FORTHCOMING PAPER ·#60E05-18-01

ESTIMATION OF THE MARSHALL-OLKIN PARETO DISTRIBUTION PARAMETERS: COMPARATIVE STUDY

Omar M. Bdair and Hanan A. Haj Ahmad*[1]
*Faculty of Engineering Technology, Al-Balqa Applied University, Amman 11134, Jordan
**Department of Mathematics, University of Hail, Hail, KSA

ABSTRACT
This article deals with different methods of point estimation for the unknown parameters of Marshall-Olkin Pareto distribution (MOP). This is a new lifetime that generalizes Pareto distribution which was introduced by Marshall-Olkin (1997). Some classical point estimation methods are considered and their asymptotic properties are discussed along with studying Bayesian estimation method. The main purpose of this work is to determine which estimation method is more efficient under MOP distribution based on minimum average relative mean square error (MSE). Real data analyses are performed and it has been shown that MOP distribution is a better fit than the original Pareto distribution. In this paper, we compare the performances of these procedures through extensive numerical simulations.

KEYWORDS: Marshall-Olkin distribution; Pareto distribution; Percentile estimation method; Least square estimation; L-moments estimation; Bayesian estimation; Lindley method; Monte Carlo simulations; Newton-Raphson method.

MSC: 60E05.

RESUMEN
Este articulo trata con diferentes métodos de estimación puntual de parámetros desconocidos de la distribución de Marshall-Olkin Pareto (MOP). Esta es una distribución nueva del tiempo de vida de Pareto que fue introducida por Marshall-Olkin (1997). Se consideran algunos métodos clásicos de estimación puntual y sus propiedades asintóticos los que son discutidos junto con el método Bayesiano. El propósito principal de este trabajo es determinar cual es el método más eficiente de estimación bajo la distribución MOP basado en el ,mínimo del promedio relativo el error cuadrático medio (MSE). Se desarrollananálisis con datos reales y se demuestra que se orbiten un mejor ajuste con la distribución MOP que con la original Pareto. En este paper, comparamos el desempeño de estos procedimientos a través de extensas simulacionesnuméricas

PALABRAS CLAVES: distribución de Marshall-Olkin distribución; distribución de Pareto; método Percentil de estimación; estimación mínimo cuadrática ; estimación L de momentos ; estimación Bayesiana; método de Lindley ; simulación de Monte Carlo; método de Newton-Raphson.

1. INTRODUCTION

Expanding family of distributions is an attractive subject to many scientists because expanding or generalizing a certain distribution will add some flexibility to the original distribution. In recent years several methods of generating new distributions from classical ones were developed. Many generalized classes of distributions have been applied to describe various phenomena. Kumaraswamy (1980) introduced a generalization of Beta distribution, where he found that the new one is much better suited than Beta distribution for computation intensive activities. Later in (1997) a new method was proposed by Marshall and Olkin, their idea of obtaining a new distribution depends on adding a parameter to the original distribution. The new family of distributions includes the original distributions as special cases, and it gives more flexibility to the original models.
Many physical and lifetime applications were discussed in literature concerning Marshall-Olkin distribution (MO), Sankaran and Jayakumar (2006) studied the physical interpretation of MO family by considering odd models. Jose (2011) considered the applications of MO family in reliability theory, while Lai (2013) studied some methods and mechanisms to construct generalizations of some life time distributions and their applications in reliability

[1]E-mail addresses: bdairmb@yahoo.com*(O. M. Bdair),* hananahm1@yahoo.com*(H. Haj Ahmad)* .

engineering, insurance and others. It was noticed that adding a parameter to the original distribution will yield a distribution with interesting hazard function, therefore it can be used to model real data in a better way than the basic distribution. That was a reason for many researchers to use MO extended family of distribution in producing new models. For more details, one may refer to Jose and Alice (2001, 2005), Ghitany et. al (2005), Ghitany and Kotz (2007) and Jose and Uma (2009).

The Pareto distribution is a well known model, it was first studied by a professor of economics "Vilfredo Pareto". Many forms of Pareto distribution appeared in the literature and it was used in a wide range of scientific applications. For instance, it was found that it is compliant in lifetime models such as actuarial sciences, finance, economic, life testing and climatology, where it usually describes the occurrence of extreme weather.

Several generalized forms of Pareto distribution were discussed in literature. Generalized Pareto (GP) distribution was first studied by Pickands (1975), and then it was studied by many authors like Gupta et. al (1998) and Hogg et. al (2005). The GP distribution was used as a model for excesses over thresholds. Its applications include environmental extreme events, ozone levels in the upper atmosphere, large insurance claims or large fluctuation in financial data, and reliability studies. Its areas of applications are successfully addressed in several books, such as those by Castillo et. al(2004), Kotz and Nadarajah(2000), and Ahsanullah (1992). Alice and Jose (2004) considered Marshall-Olkin Pareto and Marshall-Olkin semi-Pareto distributions; they developed time series models with modification structure. Ghitany (2005) considered MO of Pareto distribution and studied some of its statistical properties and its hazard rate. He also showed that the limiting distributions of the sample extremes were of exponential and Fréchet type. He used maximum likelihood estimation method to estimate the unknown parameters of MO Pareto distribution. However, there are some situations in which Pareto distribution may not be suitable from a theoretical or applied point of view. So, to obtain a more flexible family of distributions, we introduce here an extension of Pareto distribution based the MO extension method.

In this paper we study different classical point estimation methods for the unknown parameters of MO for Pareto Type I distribution (MOP) as well as the Baysian method. Some properties of the density function are discussed. Numerical methods are used to solve the obtained normal equations. Simulations are also used to make comparison between those methods, and also to determine which method is more efficient according to the mean square error. Bayesian estimation method was used and a comparison between Bayesian method and the classical methods of estimation is performed.

The rest of the paper is organized as follows: In section 2 we introduce MOP distribution. Classical point estimation methods for the unknown parameters are discussed in section 3, while in section 4, Bayesian estimation method is considered. In section 5 simulation study and real life data analysis are presented and also comparison results among all estimation methods are provided.

