over the classical empirical distribution function. The Richardson extrapolation approach had major improvement in terms of average squared errors over classical empirical estimators and had comparable results with smooth approaches such as cubic spline and constrained linear spline. Furthermore, the Richardson approach applied for real data application and used to estimate the hazardous concentration five percent. Future studies may seek to examine smoothing approaches based on Richardson extrapolation approach and investigate the superiority of non-smooth approaches in estimating the distribution function compared with smooth approaches as suggested by one of the reviewers.

ACKNOWLEDGMENTS

The author acknowledges the valuable remarks and suggestions made by the referees and editor which improved the presentation of the paper.

REFERENCES

Babu, G. J., Canty, A. J., and Chaubey, Y. P. (2002). Application of Bernstein polynomials for smooth estimation of a distribution and density function. *Journal of Statistical Planning and Inference*, 105:377–392.

Balabdaoui, F., Rufibach, K., and Santambrogio, F. (2009). Least squares estimation of two ordered monotone regression curves. *Preprint*, Hal-00417281, `https://arxiv.org/abs/0904.2052`.

Balabdaoui, F., Rufibach, K., and Santambrogio, F. (1980). Ordmonreg: Compute least squares estimates of one bounded or two ordered isotonic regression curves. *R package version 1.0.3*, `https://CRAN.R-project.org/package=OrdMonReg`.

Baszczynska, A. (2016). Kernel estimation of cumulative distribution function of a random variable with bounded support. *Statistics in Transition*, 17(3):541–556.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. John Wiley and Sons.

Borchers, H. (2021). Pracma: practical numerical math functions. *R package version 2.3.6*, `https://CRAN.R-project.org/package=pracma`.

Burden, R. and Faires, D. (2011). *Numerical Analysis*. 9th Ed., Cengage Learning.

Charles, J. K., Sun, S., and Martin, C. F. (2010). Cumulative distribution estimation via control theoretic smoothing splines. *Three Decades of Progress in Control Sciences*, 4(1):95–104.

Cheng, M. Y. and Peng, L. (2002). Regression modelling for nonparametric estimation of distribution and quantile functions. *Statistica Sinica*, 12:1043–1060.

Csaki, E. (1984). Empirical distribution function. In Krishnaiah, P. R. and Sen, P. K., eds., *Handbook of Statistics*, 4, 405–430.

Cullen, A. and Frey, H. (1999). *Probabilistic Techniques in Exposure Assessment*. 1st edition, Plenum Publishing Company.

Cumming, G. and Finch, S. (2005). Inference by eye: confidence intervals and how to read pictures of data. *The American Psychologist*, 60(2):170–180.

Cunnane, C. (1978). Unbiased plotting positions? A review. *Journal of Hydrology*, 37:205–222.

Delignette-Muller, M. and Dutang, C. (2021). Fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, 64(4):1–34.

Dodge, Y. (2002). *Statistical Data Analysis Based on the L1-norm and Related Methods*. Birkhauser Basel.

Dvoretzky, A., Kiefer, J., and Wolfowitz, A. J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Annals of Mathematical Statistics*, 27(3):642–669.

Efromovich, S. (2001). Second order efficient estimating a smooth distribution function and its applications. *Methodology and Computing in Applied Probability*, 3:179–198.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, London.

Elamir, E. A. (2012). Mean absolute deviation about median as a tool of explanatory data analysis. *Proceedings of the World Congress on Engineering*, I WCE 2012, July 4–6, 2012, London, U.K., Vol. I, 324–329, `https://www.iaeng.org/publication/WCE2012/WCE2012-pp324-329.pdf`.

Erdogan, M. S, Disibuyuk, C., and Ege Oruc, O. (2019). An alternative distribution function estimation method using rational Bernstein polynomials. *Journal of Computational and Applied Mathematics*, 353:232–242.

Fritsch, F. N. and Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17:238–246.

Gandomi, A. and Haider, M. (2015). Beyond the hype: big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2):137–144.

Geary, R. C. (1935). The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika*, 27:310–332.

Gorard, S. (2005). Revisiting a 90-year-old debate: the advantages of the mean deviation. *British Journal of Educational Studies*, 53:417–430.

Groeneveld, R. A. and Meeden, G. (1984). Measuring skewness and kurtosis. *The Statistician*, 33:391–399.

Habib, E. A. (2012). On uses of mean absolute deviation: decomposition, skewness and correlation coefficients. *Metron*, LXX(2–3):1–20.

Haddou, M. and Perron, F. (2006). Nonparametric estimation of a Cdf with mixtures of Cdf concentrated on small intervals. `http://www.crm.umontreal.ca/pub/Rapports/3200-3299/3210.pdf`.

Hazen, A. (1914). Storage to be provided in impounding reservoirs for municipal water supply. *Transactions of the American Society of Civil Engineers*, 108(77):1547–1550.

Healy, K. (2019). *Data Visualization*. Princeton University Press.

Hosking, J. R. and Wallis, J. R. (1995). A comparison of unbiased and plotting-position estimators of L moments. *Water Resources Research*, 31:2019–2025.

Hose, G. and Van den Brink, P. (2004). Confirming the species-sensitivity distribution concept for Endosulfan using laboratory, mesocosm, and field Data. *Archives of Environmental Contamination and Toxicology*, 47(4):511–520.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. 1st edition, Springer, New York.

Larson-Hall, J. (2017). Moving beyond the bar plot and the line graph to create informative and attractive graphics. *The Modern Language Journal*, 101(1):244–270.

Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*. 2nd ed., Springer, New York.

Levy, D. (2012). *Introduction to Numerical Analysis*. `https://www.math.umd.edu/~diom/courses/AMSC466/Levy-notes.pdf`.

Marron, J. S. and Wand, M. P. (1992). Exact mean integrated squared error. *The Annals of Statistics*, 20:712–736.

Matt, W. and Joshua, F. (2019). *Advanced R Statistical Programming and Data Models: Analysis, Machine Learning, and Visualization*. Apress, Berkeley, CA.

Mombeni, H., Mansouri, B., and Akhoond, M. (2021). Asymmetric kernels for boundary modification in distribution function estimation. *REVSTAT – Statistical Journal*, 19(4):463–484.

Munoz-Perez, J. and Sanchez-Gomez (1990). A characterization of the distribution function: the dispersion function. *Statistics and Probability Letters*, 10:235–239.

Pham-Gia, T. and Hung, T. L. (2001). The mean and median absolute deviations. *Mathematical and Computer Modelling*, 34:921–936.

Posthuma, L., Suter, G., and Traas, T. (2010). Species sensitivity distributions in ecotoxicology. *Environmental and Ecological Risk Assessment Series*, Lewis Publishers, A CRC Press Company.

R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria, `https://www.R-project.org`.

Runkler, A. T. (2020). *Data Analytics: Models and Algorithms for Intelligent Data Analysis*. 3rd Ed., Springer, New York.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

Tukey, J. W. (1987). *Exploratory Data Analysis*. Addison Wesley, Reading.

Xue, L. and Wang, J. (2010). Distribution function estimation by constrained polynomial spline regression. *Journal of Nonparametric Statistics*, 22(4):443–457.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 101–115
https://doi.org/10.57805/revstat.v23i1.560

Survival Copula Entropy and Dependence in Bivariate Distributions

Authors: N. Unnikrishnan Nair
– Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India
unnikrishnannair4@gmail.com

S. M. Sunoj ✉
– Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India
smsunoj@cusat.ac.in

Received: December 2022 Revised: February 2023 Accepted: February 2023

Abstract:

- In the present work we propose survival copula entropy as an alternative to Shannon entropy, cumulative residual entropy and copula entropy measures in computing the uncertainty in bivariate populations. We examine the relationships between the various measures. The properties of survival copula entropy are discussed, especially its applications to ascertain the nature and extent of uncertainty among copulas.

Keywords:

- *bivariate Shannon entropy; copula entropy; cumulative residual entropy; dependence measures.*

AMS Subject Classification:

- 94A17, 62H05.

✉ Corresponding author

1. INTRODUCTION

Knowledge on the nature and extent of the joint behaviour of random quantities is a topic of considerable interest in all fields of scientific activity, In this context the joint distribution of random variables is an indispensable tool in analysing various aspects of interrelationship among the constituent variables, Among various measures developed for understanding the amount of uncertainty prevailing in the outcomes generated by the distribution, entropy has established itself as an efficient mechanism in a variety of fields. The basic measure of uncertainty employed in the bivariate case is the Shannon entropy defined as

$$h(X,Y) = h_{X,Y} = -\int_S\int f(x,y)\log f(x,y)dxdy, \tag{1.1}$$

where $f(x,y)$ is the probability density function of the random vector (X,Y) with support S. Since the introduction of (1.1), several modifications were introduced by way of additional parameters to impart more flexibility, measures with structural changes, replacing joint density by conditional ones etc., to provide a wide range of new measures. The structural modification to obtain a new measure by replacing the density by the survival function is due to Rao *et al.* (2004) in the univariate case, motivated researchers to apply the same logic in the bivariate case resulting in the bivariate version of cumulative residual entropy

$$H(X,Y) = H_{X,Y} = -\int_S\int \bar{F}(x,y)\log \bar{F}(x,y)dxdy, \tag{1.2}$$

where $\bar{F}(x,y) = P(X > x, Y > y)$ is the survival function of (X,Y). The expression (1.2) is implied as a particular case in the definition of bivariate cumulative residual entropy in Rajesh *et al.* (2014). A critical aspect to be considered in using bivariate distribution for modelling is the dependence relation existing between X and Y. In this respect copulas are found to be more general and flexible as they provide means of obtaining the joint distribution through the marginals connected by known dependence relationships. There are three approaches to study the nature and extent of dependence in copulas. The first is through global measures that specify the association like the Pearson's correlation coefficient, Spearman's rho, Kendall's tau, Blomquist's beta, etc. A second alternative is to use dependence concepts like total positivity, quadratic dependence and stochastic increase. Finally, we have time dependent measures of association which are used when analysing data on duration variables where the time elapsed since the commencement of observation is vital. See Nair and Sankaran (2010) for a review of materials in this connection. Any one of these methods, depending on the appropriateness of the techniques chosen in the problem at hand, will enable us to know whether dependence is positive or negative and also to compare the degree of association. In view of the flexibility of the copula over distribution functions, Ma and Sun (2008) proposed the copula entropy

$$i(X,Y) = i_{X,Y} = -\int_0^1\int_0^1 c(u,v)\log c(u,v)dudv, \tag{1.3}$$

as a measure of uncertainty in the copula density $c(u,v)$ associated with (X,Y). It provides a tool to connect copulas and information theory. Since its inception the measure (1.3) was used to ascertain structural learning, dependence aspects, variable selection, casual discovery, etc in various disciplines such as hydrology (Chen *et al.*, 2013), biology (Charzyńska and Gambin, 2015), neuroscience (Kayser *et al.*, 2015) and medicine (Ma, 2019a,b; Mesiar and

Sheikhi, 2021) etc. The generality and range of application the copula entropy motivate the investigation of more refined measures of uncertainty for copulas. Further as an index of information, the negative values of (1.3) are difficult to interpret and it is more preferable to have a measure that assumes positive values only, and that too in a finite interval. The form of copula density in many standard cases is analytically and computationally more complicated to work with, when compared to the usual copula or the survival copula. To study various aspects of dependence, most concepts in that area are expressed in terms of the copula than its density. Moreover inference procedures available in literature for copulas can be made use of in estimating and testing copula-based entropies. These facts suggest proposing an uncertainty measure based on copulas and investigating its properties. The objective of the present work is thus to make a preliminary study of a new measure of uncertainty in terms of the survival copulas, in the same manner as the development of (1.2) from (1.1).

A summary of the present work is as follows. In Section 2, we define the survival copula entropy and obtain some relationships between cumulative residual entropy, copula entropy and survival copula entropy. Following this, in Section 3 the properties of the new entropy especially its role as a measure of dependence is discussed. In Section 4, application of survival copula entropy to some real situations is demonstrated. The paper ends with a brief conclusion in Section 5.

2. SURVIVAL COPULA ENTROPY

As mentioned in the introduction, let (X, Y) be a random vector with distribution function $F(x, y)$ and survival function $\bar{F}(x, y)$. Recall that a copula is a function $C : I^2 \to I$, where I is the unit interval [0, 1], such that

$$C(0, v) = C(u, 0) = 0;\;\; C(u, 1) = C(1, u) = u,$$

for all u, v in I and C is 2-increasing so that the C-volume of the rectangle $[a, b] \times [k, d]$, $V_C([a, b] \times [k, d]) \geq 0$ for all rectangles in I^2. The function C induces a probability measure on I^2 via $V_C([a, b] \times [k, d]) = C(u, v)$. When C is absolutely continuous we have the copula density $c(u, v) = \frac{\partial^2 C}{\partial u \partial v}$. A survival copula $\bar{C} : I^2 \to I$ satisfies

$$\bar{C}(u, 1) = 0 = \bar{C}(1, u) \;\text{ and }\; \bar{C}(u, 0) = u = \bar{C}(0, u),$$

for all u in I and volume $V_{\bar{C}}([a, b] \times [k, d]) \geq 0$. Further,

$$\bar{C}(u, v) = u + v - 1 + C(1 - u, 1 - v).$$

With these basic notions we define the measures of uncertainty with reference to C and $\bar{C}$.

Definition 2.1. The survival copula entropy (SCE) associated with the survival copula $\bar{C}$ of (X, Y) is defined as

$$I_{\bar{C}}(X, Y) = -\int_0^1 \int_0^1 \bar{C}(u, v) \log \bar{C}(u, v) du dv. \tag{2.1}$$

Example 2.1. Consider the Gumbel-Barnett family

$$\bar{C}(u,v) = u\ v\ \exp[-\theta \log u \log v],\ 0 \le \theta \le 1.$$

$$\begin{aligned} I_{\bar{C}}(X,Y) &= -\int_0^1\int_0^1 uve^{-\theta \log u \log v}[\log u + \log v - \theta \log u \log v]dudv \\ &= -e^{\frac{4}{\theta}} EI\Big(-\frac{4}{\theta}\Big), \end{aligned}$$

where $EI(z) = -\int_{-z}^{\infty} \frac{e^{-t}}{t}dt$. The value of $I_{\bar{C}}$ for selected values of θ are given in Table 1.

Table 1: Survival copula entropy of the Gumbel-Barnett copula.

θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$I_{\bar{C}}$	0.24404	0.23859	0.23357	0.22570	0.22455	0.22048	0.21664	0.21302	0.20960

The quantity in (2.1) is obviously a measure of uncertainty since $\bar{C}(u,v)$ is a bivariate survival function with uniform marginals and (2.1) is thus a cumulative entropy, by definition (1.2). As a measure of uncertainty in bivariate distributions, it is of interest to examine its structure in relation to the existing similar measures like cumulative entropy (1.2) and the copula entropy (1.3). For this purpose we assume that the marginal survival functions $\bar{F}_X$ and $\bar{F}_Y$ of X and Y are continuous, strictly decreasing over the half-line $[0,\infty)$ with $\bar{F}_X(0) = 1 = \bar{F}_Y(0)$. Then the bivariate cumulative entropy of (X,Y) can be written as

$$\begin{aligned} H(X,Y) &= -\int_0^\infty\int_0^\infty \bar{F}(x,y)\log \bar{F}(x,y)dxdy \\ &= -\int_0^1\int_0^1 \bar{F}\big(\bar{F}_X^{-1}(u),\bar{F}_Y^{-1}(v)\big)\log \bar{F}\big(\bar{F}_X^{-1}(u),\bar{F}_Y^{-1}(v)\big)d\bar{F}_X^{-1}(u)d\bar{F}_Y^{-1}(v) \\ &= -\int_0^1\int_0^1 \bar{C}(u,v)\log \bar{C}(u,v)d\bar{F}_X^{-1}(u)d\bar{F}_Y^{-1}(v), \end{aligned} \tag{2.2}$$

by Sklar's theorem. It is evident that the copula version of (1.2) is not in general the same as SCE. However for bivariate uniform distributions like

$$\bar{F}(x,y) = (1-x)(1-y)[1+\theta xy],\ 0 \le x,y \le 1,$$

the quantities $H_{XY}(x,y)$ and $I_{\bar{C}}(X,Y)$ remain the same. Some similar calculations reveal that

$$I_{\bar{C}}(X,Y) = -\int_0^\infty\int_0^\infty \big[\bar{F}(x,y)\log \bar{F}(x,y)\big]d\bar{F}_X d\bar{F}_Y, \tag{2.3}$$

showing that the distribution function counterpart of SCE is not identical with (1.2). Thus survival copula entropy is a different measure than the other comparable measures of uncertainty of bivariate distributions as can be seen from this and the following discussions.

We now examine the relationships between the two copula-based entropies $i(X,Y)$ and $I(X,Y)$. Although there exists nice relationship $c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$ between c and C, a simple equation connecting their entropies appears to be elusive and the same is true even for specific copulas. However, if we consider the random vector (X_E, Y_E) associated with (X, Y) through

$$f_{X_E,Y_E}(x,y) = \frac{\bar{F}(x,y)}{E(XY)}, \tag{2.4}$$

where f_{X_E,Y_E} stands for the density function of (X_E, Y_E), some relationships useful in different contexts can be derived. Note that (2.4) is often called the equilibrium distribution of (X, Y). For a detailed discussion of the properties and applications of such distributions, we refer to Nair and Preeth (2009) and Navarro and Sarabia (2010). Using (2.4) and (1.1) we can write the Shannon entropy of (X_E, Y_E) as

$$\begin{aligned} h_{X_E,Y_E} &= -\int_0^\infty \int_0^\infty \frac{\bar{F}(x,y)}{E(XY)} \left[\log \bar{F}(x,y) - E(XY)\right] dxdy \\ &= \frac{H_{X,Y}}{E(XY)} + \log E(XY), \end{aligned} \tag{2.5}$$

showing $H_{X,Y}$ as a change of origin and scale in h_{X_E,Y_E}.

From Ma and Sun (2011), the mutual information among (X, Y)

$$\begin{aligned} M(X,Y) &= -\int_0^\infty \int_0^\infty f(x,y) \log \frac{f(x,y)}{f_X(x) f_Y(y)} dxdy \\ &= h_X + h_Y - h_{X,Y}, \end{aligned}$$

is the negative of the copula entropy $i_{X,Y}$ where h_X and h_Y respectively denote the Shannon entropy of X and Y and so,

$$-i_{X,Y} = h_X + h_Y - h_{X,Y}$$

giving

$$i_{X_E,Y_E} = h_{X_E,Y_E} - h_{X_E} - h_{Y_E}. \tag{2.6}$$

From the definition of the univariate equilibrium distribution $f_{X_E} = \frac{\bar{F}_X(x)}{E(X)}$, we get

$$h_{X_E} = \frac{H_X}{E(X)} + \log E(X), \tag{2.7}$$

and similarly,

$$h_{Y_E} = \frac{H_Y}{E(Y)} + \log E(Y), \tag{2.8}$$

where $H_X = -\int_0^\infty \bar{F}_X(x) \log \bar{F}_X(x) dx$ is the cumulative residual entropy of X and H_Y is similarly defined (see Rao *et al.*, 2004). From equations (2.5) through (2.8) the expression for copula entropy is obtained in terms of cumulative entropies as

$$i_{X_E,Y_E} = \frac{H_{X,Y}}{E(XY)} - \frac{H_X}{E(X)} - \frac{H_Y}{E(Y)} + \log \frac{E(XY)}{E(X)E(Y)}. \tag{2.9}$$

Finally from (2.3),

$$\begin{aligned} I_{\bar{C}}(X,Y) &= -\int_0^\infty \int_0^\infty E(XY) f_E(x,y) [\log E(XY) f_E(x,y)] f_X(x) f_Y(y) dxdy \\ &= -\int_0^1 \int_0^1 E(XY) f_E\left(\bar{F}_X^{-1}, \bar{F}_Y^{-1}\right) \left[\log E(XY) f_E\left(\bar{F}_X^{-1}, \bar{F}_Y^{-1}\right)\right] dudv. \end{aligned} \tag{2.10}$$

From

$$f(x,y) = \frac{\partial^2 \bar{F}(x,y)}{\partial x \partial y},$$

we have

$$\begin{aligned} f\big(\bar{F}_X^{-1}, \bar{F}_Y^{-1}\big) &= \frac{\partial^2 \bar{F}\big(\bar{F}_X^{-1}, \bar{F}_Y^{-1}\big)}{\partial \bar{F}_X^{-1} \partial \bar{F}_Y^{-1}} \\ &= \frac{\partial^2 C(u,v)}{\partial u \partial v} \frac{\partial u \partial v}{\partial \bar{F}_X^{-1} \partial \bar{F}_Y^{-1}} \\ &= c(u,v) \frac{\partial u \partial v}{\partial \bar{F}_X^{-1} \partial \bar{F}_Y^{-1}}. \end{aligned}$$

Thus from (2.10), the survival copula entropy is related to the copula entropy as

$$I_{\bar{C}}(X,Y) = -\int_0^1 \int_0^1 E(XY) f_X\big(\bar{F}_X^{-1}\big) f_Y\big(\bar{F}_Y^{-1}\big) c(u,v) \log\big[E(XY) f_X\big(\bar{F}_X^{-1}\big) f_Y\big(\bar{F}_Y^{-1}\big) c(u,v)\big]\, du dv. \tag{2.11}$$

In the next example we demonstrate how the above results work out in a specific distribution.

Example 2.2. Let (X,Y) follows bivariate Pareto distribution

$$\bar{F}(x,y) = (1+x+y)^{-\theta},\ \theta > 0;\ x,y > 0.$$

Then we have

$$\bar{F}_X(x) = (1+x)^{-\theta};\ \ \bar{F}_Y(y) = (1+y)^{-\theta},$$

$$E(XY) = \int_0^\infty \int_0^\infty (1+x+y)^{-\theta} dx dy = [(\theta-1)(\theta-2)]^{-1}, \quad \theta > 2 \tag{2.12}$$

and

$$f_{X_E,Y_E}(x,y) = (\theta-1)(\theta-2)(1+x+y)^{-\theta},$$

$$\begin{aligned} H_{XY} &= -\int_0^\infty \int_0^\infty (1+x+y)^{-\theta} (-\theta \log(1+x+y)) dx dy \\ &= \frac{(2\theta-3)\theta}{(\theta-1)(\theta-2)} \end{aligned} \tag{2.13}$$

and

$$\begin{aligned} h_{X_E,Y_E} &= -\int_0^\infty \int_0^\infty (\theta-1)(\theta-2)(1+x+y)^{-\theta} \Big[\log(\theta-1)(\theta-2) + \log(1+x+y)^{-\theta}\Big] dx dy \\ &= (2\theta-3)\theta + \log(\theta-1)(\theta-2). \end{aligned} \tag{2.14}$$

The formula (2.5) is verified from (2.12), (2.13) and (2.14). Also

$$H_X = -\int_0^\infty (1+x)^{-\theta} (-\theta \log(1+x)) dx = \frac{\theta}{(\theta-1)^2}.$$

Similarly $h_Y = \frac{\theta}{(\theta-1)^2}$ and $E(X) = E(Y) = \frac{1}{(\theta-1)}$. Hence from (2.9)

$$\begin{aligned} i_{X_E,Y_E} &= (2\theta-3)\theta - \frac{\theta}{\theta-1} - \frac{\theta}{\theta-1} + \log \frac{(\theta-1)^2}{(\theta-1)(\theta-2)} \\ &= \frac{2\theta^2 - 7\theta + 1}{\theta - 1} + \log \frac{\theta-1}{\theta-2}, \qquad \theta > 2. \end{aligned}$$

The survival copula is

$$\bar{C}_{X,Y}(u,v) = \left(u^{-\frac{1}{\theta}} + v^{-\frac{1}{\theta}} - 1\right)^{-\theta}, \tag{2.15}$$

which is the Clayton family. Also

$$c(u,v) = \frac{\theta+1}{\theta} \frac{u^{-\frac{1}{\theta}-1} v^{-\frac{1}{\theta}-1}}{\left(u^{-\frac{1}{\theta}} + v^{-\frac{1}{\theta}-1}\right)^{\theta+2}} (\theta+2). \tag{2.16}$$

One can directly calculate both $I_{\bar{C}}$ and $i(X,Y)$ from (2.15) and (2.16). We may also use the fact that

$$\bar{C}(u,v) = \frac{\theta+1}{\theta} \frac{u^{-\frac{1}{\theta}-1} v^{-\frac{1}{\theta}-1}}{\left(u^{-\frac{1}{\theta}} + v^{-\frac{1}{\theta}-1}\right)^{\theta+2}} c(u,v)$$

or formula (2.11). Note that when applying (2.11), $f_X(x) = \theta(1+x)^{-\theta-1}$ so that $f_X\left(\bar{F}_X^{-1}\right) = \theta u^{-\frac{\theta+1}{\theta}}$ and $f_Y\left(\bar{F}_Y^{-1}\right) = \theta v^{-\frac{\theta+1}{\theta}}$

We have so far discussed the entropy function derived from the survival copula. One can also define the entropy based on the usual copula C.