2 .PROBABILITY DENSITY FUNCTION

The Pareto Type I distribution is a continuous statistical distribution with probability density function (pdf) of the form $f(x;\theta,\beta)=\frac{\theta\beta^{\theta}}{x^{\theta+1}}, \quad x\geq\beta,$ where $\theta>0$ is the shape parameter and $\beta>0$ is a scale parameter. The survival function of Pareto distribution is given by $\overline{F}(x;\theta,\beta)=(\frac{\beta}{x})^{\theta}$.

Marshall-Olkin (MO) distribution can be obtained by adding a parameter α, see Marshall and Olkin (1997). The survival function of MO distribution is given by $\overline{G}(x)=\frac{\alpha\overline{F}(x)}{1-\overline{\alpha}\overline{F}(x)}, \quad -\infty<x<\infty,\ \alpha>0$ where $\overline{\alpha}=1-\alpha$.

Note that, when $\alpha=1,\ \overline{G}(x)=\overline{F}(x)$. Consequently, the survival function of the three parameters Marshall-Olkin Pareto (MOP) distribution is obtained as $\overline{G}(x;\alpha,\theta,\beta)=\frac{\alpha(\frac{\beta}{x})^{\theta}}{1-\overline{\alpha}(\frac{\beta}{x})^{\theta}}, \quad x\geq\beta.$ Consequently, the corresponding pdf of MOP distribution is

$$g(x;\alpha,\theta,\beta)=\frac{\alpha\theta\beta^{\theta}x^{-(\theta+1)}}{(1-\overline{\alpha}(\frac{\beta}{x})^{\theta})^{2}}, \quad \alpha,\theta,\beta>0,\ \ x\geq\beta. \tag{1}$$

The density function of MOP distribution can be rewritten as a linear combination of Pareto distribution using the well-known of binomial expansion. First we assume that $0<\alpha<1,$ then the binomial expansion for the denominator in Eq. (1) can be applied since $\left|\bar{\alpha}(\frac{\beta}{x})^{\theta}\right|<1,$ so we can rewrite it as:

$$\begin{aligned}(1-\bar{\alpha}(\frac{\beta}{x})^{\theta})^{-2} &= \sum_{j=0}^{\infty}\frac{\Gamma(j+2)}{\Gamma(2)j!}\left(\bar{\alpha}(\frac{\beta}{x})^{\theta}\right)^{j} \\ &= \sum_{j=0}^{\infty}(j+1)\bar{\alpha}^{j}(\frac{\beta}{x})^{j\theta},\end{aligned} \tag{2}$$

where $\Gamma(.)$ is the gamma function. Applying the expansion (2) in (1), yields

$$\begin{aligned}g(x;\alpha,\theta,\beta) &= \sum_{j=0}^{\infty}\alpha\bar{\alpha}^{j}(j+1)\theta\frac{\beta^{\theta(j+1)}}{x^{\theta(j+1)+1}} \\ &= \sum_{j=0}^{\infty}\eta_j f_P(x;\theta^*,\beta),\end{aligned} \tag{3}$$

where $f_P(x;\theta^*,\beta)$ is the probability density function of Pareto distribution with parameters θ^* and β, where $\theta^*=(j+1)\theta$ and $\eta_j=\alpha\bar{\alpha}^{j}$.

For $\alpha>1$, we can use similar argument as in Eq.(2), and after some algebraic manipulations we can obtain

$$g(x;\alpha,\theta,\beta)=\sum_{j=0}^{\infty}v_j f_P(x;\theta^*,\beta), \tag{4}$$

where $v_j=\frac{(-1)^j}{\alpha(j+1)}\sum_{k=j}^{\infty}(k+1)\binom{k}{j}(1-\frac{1}{\alpha})^k$, note that $\sum_{j=0}^{\infty}\eta_j=\sum_{j=0}^{\infty}v_j=1$.

Hence the MOP density function can be expressed as infinite linear combinations of Pareto density functions. Therefore, expression (3) and (4) can be used to find and simplify many mathematical properties related to moments.

3. CLASSICAL POINT ESTIMATION METHODS

In this section we consider different methods of point estimation for the MOP parameters. The asymptotic properties are discussed for some point estimation methods. Numerical techniques are helpful in obtaining the estimated values of these parameters, then we compare between these estimation methods to decide which method is more efficient.

3.1.Maximum Likelihood Estimation

The maximum likelihood estimation (MLE) is widely used in inferential statistics as it has many nice properties, such as invariance, consistency, and normal approximation properties. It depends basically on maximizing the likelihood function of MOP distribution. Let $X_1,X_2,...,X_n$ be a random sample from MOP distribution, then the log-likelihood function for the vector of parameters $\gamma=(\alpha,\theta,\beta)$ can be expressed by

$$\ell(\gamma)=n\log(\alpha\theta)+n\theta\log\beta-(\theta+1)\sum_{i=1}^{n}\log(x_i)-2\sum_{i=1}^{n}\log(1-\bar{\alpha}(\frac{\beta}{x_i})^{\theta}). \tag{5}$$

From the above log-likelihood equation we compute the derivatives with respect to the parameter vector γ, but since $x\geq\beta$, then the MLE of the parameter β is assumed to be $x_{(1)}$, where $x_{(1)}$ is the first order statistics. Taking partial derivatives of Eq. (5) with respect to α and $\theta,$ then equate them to zero will yield two nonlinear normal equations as follows:

$$\frac{\partial \ell(\gamma)}{\partial \alpha}=\frac{n}{\alpha}-2\mu\sum_{i=1}^{n}\frac{1}{x_i^{\theta}-\bar{\alpha}\mu}=0$$

$$\frac{\partial \ell(\gamma)}{\partial \theta}=\frac{n}{\theta}+\frac{n}{\theta}\log\mu-\sum_{i=1}^{n}\log(x_i)+2\bar{\alpha}\mu\sum_{i=1}^{n}\frac{\log(\frac{\mu^{\frac{1}{\theta}}}{x_i})}{x_i^{\theta}-\bar{\alpha}\mu}=0,$$

where $\mu=\beta^{\theta}$.