Definition 2.2. The cumulative copula entropy (CCE) of (X,Y) is defined as

$$I_C(X,Y) = -\int_0^1 \int_0^1 C(u,v) \log C(u,v) du dv.$$

Example 2.3. The Farlie-Gumbel-Morgestern copula

$$C(u,v) = uv[1+\theta(1-u)(1-v)], \quad -1 \le u \le 1$$

has CCE given by

$$I_C = -\int_0^1 \int_0^1 uv(1+\theta(1-u)(1-v))[\log u + \log v + \log(1+\theta(1-u)(1-v))] du dv.$$

The integral does not converge for $\theta > 0$. For $\theta < 0$,

$$\begin{aligned} I_C = \left(168\theta^2\right)^{-1} \Big[& -690 + 84\theta^2 + 10\theta^3 - 3\pi^2 + 9\pi^2\theta + 3(17 + 9\theta + 9\theta^2 + \theta^3) \\ & - 6(-1+3\theta)\log(-\theta)\log(1+\theta) + 18(1-3\theta)\Big] PolyLog[2, 1+\theta], \end{aligned}$$

where $PolyLog(n,z) = \sum_{k=1}^{\infty} z^k/k^n$.

Values of the entropy function for some values of θ are given in Table 2.

Table 2: CCE for FGM copula.

θ	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9
$I_{\bar{C}}$	0.24811	0.24615	0.24412	0.24200	0.23981	0.23753	0.23516	0.23268	0.23013

Remark 2.1. In general I_C and $I_{\bar{C}}$ are different for the random vector (X, Y). When (X, Y) is radially symmetric, that is for any (u, v) in I^2, the rectangles $[0, u] \times [0, v]$ and $[1-u, 1] \times [1-v, 1]$ have equal C-volume, then $C = \bar{C}$ and the entropy satisfies $I_C = I_{\bar{C}}$. Since the algebra involved in deriving various results in I_C is similar to those in $I_{\bar{C}}$, we restrict our subsequent discussions to the latter case.

3. PROPERTIES OF SURVIVAL COPULA ENTROPY

An important aspect in favour of SCE among other measures is that its values lies in a finite interval which makes it easier for comparison and interpretation.

Proposition 3.1. *The SCE satisfies*

$$\frac{1}{18} \leq I_{\bar{C}}(X, Y) \leq \frac{1}{3}. \tag{3.1}$$

Proof: It is well known that for every copula C and for all (u, v) in I^2,

$$W(u, v) \leq C(u, v) \leq M(u, v),$$

where $M(u, v) = \min(u, v)$ and $W = \max(u + v - 1, 0)$ are copula versions of the Fréchet-Hoeffding bounds of a bivariate distribution in $\mathbb{R}^2$. Let $\bar{M}$ and $\bar{W}$ be the survival copulas corresponding to M and W respectively. Then

$$\begin{aligned} \bar{M}(u, v) &= u + v - 1 + \min(1-u, 1-v) \\ &= u + v - \max(u, v) \\ &= \min(u, v) = M. \end{aligned}$$

Similarly $\bar{W} = W$. Thus $W(u, v) \leq \bar{C}(u, v) \leq M(u, v)$ so that

$$\begin{aligned} I_{\bar{C}}(X, Y) &= -\int_0^1 \int_0^1 \bar{C}(u, v) \log \bar{C}(u, v) du dv \\ &\geq -\int_0^1 \int_0^1 \bar{M}(u, v) \log \bar{M}(u, v) du dv \end{aligned}$$

or

$$\begin{aligned} I_{\bar{C}}(X, Y) &\geq -\int_0^1 \int_0^1 \min(u, v) \log \min(u, v) du dv \\ &= -\int_0^1 \int_u^1 v \ \log u \ du dv - \int_0^1 \int_0^u v \ \log v \ du dv \\ &= \frac{1}{18}, \quad \text{using } \ 0 \log 0 = 0. \end{aligned}$$

Also

$$\begin{aligned}
I_{\bar{C}}(X,Y) &\leq -\int_0^1\int_0^1 \bar{W}(u,v)\log\bar{W}(u,v)dudv \\
&= \int_0^1\int_0^1 \max(u+v-1,0)\log\max(u+v-1,0)dudv \\
&= \int_0^1\left(u-\frac{1}{2}\right)\log u du - \int_0^1\int_{1-u}^1\left(\frac{v\left(u+\frac{v}{2}-1\right)}{u+v-1}dv\right)du \\
&= \int_0^1\left(u-\frac{1}{2}\right)\log u du - \int_0^1\int_{1-u}^1\left[\frac{v}{2}+\frac{3}{2}(1-u)+\frac{3}{2}\left(\frac{(1-u)^2}{u+v-1}\right)\right]dudv \\
&= \frac{1}{3}.
\end{aligned}$$

□

Remark 3.1. It is not necessary that the SCE for all copulas attain the above bounds. For example, the Clayton copula contains M and W and hence their entropies lie in $\left[\frac{1}{18},\frac{1}{3}\right]$. At the same time the Gumbel-Barnett family does not include M and W and the bounds prescribed by M and W are not attained for this family. On the other hand $C_2 = \max(\theta uv + (1-\theta)(u+v-1),0)$ contains W but not M, while $C_3 = \theta/\log\left(e^{\theta/u}+e^{\theta/v}-e^{-v}\right)$ has M as a member, but not W. Further C and $\bar{C}$ have the same entropy if and only if C is radially symmetric.

3.1. Ordering copulas via entropy

There are many situations where the data on the same random variable comes from different sources and the problem is to choose the more informative one for analysis. In such circumstances the entropies in each case has to be compared. The ordering of copulas comes handy in comparing the entropies. If C_1 and C_2 are two copulas, we say that C_1 is smaller (larger) than C_2 in concordance order, if $C_1(u,v) \leq (\geq) C_2(u,v)$ for all u,v in I, and is denoted by $C_1 \prec (\succ) C_2$. Note that $C_1 \prec (\succ) C_2 \Leftrightarrow \bar{C}_1 \prec (\succ)\bar{C}_2$. The following proposition is immediate.

Proposition 3.2.

(3.2) $$\bar{C}_1 \prec (\succ)\bar{C}_2 \implies I_{\bar{C}_1} \leq (\geq) I_{\bar{C}_2}.$$

With some additional assumptions the converse of (3.2) is also true.

Proposition 3.3. *Let $\mathcal{A}$ denote the class of copulas that are concordance ordered, that is, for elements $\bar{C}_1$ and $\bar{C}_2 \in \mathcal{A}$, we have either $\bar{C}_1(u,v) \leq \bar{C}_2(u,v)$ or $\bar{C}_1(u,v) \geq \bar{C}_2(u,v)$. Then*

(3.3) $$I_{\bar{C}_1} \leq (\geq) I_{\bar{C}_2} \implies \bar{C}_1(u,v) \leq (\geq)\bar{C}_2(u,v) \text{ for all } u,v \text{ in } I.$$

Proof: To prove the above implication assume that $I_{\bar{C}_1} \geq I_{\bar{C}_2}$ which is equivalent to

$$\int_0^1 \int_0^1 \bar{C}_1(u,v) \log \bar{C}_1(u,v) dudv \geq \int_0^1 \int_0^1 \bar{C}_2(u,v) \log \bar{C}_2(u,v) dudv. \tag{3.4}$$

Since C_1 and C_2 are ordered, if $C_1 \leq C_2$ then $ref3.3$ is violated and hence $C_1 \geq C_2$. □

As an example, from the values of $I_{\bar{C}}$ given in Tables 1 and 2, it is seen that entropies are decreasing. It is well known that the corresponding copulas are also decreasing functions of θ in their assumed ranges.

3.2. Survival copula entropy and dependence

An important use of SCE is assessing the nature of dependence between X and Y, thus making a connection between entropy and dependence. The above discussion on concordances ordering and entropy have significant implications in ascertaining the mode of dependence and the SCE. An important and perhaps mostly used dependence concept is positive (negative) quadrant dependence PQD (NQD). Recall that a copula C is PQD (NQD) if $\bar{C}(u,v) \geq (\leq) uv$, for u, v in I.

Proposition 3.4. *The vector (X, Y) is PQD (NQD), then*

$$I_{\bar{C}} \geq (\leq) \frac{1}{4}. \tag{3.5}$$

Proof:

$$(X,Y) \text{ is PQD (NQD)} \implies C(u,v) \geq (\leq) \prod(u,v) = uv.$$

From

$$\int_0^1 \int_0^1 C(u,v) \log C(u,v) dudv \geq (\leq) \int_0^1 \int_0^1 uv(\log u + \log v) dudv = \frac{1}{4},$$

the result follows. □

Remark 3.2. In view of (3.1), PQD (NQD) random vectors are sought the interval $[\frac{1}{4}, \frac{1}{3}]$ and $[\frac{1}{18}, \frac{1}{4}]$ respectively. Further, if the random variables X and Y are independent then $I_{\bar{C}} = \frac{1}{4}$. The next proposition gives a criterion to check whether which of two random variables are more positively dependent.

Proposition 3.5. *For copulas $\bar{C}_1, \bar{C}_2 \in \mathcal{A}$, $\bar{C}_1$ is more PQD than $\bar{C}_2$ if and only if $I_{\bar{C}_1} \geq I_{\bar{C}_2}$.*

Proof: We say that $\bar{C}_1$ is more PQD than $\bar{C}_2$ if $\bar{C}_1 \succ \bar{C}_2$. Thus

$$\begin{aligned} \bar{C}_1 \text{ is more PQD than } \bar{C}_2 &\Leftrightarrow C_1(u,v) \geq C_2(u,v) \text{ for all } u, v \\ &\Leftrightarrow I_{\bar{C}_1} \geq I_{\bar{C}_2}, \end{aligned}$$

by (3.2) and Remark 3.2. □

We can also compare members of a specified family of copulas $\bar{C}_\theta(u,v)$ indexed by a parameter $\theta \in \Theta$. The family $\{\bar{C}_\theta\}, \theta \in \Theta$ is positively (negatively) ordered whenever $\bar{C}_{\theta_1} \prec (\succ)\ \bar{C}_{\theta_2}$ for $\theta_1, \theta_2 \in \Theta, \theta_1 \leq (\geq)\theta_2$. In this case we have the next proposition that gives a criterion to distinguish between more positive dependent among families of copulas.

Proposition 3.6. *Let $\{\bar{C}_\theta\}$ be positively (negatively) ordered. Then $I_{\bar{C}_{\theta_1}} \leq (\geq) I_{\bar{C}_{\theta_2}}$, for all $\theta_1, \theta_2 \in \Theta, \theta_1 \leq (\geq)\theta_2$.*

Example 3.1. The Gumbel-Barnett family in Example 2.1 is negatively ordered as can be verified from Table 1 and FGM copula in Table 2 is positively ordered.

Remark 3.3. For many standard copula families, it is algebraically difficult to establish whether it is positively or negatively ordered. Proposition 3.4 gives a relative simple alternative tool to resolve this problem.

Remark 3.4. The relationship CCE has with well known measures of dependence is also worth examination. The measures in common use are the

$$\text{Kendall's tau,} \quad \tau = 4\int\int_{I^2} C(u,v)\frac{\partial^2 C}{\partial u \partial v} dudv - 1,$$

$$\text{Spearman's rho,} \quad \rho = 12\int\int_{I^2} C(u,v)dudv - 3,$$

$$\text{Blomqvist's beta,} \quad \beta = 4C\left(\frac{1}{2},\frac{1}{2}\right) - 1,$$

$$\text{Gini coefficient,} \quad \xi = 4\left[\int\int_{I^2} C(u,1-u)du - \int\int_{I^2}(u - C(u,u))du\right], \quad \text{and}$$

the product moment correlation coefficient

$$r = [D(X)D(Y)]^{-1}\int\int_{I^2}[C(u,v) - uv]dF^{-1}(u)dG^{-1}(u)$$

where $D(X), D(Y)$ are standard deviations of X and Y, and F and G are their marginal distribution functions. It is known that (Nelson, 2006) when X and Y are PQD:

(i) $3\tau \geq \rho \geq 0, \quad \beta \geq 0$ and $\xi \geq 0$.

(ii) For non-decreasing function $p(x)$ and $q(y)$ whose expectations are finite and

$$E(p(X)q(Y)) < \infty, \qquad \text{Cov}(p(X), q(Y)) \geq 0$$

and conversely, implying that $\text{Cov}(X,Y)$ and hence $r \geq 0$. Combining these with our earlier propositions we find that when the cumulative copula $I_{\bar{C}}$ lies in the interval $[\frac{1}{4}, \frac{1}{3}]$, we have positive dependence in the sense of PQD as well as τ, ρ, β, ξ and r. In general it is difficult to find expressions that connects $\bar{C}$ with the various coefficients τ, ρ, etc, but one can obtain formulas in respect of specific copulas. For example, in the Gumbel-Barnett copula, the Spearman's coefficient is $\rho = 12\Big[-\frac{e^{-\frac{4}{\theta}}}{\theta}\ EI\big(-\frac{4}{\theta}\big)\Big] - 3$ so that from Example 2.1,

$$\rho = 12\left(\frac{I_{\bar{C}}}{\theta}\right) - 3.$$

In this copula, $I_{\bar{C}}$ is a decreasing function of θ and the maximum of $I_{\bar{C}}$ occurs at $\theta = 0$, the case of independence in which case $I_{\bar{C}} = 0.25$. Hence for this copula $\rho \le 0$ so that there is negative dependence for all $0 \le \theta \le 1$. By virtue of Propositions 3.2 and 3.4, we also conclude that as the $I_{\bar{C}}$ value decreases from 0.25, so does the extent of negative dependence between X and Y.

3.3. Effect of transformations

There are occasions where transformations have to be applied to the baseline random variables to facilitate easier analysis. It is of interest to know how the SCE is affected by such transformations.

Proposition 3.7. *Let $T(X)$ and $W(Y)$ be strictly monotone transformations of X and Y. Let the corresponding SCE's be $I(T(X),W(Y))$ and $I(X,Y)$ respectively. Then*

- **(i)** *$I(T(X),W(Y)) = I(X,Y)$ when $T(X)$ and $W(Y)$ are both strictly increasing,*
- **(ii)** *$I(T(X),W(Y)) = I(X,Y)$ when $T(X)$ and $W(Y)$ are both strictly decreasing and I is the cumulative copula entropy of (X,Y) and*
- **(iii)** $$I(T(X),W(Y)) = \begin{cases} \int_0^1 \int_0^1 \big[v - \bar{C}(u,v)\big] \log\big[v - \bar{C}(u,v)\big] du dv, \\ \int_0^1 \int_0^1 \big[u - \bar{C}(u,v)\big] \log\big[u - \bar{C}(u,v)\big] du dv, \end{cases}$$

where $T(X)$ is strictly increasing (decreasing) and $W(Y)$ is strictly decreasing (increasing).

Proof: Proceeding as in Theorems 2.4.3 and 2.4.4 in (Nelson, 2006, p. 25, 26) we find that

$$\bar{C}_{T(X),W(Y)}(u,v) = \bar{C}_{X,Y}(u,v)$$

when both $T(X)$ and $T(Y)$ are increasing,

$$\bar{C}_{T(X),W(Y)}(u,v) = v - \bar{C}_{X,Y}(1-u,v)$$

when $T(X)$ is increasing and $W(Y)$ is decreasing,

$$\bar{C}_{T(X),W(Y)}(u,v) = u - \bar{C}_{X,Y}(u,1-v)$$

when $T(X)$ is decreasing and $W(Y)$ is increasing, and

$$\bar{C}_{T(X),W(Y)}(u,v) = C_{X,Y}(u,v)$$

when $T(X)$ and $W(Y)$ are strictly decreasing. Then first and last results directly establish (i) and (ii) of the Proposition and (iii) is obtained from the 2nd and 3rd results with a transformation of $(1-u)$ $(1-v)$ to u (v). It may be noted that entropy of $(T(X), W(Y))$ is independent of the form of the two functions. □

Example 3.2. The linear transformations $T(x) = \alpha x + \beta$ and $W(y) = \gamma y + \phi$ are common in data analysis. When $\alpha > 0$, $\gamma > 0$, in this case

$$\bar{C}(u, v) = u\ v\ \exp[-\theta \log u \log v], \qquad 0 \leq \theta \leq 1,$$

and

$$I_{\bar{C}}(\alpha X + \beta, \gamma Y + \phi) = -e^{\frac{4}{\theta}} EI\left(-\frac{4}{\theta}\right) = I_{\bar{C}}(X, Y), \qquad \text{for} \quad \alpha, \beta, \gamma, \phi > 0.$$

4. APPLICATIONS

In this section we demonstrate how the results obtained in the previous sections can be implemented in a practical problem. The example considered relates to an investigation on 20 individuals for isolated aortic regurgitations before and after surgery and 20 persons for isolated mitral regurgitation analysed in Kumar and Shoukri (2007). Data on pre-operative ejection fraction (X) and post-operative ejection fraction (Y) arranged in order of magnitude are

$$X:\ 0.29, 0.36, 0.39, 0.41, 0.50, 0.53, 0.54, 0.55, 0.56, 0.56, 0.56, 0.58, 0.60, 0.60, 0.62, 0.64, 0.64, 0.67, 0.80, 0.87,$$

$$Y:\ 0.17, 0.24, 0.26, 0.26, 0.27, 0.29, 0.30, 0.32, 0.33, 0.33, 0.34, 0.38, 0.47, 0.47, 0.50, 0.56, 0.58, 0.59, 0.62, 0.63.$$

The first step in the analysis is the estimation of the SCE. We consider the empirical survival copula for a random sample $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ from a continuous bivariate distribution given by

$$\bar{C}\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{(\text{number of pairs in the sample with } x > x_{(i)}, y > y_{(j)})}{n},$$

where $x_{(i)}\left(y_{(j)}\right)$ is the i-th (j-th) order statistic of the observations on $X(Y)$. Using $\bar{C}\left(\frac{i}{n}, \frac{j}{n}\right)$ as the estimator of $\bar{C}(u, v)$, the resubstitution estimator of $I_{\bar{C}}$ is

$$\hat{I}_{\bar{C}} = \frac{2}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{C}\left(\frac{i}{n}, \frac{j}{n}\right) \log \bar{C}\left(\frac{i}{n}, \frac{j}{n}\right),$$

at those points for which $\frac{i}{n} + \frac{j}{n} - 1 > 0$. The estimated value for the given sample is $\hat{I}_{\bar{C}} = 0.3049$. Kumar and Shoukri (2007) in their analysis, found that the Pearson correlation coefficient $r = 0.6870$, Kendall's rank correlation $\tau = 0.5050$ and Spearman's rank correlation $\rho = 0.6970$. Thus all the measures indicate positive dependence between the two sets of observations. Using Proposition 3.4, our nonparametric estimate also support this conclusion irrespective of the copula model, since $\hat{I}_{\bar{C}} > \frac{1}{4}$. Further, the data gives satisfactory evidence for the PQD property of the underlying copula.

5. CONCLUSION

In this work we have proposed a measure of uncertainty based on survival copula and examined some of its properties. Apart from being useful to evaluate uncertainty it can be of application in assessing copula properties like independence and their ordering. The proposed concept is more general than distribution-based counterparts and has some advantages over them and the existing copula entropy.

DATA AVAILABILITY STATEMENTS

Authors can confirm that all relevant data are included in the article.

CONFLICT OF INTEREST STATEMENT

On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

Charzyńska, A. and Gambin, A. (2015). Improvement of the k-NN entropy estimator with applications in systems biology. *Entropy*, 18(1):13.

Chen, L., Singh, V. P., and Guo, S. (2013). Measure of correlation between river flows using the copula-entropy method. *Journal of Hydrologic Engineering*, 18(12):1591–1606.

Kayser, S. J., Ince, R. A. A., Gross, J., and Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience*, 35(44):14691–14701.

Kumar, P. and Shoukri, M. M. (2007). Copula based prediction models: an application to an aortic regurgitation study. *BMC Medical Research Methodology*, 7(1):1–9.

Ma, J. (2019a). Discovering association with copula entropy. *arXiv preprint*, arXiv:1907.12268.

Ma, J. (2019b). Variable selection with copula entropy. *arXiv preprint*, arXiv:1910.12389.

Ma, J. and Sun, Z. (2008). Dependence structure estimation via copula. *arXiv preprint*, arXiv:0804.4451.

Ma, J. and Sun, Z. (2011). Mutual information is copula entropy. *Tsinghua Science & Technology*, 16(1):51–54.

Mesiar, R. and Sheikhi, A. (2021). Nonlinear random forest classification, a copula-based approach. *Applied Sciences*, 11(15):7140.

Nair, N. U. and Preeth, M. (2009). On some properties of equilibrium distributions of order n. *Statistical Methods and Applications*, 18(4):453–464.

Nair, N. U. and Sankaran, P. G. (2010). A new measure of association for bivariate survival data. *Journal of Statistical Planning and Inference*, 140(9):2569–2581.

Navarro, J. and Sarabia, J. M. (2010). Alternative definitions of bivariate equilibrium distributions. *Journal of Statistical Planning and Inference*, 140(7):2046–2056.

Nelson, R. B. (2006). *An Introduction to Copulas*. Springer, New York.

Rajesh, G., Abdul-Sathar, E. I., Nair, K. R. M., and Reshmi, K. V. (2014). Bivariate extension of dynamic cumulative residual entropy. *Statistical Methodology*, 16:72–82.

Rao, M., Chen, Y., Vemuri, B. C., and Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE transactions on Information Theory*, 50(6):1220–1228.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 117–138
https://doi.org/10.57805/revstat.v23i1.544

Bayesian and Frequentist Estimation of Stress-Strength Reliability from a New Extended Burr XII Distribution

Authors: Varun Agiwal
– Indian Institute of Public Health, Hyderabad, Telangana, India
varunagiwal.stats@gmail.com

Shikhar Tyagi
– Department of Statistics and Data Science, Christ Deemed to be University, Bangalore, India
shikhar1093tyagi@gmail.com

Christophe Chesneau ✉
– Department of Mathematics, LMNO, Université de Caen-Normandie, Caen, France
christophe.chesneau@gmail.com

Received: March 2021 Revised: March 2023 Accepted: March 2023

Abstract:

- In this article, we propose and study a new three-parameter heavy-tailed distribution that unifies the Burr type XII and power inverted Topp-Leone distributions in an original manner. This unification is made through the use of a simple 'shift parameter'. Among its interesting functionalities, it exhibits possibly decreasing and unimodal probability density and hazard rate functions. We examine its quantile function, stochastic dominance, ordinary moments, weighted moments, incomplete moments, and stress-strength reliability coefficient. Then, the classical and Bayesian approaches are developed to estimate the model and stress-strength reliability parameters. Bayes estimates are obtained under the squared error and entropy loss functions. Simulated data are considered to point out the performance of the derived estimates based on the mean squared error. In the final part, the potential of the new model is exemplified by the analysis of two engineering data sets, showing that it is preferable to other reputable and comparable models.

Keywords:

- *Burr distribution; Bayesian inference; maximum likelihood method; stress-strength reliability; data analysis.*

AMS Subject Classification:

- 62E15, 60E05, 62F10.