The above normal equations of α and μ form an implicit system, hence do not have unique root, so they can be solved analytically. The maximum likelihood estimators (MLE) have been obtained using Newton-Raphson (N-R) method or using the functions nlm or optim in R statistical package which maximize the log-likelihood function. The normal approximation of the MLE of vector parameter γ can be used to construct approximate confidence intervals and testing hypotheses on the parameters $\alpha,\ \theta$ and β. From the asymptotic property of the MLE we have $\sqrt{n}(\hat{\gamma}-\gamma)\overset{D}{\rightarrow}N_3(0,I^{-1}(\gamma)),$ where $I(\gamma)$ is the Fisher information matrix, i.e.,

$$I(\gamma)=-\frac{1}{n}\begin{bmatrix}E(\ell_{\alpha\alpha}) & E(\ell_{\alpha\theta}) & E(\ell_{\alpha\beta})\\ E(\ell_{\theta\alpha}) & E(\ell_{\theta\theta}) & E(\ell_{\theta\beta})\\ E(\ell_{\beta\alpha}) & E(\ell_{\beta\theta}) & E(\ell_{\beta\beta})\end{bmatrix}=-\begin{bmatrix}I_{11} & I_{12} & I_{13}\\ I_{21} & I_{22} & I_{23}\\ I_{31} & I_{32} & I_{33}\end{bmatrix},$$

where the second partial derivatives are as follow:

$$\ell_{\alpha\alpha}=\frac{-n}{\alpha^2}+2\mu^2\sum_{i=1}^{n}\frac{1}{(x_i^{\theta}-\bar{\alpha}\mu)^2}$$

$$\ell_{\alpha\theta}=-2\mu\sum_{i=1}^{n}\frac{x_i^{\theta}\log(\frac{\mu^{\frac{1}{\theta}}}{x_i})}{(x_i^{\theta}-\bar{\alpha}\mu)^2}$$

$$\ell_{\theta\theta}=-\frac{n}{\theta^2}+2\bar{\alpha}\mu\sum_{i=1}^{n}\frac{x_i^{\theta}(\log(\frac{\mu^{\frac{1}{\theta}}}{x_i}))^2}{(x_i^{\theta}-\bar{\alpha}\mu)^2}$$

$$\ell_{\beta\beta}=-\frac{n\theta}{\mu^{\frac{2}{\theta}}}+2\bar{\alpha}\theta\mu^{1-\frac{2}{\theta}}\sum_{i=1}^{n}\frac{(\theta-1)x_i^{\theta}+\bar{\alpha}\mu}{(x_i^{\theta}-\bar{\alpha}\mu)^2}$$

$$\ell_{\alpha\beta}=-2\theta\mu^{\frac{1}{\theta}+1}\sum_{i=1}^{n}\frac{x_i^{\theta}}{(x_i^{\theta}-\bar{\alpha}\mu)^2}$$

$$\ell_{\theta\beta}=\frac{n}{\mu^{\frac{1}{\theta}}}+\frac{2\bar{\alpha}\theta}{\mu^{\frac{1}{\theta}-1}}\sum_{i=1}^{n}\frac{x_i^{\theta}\log(\frac{\mu^{\frac{1}{\theta}}}{x_i})}{(x_i^{\theta}-\bar{\alpha}\mu)^2}+\frac{2\bar{\alpha}}{\mu^{\frac{1}{\theta}-1}}\sum_{i=1}^{n}\frac{1}{x_i^{\theta}-\bar{\alpha}\mu}$$

The expected values of these second derivatives can be evaluated using integration techniques, hence the entries of the Fisher information matrix for $\alpha\neq 1$ are:

$$I_{11} = \frac{-1}{3\alpha^2}$$

$$I_{12} = I_{21} = \frac{1}{3\bar{\alpha}^2 \alpha\theta}(\bar{\alpha} - \alpha^2 \log\alpha)$$

$$I_{13} = I_{31} = \frac{-\theta(\alpha+2)}{3\beta\alpha^2}$$

$$I_{22} = \frac{1}{\theta^2}(\frac{2}{3}(\alpha Poly\log(2,\bar{\alpha}) - \bar{\alpha}) - 1)$$

$$I_{23} = I_{32} = \frac{1}{\alpha\beta} + \frac{1}{3\beta\alpha\bar{\alpha}}(\bar{\alpha} + \alpha^2 \log\alpha)$$

$$I_{33} = \frac{\theta}{3\alpha^2\beta^2}(\bar{\alpha}^2(\theta-1)(\alpha+2) - \alpha^2),$$

where $poly\log(s,z)$ is a special function $Li_s(z)$ which is the polylogarithm function defined by the power series $Li_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}$.

The variances of the MLE of parameters can be obtained from the asymptotic property of MLE so that $V(\hat{\alpha}_{MLE}) \approx \frac{I_{22}I_{33} - I^2{}_{23}}{Det(I(\gamma))}$, $V(\hat{\theta}_{MLE}) \approx \frac{I_{11}I_{33} - I^2{}_{13}}{Det(I(\gamma))}$ and $V(\hat{\beta}_{MLE}) \approx \frac{I_{11}I_{22} - I^2{}_{12}}{Det(I(\gamma))}$ where $Det(I(\gamma))$ is the determinant of information matrix I.