✉ Corresponding author

1. INTRODUCTION

In the analysis of survival data, the researcher attempts to make predictions about the lifetime of all elements / systems by fitting a statistical distribution / model. The underlying distribution of a dataset can then be used to estimate component life characteristics, such as reliability or probability of failure at any given time, average life, failure rate, etc. Reliability is used to assess the characteristics of strength and failure, compare several different models, predict product reliability, etc. In recent years, the Burr type XII (BXII) distribution created in Burr (1942) has gained great applicability in the field of reliability / survival analysis and has been discussed by many authors. It is widely recognized as one of the most straightforward and applicable heavy-tailed distributions. The fundamental properties and estimation methods based on the BXII distribution have been derived in Wang *et al.* (1996), Zimmer *et al.* (1998), Moore and Papadopoulos (2000), Mousa and Jaheen (2002) and Wu *et al.* (2007). Due to its flexibility for data modeling, some extensions of the BXII distribution have been introduced in the literature. Among them are the beta BXII distribution (see Paranaiba *et al.*, 2011), Kumaraswamy BXII distribution (see Paranaiba *et al.*, 2013), beta exponentiated BXII distribution (see Mead, 2014), Marshall-Olkin BXII distribution (see Al-Saiari *et al.*, 2014), McDonald BXII distribution (see Gomes *et al.*, 2015), Weibull BXII distribution (see Afify *et al.*, 2018), Kumaraswamy exponentiated distribution (see Mead and Afify, 2017), generalized Burr-G distribution (see Nasir *et al.*, 2017), Topp-Leone BXII distribution (see Reyad and Othman, 2017), transmuted BXII distribution (see Afify *et al.*, 2018), generalized BXII power series distribution (see Elbatal *et al.*, 2019) and modified BXII distribution (see Jamal *et al.*, 2020).

Along with these extended BXII distributions, other successful distributions for modeling survival phenomena have been established in recent years. This is the case for the 'power inverted Topp-Leone (PITL) distribution' invented in Abushal *et al.* (2021), which also belongs to the heavy-tailed family of distributions. The first thing to know about the PITL distribution is mathematical; the PITL distribution is the distribution of $(1 - X)^{1/c}X^{-1/c}$, where $c > 0$ and X is a random variable with the classical one-parameter Topp-Leone distribution. It is also the power version of the inverted Topp-Leone (ITL) distribution proposed in Hassan *et al.* (2020). The PITL distribution is motivated in Abushal *et al.* (2021) by the following advantages: (i) it benefits from more flexibility compared with the ITL distribution on several aspects, including the shape possibilities of the associated probability density function (pdf) and hazard rate function (hrf), (ii) the inferences of the PITL model are quite manageable with the standard estimation methods, (iii) precise acceptance sampling plans can be developed without difficulty, and (iv) the PITL model is better than other competitive models, a claim illustrated with the analysis of the vinyl chloride data from Bhaumik *et al.* (2009) and the precipitation data from Hinkley (1977).

The purpose of this article is to create an original three-parameter heavy-tailed distribution that unifies the BXII and PITL distributions and to present its main statistical properties. A new tuning parameter that permits a shift between these two famous distributions largely controls this unification. It thus makes it possible to reach a wide range of intermediate distributions with equivalent interests and potentials. The proposed distribution is called the new extended BXII (NEB) distribution. In the first part of the article, we discuss the main characteristics of the NEB distribution, with an emphasis on the role of the shift parameter.

Also, some of its functionalities and distributional measures are derived. Among others, we show that the pdf and hrf may be both decreasing and unimodal, which remains a rare feature for a three-parameter heavy-tailed distributions. Then, we examine the quantile function (qf), stochastic dominance, ordinary moments, weighted moments, incomplete moments, and an important measure of system performance: the stress-strength reliability coefficient, defined on the basis of two independent random variables with the NEB distribution. The historical motivations behind this coefficient in a general setting can be found in Church and Harris (1970). The second part of the article is devoted to the inferences of the NEB model. This includes properties, estimation of the model parameters, and estimation of the stress-strength reliability coefficient through classical and Bayesian methods. We now emphasize that the problem of estimating the stress-strength reliability is widely discussed in many articles and remains a common demand in mechanical reliability systems. For the consideration of various lifetime models, we may refer to Mokhlis (2005), Lio and Tsai (2012), Rao *et al.* (2015), Laslan and Nadar (2017) and, more recently, Byrnes *et al.* (2019) and Maurya and Tripathi (2020), and the references cited therein. Following the spirit of these works, the estimation of the stress-strength reliability coefficient in the context of the NEB distribution opens some perspectives in reliability studies. In this regard, we analyze two sets of engineering data. Additionally, statistical comparisons with existing lifetime models that incorporate three or four parameters derived from the BXII model are carried out, and the results are satisfactory for the NEB model.

From the above consideration, we organize the paper as follows: Section 2 defines the NEB distribution along with a selection of its properties. Section 3 concerns the parameters and stress-strength reliability estimates via the maximum likelihood approach, with discussions on their asymptotic distributions. Then, in Section 4, the Bayes estimates are obtained under two different loss functions assuming uniform and gamma prior distributions for the parameters. Sections 5 and 6 provide the applicability of the new distribution and obtain the performance of the estimates. Last, Section 7 provides the concluding remarks.

2. PROPOSED DISTRIBUTION AND ITS PROPERTIES

2.1. Definition and motivation

At the basis of the NEB distribution, there is the following analytical result.

Proposition 2.1. *Let $a \in [0,2]$ and $c, k > 0$. Then, the following function:*

$$F(x) = \begin{cases} 1 - \dfrac{(1+ax^c)^k}{(1+x^c)^{2k}}, & x > 0, \\ 0, & x \leq 0, \end{cases} \tag{2.1}$$

has the properties of a valid cumulative distribution function (cdf).

Proof: First, it is clear that $F(x) \leq 1$ and, by the Bernoulli inequality, we have $(1 + x^c)^2 \geq 1 + 2x^c \geq 1 + ax^c$, implying that $F(x) \geq 0$. Furthermore, $\lim_{x\to 0} F(x) = 0 = F(0)$ implying that $F(x)$ is continuous in 0 and, a fortiori, in $\mathbb{R}$. It is clear that $\lim_{x\to+\infty} F(x) = 1$.

Now, for $x > 0$, since $a \in [0, 2]$, we have

$$F'(x) = ckx^{c-1}(ax^c + 2 - a)\frac{(1 + ax^c)^{k-1}}{(1 + x^c)^{2k+1}} \geq 0,$$

implying that $F(x)$ is non-decreasing. The required properties are fulfilled; the function $F(x)$ is a valid cdf. □

Based on Proposition 2.1, we are now in the position to explicit the NEB distribution. The NEB distribution with parameters a, c and k, also denoted as NEB(a, c, k), is defined either with the cdf $F(x)$ given in (2.1) or the pdf specified as

$$f(x) = \begin{cases} ckx^{c-1}(ax^c + 2 - a)\dfrac{(1 + ax^c)^{k-1}}{(1 + x^c)^{2k+1}}, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{2.2}$$

It is worth mentioning that c and k are shape parameters, whereas a is a scale parameter.

Basically, a random variable X with the NEB distribution satisfies: $P(X \in D) = \int_D f(x)dx$ for any univariate real domain D and, for any function $\phi(x)$, the expectation of the transformed variable $\phi(X)$, denoted by $E(\phi(X))$, can be expressed in the following integral form: $E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(x)f(x)dx$, provided that it converges (in the integral sense). These two formulas are the basis of measures and known distributional functions based on the moments.

Thus defined, thanks to the parameter a, the NEB distribution constitutes a new lifetime distribution with three parameters extending both the BXII and PITL distributions. More precisely, a can be viewed as a 'shift parameter' that allows a slip between the BXII and PITL distributions in the following sense: when $a = 0$, the NEB distribution becomes the BXII distribution, when $a = 2$, the NEB distribution becomes the PITL distribution, naturally, when $a = 2$ and $c = 1$, the power transformation of the PITL distribution disappears and the NEB distribution becomes the ITL distribution, and, to our knowledge, all the intermediary cases $a \in (0, 2)$ bring new distributions.

To realize the possibilities of the NEB distribution modeling, let us now investigate some analytical properties of its pdf. First, when $x \to 0$, the following equivalence holds: $f(x) \sim (2 - a)ckx^{c-1}$ and, when $x \to +\infty$, we get $f(x) \sim cka^k x^{-ck-1}$. From these results, we derive the following nuanced limits:

$$\lim_{x \to 0} f(x) = \begin{cases} 0, & c > 1, \\ (2 - a)k, & c = 1, \\ +\infty, & c \in (0, 1), \end{cases}$$

and $\lim_{x \to +\infty} f(x) = 0$ for all the values of the parameters, the rate of convergence having a polynomial decay governed by the parameter c. Further investigations show that $f(x)$ is a decreasing function for $c \leq 1$ and is unimodal for $c > 1$. The mode can be determined numerically.

Furthermore, using the Riemann integral criteria, we get $\int_0^{+\infty} e^{tx} f(x)dx = +\infty$ for all $t > 0$, meaning that the NEB distribution is heavy right-tailed. It thus keeps the heavy-tailed nature of its parental distributions: the BXII and PITL distributions.

For more remarks, Figure 1 shows some possible shapes of the pdf with diverse values for a, c and k.

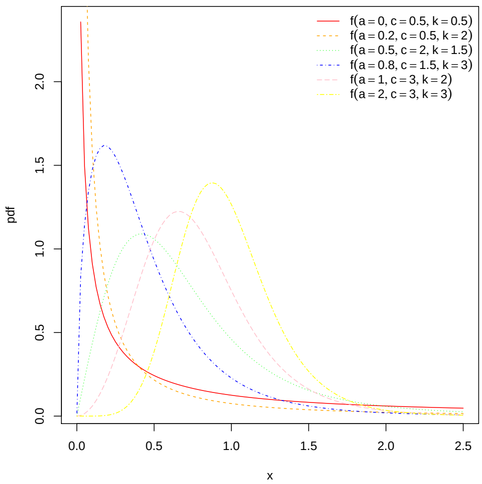


Figure 1: Panel of shapes of the pdf of the NEB distribution.

Figure 1 illustrates the decreasing and unimodal nature of $f(x)$. It is also shown that $f(x)$ has a versatile mode which is greatly affected by the parameter a. Almost symmetrical shapes can be seen, as in the yellow curve, also corresponding to the case $a = 2$ referring to the PTIL distribution. Moreover, Figure 1 illustrates the compromise that the NEB distribution made between the BXII and PITL distributions.

2.2. Complementary functions

We now focus on important reliability functions that may appear in various aspects of the NEB distribution analysis. The survival function (sf) and hrf of the NEB distribution are inscribed as

$$\bar{F}(x) = 1 - F(x) = \begin{cases} \dfrac{(1 + ax^c)^k}{(1 + x^c)^{2k}}, & x > 0, \\ 1, & x \leq 0, \end{cases}$$

and

$$h(x) = \frac{f(x)}{\bar{F}(x)} = \begin{cases} ckx^{c-1} \dfrac{ax^c + 2 - a}{(1 + ax^c)(1 + x^c)}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

respectively. An asymptotic study of $h(x)$ is now provided. First, when $x \to 0$, the following equivalence holds: $h(x) \sim (2 - a)ckx^{c-1}$, and when $x \to +\infty$, we obtain $h(x) \sim ckx^{-1}$. From these results, we derive the following limits:

$$\lim_{x \to 0} h(x) = \begin{cases} 0, & c > 1, \\ (2 - a)k, & c = 1, \\ +\infty, & c \in (0, 1), \end{cases}$$

and $\lim_{x \to +\infty} h(x) = 0$ for the values of the parameters. Since the variety of shapes is an important indicator on the modeling flexibility of a distribution (see Aarset, 1987), we provide a graphical analysis of $h(x)$ in Figure 2 with diverse values for a, c and k.

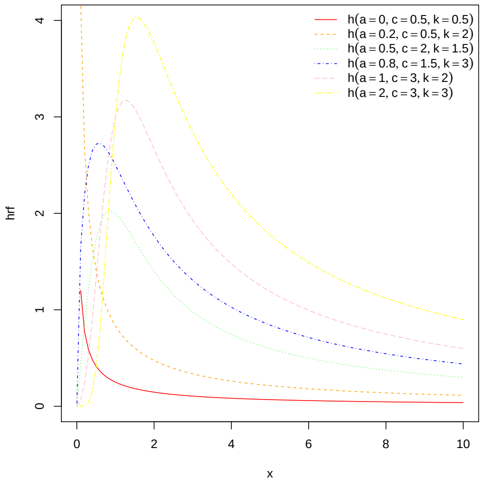


Figure 2: Panel of shapes of the hrf of the NEB distribution.

From Figure 2, we see that $h(x)$ has the same global shapes properties than $f(x)$, only varying on the weights of the tails: it is decreasing for $c \leq 1$ and has only one maximal point for $c > 1$. The parameter a mainly affects the value of the maximal point. Hence, the so-called decreasing and bathtub upside-down hazard rates of survival data can be reached by the NEB model.

We complete the presentation of the NEB distribution by expressing its qf. The notion of qf is very useful on various aspects in probability and statistics; it is at the same level of importance as the cdf to define a distribution (see Gilchrist, 2000). The expression of the qf of the NEB distribution follows through the solution of the following nonlinear equation: $F(x) = u$ with respect to x. After a step-by-step development, we come to

$$Q(u) = \frac{1}{2^{1/c}} \left\{ -[2 - a(1-u)^{-1/k}] + \sqrt{[2 - a(1-u)^{-1/k}]^2 - 4[1 - (1-u)^{-1/k}]} \right\}^{1/c},$$

where $u \in (0, 1)$. As a basic application, the three quartiles of the NEB distribution are given by $Q_1 = Q(1/4)$, $Q_2 = Q(1/2)$ and $Q_3 = Q(3/4)$, respectively. Also, among the possible uses of this qf, one can use it to generate values from any random variable with the NEB distribution, define diverse distributional functions analogous to the pdf and hrf, and various measures on skewness and kurtosis.

2.3. Stochastic dominance

The NEB distribution has several stochastic dominance properties involving $F(x)$ which are of interest in understanding the roles of the parameters a, c and k for distributional comparison. Here, we focus on the notion of first-order stochastic (fos) dominance as presented in Shaked and Shanthikumar (2007).

Proposition 2.2. *The following stochastic order properties hold: if $a_2 \geq a_1$, the NEB distribution defined with $a = a_2$ fos dominates the NEB distribution defined with $a = a_1$; if $k_2 \geq k_1$, the NEB distribution defined with $k = k_1$ fos dominates the NEB distribution defined with $k = k_2$.*

Proof: The proof is based on the monotonicity of $F(x) = F(x; a, c, k)$ with respect to the parameters. We have

$$\frac{\partial}{\partial a} F(x; a, c, k) = -kx^c \frac{(1 + ax^c)^{k-1}}{(1 + x^c)^{2k}} \leq 0,$$

which means that $F(x)$ is a decreasing function with respect to a, implying that, if $a_2 \geq a_1$, the NEB distribution defined with $a = a_2$ fos the NEB distribution defined with $a = a_1$. Now, we have

$$\frac{\partial}{\partial k} F(x; a, c, k) = \frac{(1 + ax^c)^k}{(1 + x^c)^{2k}} [2 \log(1 + x^c) - \log(1 + ax^c)] \geq 0,$$

which means that $F(x)$ is an increasing function with respect to k, implying that, if $k_2 \geq k_1$, the NEB distribution defined with $k = k_1$ fos dominates the NEB distribution defined with $k = k_2$. This ends the proof of the three items of the proposition. □

Thus, based on Proposition 2.2, we see that the parameter c has the most complex role for the comparison of NEB distributions differing with their parameters. Moreover, the first result and the expression of $F(x)$ justify the naming of 'shift parameter' for a.

2.4. Moment properties

The following result concerns the ordinary moments of the NEB distribution.

Proposition 2.3. *Let X be a random variable with the NEB distribution and r be an integer. Then, X admits an r-th ordinary moment, i.e., $\mu'_r = E(X^r)$, if and only if $r < ck$. In this case, μ'_r can be expressed as the following infinite sum expansion:*

$$\mu'_r = k \sum_{\ell=0}^{+\infty} \binom{k-1}{\ell} (a-1)^\ell \Big[aB\Big(\frac{r}{c} + \ell + 2, k - \frac{r}{c}\Big) + (2 - a)B\Big(\frac{r}{c} + \ell + 1, k + 1 - \frac{r}{c}\Big)\Big],$$

where $B(u, v)$ is the beta function: $B(u, v) = \int_0^1 t^{u-1}(1 - t)^{v-1} dt$ with $u, v > 0$.

Proof: Provided that it exists, we have $\mu'_r = \int_{-\infty}^{+\infty} x^r f(x)dx$. In view of the definition of $f(x)$ in (2.2), only the neighborhoods of $x = 0$ and $+\infty$ of the function $x^r f(x)$ need processing, and we can invoke the integral Riemann criteria in this regard. In the neighborhood of $x = 0$, we have $x^r f(x) \sim ck(2 - a)x^{r+c-1}$, which is the main term of a convergent integral over $x \in (0, d)$ with $d > 0$ if and only if $r + c > 0$, which is always fulfilled. Also, in the neighborhood of $x = +\infty$, we have $x^r f(x) \sim cka^k x^{r-ck-1}$ which is the main term of a convergent integral over $x \in (d, +\infty)$ if and only if $r - ck < 0$, which is satisfied if $r < kc$. In the end, μ'_r exists if and only $r < ck$.

In this case, in order to express $\mu'_r = \int_{-\infty}^{+\infty} x^r f(x)dx$ as desired, for $x > 0$, we set $f(x) = f_1(x) + f_2(x)$, where

$$f_1(x) = ckax^{2c-1}\frac{(1+ax^c)^{k-1}}{(1+x^c)^{2k+1}}, \quad f_2(x) = ck(2-a)x^{c-1}\frac{(1+ax^c)^{k-1}}{(1+x^c)^{2k+1}},$$

which can be also written as

$$f_1(x) = ckax^{2c-1}\frac{[1+(a-1)x^c/(1+x^c)]^{k-1}}{(1+x^c)^{k+2}}$$

and

$$f_2(x) = ck(2-a)x^{c-1}\frac{[1+(a-1)x^c/(1+x^c)]^{k-1}}{(1+x^c)^{k+2}}.$$

Since $a \in [0,2]$ and $x > 0$, it is clear that $|(a-1)x^c/(1+x^c)| < 1$. Therefore, the generalized version of the binomial formula gives

$$\left[1+(a-1)\frac{x^c}{1+x^c}\right]^{k-1} = \sum_{\ell=0}^{+\infty}\binom{k-1}{\ell}(a-1)^\ell\frac{x^{c\ell}}{(1+x^c)^\ell}.$$

Note that the limit $+\infty$ can be replaced by $k-1$ if k is an integer greater to 1. So $f_1(x)$ and $f_2(x)$ can be expressed as

$$f_1(x) = cka\sum_{\ell=0}^{+\infty}\binom{k-1}{\ell}(a-1)^\ell\frac{x^{c(\ell+2)-1}}{(1+x^c)^{\ell+k+2}}$$

and

$$f_2(x) = ck(2-a)\sum_{\ell=0}^{+\infty}\binom{k-1}{\ell}(a-1)^\ell\frac{x^{c(\ell+1)-1}}{(1+x^c)^{\ell+k+2}},$$

respectively. By invoking the dominated convergence theorem to justify the exchange of the signs $\sum$ and $\int$, we obtain

$$\begin{aligned}\mu'_r &= \int_0^{+\infty} x^r f_1(x)dx + \int_0^{+\infty} x^r f_2(x)dx \\ &= cka\sum_{\ell=0}^{+\infty}\binom{k-1}{\ell}(a-1)^\ell\int_0^{+\infty}\frac{x^{r+c(\ell+2)-1}}{(1+x^c)^{\ell+k+2}}dx \\ &\quad + ck(2-a)\sum_{\ell=0}^{+\infty}\binom{k-1}{\ell}(a-1)^\ell\int_0^{+\infty}\frac{x^{r+c(\ell+1)-1}}{(1+x^c)^{\ell+k+2}}dx. \end{aligned} \tag{2.3}$$

With the change of variable $y = x^c$, the two integral terms can be expressed as

$$\int_0^{+\infty}\frac{x^{r+c(\ell+2)-1}}{(1+x^c)^{\ell+k+2}}dx = \frac{1}{c}\int_0^{+\infty}\frac{y^{r/c+\ell+1}}{(1+y)^{\ell+k+2}}dy = \frac{1}{c}B\Big(\frac{r}{c}+\ell+2, k-\frac{r}{c}\Big)$$

and

$$\int_0^{+\infty}\frac{x^{r+c(\ell+1)-1}}{(1+x^c)^{\ell+k+2}}dx = \frac{1}{c}\int_0^{+\infty}\frac{y^{r/c+\ell}}{(1+y)^{\ell+k+2}}dy = \frac{1}{c}B\Big(\frac{r}{c}+\ell+1, k+1-\frac{r}{c}\Big).$$

By putting these equations into (2.3), we obtain the stated result. □

In any case, if $r < ck$, μ'_r can be evaluated in a numerical way by using any standard mathematical software.

With one of these approaches, we are able to evaluate standard moment measures such as the mean of X specified by $\mu = \mu'_1$ and the variance of X given as $V = \mu'_2 - \mu^2$, as well as moment measures of skewness and kurtosis.

The incomplete moments of X taken at a specific value $t \geq 0$ is also of interest. It can be expanded as described in the new result.

Proposition 2.4. *Let X be a random variable with the NEB distribution, r be an integer and $t \geq 0$. Then, X admits incomplete moments of all orders and the r-th incomplete moment of X at the level t, i.e., $\mu'_r(t) = E(X1_{\{X \leq t\}})$, can be expressed as the following infinite sum expansion:*

$$\begin{aligned} \mu'_r(t) = \; & k \sum_{\ell=0}^{+\infty} \binom{k-1}{\ell} (a-1)^{\ell} \\ & \times \left[aB_{t^c/(1+t^c)}\left(\frac{r}{c} + \ell + 2, k - \frac{r}{c}\right) + (2-a) B_{t^c/(1+t^c)}\left(\frac{r}{c} + \ell + 1, k + 1 - \frac{r}{c}\right)\right], \end{aligned}$$

where $B_x(u, v)$ denotes the incomplete beta function taken at x: $B_x(u, v) = \int_0^x t^{u-1}(1-t)^{v-1} dt$ with $x \in [0, 1]$ and $u, v > 0$.

Proof: The proof is almost identical to the one of Proposition 2.3, we thus omit it. □

Following the spirit of Abushal *et al.* (2021), we can use the incomplete moments of X to define several inequality measures, and various residual life functions, as well as the related moments. We end this part with a generalization of the ordinary moments by investigating the weighted probability moments.

Proposition 2.5. *Let X be a random variable with the NEB distribution, and r and s be integers. Then, X admits an (r, s)-th probability weighted moment, i.e., $\mu'_{r,s} = E(X^r \bar{F}(X)^s)$, if and only if $r < ck$. In this case, $\mu'_{r,s}$ can be expressed as the following infinite sum expansion:*

$$\begin{aligned} \mu'_{r,s} = \; & \frac{k}{1+s} \sum_{\ell=0}^{+\infty} \binom{k(1+s)-1}{\ell} (a-1)^{\ell} \\ & \times \left[aB\left(\frac{r}{c} + \ell + 2, k(1+s) - \frac{r}{c}\right) + (2-a) B\left(\frac{r}{c} + \ell + 1, k(1+s) + 1 - \frac{r}{c}\right)\right]. \end{aligned}$$

Proof: First, let us notice that, for $x > 0$,

$$f(x)\bar{F}(x)^s = \frac{1}{1+s} ck(1+s)x^{c-1}(ax^c + 2 - a)\frac{(1+ax^c)^{k(1+s)-1}}{(1+x^c)^{2k(1+s)+1}} = \frac{1}{1+s} f_\circ(x),$$

where $f_\circ(x)$ denotes the pdf of the NEB distribution with parameters a, c and $k(1+s)$. Therefore, we have

$$\mu'_{r,s} = \int_{-\infty}^{+\infty} x^r f(x)\bar{F}(x)^s dx = \frac{1}{1+s}\int_0^{+\infty} x^r f_\circ(x) dx = \frac{1}{1+s}\mu'^{\circ}_r,$$

where $\mu_r'^{\circ}$ denotes the r-th ordinary moment of a random variable with the NEB distribution with parameters a, c and $k(1+s)$. Hence, the desired result follows from Proposition 2.3 with adjustment on the definition of the parameters. □

Probability-weighted moments can be considered as extended versions of the ordinary moments. Also, they appear in the theory of order statistics, and remain standard in several branches of statistics. On this topic, we may refer to Hosking (1989).

2.5. Stress-strength reliability coefficient

Let X and Y be two independent random variables following the NEB distributions with parameters a, c and k_1, and a, c and k_2, respectively. We are interested in the determination of the common stress-strength reliability coefficient defined by

$$R = P(Y < X). \tag{2.4}$$

This coefficient is a measure of reliability of a component with strength modeled by X, subject to a stress modeled by Y. Further details on this special coefficient can be found in Church and Harris (1970).

Proposition 2.6. *The coefficient R precised in (2.4) is*

$$R = \frac{k_2}{k_1 + k_2}.$$

Proof: Let $F_2(x)$ be the cdf of Y and $f_1(x)$ be the pdf of X. Then, based on (2.1) and (2.2), after a linear integral development, we get

$$\begin{aligned}
R &= \int_{-\infty}^{+\infty} F_2(x) f_1(x)dx = \int_0^{+\infty} \left[1 - \frac{(1+ax^c)^{k_2}}{(1+x^c)^{2k_2}}\right] \\
&\quad \times\ ck_1 x^{c-1}(ax^c + 2 - a)\frac{(1+ax^c)^{k_1-1}}{(1+x^c)^{2k_1+1}}dx \\
&= 1 - \int_0^{+\infty} ck_1 x^{c-1}(ax^c + 2 - a)\frac{(1+ax^c)^{k_1+k_2-1}}{(1+x^c)^{2(k_1+k_2)+1}}dx \\
&= 1 - \frac{k_1}{k_1+k_2}\int_0^{+\infty} c(k_1+k_2)x^{c-1}(ax^c + 2 - a)\frac{(1+ax^c)^{k_1+k_2-1}}{(1+x^c)^{2(k_1+k_2)+1}}dx.
\end{aligned}$$

Note that the last integral term is equal to one since it corresponds to the integral of a pdf over its whole support; it is the pdf of the NEB distribution with parameters a, c and $k_1 + k_2$. Hence $R = 1 - k_1/(k_1 + k_2) = k_2/(k_1 + k_2)$. This ends the proof. □

Thus, in the configuration of Proposition 2.6, R has a quite simple expression. It is decreasing with respect to k_1, whereas it is increasing with respect to k_2. If $k_1 = k_2$, we get $R = 1/2$ meaning that there is a equal chance of Y to be greater than X, and vice-versa.

The rest of the article is devoted to the inferences of the NEB model, beginning with the estimation of the model parameters through the maximum likelihood approach.

3. MAXIMUM LIKELIHOOD ESTIMATION

3.1. Estimation of the parameters

Let n be a positive parameter. Let us denote by $x_1, ..., x_n$ n independent observations from the NEB distribution. Then, the maximum likelihood method proposes to used the maximum likelihood estimates (MLEs) $\hat{a}$, $\hat{c}$ and $\hat{k}$ of a, c and k, respectively, defined by $(\hat{a}, \hat{c}, \hat{k}) = \mathrm{argmax}_{(a,c,k)\in[0,2]\times(0,+\infty)^2}\, \ell(a, c, k)$, where $\ell(a, c, k)$ denotes the log-likelihood function defined by

$$\begin{aligned}\ell(a, c, k) \;=\;& n\log c + n\log k + (c-1)\sum_{i=1}^{n}\log(x_i) + \sum_{i=1}^{n}\log(a x_i^c + 2 - a)\\ &+\; (k-1)\sum_{i=1}^{n}\log(1 + a x_i^c) - (2k+1)\sum_{i=1}^{n}\log(1 + x_i^c).\end{aligned}$$

The MLEs $\hat{a}$, $\hat{c}$ and $\hat{k}$ can be determined through the score equations. Now let $\hat{V}_a$, $\hat{V}_c$ and $\hat{V}_k$ defined by $(\hat{V}_a, \hat{V}_c, \hat{V}_k) = \mathrm{diag}\big[I(a, c, k)^{-1}\big]\,|_{\,a=\hat{a},c=\hat{c},k=\hat{k}}$, where

$$I(a, c, k) = \left(-\frac{\partial^2}{\partial u \partial v}\ell(a, c, k)\right)_{(u,v)=(a,c,k)^2}.$$

By applying a the well-known asymptotic property of the MLEs, as m and n tends to $+\infty$, the underlying distribution of

$$\left\{(1/\sqrt{\hat{V}_a})(\hat{a} - a), (1/\sqrt{\hat{V}_c})(\hat{c} - c), (1/\sqrt{\hat{V}_k})(\hat{k} - k)\right\}$$

can be approximated by the standard trivariate normal distribution. As an immediate consequence, a two-sided asymptotic $100(1-\alpha)\%$ confidence interval of a with $\alpha \in (0, 1)$ is given as $I_a = \left[\hat{a} - u_\alpha\sqrt{\hat{V}_a}, \hat{a} + u_\alpha\sqrt{\hat{V}_a}\right]$, where $u_\alpha = Q_U(1 - \alpha/2)$, $Q_U(x)$ denoting the qf of the standard univariate normal distribution. Analogous two-sided asymptotic $100(1-\alpha)\%$ confidence intervals for c and k can be presented in a similar way. The general theory and formulas of the maximum likelihood approach can be found in Casella and Berger (1990).

3.2. Estimation of R

We now focus on the estimation of the stress-strength reliability coefficient R as described in Subsection 2.5, recalling that $R = k_2/(k_1 + k_2)$. Such estimation problem is of interest in various applied studied, as motivated in Mokhlis (2005), Lio and Tsai (2012), Rao *et al.* (2015), Laslan and Nadar (2017), Byrnes *et al.* (2019), Maurya and Tripathi (2020) and Agiwal (2021). We follow the same methodology as the one employed in Agiwal (2021).

Let n and m be two positive integers. Let us denote by $x_1, ..., x_n$ n independent observations from the NEB distribution with parameters a, c and k_1, and $y_1, ..., y_m$ m independent observations from the NEB distribution with parameters a, c and k_2, assuming that a and c are known. Then, the log-likelihood function based on these two samples is given by

$$\begin{aligned}\ell(k_1,k_2) =\ & (n+m)\log c + n\log k_1 + (c-1)\sum_{i=1}^{n}\log(x_i) + \sum_{i=1}^{n}\log(ax_i^c+2-a)\\ &+ (k_1-1)\sum_{i=1}^{n}\log(1+ax_i^c) - (2k_1+1)\sum_{i=1}^{n}\log(1+x_i^c) + m\log k_2 + (c-1)\sum_{i=1}^{m}\log(y_i)\\ &+ \sum_{i=1}^{m}\log(ay_i^c+2-a) + (k_2-1)\sum_{i=1}^{m}\log(1+ay_i^c) - (2k_2+1)\sum_{i=1}^{m}\log(1+y_i^c).\end{aligned}$$

The MLEs $\hat{k}_1$ and k_2 of k_1 and k_2, respectively, are obtained as

$$(\hat{k}_1,\hat{k}_2) = \text{argmax}_{(k_1,k_2)\in(0,+\infty)^2}\,\ell(k_1,k_2).$$

Classically, they satisfy the score equations defined by $\partial\ell(k_1,k_2)/\partial k_1\,|_{\,k_1=\hat{k}_1,k_2=\hat{k}_2}=0$ and $\partial\ell(k_1,k_2)/\partial k_2\,|_{\,k_1=\hat{k}_1,k_2=\hat{k}_2}=0$, which give

$$\hat{k}_1 = \left\{-\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{1+ax_i^c}{(1+x_i^c)^2}\right)\right\}^{-1}, \quad \hat{k}_2 = \left\{-\frac{1}{m}\sum_{i=1}^{m}\log\left(\frac{1+ay_i^c}{(1+y_i^c)^2}\right)\right\}^{-1}.$$

Now, we have

$$\frac{\partial^2}{\partial k_1^2}\ell(k_1,k_2) = -\frac{n}{k_1^2}, \quad \frac{\partial^2}{\partial k_2^2}\ell(k_1,k_2) = -\frac{m}{k_2^2}, \quad \frac{\partial^2}{\partial k_1 k_2}\ell(k_1,k_2) = 0.$$

By applying a the well-known asymptotic property of the MLEs, as m and n tends to $+\infty$, the underlying distribution of $\left\{(\hat{k}_1/\sqrt{n})(\hat{k}_1-k_1), (\hat{k}_2/\sqrt{m})(\hat{k}_2-k_2)\right\}$ can be approximated by the standard bivariate normal distribution. On the other side, by substitution, a point estimate for R is obtained as

$$\hat{R} = \frac{\hat{k}_2}{\hat{k}_1+\hat{k}_2}. \tag{3.1}$$

By applying the multivariate delta method (see Klein, 1953), since the underlying random estimates of k_1 and k_2 are independent, an estimate for the variance of the underlying random estimate of R is inscribed as

$$\begin{aligned}\hat{V}_R &= \left(-\frac{\partial^2}{\partial k_1^2}\ell(k_1,k_2)\right)^{-1}\left(\frac{\partial}{\partial k_1}R\right)^2 + \left(-\frac{\partial^2}{\partial k_2^2}\ell(k_1,k_2)\right)^{-1}\left(\frac{\partial}{\partial k_2}R\right)^2\Bigg|_{k_1=\hat{k}_1,k_2=\hat{k}_2}\\ &= \frac{\hat{k}_1^2\hat{k}_2^2}{(\hat{k}_1+\hat{k}_2)^4}\left(\frac{1}{n}+\frac{1}{m}\right).\end{aligned}$$

Therefore, as m and n tends to $+\infty$, the underlying distribution of $(1/\sqrt{\hat{V}_R})(\hat{R}-R)$ can be approximated by the standard univariate normal distribution. As an immediate consequence, a two-sided asymptotic $100(1-\alpha)\%$ confidence interval of R is given as

$$I_R = \left[\hat{R} - u_\alpha\frac{\hat{k}_1\hat{k}_2}{(\hat{k}_1+\hat{k}_2)^2}\sqrt{\frac{1}{n}+\frac{1}{m}}, \hat{R} + u_\alpha\frac{\hat{k}_1\hat{k}_2}{(\hat{k}_1+\hat{k}_2)^2}\sqrt{\frac{1}{n}+\frac{1}{m}}\right].$$

The rest of the study focuses on the Bayesian inferences of the NEB model, with applications.

4. BAYESIAN INFERENCE

In the Bayesian framework, not only data but also prior information about the unknown parameter is used to analyze the data and draw conclusions. In this way, Bayesian inference incorporates the prior distribution of the model parameters with the likelihood function to produce the posterior distribution that gathers more quality inferences and controls the uncertainty. However, the choice of a suitable prior has a significant role in changing the result. If sufficient information is available about the parameter, then an informative prior is considered; otherwise, one can use a non-informative prior.

Here, we consider both informative and non-informative priors for the Bayesian analysis of the unknown model parameters and stress-strength reliability coefficient of the NEB distribution. Since the shape of the proposed distribution is skewed to the right, we use a gamma prior as a skewed distribution for the independent parameters k_1, k_2 and c, whereas a follows a uniformly distributed prior. Indeed, we know that the gamma distribution is very flexible and is used frequently everywhere. A slight change in the parameters is also observed, as are changes in the shape of the distributions. So, we consider this prior for the Bayesian computation in our manuscript. Because a is the scale parameter, it has little effect on the distribution's shape. As a result, we can easily consider the improper prior in place for uniform distribution. The description of the said priors can be summarized as follows: $\pi(k_1) = \text{Gamma}(r_1, s_1),\ r_1 > 0, s_1 > 0,\ \pi(k_2) = \text{Gamma}(r_2, s_2),\ r_2 > 0, s_2 > 0$, $\pi(c) = \text{Gamma}(r_3, s_3),\ r_3 > 0, s_3 > 0$ and $\pi(a) \propto 1,\ a \in [0, 2]$, where $\text{Gamma}(r, s)$ denotes the standard gamma distribution with 'shape parameter' r and 'scale parameter' s, and $(r_1, s_1, r_2, s_2, r_3, s_3)$ are called the hyper-parameters. One can notice that, if $r_1 = s_1 = r_2 = s_2 = r_3 = s_3 = 0$, the prior is reduced to a non-informative form of gamma prior. Consequently, the joint prior $\pi(\Theta = (c, a, k_1, k_2))$ is defined as follows:

$$\pi(\Theta) = \frac{s_1^{r_1} s_2^{r_2} s_3^{r_3}}{\Gamma(r_1)\Gamma(r_2)\Gamma(r_3)} k_1^{r_1-1} k_2^{r_2-1} c^{r_3-1} e^{-k_1 s_1 - k_2 s_2 - c s_3},$$

where $\Gamma(u)$ is the gamma function, i.e., $\Gamma(u) = \int_0^{+\infty} t^{u-1} e^{-t} dt$, $u > 0$.

The posterior distribution $\pi(\Theta|\mathbf{data})$ of the parametric space (Θ) is obtained by incorporation of likelihood function $(L(\Theta|\mathbf{data}))$ with the joint prior distribution $\pi(\Theta)$, that is

$$\begin{aligned}
\pi(\Theta|\mathbf{data}) &= K L(\Theta|\mathbf{data})\pi(\Theta) \\
&= K c^{n+m+r_3-1} k_1^{n+r_1-1} k_2^{m+r_2-1} \frac{s_1^{r_1} s_2^{r_2} s_3^{r_3} e^{-k_1 s_1 - k_2 s_2 - c s_3}}{\Gamma(r_1)\Gamma(r_2)\Gamma(r_3)} \\
&\times \prod_{i=1}^{n} \frac{x_i^{c-1}(a x_i^c + 2 - a)(1 + a x_i^c)^{k_1-1}}{(1 + x_i^c)^{2k_1+1}} \prod_{j=1}^{m} \frac{y_j^{c-1}(a y_j^c + 2 - a)(1 + a y_j^c)^{k_2-1}}{(1 + y_j^c)^{2k_2+1}},
\end{aligned}$$

where K is a constant such that $K^{-1} = \int L(\Theta|\mathbf{data})\pi(\Theta) d\Theta$.

Based on decision theory, it is a well known discussion that the best estimate decision depends on the pattern of the loss function adopted for a particular situation and the resulting outcome may be under or / and over estimation. If the amount of loss is equal in under and over estimation then the symmetric loss function is considered. On the other situations, the asymmetric loss function is useful when positive loss may be more serious than a given negative loss of the same magnitude or vice-versa. Here, we employ both asymmetric and symmetric loss functions to investigate the suitability of the loss functions for the model. More precisely, we use the squared error (symmetric) loss function (SELF) and entropy (asymmetric) loss function (ELF). The SELF and ELF are inscribed as $L_{SELF}(\theta,\hat{\theta}) = \left(\hat{\theta}-\theta\right)^2$ and $L_{ELF}(\theta,\hat{\theta}) = \hat{\theta}/\theta - \log\left(\hat{\theta}/\theta\right) - 1$, respectively. Under the SELF and ELF, for any parametric function, say $\phi(\Theta)$, the Bayes estimate is obtained as follows:

$$\phi^{\star}_{self}(\Theta|\mathbf{data}) = K\int \phi(\Theta)\pi(\Theta|\mathbf{data})d\Theta \tag{4.1}$$

and

$$\phi^{\star}_{elf}(\Theta|\mathbf{data}) = \left(K\int \phi^{-1}(\Theta)\pi(\Theta|\mathbf{data}\ )d\Theta\right)^{-1}, \tag{4.2}$$

respectively. To obtain (4.1) and (4.2), we get the Bayesian estimates of the model parameters as well as the stress-strength reliability coefficient, where the Bayes estimate under the SELF is the posterior mean and under the ELF is the inverse of the harmonic mean. Due to the presence of multiple integrations in equations (4.1) and (4.2), they are very difficult to solve in an exact manner. Therefore, an iterative numerical procedure is required to solve these equations. For this situation, the Markov Chain Monte Carlo (MCMC) technique is suggested to generate a sequence of random draws from posteriors of interest. Using the MCMC method, a stochastic chain is produced that contains a sequence of random samples. The Gibbs sampling and the Metropolis-Hastings (MH) algorithm are two approaches in MCMC to computing the posterior distribution. To implement these approaches, the full conditional posterior distribution is derived for the study parameters. By putting $\underline{\mathbf{x}} = (x_1,...,x_n)$ and $\underline{\mathbf{y}} = (y_1,...,y_n)$, they are given as follows:

$$\begin{aligned}
\pi_1(c|a,k_1,k_2,\underline{\mathbf{x}},\underline{\mathbf{y}}) &\propto c^{n+m+r_3-1}e^{-cs_3}\prod_{i=1}^{n}\frac{x_i^{c-1}(ax_i^c+2-a)(1+ax_i^c)^{k_1-1}}{(1+x_i^c)^{2k_1+1}} \\
&\quad\times\prod_{j=1}^{m}\frac{y_j^{c-1}(ay_j^c+2-a)(1+ay_j^c)^{k_2-1}}{(1+y_j^c)^{2k_2+1}},
\end{aligned}$$

$$\pi_2(a|c,k_1,k_2,\underline{\mathbf{x}},\underline{\mathbf{y}}) \propto \prod_{i=1}^{n}(ax_i^c+2-a)(1+ax_i^c)^{k_1-1}\prod_{j=1}^{m}(ay_j^c+2-a)(1+ay_j^c)^{k_2-1},$$

$$\pi_3(k_1|a,c,\underline{\mathbf{x}},\underline{\mathbf{y}}) \propto k_1^{n+r_1-1}e^{-k_1s_1}\prod_{i=1}^{n}\frac{(1+ax_i^c)^{k_1-1}}{(1+x_i^c)^{2k_1+1}}$$

and

$$\pi_4(k_2|a,c,\underline{\mathbf{x}},\underline{\mathbf{y}}) \propto k_2^{m+r_2-1}e^{-k_2s_2}\prod_{j=1}^{m}\frac{(1+ay_j^c)^{k_2-1}}{(1+y_j^c)^{2k_2+1}}.$$

Based on estimated values of the parameters $(\hat{k}_1,\hat{k}_2)$, the estimated value of stress-strength reliability coefficient is obtained. To evaluate the above conditional posterior distribution, the following steps are considered:

Step 1: Starting with an initial value vector $\Theta^0 = (c^0, a^0, k_1^0, k_2^0)$ and set $l = 1$.

Step 2: Generate the point vector $\Theta^p = (c^p, a^p, k_1^p, k_2^p)$ from the candidate proposal density $q(\Theta^p|\Theta^0)$ where $q(\Theta^p|\Theta^0)$ proposes a probability with a move Θ^p, having conditional probability density given Θ^0.

Step 3: Determine the Hastings-ratio using Θ^p and Θ^0 as specified by

$$\rho(\Theta^p|\Theta^0) = \frac{\pi_1(c^p|a^0, k_1^0, k_2^0, \mathbf{data})q(\Theta^p|\Theta^0)}{\pi_1(c^0|a^0, k_1^0, k_2^0, \mathbf{data})q(\Theta^0|\Theta^p)}.$$

Similarly for the remaining parameters, the Hastings-ratio is obtained.

Step 4: Take into account Θ^p with probability $\gamma \leq \min[1, \rho(\Theta^p|\Theta^0)]$, otherwise $\Theta = \Theta^0$ with rejection probability $1 - \gamma$, where γ is generated from the uniform $U(0,1)$ distribution.

Step 5: Repeat Steps 2-4, $K = 5000$ times and record the sequence of parameter observations. Next, we get the Bayes estimate under different loss functions.

5. SIMULATION STUDY

This section performs a simulation experiment to determine the effectiveness of the proposed method in the model parameters as well as the stress-strength reliability coefficient for the NEB distribution. For this, various sample sizes, along with different sets of parameter values, are considered for making better inferences. We take the following sample size combinations, namely, $(n, m) = \{(20, 20), (30, 50), (50, 30), (50, 50), (40, 60), (60, 40), (40, 40)\}$ and different sets of stress-strength reliability coefficient values, namely $(k_1, k_2) = \{(2, 1), (2, 2), (1, 2)\}$ so that the true reliability parameter values are small (0.33), moderate (0.50) and high (0.67), respectively. The remaining parameter values are $a = 2.5$ and $c = 1.5$. We evaluate the performance of the stress-strength reliability coefficient on the basis of simulated samples with diverse sample sizes and combinations using the R software. To this end, we simulate a random sample of different sizes from the NEB distribution. In this regard, we use the Newton steps to generate a sample of size n from the NEB(a, c, k) distribution by following the steps below:

Step 1: Set n, a, c and k.

Step 2: Set initial value x_0.

Step 3: Set $j = 1$.

Step 4: Generate a value u from the uniform $U(0,1)$ distribution.

Step 5: Update x_0 through the Newton formula for solving $F(x) = u$ such as $x_{new} = x_0 - \dfrac{F(x_0) - u}{f(x_0)}$, with the defined with the used parameters a, c and k.

Step 6: If $|x_0 - x_{new}| \leq \epsilon$ with $\epsilon > 0$ chosen as small, then x_{new} will be the desired value from $F(x)$.

Step 7: If $|x_0 - x_{new}| > \epsilon$, then, set $x_0 = x_{new}$ and go to Step 5.

Step 8: Repeat Steps 4-7, for $j = 1, 2, ..., n$ and obtained $x_1, x_2, ..., x_n$.

Using the generated samples, the maximum likelihood and Bayes estimates are obtained based on derived estimates of the parameters and reliability function. For the Bayes estimates, we use different loss functions under different priors and the hyper-parameters of the gamma prior are taken as follows:

1. When $r_1 = s_1 = r_2 = s_2 = r_3 = s_3 = 0$ (the non-informative prior case), the Bayes estimates are denoted as $SELF_0$ and ELF_0.
2. When the prior means are equal to the true value of parameters and the prior variances are equal to 1, the Bayes estimates are denoted as $SELF_1$ and ELF_1.

The results are based on 5000 replications. We vary the sample sizes with fixed values of the stress-strength reliability coefficient and for various combinations of the model parameters with fixed samples sizes. For different parameter values, different sample sizes and different priors under both SELF and ELF, we report the average estimates (AVs) and the corresponding mean squared errors (MSEs) of the MLEs and Bayes estimates of the model parameters and stress-strength reliability coefficient. The simulation results are postponed in Tables 1–5.

We deduce the following findings from the results:

1. The MSE of all estimates, obtained with different parameter values, decreases as the sample sizes increase.
2. For the distribution parameters and reliability function, the MSE based on the MLEs is higher as compared to the one of the Bayes estimates.
3. For gamma priors in comparison with informative and non-informative forms, the MSE of informative priors is smaller.
4. For reliability function and stress-strength reliability coefficients, the ELF performs better than the SELF in terms of the lesser value of the MSE.
5. For varying n and m, the MSE of k_1 is mostly greater than k_2 when $k_1 < k_2$ and $k_2 < k_1$.

Table 1: AVs and MSEs of the estimates of R with varying n and m.