Now without loss of generality we may assume that $\beta = 1$ since it is a scale parameter. Then the MLE of α, say $\hat{\alpha}_{MLE}$, when the other parameters are known, can be obtained by solving the non linear equation

$$\hat{\alpha}_{MLE} = \frac{n}{2\sum_{i=1}^{n} \frac{1}{x_i^{\theta} - (1-\hat{\alpha}_{MLE})}}. \tag{6}$$

Let us consider the MLEs of β and θ, say $\hat{\beta}_{MLE}$ and $\hat{\theta}_{MLE}$ when the shape parameter α is known and another parameter is also known. For known α and θ, $\hat{\beta}_{MLE}$ and for known α and β, $\hat{\theta}_{MLE}$ can be obtained by maximizing

$$w(\beta) = n\theta\log\beta - 2\sum_{i=1}^{n}\log(1-\bar{\alpha}(\frac{\beta}{x_i})^{\theta}), \tag{7}$$

with respect to β and by maximizing

$$w(\theta) = n\log(\theta) + n\theta\log\beta - (\theta+1)\sum_{i=1}^{n}\log(x_i) - 2\sum_{i=1}^{n}\log(1-\bar{\alpha}(\frac{\beta}{x_i})^{\theta}), \tag{8}$$

with respect to θ, respectively. Therefore, $\hat{\beta}_{MLE}$ can be obtained by a numerical solution using the Newton Raphson method of the non-linear equation $\frac{n\theta}{\beta} = -2\sum_{i=1}^{n}\frac{\bar{\alpha}\theta(\beta/x_i)^{\theta-1}}{1-\bar{\alpha}(\beta/x_i)^{\theta}}$, while $\hat{\theta}_{MLE}$ can be obtained also by a numerical solution using the Newton Raphson method of the non-linear equation

$$\frac{n}{\theta} + n\log\beta = \sum_{i=1}^{n}\log x_i - 2\sum_{i=1}^{n}\frac{\bar{\alpha}(\beta/x_i)^{\theta}\log(\beta/x_i)}{1-\bar{\alpha}(\beta/x_i)^{\theta}}.$$

3.2 Moment Estimation

In order to find the method of moment estimators (MME) it is necessary to compute the moments of MOP distribution. So we use the idea of binomial expansion in Eqs. (3) and (4) in order to obtain the following:

$$E(X^r) = \sum_{j=0}^{\infty} w_j E(Y_j^r), \tag{9}$$

where

$$w_j = \begin{cases} \eta_j & 0 < \alpha < 1 \\ \nu_j & \alpha > 1 \end{cases}$$

and $Y_j \sim f_P(x;\theta^*,\beta)$.

Since the r^{th} moment for Pareto distribution is given by $E(Y_j^r) = \frac{(j+1)\theta\beta^r}{(j+1)\theta - r}$, the expected value in Eq.(9) can be written as

$$E(X^r) = \sum_{j=0}^{\infty} w_j \frac{(j+1)\theta\beta^r}{(j+1)\theta - r}. \tag{10}$$

Using this formula, we can find the first three sample moments. Equating the sample moments with the population moments of MOP distribution, we obtain the following three equations:

$$m_1 = \overline{X} = \theta\beta \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 1} \tag{11}$$

$$m_2 = E(X^2) = \theta\beta^2 \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 2} \tag{12}$$

$$m_3 = E(X^3) = \theta\beta^3 \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 3}. \tag{13}$$

Numerical methods are used to solve the three equations (11), (12) and (13) in order to estimate the needed parameters.

Population and sample variances can be used to obtain the second moment instead of equation(12) that is

$$S^2 = \sigma^2 = \theta\beta^2 \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 2} - \theta^2\beta^2 \Big(\sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 1}\Big)^2. \tag{14}$$

Now we discuss the asymptotic distribution properties of the MME's of α, θ and β. Let $\gamma = (\alpha, \theta, \beta)$, and let

$$f_1(\gamma) = \overline{X} - \theta\beta \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 1}$$

$$f_2(\gamma) = S^2 - \theta\beta^2 \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 2} + \theta^2\beta^2 \Big(\sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 1}\Big)^2$$

$$f_3(\gamma) = E(X^3) - \theta\beta^3 \sum_{j=1}^{\infty} \frac{w_j(j+1)}{(j+1)\theta - 3}$$

Using Taylor expansion of $f(\hat{\gamma}_{MME})$ about the true value of $\gamma = (\alpha, \theta, \beta)$, where $f(\gamma) = (f_1(\gamma), f_2(\gamma), f_3(\gamma))$, we obtain

$$f(\hat{\gamma}_{MME})-f(\gamma)=[\hat{\alpha}_{MM}-\alpha,\hat{\theta}_{MM}-\theta,\hat{\beta}_{MM}-\beta]\begin{bmatrix}(\frac{\partial f_1}{\partial\alpha}) & (\frac{\partial f_2}{\partial\alpha}) & (\frac{\partial f_3}{\partial\alpha})\\ (\frac{\partial f_1}{\partial\theta}) & (\frac{\partial f_2}{\partial\theta}) & (\frac{\partial f_3}{\partial\theta})\\ (\frac{\partial f_1}{\partial\beta}) & (\frac{\partial f_2}{\partial\beta}) & (\frac{\partial f_3}{\partial\beta})\end{bmatrix}_{\gamma=\bar{\gamma}},$$

where $\bar{\gamma}$ is a point between $\hat{\gamma}_{MME}$ and γ. It is clear that as $n\to\infty$, $\hat{\gamma}_{MME}\to\gamma$, and $\bar{\gamma}\to\gamma$. Using central limit theorem we obtain $\sqrt{n}(\bar{X}-E(\bar{X}))\to N(0,\sigma^2)$, and $\sqrt{n}(S^2-E(S^2))\to N(0,\frac{(n-1)^2m_4-(n-1)(n-3)m_2^2}{n^3})$, where $m_4=E(X^4)=\theta\beta^4\sum_{j=1}^{\infty}\frac{w_j(j+1)}{(j+1)\theta-4}$. Therefore $[\sqrt{n}(\bar{X}-E(\bar{X})),\sqrt{n}(S^2-E(S^2))]\to N(0,\begin{bmatrix}\sigma^2 & Cov(\bar{X},S^2)\\ Cov(\bar{X},S^2) & \frac{(n-1)^2m_4-(n-1)(n-3)m_2^2}{n^3}\end{bmatrix}$,

where $Cov(\bar{X},S^2)=\frac{n}{n-1}\left(m_3-2m_1m_2+m_1\sigma^2+m_1^3\right)-m_1(m_2-m_1^2)$, is the covariance between $\bar{X}$ and S^2.

3.3.Estimators Based on Percentiles

This method was introduced by Kao (1958, 1959), which can be used when the data has a distribution function with closed form. The idea depends on estimating the unknown parameters by fitting straight line to the theoretical percentile points obtained from the distribution function and the sample percentile points. Kao (1958, 1959) found that this method can be useful in Weibull and exponential distributions. In this section we use the same technique for the MOP distribution.

Consider the cumulative distribution function (cdf) of MOP distribution $G(x;\gamma)=1-\frac{\alpha(\frac{\beta}{x})^{\theta}}{1-\bar{\alpha}(\frac{\beta}{x})^{\theta}}=1-\frac{\alpha}{(\frac{\beta}{x})^{-\theta}-\bar{\alpha}}$,

where $\gamma=(\alpha,\theta,\beta)$ therefore, $x=\beta\left[\frac{\alpha}{1-G(x;\gamma)}+\bar{\alpha})\right]^{1/\theta}$.

Let $X_{(i)}$ denote the i^{th} order statistics from a sample of size n. If p_i denote some estimate of $G(x_{(i)};\gamma)$, then the estimate of $\gamma=(\alpha,\theta,\beta)$ can be obtained by minimizing

$$\lambda=\sum_{i=1}^{n}\left(x_{(i)}-\beta\left[\frac{\alpha}{1-p_i}+\bar{\alpha}\right]^{1/\theta}\right)^2. \tag{15}$$

This method is used by several authors, see for example Bdair (2012), Gupta and Kundu (2001) and Kundu and Raqab (2005). Since Eq. (15) is a non-linear function, non-linear optimization technique can be used to find the minimum values of the needed estimators, these estimators are called percentile estimators (PCE's). It is possible to use several p_i's as estimators of $G(x_{(i)})$. For example $p_i=\frac{i}{n+1}$ is the most used estimator of $G(x_{(i)})$, as $\frac{i}{n+1}$ is the expected value of $G(x_{(i)})$. In this paper we also use $p_i=\frac{i}{n+1}$. Some of the other choices of p_i's are $p_i=(i-(3/8))/(n+(1/4))$ or $p_i=(i-(1/2))/n$, see Mann et. al (1974).

If we assume that α is known then the PCE of (θ,β) can be obtained by taking partial derivative with respect to θ and β. Let $\pi_i=\frac{\alpha}{1-p_i}+\bar{\alpha}$, then the equations that should be solved are:

$$\frac{\partial\lambda}{\partial\theta}=\frac{2\beta}{\theta^2}\sum_{i=1}^{n}\left(x_{(i)}-\beta(\pi_i)^{1/\theta}\right)(\pi_i)^{1/\theta}\times\ln(\pi_i)$$

$$\frac{\partial\lambda}{\partial\beta}=-2\sum_{i=1}^{n}\left(x_{(i)}-\beta(\pi_i)^{1/\theta}\right)(\pi_i)^{1/\theta}.$$

Solving these normal equations after equating them to zero will give

$$\hat{\beta}_{PCE} = \frac{\sum_{i=1}^{n} x_{(i)}}{\sum_{i=1}^{n} (\pi_i)^{1/\theta}} \qquad (16)$$

When θ is known, this gives the PCE of β denoted by $\hat{\beta}_{PCE}$. Substituting $\hat{\beta}_{PCE}$ in one of the above normal equations will give the PCE of θ ($\hat{\theta}_{PCE}$)

Now, if α is unknown then without loss of generality we may assume $\beta = 1,$ since it is a scale parameter. Let $G(x) = G(x;\alpha,\theta,1),$ then $\alpha = \frac{(1-G(x))}{G(x)}(x^{\theta}-1).$ The PCE of α ($\hat{\alpha}_{PCE}$) can be obtained by minimizing $\sum_{i=1}^{n}(\alpha - \frac{1-p_i}{p_i}(x_{(i)}^{\theta}-1))^2,$ with respect to α . So the PCE of α ($\hat{\alpha}_{PCE}$) is given by

$$\hat{\alpha}_{PCE} = \frac{\sum_{i=1}^{n} \frac{(x_{(i)}^{\theta}-1)(1-p_i)}{p_i}}{n} \qquad (17)$$

Similarly, when α is known, the PCE of θ is found to be

$$\hat{\theta}_{PCE} = \frac{\sum_{i=1}^{n} \frac{\ln(\frac{\alpha}{1-p_i}+\overline{\alpha})}{\ln x_i}}{n}. \qquad (18)$$

Interestingly, all the PCE's estimators have closed forms when assuming that the other parameters are known.

3.4.Least Squares Estimators and Weighted Least Squares Estimators

The method of least squares estimate or regression estimate was first suggested by Swain et. al (1988). It was used to estimate the parameters of beta distribution. The method can be described as follows: Suppose $Y_1, Y_2, ..., Y_n$ is a random sample of size n from a distribution function $G(.)$ and $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$ denote the order statistics of the observed sample. Let $G(Y_{(i)})$ be the distribution function of the i^{th} order statistics of the observed sample, then $G(Y_{(i)})$ has $Uniform(0,1)$ distribution. Therefore, we have

$$E(G(Y_{(i)})) = \frac{i}{n+1}, \qquad Var(G(Y_{(i)})) = \frac{i(n-i+1)}{(n+1)^2(n+2)}, \qquad (19)$$

and

$$Cov(G(Y_{(i)}), G(Y_{(j)})) = \frac{i(n-j+1)}{(n+1)^2(n+2)}, \quad for \quad i < j. \qquad (20)$$

One may refer to Johnson et. al (1995) for more details. The least squares estimators of the unknown parameters can be obtained by minimizing $\sum_{i=1}^{n}\left(G(Y_{(i)}) - \frac{i}{n+1}\right)^2,$ with respect to the unknown parameters $\alpha,\ \theta,\ \beta$, and hence minimizing

$$\sum_{i=1}^{n}\left(\frac{(\frac{\beta}{Y_{(i)}})^{-\theta}-1}{(\frac{\beta}{Y_{(i)}})^{-\theta}-\overline{\alpha}} - \frac{i}{n+1}\right)^2. \qquad (21)$$

Numerical solution of Eq. (21) can give the least squares estimate of the three unknown parameters denoted by $\hat{\alpha}_{LSE}$,

$\hat{\theta}_{LSE}$ and $\hat{\beta}_{LSE}$.