(k_1,k_2)	R	(n,m)	MLE		$SELF_0$		ELF_0		$SELF_1$		ELF_1	
			AV	MSE	AV	MSE	AV	MSE	AV	MSE	AV	MSE
(2,1)	0.33	(20,20)	0.3919	0.0031	0.3656	0.0027	0.3772	0.0026	0.3559	0.0026	0.3711	0.0025
		(30,50)	0.3674	0.0024	0.3381	0.0016	0.3597	0.0019	0.3351	0.0014	0.3332	0.0014
		(50,30)	0.3670	0.0025	0.3568	0.0023	0.3592	0.0021	0.3427	0.0013	0.3378	0.0011
		(50,50)	0.3710	0.0022	0.3441	0.0016	0.3417	0.0015	0.3560	0.0015	0.3651	0.0018
		(40,60)	0.3677	0.0023	0.3458	0.0015	0.3616	0.0019	0.3436	0.0014	0.3407	0.0014
		(60,40)	0.3683	0.0019	0.3585	0.0016	0.3543	0.0014	0.3584	0.0014	0.3621	0.0015
		(40,40)	0.3727	0.0029	0.3439	0.0015	0.3399	0.0014	0.3550	0.0021	0.3654	0.0024
(2,2)	0.5	(20,20)	0.4715	0.0036	0.4789	0.0033	0.4795	0.0033	0.4843	0.0029	0.4859	0.003
		(30,50)	0.4783	0.0021	0.4827	0.0020	0.4851	0.0018	0.4954	0.0011	0.4893	0.0012
		(50,30)	0.4892	0.0031	0.4977	0.0030	0.4960	0.0029	0.5128	0.0020	0.5070	0.0019
		(50,50)	0.4827	0.0020	0.5042	0.0011	0.5041	0.0010	0.4848	0.0016	0.4800	0.0018
		(40,60)	0.5064	0.0021	0.4916	0.0019	0.4926	0.0018	0.5082	0.0017	0.5035	0.0017
		(60,40)	0.4886	0.0015	0.4951	0.0014	0.4940	0.0014	0.5014	0.0009	0.4965	0.0009
		(40,40)	0.4889	0.0036	0.5045	0.0016	0.4983	0.0016	0.4913	0.0026	0.4855	0.0028
(1,2)	0.67	(20,20)	0.6238	0.0029	0.6324	0.0027	0.6373	0.0026	0.6405	0.0029	0.6467	0.0027
		(30,50)	0.6428	0.0018	0.6434	0.0016	0.6412	0.0015	0.6568	0.0015	0.6637	0.0014
		(50,30)	0.6316	0.0028	0.6587	0.0026	0.6526	0.0022	0.6459	0.0025	0.6550	0.0021
		(50,50)	0.6416	0.0019	0.6536	0.0018	0.6507	0.0016	0.6548	0.0013	0.6602	0.0011
		(40,60)	0.6443	0.0016	0.6516	0.0015	0.6594	0.0015	0.6575	0.0015	0.6625	0.0013
		(60,40)	0.6419	0.0017	0.6547	0.0015	0.6502	0.0013	0.6552	0.0015	0.6624	0.0013
		(40,40)	0.6277	0.0028	0.6561	0.0022	0.6526	0.0020	0.6318	0.0025	0.6488	0.0020

Table 2: AVs and MSEs of the estimates of k_1 with varying n and m.

k_1	(n,m)	MLE AV	MLE MSE	$SELF_0$ AV	$SELF_0$ MSE	ELF_0 AV	ELF_0 MSE	$SELF_1$ AV	$SELF_1$ MSE	ELF_1 AV	ELF_1 MSE
1	(20,20)	1.3978	0.3245	1.3142	0.2019	1.2756	0.1689	1.2314	0.1632	1.2249	0.1552
	(30,50)	1.3297	0.2966	1.2527	0.1728	1.2239	0.1588	1.1914	0.1527	1.1977	0.1290
	(50,30)	1.3540	0.2815	1.2803	0.1997	1.2533	0.1354	1.1591	0.1617	1.1445	0.1403
	(50,50)	1.3446	0.2448	1.2695	0.1590	1.2397	0.1370	1.1554	0.1400	1.1465	0.1498
	(40,60)	1.3789	0.2644	1.2936	0.1723	1.2681	0.1488	1.1498	0.1577	1.1430	0.1211
	(60,40)	1.3394	0.2394	1.2499	0.1634	1.2831	0.1584	1.1549	0.1329	1.1459	0.1238
	(40,40)	1.3080	0.2467	1.2877	0.1541	1.2020	0.1276	1.1594	0.1386	1.1485	0.1265
2	(20,20)	2.4732	0.312	2.2916	0.2235	2.2478	0.1481	2.2655	0.1567	2.2213	0.1235
	(30,50)	2.3296	0.3117	2.2513	0.2102	2.1981	0.1522	2.1634	0.1250	2.0348	0.0483
	(50,30)	2.4364	0.2752	2.1203	0.1041	1.9262	0.1312	2.1488	0.0906	2.1041	0.0378
	(50,50)	2.4488	0.2502	2.2601	0.2098	2.2046	0.1383	2.2330	0.1419	2.0907	0.1185
	(40,60)	2.2553	0.2214	2.2078	0.1740	2.1927	0.1556	2.1828	0.1167	2.1766	0.0825
	(60,40)	2.4525	0.2404	2.1537	0.0815	2.1444	0.1025	2.1106	0.0694	2.0210	0.0758
	(40,40)	2.4793	0.2340	2.2378	0.1422	2.2142	0.1149	2.1208	0.1073	1.9659	0.1097

Table 3: AVs and MSEs of the estimates of k_2 with varying n and m.

k_2	(n,m)	MLE AV	MLE MSE	$SELF_0$ AV	$SELF_0$ MSE	ELF_0 AV	ELF_0 MSE	$SELF_1$ AV	$SELF_1$ MSE	ELF_1 AV	ELF_1 MSE
1	(20,20)	1.4710	0.2121	1.3760	0.1867	1.3076	0.1203	1.3129	0.1325	1.2739	0.1091
	(30,50)	1.4542	0.1936	1.3164	0.1181	1.2064	0.0564	1.2502	0.0761	1.1671	0.0459
	(50,30)	1.4152	0.1874	1.3290	0.1259	1.1942	0.0524	1.2036	0.0604	1.0937	0.0340
	(50,50)	1.4042	0.1855	1.3683	0.1581	1.2797	0.0952	1.3013	0.1046	1.2181	0.0665
	(40,60)	1.4653	0.2018	1.2885	0.1717	1.2183	0.1044	1.3210	0.1129	1.2558	0.0763
	(60,40)	1.4996	0.2322	1.3121	0.2061	1.2150	0.1168	1.2499	0.0843	1.1669	0.0532
	(40,40)	1.4183	0.1778	1.2856	0.1583	1.1743	0.0628	1.2454	0.0865	1.1539	0.0533
2	(20,20)	2.5262	0.3023	2.3506	0.2678	2.3008	0.2022	2.3265	0.2231	2.2753	0.1347
	(30,50)	2.3960	0.2065	2.3144	0.1845	2.2561	0.1528	2.2813	0.1113	2.1318	0.0483
	(50,30)	2.4262	0.2112	2.3415	0.2393	2.2966	0.1500	2.3161	0.1927	2.2163	0.1116
	(50,50)	2.4151	0.2259	2.3351	0.1966	2.2297	0.1231	2.3016	0.1286	2.1652	0.0655
	(40,60)	2.4631	0.3068	2.3378	0.2370	2.3048	0.1917	2.3079	0.2030	2.2521	0.1085
	(60,40)	2.4371	0.2929	2.2694	0.1812	2.2513	0.1236	2.2637	0.1621	2.2191	0.0777
	(40,40)	2.4664	0.3095	2.2665	0.1609	2.2678	0.1553	2.2918	0.1370	2.1326	0.0678

Table 4: AVs and MSEs of the estimates of c with varying n and m.

c	(n,m)	MLE AV	MLE MSE	$SELF_0$ AV	$SELF_0$ MSE	ELF_0 AV	ELF_0 MSE	$SELF_1$ AV	$SELF_1$ MSE	ELF_1 AV	ELF_1 MSE
2.5	(20,20)	2.5632	0.1342	2.4183	0.1268	2.4678	0.1176	2.5105	0.1192	2.4747	0.1081
	(30,50)	2.5316	0.1100	2.4638	0.0992	2.4925	0.0821	2.4755	0.0955	2.4908	0.0848
	(50,30)	2.5489	0.1006	2.4300	0.1235	2.4681	0.0980	2.4732	0.0944	2.4830	0.0843
	(50,50)	2.5660	0.1182	2.4469	0.1140	2.4548	0.1110	2.4781	0.1060	2.4843	0.0961
	(40,60)	2.5683	0.1052	2.4835	0.0905	2.4977	0.0881	2.4930	0.0895	2.4971	0.0894
	(60,40)	2.5479	0.0984	2.4797	0.0997	2.5170	0.9567	2.4908	0.0894	2.4809	0.0803
	(40,40)	2.5647	0.1067	2.5247	0.0990	2.5625	0.1135	2.4862	0.0973	2.4858	0.0865

Table 5: AVs and MSEs of the estimates of a with varying n and m.

a	(n,m)	MLE AV	MLE MSE	$SELF_0$ AV	$SELF_0$ MSE	ELF_0 AV	ELF_0 MSE	$SELF_1$ AV	$SELF_1$ MSE	ELF_1 AV	ELF_1 MSE
1.5	(20,20)	1.5378	0.0231	1.5442	0.0156	1.5256	0.0162	1.5114	0.0148	1.5249	0.0151
	(30,50)	1.4814	0.0189	1.5063	0.0102	1.5108	0.0115	1.4938	0.0103	1.4970	0.0100
	(50,30)	1.4918	0.0171	1.4958	0.0104	1.5009	0.0108	1.4952	0.0101	1.4947	0.0126
	(50,50)	1.4894	0.0184	1.4937	0.0116	1.5205	0.0127	1.5221	0.0113	1.4845	0.0120
	(40,60)	1.5205	0.0173	1.4836	0.0116	1.5386	0.0124	1.5028	0.0102	1.5155	0.0102
	(60,40)	1.5075	0.0178	1.5340	0.0128	1.4709	0.0147	1.4918	0.0106	1.4979	0.0106
	(40,40)	1.4955	0.0188	1.5366	0.0120	1.4865	0.0123	1.4996	0.0108	1.4987	0.0108

6. APPLICATION

In this section, we work with two engineering data sets, initially reported in Caramanis *et al.* (1983) and Mazumdar and Gaver (1984), to demonstrate that the proposed methodologies can be used in practice quite effectively. These data sets represent two different algorithms, called SC16 and P3, used by the electric utility industry to compare and estimate unit capacity factors. More precisely, SC16 represents the Southern Company's program using a piecewise linear representation of equivalent charging duration (ELDC) curves in 16 megawatt increments to represent the original charging duration curve. On his side, P3 represents the ELDC using the Gram-Charlier series involving all cumulative power in megawatts. The data sets considered are detailed as follows:

SC16(X), $n = 23$: 0.853, 0.759, 0.866, 0.809, 0.717, 0.544, 0.492, 0.403, 0.344, 0.213, 0.116, 0.116, 0.092, 0.070, 0.059, 0.048, 0.036, 0.029, 0.021, 0.014, 0.011, 0.008, 0.006.

P3(Y), $m = 22$: 0.853, 0.759, 0.874, 0.800, 0.716, 0.557, 0.503, 0.399, 0.334, 0.207, 0.118, 0.118, 0.097, 0.078, 0.067, 0.056, 0.044, 0.036, 0.026, 0.019, 0.014, 0.010.

We remove the value 0.000 from the P3 algorithm so that it does not make the parameter likelihood estimates meaningless. First, we check the validity of the proposed distribution using the negative log-likelihood (-logL), Kolmogorov-Smirnov (K-S) statistic, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). We compare the fits of the NEB distribution with those of the Topp-Leone BXII (TLBXII) distribution, Marshall-Olkin Extended BXII (MOEBXII) distribution, Weibull BXII (WBXII) distribution, and transmuted BXII (TBXII) distribution, as referenced in the introductory part. The expressions of the pdfs of the competitor distributions are briefly presented below:

$$\text{TLBXII}: \ f(x; l, c, k) = 2lckx^{c-1}(1+x^c)^{-(2k+1)}\left[1-(1+x^c)^{-2k}\right]^{l-1},$$

$$\text{MOEBXII}: \ f(x; a, c, k) = ack\frac{x^{c-1}(1+x^c)^{-(k+1)}}{[1-(1-a)(1+x^c)^{-k}]^2},$$

$$\text{WBXII}: \ f(x; a, l, c, k) = \frac{alckx^{c-1}}{1+x^c}\{k\log(1+x^c)\}^{l-1}\exp\left[-a\{k\log(1+x^c)\}^l\right],$$

$$\text{TBXII}: \ f(x; c, k, \theta, \lambda) = \frac{ck}{\theta^c}x^{c-1}\left[1+\left(\frac{x}{\theta}\right)^c\right]^{-(k+1)}\left[1-\lambda+2\lambda\left\{1+\left(\frac{x}{\theta}\right)^c\right\}^{-k}\right],$$

All the involved parameters are supposed to be strictly positive, except $\lambda \in [-1, 1]$ for the last distribution. It is supposed that $x > 0$, the standard completion applied on these pdfs for $x \leq 0$. We use the maximum likelihood estimation and the K-S test to fit the two data sets separately for the proposed and the above competitor distributions. We discover that the NEB distribution provides a better fit. We also use both information criteria to find the best model in two data sets that have a good fit based on the minimum values of AIC and BIC, and conclude that the NEB distribution fits both data sets better than the others distributions. The values of MLE, K-S test, AIC, and BIC are collected in Tables 6 and 7.

Table 6: MLEs, AIC, BIC and KS statistic for the SC16 data.

Model	MLEs	-logL	K-S	AIC	BIC
NEB(a, c, k)	$(1.3907, 0.7741, 5.3351)$	-7.0094	0.9400	-8.0187	-4.6123
TLBXII(l, c, k)	$(234.8565, 0.1510, 5.4960)$	-6.1992	0.9739	-6.3985	-2.9920
MOEBXII(a, c, k)	$(0.8894, 0.8332, 3.7239)$	-6.9174	0.9433	-7.8348	-4.4283
WBXII(a, l, c, k)	$(1.1192, 13.1646, 0.0756, 1.582)$	-7.4218	0.9500	-6.8437	-2.3017
TBXII(c, k, θ, λ)	$(0.7533, 337.0553, 573.6415, 0.0924)$	-7.4215	1.0206	-6.8308	-2.2888

Table 7: MLEs, AIC, BIC and KS statistic for the P3 data.

Model	MLEs	-logL	K-S	AIC	BIC
NEB(a, c, k)	$(1.3630, 0.8704, 5.2907)$	-4.4913	0.9407	-2.9825	0.2906
TLBXII(l, c, k)	$(255.2371, 0.1688, 5.5850)$	-3.9032	0.9767	-1.8063	1.4668
MOEBXII(a, c, k)	$(0.7887, 0.9515, 3.6233)$	-4.4305	0.9448	-2.8611	0.4120
WBXII(a, l, c, k)	$(4.4466, 9.8049, 0.1123, 1.3798)$	-4.8735	0.9413	-1.7470	2.6171
TBXII(c, k, θ, λ)	$(0.8467, 474.7816, 428.0671, 0.1096)$	-4.8649	1.0368	-1.7298	2.6344

From Tables 6 and 7, we can note that the parameter a is estimated in an intermediate way between 0 and 2, justifying the alternative identity of the distribution NEB compared to the BXII and PITL distributions.

For both data sets, the MLEs and Bayes estimates of the model parameters are given along with their standard errors (SEs), and the stress-strength reliability coefficient values are obtained in Table 8. As we had no prior information apart from a few observations, we only use non-informative values for the gamma prior.

Table 8: Maximum likelihood and Bayes estimates of R and distribution parameters with SEs based on the considered data-sets.

Estimates	R	a		c		k_1		k_2	
		AV	SE	AV	SE	AV	SE	AV	SE
MLE	0.4861	1.1371	0.3581	0.8254	0.1494	4.1941	0.8725	3.9586	0.9018
$SELF_0$	0.4878	1.0790	0.2190	0.8488	0.1502	4.5912	0.7115	4.3966	0.1093
ELF_0	0.4747	0.9715	0.3127	0.8448	0.1684	3.9384	0.8198	3.7017	0.1121

Based on Table 8, an estimate of R is approximately obtained as 0.48. We conclude that the P3 algorithm has slightly more storage capacity for the electric utility industry.

7. CONCLUSION

This article emphasized a new three-parameter heavy right-tailed distribution that consolidates, in a certain sense, the "popular Burr type XII distribution" and the "promising power inverted Topp-Leone distribution". The slip between these two well-established distributions was made by a special shift parameter. The new distribution benefits from notable advantages, including a flexible decreasing and unimodal probability density function, a decreasing upside-down bathtub-shaped hazard rate function, as well as a manageable quantile function, (first-order) stochastic ordering properties, moments, incomplete moments, and probability weighted moments. The classical and Bayesian approaches were developed to estimate the model and stress-strength reliability parameters. The effectiveness and potential of the new model were highlighted using both simulated and actual data, demonstrating that it can be a superior replacement for other lifetime models in the literature.

ACKNOWLEDGMENTS

We thank the two referees for the constructive comments on the paper.

REFERENCES

Aarset, M. V. (1987). How to identify bathtub hazard rate. *IEEE Transactions on Reliability*, 36:106–108.

Abushal, T. A., Hassan, A. S., El-Saeed, A. R., and Nassr, S. G. (2021). Power inverted Topp-Leone distribution in acceptance sampling plans. *Computers, Materials & Continua*, 67(1):991–1011.

Afify, A. Z., Cordeiro, G. M., Bourguignon, M., and Ortega, E. M. M. (2018). Properties of the transmuted Burr XII distribution, regression and its applications. *Journal of Data Science*, 16(3):485–510.

Afify, A. Z., Cordeiro, G. M., Ortega, E. M., Yousof, H. M., and Butt, N. S. (2018). The four-parameter Burr XII distribution: Properties, regression model, and applications. *Communications in Statistics – Theory and Methods*, 47(11):2605–2624.

Agiwal, V. (2021). Bayesian estimation of stress strength reliability from inverse Chen distribution with application on failure time data. *Annals of Data Science*, `https://doi.org/10.1007/s40745-020-00313-w`.

Al-Saiari, A. Y., Baharith, L. A., and Mousa, S. A. (2014). Marshall-Olkin extended Burr type XII distribution. *International Journal of Statistics and Probability*, 3:78–84.

Bhaumik, D. K., Kapur, K., and Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples. *Technometrics*, 51(3):326–334.

Burr, I. W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, 13:215–232.

Byrnes, J. M., Lin, Y. J., Tsai, T. R., and Lio, Y. (2019). Bayesian inference of $\delta = P(X < Y)$ for Burr type XII distribution based on progressively first failure-censored samples. *Mathematics*, 7(9):794.

Caramanis, M., Stremmel, J. V., Fleck, W., and Daniel, S. (1983). Probabilistic production costing. *International Journal of Electrical Power & Energy Systems*, 5:75–86.

Casella, G. and Berger, R. L. (1990). *Statistical Inference*. Brooks/Cole Publishing Company: Bel Air, CA, USA.

Church, J. D. and Harris, B. (1970). The estimation of reliability from stress-strength relationships. *Technometrics*, 12(1):49–54.

Elbatal, I., Altun, E., Afify, A. Z., and Ozel, G. (2019). The generalized Burr XII power series distributions with properties and applications. *Annals of Data Science*, 6(3):571–597.

Gilchrist, W. (2000). *Statistical Modelling with Quantile Functions*. CRC Press, Abingdon.

Gomes, A. E., da-Silva, C. Q. and Cordeiro, G. M. (2015). Two extended Burr models: Theory and practice. *Communication in Statistics Theory – Methods*, 44(8):1706–1734.

Hassan, A. S., Elgarhy, M., and Ragab, R. (2020). Statistical properties and estimation of inverted Topp-Leone distribution. *Journal of Statistics Applications & Probability*, 9(2):319–331.

Hinkley, D. (1977). On quick choice of power transformation. *Applied Statistics*, 26(1):67–69.

Hosking, J. R. (1989). *The Theory of Probability Weighted Moments*. Technical Report RC-12210, IBM Thomas J. Watson Research Center, Yorktown Heights, NY.

Jamal, F., Chesneau, C., Nasir, M. A., Saboor, A., Altun, E., and Khan, M. A. (2020). On a modified Burr XII distribution having flexible hazard rate shapes. *Mathematica Slovaca*, 70(1):193–212.

Klein, L. R. (1953). *A Textbook of Econometrics*. Prentice Hall, New York, NY, USA.

Laslan, F. K. I. and Nadar, M. (2017). Statistical inference of $P(X < Y)$ for the Burr type XII distribution based on records. *Hacettepe Journal of Mathematics and Statistics*, 46(4):713–742.

Lio, Y. L. and Tsai, T. R. (2012). Estimation of $\delta = P(X < Y)$ for Burr XII distribution based on the progressively first failure-censored samples. *Journal of Applied Statistics*, 39(2):309–322.

Maurya, R. K. and Tripathi, Y. M. (2020). Reliability estimation in a multicomponent stress-strength model for Burr XII distribution under progressive censoring. *Brazilian Journal of Probability and Statistics*, 34(2):345–369.

Mazumdar, M. and Gaver, D. P. (1984). On the computation of power-generating system reliability indexes. *Technometrics*, 26:173–185.

Mead, M. E. (2014). The beta exponentiated Burr XII distribution. *Journal of Statistics: Advances in Theory and Applications*, 12:53–73.

Mead, M. E. and Afify, A. Z. (2017). On five parameter Burr XII distribution: properties and applications. *South African Statistical Journal*, 51:67–80.

Mokhlis, N. A. (2005). Reliability of a stress-strength model with Burr type III distributions. *Communications in Statistics – Theory and Methods*, 34(7):1643–1657.

Moore, D. and Papadopoulos, A. S. (2000). The Burr type XII distribution as a failure model under various loss functions. *Microelectronics Reliability*, 40(12):2117–2122.

Mousa, M. A. and Jaheen, Z. F. (2002). Bayesian prediction for progressively censored data from the Burr model. *Statistical Papers*, 43(4):587–593.

Nasir, M. A., Tahir, M. H., Jamal, F., and Ozel, G. (2017). A new generalized Burr family of distributions for the lifetime data. *Journal of Statistics Applications and Probability*, 6(2):401–417.

Paranaiba, P. F., Ortega, E. M. M., Cordeiro, G. M., and de Pascoa, M. (2013). The Kumaraswamy Burr XII distribution: theory and practice. *Journal of Statistical Computation and Simulation*, 83(11):2117–2143.

Paranaiba, P. F., Ortega, E. M. M., Cordeiro, G. M., and Pescim, R. R. (2011). The beta Burr XII distribution with application to lifetime data. *Computation Statistics and Data Analysis*, 55(2):1118–1136.

Rao, G. S., Aslam, M., and Kundu, D. (2015). Burr-XII distribution parametric estimation and estimation of reliability of multicomponent stress-strength. *Communications in Statistics – Theory and Methods*, 44(23):4953–4961.

Reyad, H. M. and Othman, S. A. (2017). The Topp-Leone Burr-XII distribution: Properties and applications. *Journal of Advances in Mathematics and Computer Science*, 21(5):1–15.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Wiley, NY, USA.

Wang, F. K., Keats, J. B., and Zimmer, W. J. (1996). Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data. *Microelectronics Reliability*, 36(3):359–362.

Wu, S. J., Chen, Y. J., and Chang, C. T. (2007). Statistical inference based on progressively censored samples with random removals from the Burr type XII distribution. *Journal of Statistical Computation and Simulation*, 77(1):19–27.

Zimmer, W. J., Keats, J. B., and Wang, F. K. (1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, 30(4):386–394.

REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 139–160
https://doi.org/10.57805/revstat.v23i1.469

Stochastic Generator of a New Family of Lifetime Distributions with Illustration

Authors: Amjad D. Al-Nasser
– Department of Statistics, Faculty of Science, Yarmouk University, Jordan
amjadyu@yahoo.com

Ahmad A. Hanandeh ✉
– Department of Mathematics, Faculty of Science, Islamic University of Madinah, Saudi Arabia
hanandeh@iu.edu.sa

Received: May 2022 Revised: March 2023 Accepted: March 2023

Abstract:

- In this article, a new family of lifetime distributions, referred to as exponential transformation (ET) family, is introduced to help researchers model different types of data sets. Furthermore, the maximum likelihood method was used in estimating the model's parameters. Some structural properties of the distribution have been derived and studied. These are density function, distribution function, and reliability function, hazard rate function, moments, moment generating function, entropies, order statistics, Bonferroni and Lorenz curves. Simulation method was used to investigate the behaviors of the parameters of the proposed distribution; the results showed that the mean square error and standard error for the chosen parameter values decrease as the sample size increases. The proposed distribution was tested on real-life data, the results showed that the ET-exponential distribution performed better than other well-known distributions in modeling data. The results also showed that the distribution can be used as an alternative model in modeling lifetime processes.

Keywords:

- *lifetime distribution; maximum likelihood; means square error; parameters; quantile function.*

AMS Subject Classification:

- 60E05, 62F10.

✉ Corresponding author

1. INTRODUCTION

With increasing diversity of real-life problems and applications that includes complicated phenomena, there is a growing interest by researchers in developing new lifetime distributions to overcome complicated models. Consequently, significant progress has been made towards constructing numerous classes of new distributions to generate more flexible distributions instead of the classical ones to provide more accurate data modeling. Kumaraswamy (1980) was the first who suggested proposing new distribution by taking baseline distribution, and then Azzalini (1985) introduced a system for generating new distributions by adding a skewing parameter to a symmetric distribution.

The ideas of generating a new class of distributions graduated and can be classified into five schemes; the first one is by using differential equations, the second one is by generating a weighted form of the baseline distribution. The third one concerns adding additional parameter(s) to the baseline distribution. The fourth scheme is to discretize the continuous density function. The last scheme is a distribution transformation scheme that modifies a probability distribution function by forming a stochastic representation of baseline distribution such that the new relationship must satisfy the distribution theory assumptions.

In this article, we are interested in the last scheme to regenerate a new class of lifetime distribution. In the literature, there are several generators proposed based on different mathematical functional relationships. For instance, Mudholkar and Srivastava (1993) defined the exponentiated class of distributions by exponentiating a given baseline distribution with a positive parameter. Marshall and Olkin (1997), applied the transformation scheme to the survival function by adding an additional shape parameter to the transformation scheme. Eugene *et al.* (2002), used beta as a generator to develop Beta-G class of distributions. Zografos and Balakrishnan (2009) suggested the gamma-G class of distributions. Transmuted family of distributions was developed by Shaw and Buckley (2007); then later Shaw and Buckley (2009) proposed the quadratic rank transmutation map, while Cordeiro and de Castro (2011) proposed the Kumaraswamy-generated family.

Recently, Alzaatreh *et al.* (2013) defined and studied a new T-X family. Logarithmic transformation was proposed by Maurya *et al.* (2016) and an extension of this generator was proposed by Aslam *et al.* (2020). Moreover, many used trigonometry functions to provide distribution generators, for instance, Kumar *et al.* (2015) used Sine function to develop a new class of distributions while modification of this scheme by using Cosine-Sine (CS) transformation proposed by Chesneau *et al.* (2019). Vast modification has been made by many authors to identify a new generator of family of distributions (Kumar *et al.*, 2021; Goldoust *et al.*, 2019; Zaidi *et al.*, 2021; Altun *et al.*, 2021; Zichuan *et al.*, 2020).

In similar fashion, this article will propose a new distribution generator based on the exponential function to provide new class of parameter lifetime distribution. Let Y be a non-negative continuous random variable with baseline cumulative distribution function (CDF) $F(y,\theta)$ and probability distribution function (PDF) $f(y,\theta)$; where $\theta \in (0,\infty)$ is real valued represents the distribution parameter, then the stochastic presentation of the proposed CDF for generating a new class of distributions can be defined as:

$$G(y,\theta,\alpha) = F(y,\theta)e^{-\alpha\,\overline{F}(y,\theta)} \quad y > 0;\ \alpha \geq 0 \tag{1.1}$$

where $\overline{F}(y,\theta) = 1 - F(y,\theta)$. Noting that when $\alpha = 0$, then the proposed distribution is exactly the same as the baseline distribution.

This family will be called as **exponential transformation (ET)**, *i.e.*, $ET(y,\theta,\alpha)$. Now, the PDF of ET family can be obtaining by finding the first derivative of equation (1.1):

$$g(y,\theta,\alpha) = f(y,\theta)e^{-\alpha\ \overline{F}(y,\theta)}\left(1 + \ \alpha\ F(y,\theta)\right)\ ;\quad y > 0; \alpha \geq 0. \tag{1.2}$$

This family can be joined to T-X family by (Alzaatreh *et al.*, 2013) as follows: For a general baseline CDF of a continuous probability distribution denoted by $F(y,\theta)$, a new CDF having the form

$$G(y,\theta,\alpha) = \int_0^{F(y,\theta)} r(t)dt$$

where $r(t)$ is a PDF defined over $(0,1)$, and $R(t)$ is the associated CDF. Accordingly, the PDF $g(y,\theta,\alpha)$ can be obtained as

$$g(y,\theta,\alpha) = f(y,\theta)\ r(F(y,\theta)).$$

If we used $r(t)$ in the following functional form:

$$r(t) = (1+\alpha t)\ e^{-\alpha(1-t)}.$$

Then, the proposed family could be considered as member of the T-X family. As an illustration of the proposed family, the exponential distribution will be considered as baseline distribution. In this article, the third scheme will be used to generate new class of lifetime distribution.

The remainder of this article proceeds as follows. Section 2 provides some characterizations of the ET family, including shapes of CDF, PDF, reliability measures such as survival and hazard rate. Section 3 is dedicated to the mathematical properties of the ET family such as moments, quantiles, order statistics and entropies. In Section 4, the estimation of parameters is studied. Section 5 offers detailed simulation experiments on model performance and assessment. Section 6 is devoted to studying illustrative examples based on real data. Finally, Section 7 concludes the manuscript with a summary and an eye toward future work to close the paper.

2. CHARACTERIZATIONS

2.1. Asymptotic properties of the CDF and PDF

Suppose that X is a continuous random variable of ET family as given in (1.1), it can be easily seen that this family of distribution satisfies the Kolmogorov axioms of the distribution functions. For instance, it is easily seen that the limit property of $G(x,\theta,\ \alpha)$ satisfy the property of CDF:

$$\lim_{x\to\infty} G(x,\theta,\alpha) = \lim_{x\to\infty} F(x,\theta)e^{-\alpha\ \overline{F}(x,\theta)} = 1$$

and

$$\lim_{x\to 0} G(x,\theta,\alpha) = \lim_{x\to 0} F(x,\theta)e^{-\alpha\ \overline{F}(x,\theta)} = 0.$$

Hence, the total probability is equal to one. Also, it is monotone right increasing function of x, and $0 \leq G(x,\theta,\alpha) \leq 1;\ \forall\ x$. Therefore, $G(x,\theta,\alpha)$ is an absolute continuous distribution function.

Similarly, it is easily can be noted that $g(x,\theta,\alpha)$ is non-negative real valued PDF for all x. For instance in the exponential case:

$$\lim_{x\to\infty} g(x,\theta,\alpha) = \ 0$$

and

$$\lim_{x\to 0} g(x,\theta,\alpha) = \ \theta\ e^{-\alpha}.$$

Since both parameters are positive this indicates that $g(x,\theta,\alpha)$ is a unimodal distribution. Now, the functional form given in (1.2) satisfied the PDF property:

$$\int_0^\infty g(x,\theta,\alpha)dx = \int_0^\infty f(x,\theta)e^{-\alpha\ \overline{F}(x,\theta)}\ \ (1+\alpha\ F(x,\theta))\ dx.$$

To illustrate the usefulness of the new stochastic representation given in (1.1) and the associated PDF given in (1.2), suppose that the baseline distribution is the exponential distribution with mean $\frac{1}{\theta}$, then we have ET-Exp distribution.

Corollary 2.1. *Suppose that X is a random variable of ET-Exp, then the CDF and PDF of X are given in equations (2.1) and (2.2), respectively:*

$$G(x,\theta,\alpha) = e^{-\alpha e^{-\theta x}}\left(1-e^{-\theta x}\right);\quad x>0;\theta>0,\alpha\geq 0, \tag{2.1}$$

$$g(x,\theta,\alpha) = \theta e^{-\left(\theta x+\alpha e^{-\theta x}\right)}\left(1+\alpha(1-e^{-\theta x})\right);\quad x>0;\theta>0,\alpha\geq 0. \tag{2.2}$$

Figure 1 give a good representation of the new distribution PDF with selected set of parameters, in different cases by assuming both parameters are larger than 1, less than 1 or one of them is less than one and the other parameter is more than one.

2.2. The CDF and PDF expansion

The following result proposes Taylor series expansions of the exponential function that given in ET family. Accordingly, using the exponential series, we get:

$$e^{-x} = \sum_{j=0}^{\infty}(-1)^j\ \frac{x^j}{j!}.$$

Then the CDF can be written as, respectively:

$$G(x,\theta,\alpha) = F(x,\theta)\ \sum_{j=0}^{\infty}(-1)^j\frac{\alpha^j}{j!}\ \overline{F}(x,\theta)^j,$$

$$G(x,\theta,\alpha) = F(x,\theta)\ \sum_{j=0}^{\infty}(-1)^j\frac{\alpha^j}{j!}\ \{1-F(x,\theta)\}^j.$$

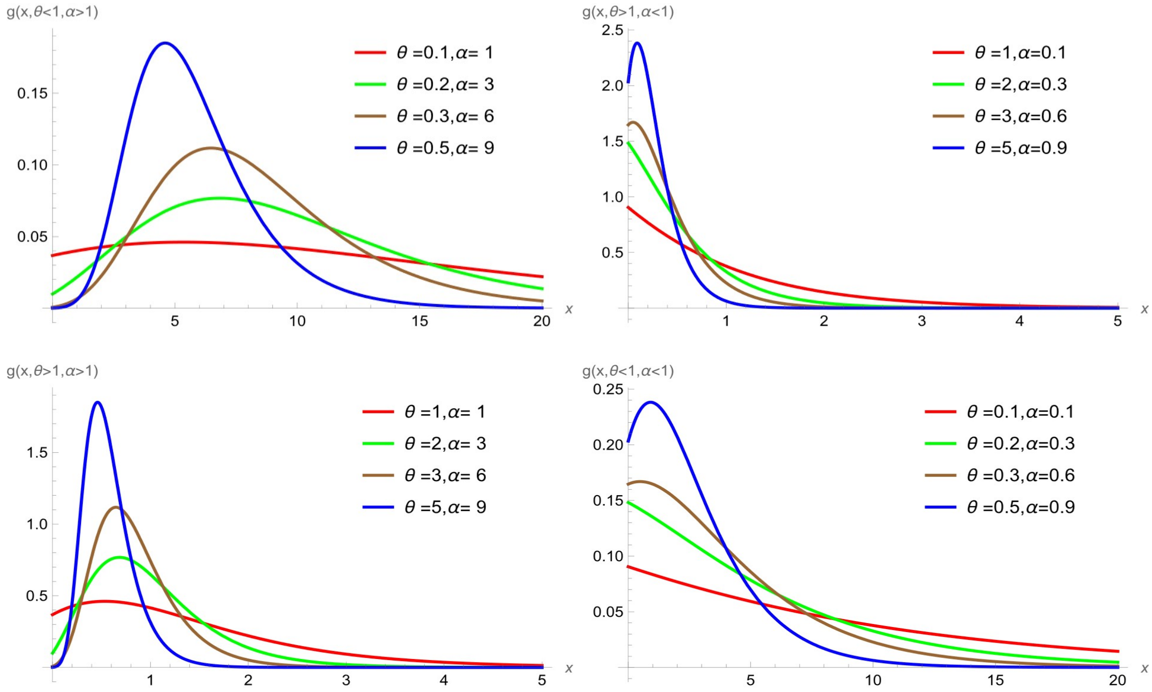


Figure 1: PDF of ET-Exp with selected parameter values.

Now, using the power series $(1-z)^m = \sum_{k=1}^{\infty} (-1)^k \binom{m}{k} z^k$, the expansion yields to:

$$G(x,\theta,\alpha) = F(x,\theta) \sum_{j=0}^{\infty} (-1)^j \frac{\alpha^j}{j!} \sum_{k=1}^{\infty} (-1)^k \binom{j}{k} F(x,\theta)^k$$

which can be simplified to

$$G(x,\theta,\alpha) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} (-1)^{j+k} \frac{\alpha^j}{j!} \binom{j}{k} F(x,\theta)^{k+1}. \tag{2.3}$$

In similar fashion, the PDF can be expanded as follows:

$$g(x,\theta,\alpha) = f(x,\theta)(1+\alpha\ F(x,\theta)) \left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} (-1)^{j+k} \frac{\alpha^j}{j!} \binom{j}{k} F(x,\theta)^k \right). \tag{2.4}$$

Based on (2.3) and (2.4), several mathematical properties of the ET family can be derived for any lifetime distribution.

2.3. Reliability measures of ET family of distributions

Reliability measures are widely used to analyze lifetime models. The most well-known measures are the survival, hazard or faultier rate and cumulative hazard functions; the following theorem presents these measures of the proposed family of distributions.

Theorem 2.1. *Let X be a random variable that follows the ET family of distributions, with PDF and CDF as defined in (1.1) and (1.2), then:*

1. *The survival function is given by*

$$\begin{aligned} S(x,\theta,\alpha) &= 1 - G(x,\theta,\alpha); \qquad x>0 \\ &= 1 - F(x,\theta)e^{-\alpha\, S_F(x,\,\theta)} \end{aligned}$$

where $S_F(x,\theta,\alpha)$ is the survival function of the baseline distribution. It is obvious that $\lim_{x\to\infty} S(x,\theta,\alpha)=0$ and $\lim_{x\to 0} S(x,\theta,\alpha)=1$.

2. *The hazard function is given by*

$$h(x,\theta,\alpha) = \frac{g(x,\theta,\alpha)}{S(x,\theta,\alpha)} = \frac{f(x,\theta)e^{-\alpha S_F(x,\,\theta,\alpha)}\,(1+\alpha\,F(x,\theta))}{1-\,F(x,\theta)e^{-\alpha S_F(x,\theta,\alpha)}}. \tag{2.5}$$

3. *The reversed hazard function is given by*

$$hr(x,\theta,\alpha) = \frac{g(x,\theta,\alpha)}{G(x,\theta,\alpha)} = \frac{f(x,\theta)e^{-\alpha S_F(x,\theta,\alpha)}\,(1+\alpha F(x,\theta))}{F(x,\theta)e^{-\alpha S_F(x,\theta,\alpha)}}. \tag{2.6}$$

Assuming the baseline distribution is the exponential distribution, then Corollary 2.2 is hold.

Corollary 2.2. *Suppose that X is a random variable of ET-Exp, then*

1. *The survival function is given by:*

$$S(x,\theta,\alpha) = 1-\ e^{-\alpha e^{-\theta x}}\Big(1-e^{-\theta x}\Big).$$

2. *The hazard rate function is given by:*

$$h(x,\theta,\alpha) = \frac{e^{-\left(\theta x+\alpha e^{-\theta x}\right)}\big(1+\alpha(1-e^{-\theta x})\big)}{1-\ e^{-\alpha e^{-\theta x}}(1-e^{-\theta x})}\ \theta.$$

3. *The reversed hazard function is given by:*

$$hr(x,\theta,\alpha) = \frac{e^{-\theta x}\big(1+\alpha(1-e^{-\theta x})\big)}{(1-e^{-\theta x})}\ \theta.$$

Figure 2 shows comparisons between the hazard rate functions of the baseline distribution and the proposed distribution.

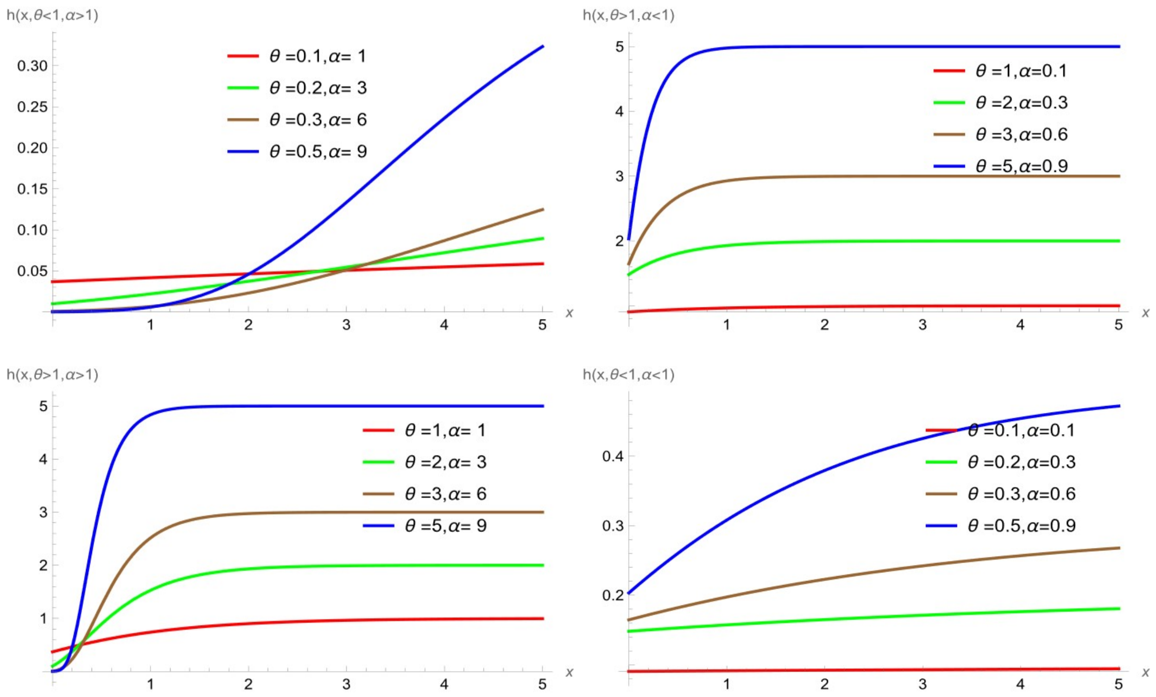


Figure 2: Hazard rate of ET-Exp with selected parameter values.

3. MATHEMATICAL PROPERTIES

Some basic mathematical properties such as ordinary moments, quantile function and moment generating function are derived in this section.

3.1. Moments

Some of the most important characteristics (tendency, dispersion, skewness and kurtosis) of a statistical distribution can be studied through moments. Suppose that the moments of $ET(x,\theta,\ \alpha)$ can be obtained by finding the expected value of $k(x)$; where

$$k(x) \ = \begin{cases} x^r, & \text{for moment of order } r, \\ e^{tx}, & \text{for moment of generating function,} \\ e^{itx}, & \text{for characteristic function.} \end{cases}$$

Hence,

$$\begin{aligned} E(k(x)) \ &= \int_0^\infty k(x)\ f(x,\theta)e^{-\alpha \overline{F}(x,\theta)}\ (1+\alpha\ F(x,\theta))dx \\ &= \int_0^\infty k(x)\ f(x,\theta)(1+\alpha\ F(x,\theta)) \left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} (-1)^{j+k}\frac{\alpha^j}{j!} \begin{pmatrix} j \\ k \end{pmatrix} F(x,\theta)^k \right)\ dx, \end{aligned}$$

which is equivalent to the expected value based on the baseline distribution

$$E_F\Big(k(x)\ e^{-\alpha\overline{F}(x,\theta)}\ (1+\alpha F(x,\theta))\Big).$$

Then the expected value can be obtained using expansion technique or by using integral estimation.

Corollary 3.1. *Suppose that X is a random variable of ET-Exp, then the r-th moment is given by:*

$$\begin{aligned} E(x^r) &= \int_0^\infty x^r\ \theta e^{-\left(\theta x+\alpha e^{-\theta x}\right)}\Big(1+\alpha\Big(1-e^{-\theta x}\Big)\Big)\ dx \\ &= \int_0^\infty x^r\ e^{-\theta x}\Big(1+\alpha\ \Big(1-e^{-\theta x}\Big)\Big)\left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{j+k}\frac{\alpha^j}{j!}\ \binom{j}{k}\ \Big(1-e^{-\theta x}\Big)^k\right)\ dx. \end{aligned}$$

Using the first fourth moments one can compute numerically the population mean, variance, standard deviation, skewness and kurtosis coefficients for some give parameter's values.

Corollary 3.2. *Suppose that X is a random variable of ET-Exp, then the moment generating function and characteristic function are, respectively, given by:*

$$\begin{aligned} E\big(e^{tx}\big) &= \int_0^\infty e^{tx}\ \theta e^{-\left(\theta x+\alpha e^{-\theta x}\right)}\Big(1+\alpha\Big(1-e^{-\theta x}\Big)\Big)\ dx \\ &= \int_0^\infty e^{(t-\theta)x}\Big(1+\alpha\ \Big(1-e^{-\theta x}\Big)\Big)\left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{j+k}\frac{\alpha^j}{j!}\ \binom{j}{k}\ \Big(1-e^{-\theta x}\Big)^k\right)\ dx. \end{aligned}$$

Similarly,

$$\begin{aligned} E\big(e^{itx}\big) &= \int_0^\infty e^{itx}\ \theta e^{-\left(\theta x+\alpha e^{-\theta x}\right)}\Big(1+\alpha\Big(1-e^{-\theta x}\Big)\Big)\ dx \\ &= \int_0^\infty e^{(it-\theta)x}\Big(1+\alpha\ \Big(1-e^{-\theta x}\Big)\Big)\left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{j+k}\frac{\alpha^j}{j!}\ \binom{j}{k}\ \Big(1-e^{-\theta x}\Big)^k\right)\ dx. \end{aligned}$$

3.2. Quantile function

The quantile function of X, say $F^{-1}(y)$, is given by the inverse function of $F(x)$. Let X follow $ET(x,\theta,\ \alpha)$ family, the quantile function of X is given by:

$$X = Q(u) = Q_F(u\ e^{-\ \alpha(1-u)}\ ;\theta)$$

where Q_F is the quantile function of the baseline distribution. Therefore, if U follow the standard uniform distribution, then $X = Q(u)$ follows the $ET(x,\theta,\ \alpha)$ family.

Now assuming that our baseline function is exponential, then, after some algebra, it follows that the Quantile function for ET-Exp distribution can be written as:

$$X = -\frac{\log(-\frac{W(u\alpha\ e^{\alpha})-\alpha}{\alpha})}{\theta}$$

where $W(.)$ is the Lambert W function.

3.3. Order statistics

Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$, be the order statistics of a random sample $X_1, X_2, ..., X_n$ from the distribution with PDF $g(x)$ and CDF $G(x)$. Then, the PDF of the i-th order statistics $X_{(i)}$ is given by:

$$g_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!}\, g(x)[\, G(x)]^{(i-1)} \times [1 - G(x)]^{(n-i)}. \tag{3.1}$$

By substituting equations (1.1) and (1.2) into equation (3.1), it follows that the PDF of the i-th order statistics $X_{(i)}$ of the ET-family is given by:

$$\begin{aligned} g_{X_{(i)}}(x) &= \frac{n!}{(i-1)!(n-i)!}\, f(x,\theta)e^{-\alpha S_F(x,\,\theta)}\, (1+\alpha F(x,\theta))\Big[F(x,\theta)e^{-\alpha S_F(x,\,\theta)}\Big]^{(i-1)} \\ & \times \Big[1 - F(x,\theta)e^{-\alpha S_F(x,\,\theta)}\Big]^{(n-i)}. \end{aligned} \tag{3.2}$$

Assuming the baseline distribution is the exponential distribution, then the equation (3.2) will be:

$$\begin{aligned} g_{X_{(i)}}(x) &= \frac{n!}{(i-1)!(n-i)!}\, \theta e^{-\left(\theta x + \alpha e^{-\theta x}\right)}\Big(1+\alpha(1-e^{-\theta x})\Big)\, \Big[e^{-\alpha e^{-\theta x}}\Big(1-e^{-\theta x}\Big)\Big]^{(i-1)} \\ & \times \Big[1 - e^{-\alpha e^{-\theta x}}\Big(1-e^{-\theta x}\Big)\Big]^{(n-i)}. \end{aligned} \tag{3.3}$$

3.4. Entropy measure

Entropy of a variable is a measure of variation of the uncertainty and it is widely used in science, e.g., physics and engineering. Here, we focus our attention on two types of entropy, namely Rényi and Tsallis.