For the weighted least square estimators of the unknown parameters we need to minimize

$$\sum_{i=1}^{n} \frac{1}{Var(G(Y_{(i)}))} \left(G(Y_{(i)}) - \frac{i}{n+1} \right)^2,$$

and hence minimizing

$$\sum_{i=1}^{n} \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left(\frac{(\frac{\beta}{Y_{(i)}})^{-\theta} - 1}{(\frac{\beta}{Y_{(i)}})^{-\theta} - \bar{\alpha}} - \frac{i}{n+1} \right)^2, \tag{22}$$

with respect to α, θ and β . We can obtain $\hat{\alpha}_{WLSE}$, $\hat{\theta}_{WLSE}$ and $\hat{\beta}_{WLSE}$ respectively by numerical solution.

3.5. L-moment Estimators

In this section we use a method of estimating the unknown parameters of a MOP distribution based on the linear combination of order statistics, see David and Nagaraja (2003) and Hosking (1990). The estimators obtained by this method are well known as L-moment estimators (LME's). The LME's are similar to the usual moment estimators but can be estimated by linear combinations of order statistics. The LME's have certain advantages over usual moment estimators. It is observed that LME's are less subject to bias in estimation and sometimes more accurate in small samples than even the MLE's. Here, the idea is to equate the first three sample L-moments with the corresponding population L-moments. From Hosking (1990), we obtain the first, second and third sample L-moments as:

$$L_1 = \frac{1}{n} \sum_{i=1}^{n} x_{(i)}, \quad L_2 = \frac{2}{n(n-1)} \sum_{i=1}^{n} (i-1) x_{(i)} - L_1$$

$$\text{and } L_3 = \frac{1}{3} \binom{n}{3}^{-1} \sum_{i=1}^{n} \left\{ \binom{i-1}{2} - 2 \binom{i-1}{1} \binom{n-i}{1} + \binom{n-i}{2} \right\} X_{(i)}.$$

Similarly, the first three population L-moments are:

$$l_1 = E(X)$$

$$l_2 = \int_0^1 x(F)(2F-1)dF$$

$$l_3 = \int_0^1 x(F)(6F^2 - 6F + 1)dF,$$

where F denotes the (cdf) of the distribution under study. Now under MOP distribution and after using some integration techniques, the above population moments will reduce to:

$$l_1 = \alpha\theta\beta \sum_{j=1}^{n} \frac{\bar{\alpha}^{j} (j+1)}{(j+1)\theta - 1}$$

$$l_2 = 2\alpha\beta \left(\frac{\bar{\alpha}\alpha}{(\beta^{-\theta} - \bar{\alpha})^2} - \frac{1}{\bar{\alpha}} \right) - \alpha\beta \left(\log\left(\frac{\bar{\alpha} - \beta^{-\theta}}{\bar{\alpha}} \right) - \frac{\bar{\alpha}}{\beta^{-\theta} - \bar{\alpha}} \right) + 3\alpha\beta$$

$$l_3 = 6\alpha\beta \{ A \log\left(\frac{\bar{\alpha} - \beta^{-\theta}}{\bar{\alpha}} \right) + B \left(\frac{1}{\bar{\alpha} - \beta^{-\theta}} - \frac{1}{\bar{\alpha}} \right) + C \left(\frac{1}{2(\bar{\alpha} - \beta^{-\theta})^2} + \frac{1}{2\bar{\alpha}^{2}} \right) +$$

$$D \left(\frac{1}{3(\bar{\alpha} - \beta^{-\theta})^3} + \frac{1}{3\bar{\alpha}^{3}} \right) \} - 5\alpha\beta \{ \bar{\alpha}^{\frac{1}{\theta}-1} \log\left(\frac{\bar{\alpha} - \beta^{-\theta}}{\bar{\alpha}} \right) +$$

$$(\bar{\alpha}^{\frac{1}{\theta}-1} + 2\bar{\alpha}^{\frac{1}{\theta}}) \left(\frac{1}{\bar{\alpha} - \beta^{-\theta}} - \frac{1}{\bar{\alpha}} \right) + \frac{\bar{\alpha}^{\frac{1}{\theta}} \alpha}{2(\bar{\alpha} - \beta^{-\theta})^2} - \frac{\bar{\alpha}^{\frac{1}{\theta}} \alpha}{2\bar{\alpha}^{2}} \} +$$

$$\alpha\beta \{ \bar{\alpha}^{\frac{1}{\theta}-1} \log\left(\frac{\bar{\alpha} - \beta^{-\theta}}{\bar{\alpha}} \right) - \bar{\alpha}^{\frac{1}{\theta}-1} + \frac{\bar{\alpha}^{\frac{1}{\theta}}}{\bar{\alpha} - \beta^{-\theta}} \}$$

where

$$A=\frac{-2^{\frac{1}{\theta}}\bar{\alpha}+10\bar{\alpha}^{-\frac{1}{\theta}+1}-12\bar{\alpha}^{-\frac{1}{\theta}+2}+4\bar{\alpha}^{-\frac{1}{\theta}+3}-2\bar{\alpha}^{-\frac{1}{\theta}}}{2\bar{\alpha}(\bar{\alpha}-2)},$$

$$B=\frac{(2\bar{\alpha}-1)(-2^{\frac{1}{\theta}}\bar{\alpha}+16\bar{\alpha}^{-\frac{1}{\theta}+1}-14\bar{\alpha}^{-\frac{1}{\theta}+2}+4\bar{\alpha}^{-\frac{1}{\theta}+3}-6\bar{\alpha}^{-\frac{1}{\theta}}}{2\bar{\alpha}(\bar{\alpha}-2)},$$