3.4.1. Rényi entropy

The Rényi entropy of a random variable X with distribution $g(x)$ of order δ, where $\delta > 0$ and $\delta \neq 1$, can be obtained as follows:

$$R(\delta) = \frac{1}{1-\delta}\log\Big(\int g^{\delta}(x)dx\Big). \tag{3.4}$$

By substituting equations (1.2) into equation (3.4) leads to

$$R(\delta) = \frac{1}{1-\delta}\log\Big(\int \Big(f(x,\theta)e^{-\alpha\ \overline{F}(x,\theta)}(1+\ \alpha\ F(x,\theta))\Big)^{\delta} dx\Big).$$

Moreover, The Rényi entropy for ET-Exp distribution is:

$$\begin{aligned} R(\delta) &= \frac{1}{1-\delta}\log\Big(\int_0^{\infty}\Big(\theta e^{-\left(\theta x+\alpha e^{-\theta x}\right)}\Big(1+\alpha(1-e^{-\theta x})\Big)\Big)^{\delta} dx\Big) \\ &= \frac{1}{1-\delta}\log\left(\int_0^{\infty}\left(e^{-\theta x}\Big(1+\alpha\Big(1-e^{-\theta x}\Big)\Big)\left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{j+k}\frac{\alpha^j}{j!}\binom{j}{k}\Big(1-e^{-\theta x}\Big)^k\right)\right)^{\delta} dx\right). \end{aligned}$$

3.4.2. Tsallis entropy

The Tsallis entropy of a random variable X with distribution $g(x)$ of order λ, where $\lambda > 0$ and $\lambda \neq 1$, can be obtained as follows:

$$S(\lambda) = \frac{1}{1-\lambda}\left(1 - \int g^{\lambda}(x)dx\right). \tag{3.5}$$

By substituting equations (1.2) into equation (3.5) leads to

$$S(\lambda) = \frac{1}{1-\lambda}\left(1 - \int \left(f(x,\theta)e^{-\alpha\ \overline{F}(x,\theta)}(1+\ \alpha\ F(x,\theta))\right)^{\lambda} dx\right).$$

Moreover, the Tsallis entropy for ET-Exp distribution is:

$$S(\lambda) = \frac{1}{1-\lambda}\left(1 - \int_0^{\infty} \left(\theta e^{-\left(\theta x + \alpha e^{-\theta x}\right)}\left(1+\alpha(1-e^{-\theta x})\right)\right)^{\lambda} dx\right),$$

$$S(\lambda) = \frac{1}{1-\lambda}\left(1 - \int_0^{\infty}\left(e^{-\theta x}\left(1+\alpha\left(1-e^{-\theta x}\right)\right)\left(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{j+k}\frac{\alpha^j}{j!}\binom{j}{k}\left(1-e^{-\theta x}\right)^k\right)\right)^{\lambda} dx\right).$$

4. PARAMETER ESTIMATION

In this section, estimation of the unknown parameters of the $ET(x,\theta,\alpha)$ family of distributions based on complete samples are determined using method of moment (MOM) and maximum likelihood estimation (MLE) method. Let $x_1,\ x_2, ...,\ x_n$ be the observed values from $ET(x,\theta,\alpha)$ family.

4.1. Method of moment

The MOM estimator can be obtained by solving the following equations:

$$E_F\left(x\ e^{-\alpha\overline{F}(x,\theta)}\ (1+\ \alpha\ F(x,\theta))\right) = \frac{\sum_{i=1}^{n} x_i}{n},$$

$$E_F\left(x^2\ e^{-\alpha\overline{F}(x,\theta)}\ (1+\ \alpha\ F(x,\theta))\right) = \frac{\sum_{i=1}^{n} x^2{}_i}{n}.$$

Using Mathematica, we may replace the first moment of ET-Exp family by:

$$\begin{aligned} E(X) &= \frac{1}{\alpha\theta}\left(1 - e^{-\alpha} + \alpha\left(\int_0^{\alpha}\frac{\sinh(t)}{t}dt - \int_0^{\alpha}\frac{\cosh(t)-1}{t}dt\right)\right) \\ &= \frac{1}{\alpha\theta}\left(1 - e^{-\alpha} + \alpha\left(\log(\alpha) - \mathrm{Chi}(\alpha) + \mathrm{Shi}(\alpha) + \boldsymbol{\gamma}\right)\right), \end{aligned}$$

while the second moment can be replaced by the following formula:

$$E(X^2) = \frac{2\big(\alpha^2{}_3F_3(1,1,1;2,2,2;-\alpha) + \log(\alpha) + \Gamma(0,\alpha) + \boldsymbol{\gamma}\big)}{\alpha\theta^2}$$

where $\boldsymbol{\gamma}$ is Euler's constant, with numerical value ≈ 0.577216, the incomplete gamma function satisfies

$$\Gamma(0,\alpha) = \int_\alpha^\infty \frac{e^{-t}}{t}dt$$

and ${}_3F_3(1,1,1;2,2,2;-\alpha)$ is the generalized hypergeometric function.

4.2. Maximum likelihood estimation method

Using the MLE, the point estimator of the unknown parameter can be obtained by solving the following likelihood function:

$$L = \prod_{i=1}^{n} f(x_i,\theta)\ e^{-\alpha\overline{F}(x,\theta)}\ (1+\alpha\ F(x_i,\theta)).$$

Taking the Log of the likelihood function will simplify the estimation problem:

$$\text{Log}\,L\ = \left\{\sum_{i=1}^{n}\text{Log}(f(x_i,\theta)) - \sum_{i=1}^{n}\alpha\overline{F}(x_i,\theta) + \sum_{i=1}^{n}\text{Log}(1+\alpha\ F(x_i,\theta))\right\}.$$

Now, we have to find the first order condition:

$$\frac{d\,\text{Log}\,L}{d\theta} = \sum_{i=1}^{n}\frac{df(x_i,\theta)/d\theta}{f(x_i,\theta)} + \sum_{i=1}^{n}\alpha\ f(x_i,\theta) + \sum_{i=1}^{n}\frac{\alpha\ f(x_i,\theta)}{1+\alpha\ F(x_i,\theta)},$$

$$\frac{d\,\text{Log}\,L}{d\alpha} = -\sum_{i=1}^{n}\overline{F}(x_i,\theta) + \sum_{i=1}^{n}\frac{F(x_i,\theta)}{1+\alpha\ F(x_i,\theta)}.$$

Then setting each of the first order conditions to zero and using a numerical method we can find the optimal estimator of the unknown parameters.

Similarly, taking the exponential case, the MLE the point estimator of the unknown parameter can be obtained by solving the following likelihood function:

$$L = \prod_{i=1}^{n}\theta e^{-\left(\theta x_i+\alpha e^{-\theta x_i}\right)}\left(1+\alpha(1-e^{-\theta x_i})\right).$$

Taking the Log of the likelihood function will simplify the estimation problem:

$$\text{Log}\,L = \left\{\sum_{i=1}^{n}\text{Log}\left(\theta e^{-\theta x_i}\right) - \sum_{i=1}^{n}\alpha e^{-\theta x_i} + \sum_{i=1}^{n}\text{Log}(1+\alpha(1-e^{-\theta x_i}))\right\}.$$

Now, we have to find the first order condition:

$$\frac{d\,\text{Log}\,L}{d\theta} = \sum_{i=1}^{n}\left(\frac{1}{\theta}-x_i\right) + \sum_{i=1}^{n}\alpha\ \theta e^{-\theta x_i} + \sum_{i=1}^{n}\frac{\alpha\ \theta e^{-\theta x_i}}{1+\alpha(1-e^{-\theta x_i})},$$

$$\frac{d\,\text{Log}\,L}{d\alpha} = -\sum_{i=1}^{n}e^{-\theta x} + \sum_{i=1}^{n}\frac{(1-e^{-\theta x_i})}{1+\alpha(1-e^{-\theta x_i})}.$$

The non-linear equations above can not be solved analytically, and thus we have used an R-code to solve it analytically on R-software (R Core Team, 2021).

5. SIMULATION

In this section, we study the performance of ML estimators for different sample sizes, i.e, n= 50, 75, 100, 250, and 400. We have employed the inverse CDF technique for data simulation for ET-Exp distribution using R software. Bias, Variance and MSE for the ET-Exp distribution are observed. As it was expected, Table 1 shows that as the sample size increase, the values of MSE are getting smaller for the parameter estimate.

Table 1: Estimate, Bias and Mean Square Error of MLEs of parameters α and θ.

Sample Size	$\boldsymbol{\theta = 0.1}$			$\boldsymbol{\alpha = 3}$		
$\boldsymbol{n}$	**Estimate**	**Bias**	**MSE**	**Estimate**	**Bias**	**MSE**
50	0.10264	−0.00264	0.00020	3.28413	−0.28412	1.26297
75	0.10165	−0.00165	0.00013	3.17688	−0.17688	0.73347
100	0.10132	−0.00132	0.00009	3.13598	−0.13598	0.53885
250	0.10042	−0.00042	0.00004	3.05039	−0.05039	0.18580
400	0.10031	−0.00031	0.00002	3.03191	−0.03191	0.11652
Sample Size	$\boldsymbol{\theta = 3.1}$			$\boldsymbol{\alpha = 0.2}$		
$\boldsymbol{n}$	**Estimate**	**Bias**	**MSE**	**Estimate**	**Bias**	**MSE**
50	3.35913	−0.25913	0.51010	0.38221	−0.18221	0.20171
75	3.27712	−0.17712	0.32454	0.32181	−0.12181	0.12360
100	3.23269	−0.13269	0.23782	0.29313	−0.09313	0.09281
250	3.15159	−0.05159	0.09936	0.23775	−0.03775	0.03938
400	3.12687	−0.02686	0.06317	0.22022	−0.02022	0.02522
Sample Size	$\boldsymbol{\theta = 6}$			$\boldsymbol{\alpha = 3}$		
$\boldsymbol{n}$	**Estimate**	**Bias**	**MSE**	**Estimate**	**Bias**	**MSE**
50	6.15306	−0.15306	0.73829	3.27790	−0.27789	1.26813
75	6.09756	−0.09756	0.45751	3.17543	−0.17543	0.73278
100	6.09033	−0.09033	0.34335	3.14626	−0.14626	0.53472
250	6.03441	−0.03441	0.13502	3.05740	−0.05740	0.19413
400	6.01596	−0.01596	0.08147	3.03188	−0.03188	0.11481
Sample Size	$\boldsymbol{\theta = 0.6}$			$\boldsymbol{\alpha = 0.3}$		
$\boldsymbol{n}$	**Estimate**	**Bias**	**MSE**	**Estimate**	**Bias**	**MSE**
50	0.63933	−0.03933	0.01777	0.45488	−0.15488	0.22516
75	0.62767	−0.02767	0.01160	0.40929	−0.10929	0.14199
100	0.62057	−0.02057	0.00878	0.38447	−0.08447	0.11042
250	0.60632	−0.00632	0.00367	0.32713	−0.02713	0.04553
400	0.60414	−0.00414	0.00242	0.31588	−0.01588	0.03049

6. APPLICATION

In this Section, we demonstrate the capability of the ET-Exp distribution by fitting the model to four datasets, namely over the Gompertz, Exponential, Lindley, Weibull, and Generalized Exponential (GE) distributions. For theses four datasets, the maximum likelihood procedure is used to estimate the parameters of each distribution. Using the obtained estimates, we get the values of Akaike information criterion (AIC), Bayesian information criterion (BIC) and $-\log L$.

Moreover, we find the Kolmogorov-Smirnov (K-S) statistic with its corresponding P-value (P-Val), and Anderson-Darling (AD) statistics. Basic descriptive statistics are calculated for all datasets, including the five number summary, mean, variance, skewness and kurtosis. The distribution with the lowest AIC, BIC, and $-\operatorname{Log} L$ is considered as the most flexible distribution for a given dataset.

Growth hormone data: The first set of data consists of 40 observations represents the estimated time since given growth hormone medication until the children reached the target age in the Programa Hormonal de Secretaria de Saude de Minas Gerais (Lemos de Morais, 2009). The dataset was analyzed by Alizadeh *et al.* (2018). The datasets are: 2.15, 2.20, 2.55, 2.56, 2.63, 2.74, 2.81, 2.90, 3.05, 3.41, 3.43, 3.43, 3.84, 4.16, 4.18, 4.36, 4.42, 4.51,4.60, 4.61, 4.75, 5.03, 5.10, 5.44, 5.90, 5.96, 6.77, 7.82,8.00, 8.16, 8.21, 8.72, 10.40, 13.20, 13.70.

Table 2 provides the descriptive statistics for this data set and Table 3 presents the results of MLEs and goodness of fit tests for this data set using each distribution.

Table 2: The descriptive statistics for the growth hormone medication data set.

Parameters	N	**Min**	Q_1	**Median**	Q_3	**Mean**	**Max**	**Skewness**	**Kurtosis**	**Variance**
Data set I	35	2.15	3.23	4.51	6.365	5.306	13.7	1.3706	4.4008	8.4754

Table 3: MLEs and goodness of fit statistics for the growth hormone medication data set.

Distribution	α	θ	**−Log L**	**K-S**	**P-Val**	**A-D**	**P-Val**	**AIC**	**BIC**
Gompertz	0.18	0.50	87.10	0.21	0.10	1.81	0.12	178.20	181.30
Lindley	0.33		87.47	0.25	0.03	2.41	0.055	176.95	178.50
Exponential	0.19		93.41	0.33	0.0008	4.49	0.005	188.81	190.37
ET-Exp	7.18	0.52	79.84	0.11	0.83	0.63	0.62	163.68	166.79
Weibull	1.99	6.03	82.49	0.15	0.45	0.98	0.37	168.98	172.09
GE	6.51	0.48	79.10	0.10	0.86	0.53	0.72	162.20	165.31

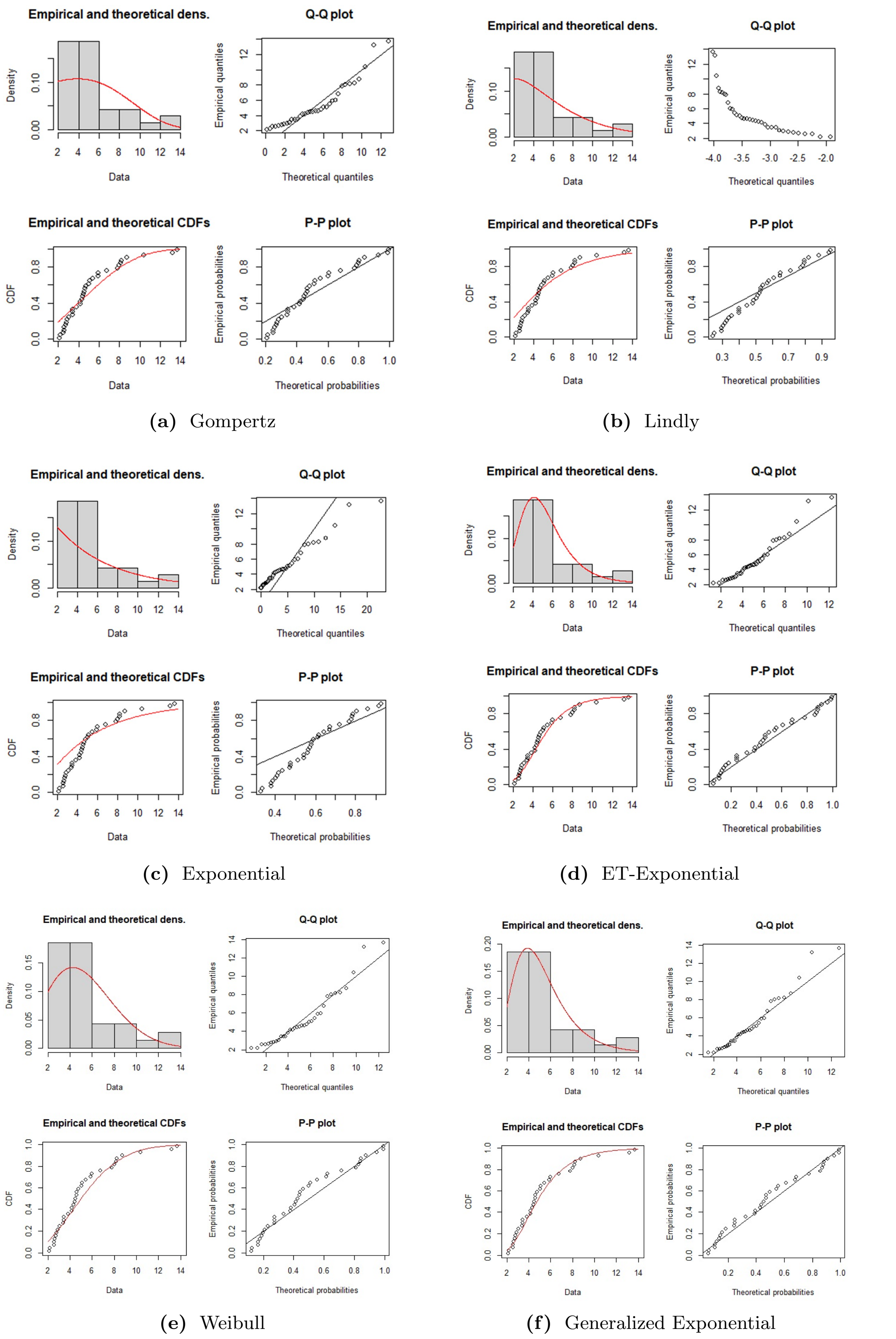


Figure 3: The empirical and theoretical PDFs, empirical and theoretical CDFs, Q-Q plots and p-p plot for (a) Gompertz, (b) Lindly, (c) Exponential, and (d) ET-Exponential, (e) Weibull, and (f) Generalized-Exponential for the growth hormone medication dataset.

Flood data: The second set of data has been presented by Dumonceaux and Antle (1973) and acts the maximum flood levels (in million of cubic feet/s) of the Susquehanna River at Harrisburg, Pennsylvania from 1890 to 1969, and its values are: 0.645, 0.613, 0.315, 0.449, 0.297, 0.402, 0.379, 0.423, 0.379, 0.324, 0.269, 0.740, 0.218, 0.412, 0.494, 0.416, 0.338, 0.392, 0.484 and 0.265.

Table 4 provides the descriptive statistics for this data set and Table 5 presents the results of MLEs and goodness of fit tests for this data set using each distribution.

Table 4: The descriptive statistics for the maximum flood levels of the Susquehanna River data set.

Parameters	N	**Min**	Q_1	**Median**	Q_3	**Mean**	**Max**	**Skewness**	**Kurtosis**	**Variance**
Data set II	20	0.218	0.3217	0.397	0.4577	5.4127	0.74	0.9116	3.368	0.0176

Table 5: MLEs and goodness of fit statistics for the maximum flood levels of the Susquehanna River data set.

Distribution	α	θ	$-$**Log L**	**K-S**	**P-Val**	**A-D**	**P-Val**	**AIC**	**BIC**
Gompertz	6.08	0.06	−9.71	0.19	0.47	1.06	0.32	−15.42	−13.43
Lindley	3.02		1.72	0.41	0.002	4.42	0.006	5.42	6.42
Exponential	2.42		2.3	0.42	0.0015	4.66	0.004	6.6	7.59
ET-Exp	33.65	10.01	−14.36	0.11	0.97	0.19	0.99	−24.71	−22.72
Weibull	3.31	0.46	−12.43	0.17	0.63	0.58	0.67	−20.86	−18.87
GE	31.81	9.83	−14.38	0.11	0.97	0.19	0.99	−24.75	−22.76

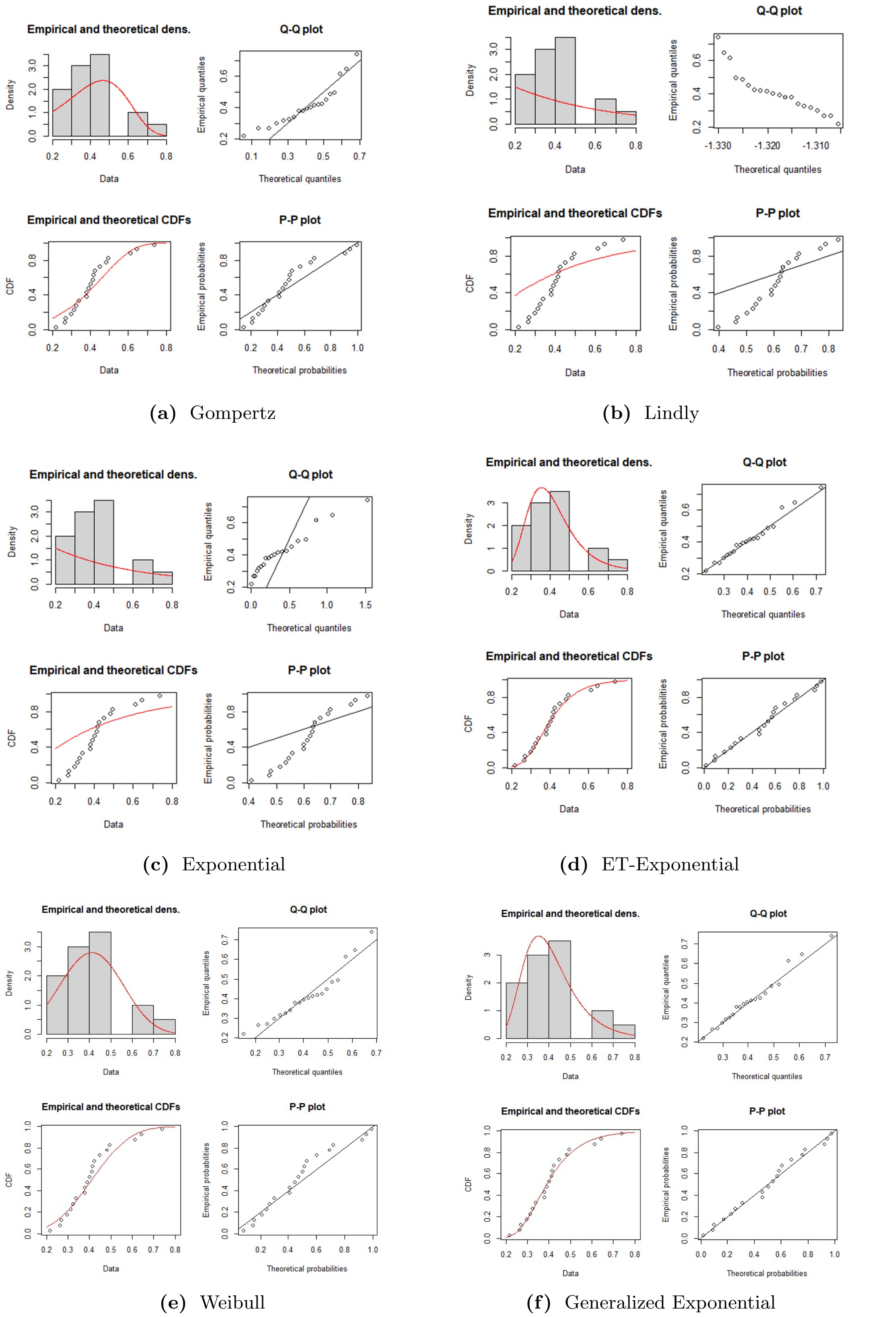


Figure 4: The empirical and theoretical PDFs, empirical and theoretical CDFs, Q-Q plots and p-p plot for (a) Gompertz, (b) Lindly, (c) Exponential, (d) ET-Exponential, (e) Weibull, and (f) Generalized-Exponential for the flood dataset.

Rock samples data: The third set of data is given by Cordeiro and dos Santos Brito (2012) consists of the shape perimeter by squared (area) from measurements of 48 rock samples from a petroleum reservoir. The data are listed as follows: 0.0903296, 0.2036540, 0.2043140, 0.2808870, 0.1976530, 0.3286410, 0.1486220, 0.1623940, 0.2627270, 0.1794550, 0.3266350, 0.2300810, 0.1833120, 0.1509440, 0.2000710, 0.1918020, 0.1541920, 0.4641250, 0.1170630, 0.1481410, 0.1448100, 0.1330830, 0.2760160, 0.4204770, 0.1224170, 0.2285950, 0.1138520, 0.2252140, 0.1769690, 0.2007440, 0.1670450, 0.2316230, 0.2910290, 0.3412730, 0.4387120, 0.2626510, 0.1896510, 0.1725670, 0.2400770, 0.3116460, 0.1635860, 0.1824530, 0.1641270, 0.1534810, 0.1618650, 0.2760160, 0.2538320, 0.2004470

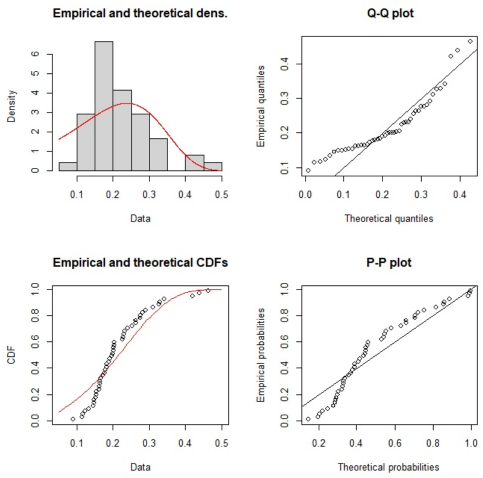
Table 6 provides the descriptive statistics for this data set and Table 7 presents the results of MLEs and goodness of fit tests for this data set using each distribution.

Table 6: The descriptive statistics for the rock samples from a petroleum reservoir.

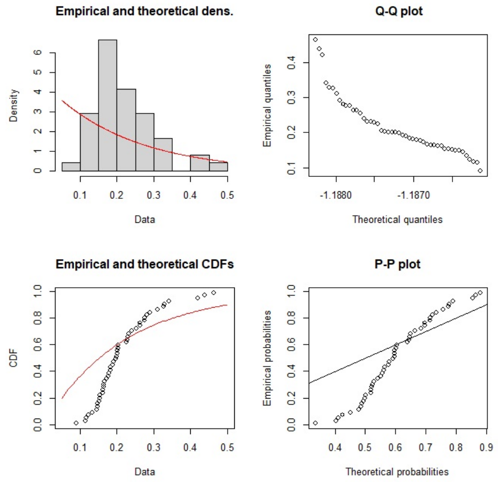
Parameters	N	**Min**	Q_1	**Median**	Q_3	**Mean**	**Max**	**Skewness**	**Kurtosis**	**Variance**
Data set III	48	0.0903	0.1623	0.1989	0.2627	0.2181	0.4641	1.1693	4.1098	0.00697

Table 7: MLEs and goodness of fit statistics for the rock samples from a petroleum reservoir data set.

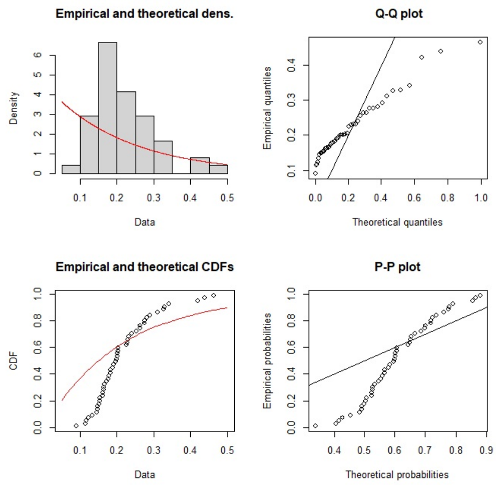
Distribution	α	θ	**−Log L**	**K-S**	**P-Val**	**A-D**	**P-Val**	**AIC**	**BIC**
Gompertz	0.14	8.22	−45.25	0.18	0.10	2.57	0.05	−86.50	−82.75
Lindley	5.31		−25.63	0.38	0.00	9.33	0.00	−49.26	−47.39
Exponential	4.58		−25.09	0.39	0.00	9.57	0.00	−48.18	−46.31
ET-Exp	19.41	16.60	−57.94	0.10	0.71	0.31	0.93	−111.89	−108.15
Weibull	2.75	0.25	−52.74	0.15	0.23	1.23	0.26	−101.48	−97.74
GE	17.84	16.06	−58.10	0.10	0.74	0.27	0.96	−112.14	−108.45



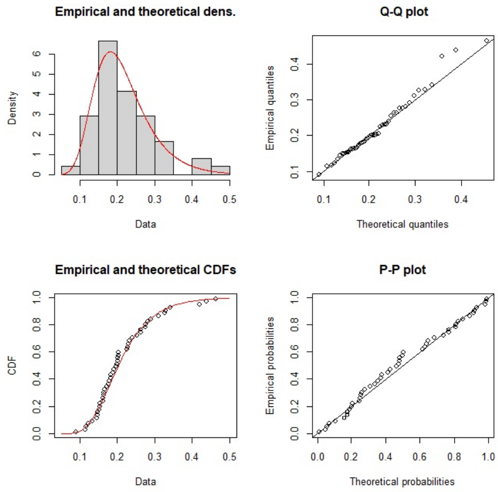
(a) Gompertz



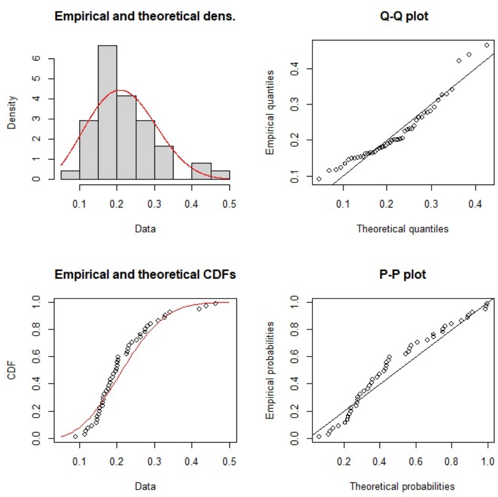
(b) Lindly



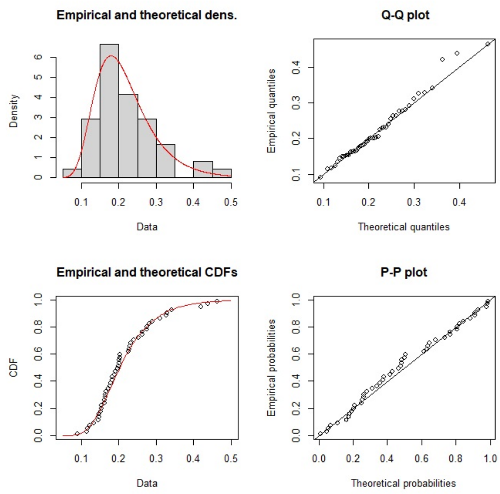
(c) Exponential



(d) ET-Exponential



(e) Weibull



(f) Generalized Exponential

Figure 5: The empirical and theoretical PDFs, empirical and theoretical CDFs, Q-Q plots and p-p plot for (a) Gpmpertz, (b) Lindly, (c) Exponential, (d) ET-Exponential, (e) Weibull, and (f) Generalized-Exponential for the rock sample dataset.

Ball bearings failure time data: The fourth set of data is obtained from Gupta and Kundu (1999) represents the number of million revolutions before failure of 23 endurance of deep-groove ball bearings in the life test. These failure times are: 17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.40, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04 and 173.40.

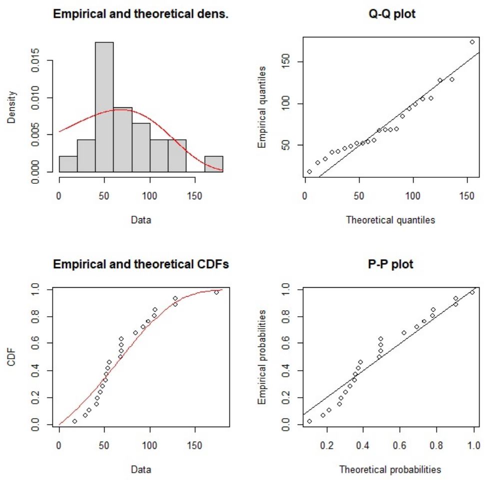
Table 8 reveals certain descriptive statistics regarding this data set and Table 9 provides the results of MLEs and goodness of fit tests for this data set using each distribution.

Table 8: The descriptive statistics for the ball bearings failure time data.

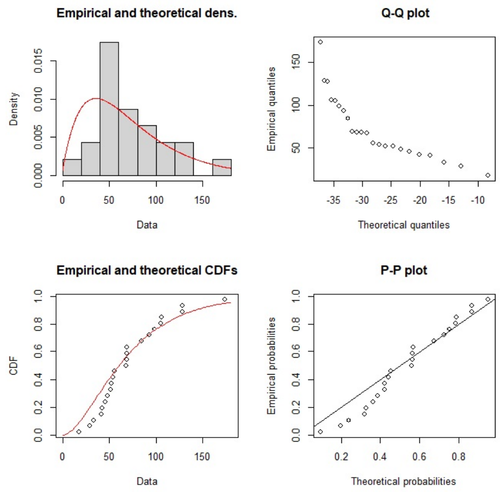
Parameters	N	**Min**	Q_1	**Median**	Q_3	**Mean**	**Max**	**Skewness**	**Kurtosis**	**Variance**
Data set IV	23	17.88	47.00	67.80	95.88	72.22	173.40	0.94	3.49	1405.58

Table 9: MLEs and goodness of fit statistics for the ball bearings failure time data.

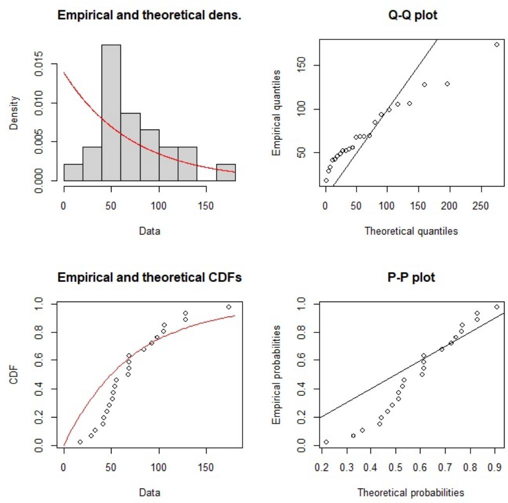
Distribution	α	θ	**−Log L**	**K-S**	**P-Val**	**A-D**	**P-Val**	**AIC**	**BIC**
Gompertz	0.33	0.02	115.98	0.15	0.65	0.73	0.53	235.96	238.23
Lindley	0.027		115.74	0.19	0.36	0.93	0.34	233.47	234.61
Exponential	0.014		121.43	0.31	0.03	0.31	0.03	244.87	246.00
ET-Exp	5.99	0.035	113.11	0.11	0.93	0.22	0.98	230.22	232.50
Weibull	2.10	81.87	113.69	0.15	0.67	0.32	0.91	231.38	233.65
GE	5.28	0.032	112.98	0.11	0.96	0.19	0.99	229.96	232.23



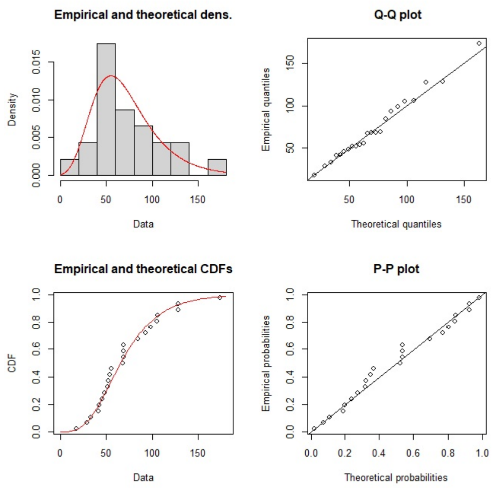
(a) Gompertz



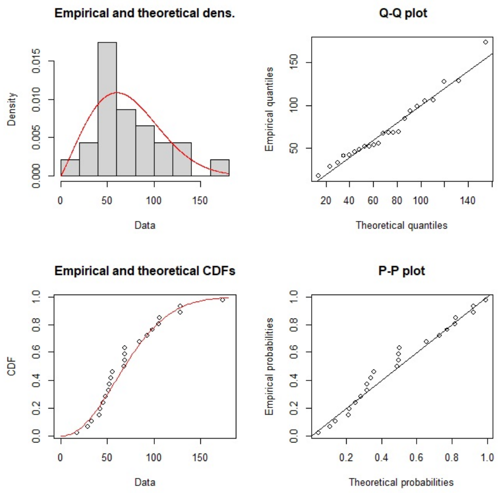
(b) Lindly



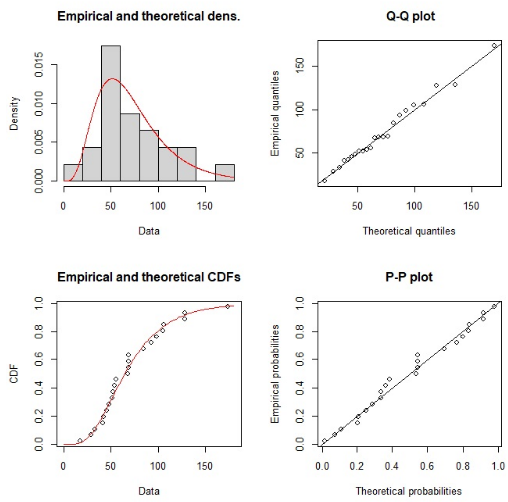
(c) Exponential



(d) ET-Exponential



(e) Weibull



(f) Generalized Exponential

Figure 6: The empirical and theoretical PDFs, empirical and theoretical CDFs, Q-Q plots and p-p plot for (a) Gpmpertz, (b) Lindly, (c) Exponential, (d) ET-Exponential, (e) Weibull, and (f) Generalized-Exponential for the ball bearings failure time data.

It can be seen that for the four datasets, GE and ET-Exp distributions have the smallest values of the Kolmogorov-Smirnov (largest P-values), Anderson-Darling, AIC and BIC goodness-of-fit tests statistics which indicate that the best fit is provided by the GE and ET-Exp distributions for these data sets. Based on the values of these statistics, we conclude that the GE and ET-Exp distributions provide the best fit in all the data sets examined. In the cases considered, the ET-Exp and GE performed far better than the Gompertz, Lindly and Exponential distributions while the Weibull distribution performed better than Gompertz and Lindly but not as good as ET-Exp and GE.

7. CONCLUDING REMARKS

A new family of lifetime distributions referred to as exponential transformation (ET) with flexible and desirable properties is proposed. Properties of the ET distribution and a sub-distribution were presented. The PDF, CDF, moments, hazard function, reliability and quantile function were presented. Entropy measures including rényi entropy, tsallis entropy for ET distribution were also derived. Estimate of the model parameters via the method of maximum likelihood obtained and applications to illustrate the usefulness of the proposed model to real data given. The applications provided show that ET distribution performs better than other several models in the literature.

ACKNOWLEDGMENTS

We would like to thank the two referees, the Editor and the associate editor for the constructive comments on the paper.

REFERENCES

Alizadeh, A., Bagheri, S., Samani, E. B., Ghobadi, S., and Nadarajah, S. (2018). Exponentiated power Lindley power series class of distributions: theory and applications. *Communications in Statistics – Simulation and Computation*, 47(9):2499–2531.

Altun, E., Korkmaz, M. Ç., El-Morshedy, M., and Eliwa, M. (2021). A new flexible family of continuous distributions: the additive odd-g family. *Mathematics*, 9(16):1837.

Alzaatreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1):63–79.

Aslam, M., Ley, C., Hussain, Z., Shah, S. F., and Asghar, Z. (2020). A new generator for proposing flexible lifetime distributions and its properties. *PloS one*, 15(4):e0231908.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2):171–178.

Chesneau, C., Bakouch, H., and Hussain, T. (2019). A new class of probability distributions via cosine and sine functions with applications. *Communications in Statistics-Simulation and Computation*, 48(8):2287–2300.

Cordeiro, G. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81(7):883–898.

Cordeiro, G. and dos Santos Brito, R. (2012). The beta power distribution. *Brazilian Journal of Probability and Statistics*, 26(1):88–112.

Dumonceaux, R. and Antle, C. E. (1973). Discrimination between the log-normal and the Weibull distributions. *Technometrics*, 15(4):923–926.

Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics – Theory and Methods*, 31(4):497–512.

Goldoust, M., Rezaei, S., Alizadeh, M., and Nadarajah, S. (2019). The odd log-logistic power series family of distributions: properties and applications. *Statistica*, 79(1):77–107.

Gupta, R. D. and Kundu, D. (1999). Theory & methods: generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

Kumar, D., Singh, U., and Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients data. *Journal of Statistics Applications & Probability*, 4(3):417-427.

Kumar, D., Kumar, P., Kumar, P., Singh, S. K., and Singh, U. (2021). PCM transformation: properties and their estimation. *Journal of Reliability and Statistical Studies*, 14(2):373–392.

Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1-2):79–88.

Lemos de Morais, A. (2009). *A class of generalized beta distributions, Pareto power series and Weibull power series*. Universidade Federal de Pernambuco.

Marshall, A. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3):641–652.

Maurya, S., Kaushik, A., Singh, R., Singh, S. K., and Singh, U. (2016). A new method of proposing distribution and its application to real data. *Imperial Journal of Interdisciplinary Research*, 2(6):1331–1338.

Mudholkar, G. and Srivastava, D. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2):299–302.

Shaw, W. and Buckley, I. (2007). The alchemy of probability distributions: Beyond Gram-Charlier & Cornish-Fisher expansions, and skew-normal or kurtotic-normal distributions. *Map, Research report, Financial Mathematics Group, King's College*, 1–18.

Shaw, W. and Buckley, I. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint*, arXiv:0901.0434.

R Core Team (2021). R: A language and environment for statistical computing. *R Version 4.0.3*, R Foundation for Statistical Computing, Vienna, Austria, Google Scholar – there is no corresponding record for this reference.

Zaidi, S., Sobhi, M., El-Morshedy, M., and Afify, A. (2021). A new generalized family of distributions: Properties and applications. *Aims Math*, 6(1):456–476.

Zichuan, M., Hussain, S., Iftikhar, A., Ilyas, M., Ahmad, Z., Khan, D. M., and Manzoor, S. (2020). A new extended-family of distributions: properties and applications. *Computational and Mathematical Methods in Medicine*, 2020:4650520.

Zografos, K. and Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6(4):344–362.

REVSTAT — Statistical journal

AIMS AND SCOPE

The aim of REVSTAT — Statistical Journal is to publish articles of high scientific content, developing Statistical Science focused on innovative theory, methods, and applications in different areas of knowledge. Important survey/review contributing to Probability and Statistics advancement is also welcome.

BACKGROUND

Statistics Portugal started in 1996 the publication of the scientific statistical journal *Revista de Estatística*, in Portuguese, a quarterly publication whose goal was the publication of papers containing original research results, and application studies, namely in the economic, social, and demographic fields. Statistics Portugal was aware of how vital statistical culture is in understanding most phenomena in the present-day world, and of its responsibilities in disseminating statistical knowledge.

In 1998 it was decided to publish papers in English. This step has been taken to achieve a larger diffusion, and to encourage foreign contributors to submit their work. At the time, the editorial board was mainly composed by Portuguese university professors, and this has been the first step aimed at changing the character of *Revista de Estatística* from a national to an international scientific journal. In 2001, the *Revista de Estatística* published a three volumes special issue containing extended abstracts of the invited and contributed papers presented at the 23rd European Meeting of Statisticians (EMS). During the EMS 2001, its editor-in-chief invited several international participants to join the editorial staff.

In 2003 the name changed to REVSTAT — Statistical Journal, published in English, with a prestigious international editorial board, hoping to become one more place where scientists may feel proud of publishing their research results.

EDITORIAL POLICY

REVSTAT — Statistical Journal is an open access peer-reviewed journal published quarterly, in English, by Statistics Portugal.

The editorial policy of REVSTAT is mainly placed on the originality and importance of the research. The whole submission and review processes for REVSTAT are conducted exclusively online on the journal's webpage revstat.ine.pt based in Open Journal System (OJS). The only working language allowed is English. Authors intending to submit any work must register, login and follow the guidelines.

There are no fees for publishing accepted manuscripts that will be made available in open access.

All articles consistent with REVSTAT aims and scope will undergo scientific evaluation by at least two reviewers, one from the Editorial Board and another external. Authors can suggest an editor or reviewer who is expert on the paper subject providing her/his complete information, namely: name, affiliation, email and, if possible, personal URL or ORCID number.

Journal prints may be ordered at expenses of the author(s), and prior to publication.

ABSTRACT AND INDEXING SERVICES

REVSTAT — *Journal Citation Reports - JCR (Clarivate); DOAJ-Directory of Open Access Journals; Current Index to Statistics; Google Scholar; Mathematical Reviews® (MathSciNet®); Zentralblatt für Mathematic; Scimago Journal & Country Rank; Scopus*

AUTHOR GUIDELINES

The whole submission and review processes for REVSTAT are conducted exclusively online on the journal's webpage https://revstat.ine.pt/ based in Open Journal System (OJS). Authors intending to submit any work must ***register, login*** and follow the indications choosing ***Submissions***.

REVSTAT — Statistical Journal adopts the COPE guidelines on publication ethics.

Work presentation

- the only working language is English;
- the first page should include the name, ORCID iD (optional), Institution, country, and email-address of the author(s);
- a summary of fewer than one hundred words, followed by a maximum of six keywords and the MSC 2020 subject classification should be included also in the first page;
- manuscripts should be typed only in black, in double-spacing, with a left margin of at least 3 cm, with numbered lines, and a maximum of 25 pages, in .pdf format;
- the title should be with no more than 120 characters (with spaces);
- figures must be a minimum of 300dpi and will be reproduced online as in the original work, however, authors should take into account that the printed version is always in black and grey tones;
- authors are encouraged to submit articles using LaTeX which macros are available at *REVSTAT style*;
- citations in text should be included in the text by name and year in parentheses, as in the following examples: § article title in lowercase (Author 1980); § This theorem was proved later by AuthorB and AuthorC (1990); § This subject has been widely addressed (AuthorA 1990; AuthorB et al. 1995; AuthorA and AuthorB 1998);
- references should be listed in alphabetical order of the author's scientific surname at the end of the article;
- acknowledgments of people, grants or funds should be placed in a short section before the References title page. Note that religious beliefs, ethnic background, citizenship and political orientations of the author(s) are not allowed in the text;
- authors are welcome to suggest one of the Editors or Associate Editors or yet other reviewer expert on the subject providing a complete information, namely: name, affiliation, email, and personal URL or ORCID number in the Comments for the Editor (submission form).

Accepted papers

After final revision and acceptance of an article for publication, authors are requested to provide the corresponding LaTeX file, as in REVSTAT style.

Supplementary files may be included and submitted separately in .tiff, .gif, .jpg, .png .eps, .ps or .pdf format. These supplementary files may be published online along with an article, containing data, programming code, extra figures, or extra proofs, etc; however, REVSTAT is not responsible for any supporting information supplied by the author(s).

Copyright Notice

Upon acceptance of an article, the author(s) will be asked to transfer copyright of the article to the publisher, Statistics Portugal, to ensure the widest possible dissemination of information.

According to REVSTAT's archiving policy, after assigning the copyright form, authors may cite and use limited excerpts (figures, tables, etc.) of their works accepted/published in REVSTAT in other publications and may deposit only the published edition in scientific repositories providing its source as REVSTAT while the original place of publication. The Executive Editor of the Journal must be notified in writing in advance.