$$C=\frac{(\bar{\alpha}-1)(-2^{\frac{1}{\theta}}\bar{\alpha}^{2}-12\bar{\alpha}^{-\frac{1}{\theta}+1}+22\bar{\alpha}^{-\frac{1}{\theta}+2}-164\bar{\alpha}^{-\frac{1}{\theta}+3}+4\bar{\alpha}^{-\frac{1}{\theta}+4}+4\bar{\alpha}^{-\frac{1}{\theta}}}{2\bar{\alpha}(\bar{\alpha}-2)},$$

and

$$D=\bar{\alpha}^{-\frac{1}{\theta}}(\bar{\alpha}-1)^{2}.$$

Therefore, the L-moment estimators can be obtained by numerical solution of the following system of equations:

$$L_1=l_1,\ L_2=l_2,\ L_3=l_3$$

4. BAYESIAN ESTIMATION

Bayesian estimation procedure of the parameters of a lifetime model has been studied in the literature extensively. Bayes estimation of parameters under generalized distributions was also studied. Singh et. al (2008) considered generalized exponential parameters, Preda et. al (2010) studied modified Weibull distribution and Singh et. al (2014) considered MO extended exponential parameters. Abdul Haq and Al-Omari (2016) considered the three component mixture of Rayleigh distribution.

In Bayesian method all parameters are considered as random variables with certain distribution called prior distribution. If prior information is not available which is usually the case, we need to select a prior distribution. Since the selection of prior distribution plays an important role in estimation of the parameters, our choice for the prior of α, θ and β are the independent gamma distributions i.e. $G(a_1,b_1)$, $G(a_2,b_2)$ and $G(a_3,b_3)$, respectively. The reason for choosing this prior density is that Gamma prior has flexible nature as a non-informative prior, especially when the values of the hyperparameters are assumed to be zero. Thus the suggested prior for α, θ and β are independent gamma $G(a_1,b_1)$, $G(a_2,b_2)$ and $G(a_3,b_3)$ distributions, respectively, which have the following densities:

$$g_{\alpha}(a_1,b_1)=\frac{b_1^{a_1}}{\Gamma(a_1)}\alpha^{a_1-1}e^{-b_1\alpha},$$

$$g_{\theta}(a_2,b_2)=\frac{b_2^{a_2}}{\Gamma(a_2)}\theta^{a_2-1}e^{-b_2\theta},$$

$$g_{\beta}(a_3,b_3)=\frac{b_3^{a_3}}{\Gamma(a_3)}\beta^{a_3-1}e^{-b_3\beta},$$

where a_1, a_2, a_3, b_1, b_2 and b_3, the hyperparameters of prior distributions, are positive real constants that reflect prior knowledge about α, θ and β.

The joint prior of α, θ and β is $g(\alpha,\theta,\beta)\propto\alpha^{a_1-1}\theta^{a_2-1}\beta^{a_3-1}e^{-b_1\alpha-b_2\theta-b_3\beta}$, $\alpha,\ \theta,\ \beta,\ a_1,\ a_2,\ a_3,\ b_1,\ b_2,\ b_3>0$.

The joint posterior density of α, θ and β is given by

$$p(\alpha,\theta,\beta\mid\underline{x})=\frac{L(\underline{x}\mid\alpha,\theta,\beta)g(\alpha,\theta,\beta)}{\int_{\alpha}\int_{\theta}\int_{\beta}L(\underline{x}\mid\alpha,\theta,\beta)g(\alpha,\theta,\beta)d\alpha d\theta d\beta},\qquad(23)$$

where $L(\underline{x}/\alpha,\theta,\beta)$ is the likelihood function of MOP distribution. Substituting $L(\underline{x}/\alpha,\theta,\beta)$ and $g(\alpha,\theta,\beta)$

for MOP distribution, the joint posterior density can be written as

$$p(\alpha,\theta,\beta\,|\,\underline{x}) \propto g_\alpha(n+a_1,b_1)g_\theta(n+a_2,b_2+\sum_{i=1}^{n}\ln x_i)g_\beta(n\theta+a_3,b_3)e^{-\Delta(\alpha,\theta,\beta)},$$

where $\Delta(\alpha,\theta,\beta)=\sum_{i=1}^{n}\ln x_i+2\sum_{i=1}^{n}\ln(1-\bar{\alpha}\left(\frac{\beta}{x_i}\right)^{\theta}$.

In the literature there are several approximation methods available to solve this kind of integration. Here we consider Lindley's and Monte Carlo Markov Chain (MCMC) approximation methods, see Lindley (1980) and Karandikar (2006), respectively. These approximation methods reduce the ratio of integrals into a whole and produce a single numerical result. The two methods are described below:

4.1 Bayes Estimation under Lindley's Approximation

This method is used to obtain the Bayes estimates (BEs) of the unknown parameters. In this method the posterior mean or expectation is expressed as the ratio of integral as follow:

$$E(\alpha,\theta,\beta\,|\,\underline{x})=\frac{\int_\alpha\int_\theta\int_\beta u(\alpha,\theta,\beta)e^{\ell(\alpha,\theta,\beta|\underline{x})+G(\alpha,\theta,\beta)}d\alpha d\theta d\beta}{\int_\alpha\int_\theta\int_\beta e^{\ell(\alpha,\theta,\beta|\underline{x})+G(\alpha,\theta,\beta)}d\alpha d\theta d\beta},$$

where $u(\alpha,\theta,\beta)$ is a function of α, θ and β only, $\ell(\alpha,\theta,\beta\,|\,\underline{x})$ is the log likelihood function and $G(\alpha,\theta,\beta)$ is the log of joint prior density.

According to Lindley (1980), if the maximum likelihood estimation of the parameters are available and n is sufficiently large, then the above ratio of integrals can be approximated as:

$$E(\alpha,\theta,\beta\setminus\underline{x})\approx u(\hat{\alpha},\hat{\theta},\hat{\beta})+\frac{1}{2}\sum_{i,j}^{m}(\hat{u}_{ij}+2\hat{u}_i\hat{e}_j)\hat{\sigma}_{ij}+\frac{1}{2}\sum_{i,j,k}^{m}\hat{\ell}_{ijk}\hat{u}_i\hat{\sigma}_{ij}\hat{\sigma}_{ki}, \tag{24}$$

where $\hat{\alpha}$, $\hat{\theta}$ and $\hat{\beta}$ are the MLEs of α, θ and β, respectively. $\hat{u}_i=\frac{\partial u(\hat{\alpha},\hat{\theta},\hat{\beta})}{\partial i}$, $\hat{u}_{ij}=\frac{\partial^2 u(\hat{\alpha},\hat{\theta},\hat{\beta})}{\partial i\partial j}$,

$\hat{e}_j=\frac{\partial G(\hat{\alpha},\hat{\theta},\hat{\beta})}{\partial j}$ and $\hat{\ell}_{ijk}=\frac{\partial^3\ell(\hat{\alpha},\hat{\theta},\hat{\beta})}{\partial i\partial j\partial k}$ are the partial derivatives with $i,j,k=1,2,3$, which indicates the number of parameters of distribution under study. The values of i,j and k refer to the parameters α, θ and β, respectively. In other words, if i,j or $k=1$ means that we take partial derivative with respect to α. For i,j or $k=2$, this means that we take partial derivative with respect to θ and finally if i,j or $k=3$ then take partial derivative with respect to β.

If α, θ and β are pair wise orthogonal then $\sigma_{ij}=0$ for $i\neq j$ and $\sigma_{ij}=(-\frac{1}{\ell_{ij}})$ for $i=j$. In MOP distribution if we assume the parameters are orthogonal then Eq. (24) reduces to

$$E(\alpha,\theta,\beta\setminus\underline{x})\approx u(\hat{\alpha},\hat{\theta},\hat{\beta})+\frac{1}{2}\sum_{i}^{m}(\hat{u}_{ii}+2\hat{u}_i\hat{e}_i)\hat{\sigma}_{ii}+\frac{1}{2}\sum_{i,j,k}^{m}\hat{\ell}_{iii}\hat{u}_i\hat{\sigma}_{ii}^2 \tag{25}$$

If $u(\alpha,\theta,\beta)=\alpha(\text{or }\theta,\beta)$, then $u_{\alpha(\text{or }\theta,\beta)}=1$ and the remaining partial derivatives are zeros. Consequently, it can be easily verified that:

$e_\alpha=\frac{a_1-1}{\alpha}-b_1$, $e_\theta=\frac{a_2-1}{\theta}-b_2$, $e_\beta=\frac{a_3-1}{\beta}-b_3$, $\ell_\alpha=\frac{n}{\alpha}-2\mu\sum_{i=1}^{n}\frac{1}{x_i^\theta-\bar{\alpha}\mu}$, $\ell_{\alpha\alpha}=\frac{-n}{\alpha^2}+2\mu^2\sum_{i=1}^{n}\frac{1}{(x_i^\theta-\bar{\alpha}\mu)^2}$,

$\ell_{\alpha\alpha\alpha}=\frac{2n}{\alpha^3}-4\mu^3\sum_{i=1}^{n}\frac{1}{(x_i^\theta-\bar{\alpha}\mu)^3}$, $\ell_\theta=\frac{n}{\theta}+\frac{n}{\theta}\log\mu-\sum_{i=1}^{n}\log(x_i)+2\bar{\alpha}\mu\sum_{i=1}^{n}\frac{\log(\frac{\mu^{\frac{1}{\theta}}}{x_i})}{x_i^\theta-\bar{\alpha}\mu}$, $\ell_{\theta\theta}=-\frac{n}{\theta^2}+2\bar{\alpha}\mu\sum_{i=1}^{n}\frac{x_i^\theta(\log(\frac{\mu^{\frac{1}{\theta}}}{x_i}))^2}{(x_i^\theta-\bar{\alpha}\mu)^2}$,

$$\ell_{\theta\theta\theta} = \tfrac{2n}{\theta^3} + 2\bar{\alpha}\mu \sum_{i=1}^{n} \frac{x_i^{\theta}(\log(\frac{\mu^{\frac{1}{\theta}}}{x_i}))^3 (x_i^{\theta} + \bar{\alpha}\mu)}{(x_i^{\theta} - \bar{\alpha}\mu)^3}, \quad \ell_{\beta} = n\theta\mu^{-\frac{1}{\theta}} + 2\bar{\alpha}\theta\mu^{1-\frac{1}{\theta}} \sum_{i=1}^{n} \frac{1}{x_i^{\theta} - \bar{\alpha}\mu},$$

$$\ell_{\beta\beta} = -n\theta\mu^{-\frac{2}{\theta}} + 2\bar{\alpha}\theta\mu^{1-\frac{2}{\theta}} \sum_{i=1}^{n} \frac{(\theta-1)x_i^{\theta} + \bar{\alpha}\mu}{(x_i^{\theta} - \bar{\alpha}\mu)^2},$$

and

$$\ell_{\beta\beta\beta} = 2n\theta\mu^{-\frac{3}{\theta}} + 4\bar{\alpha}\theta\mu^{1-\frac{3}{\theta}} \sum_{i=1}^{n} \frac{(\theta-1)x_i^{\theta} + \bar{\alpha}\mu}{(x_i^{\theta} - \bar{\alpha}\mu)^2} + 2\bar{\alpha}\theta^2\mu^{1-\frac{3}{\theta}} \sum_{i=1}^{n} \frac{(\theta-1)x_i^{\theta}(x_i^{\theta} + \bar{\alpha}\mu)}{(x_i^{\theta} - \bar{\alpha}\mu)^3}, \text{ where } \mu = \beta^{\theta}.$$

Using the above partial derivatives and substituting them in Eq. (25), we obtain the Bayes estimate for MOP parameters as follow:

$$\hat{\alpha}_L = \hat{\alpha} + \hat{u}_{\alpha}\hat{e}_{\alpha}\hat{\sigma}_{\alpha\alpha} + 0.5(\hat{u}_{\alpha}\hat{\sigma}^2{}_{\alpha\alpha}\hat{\ell}_{\alpha\alpha\alpha})$$

$$\hat{\theta}_L = \hat{\theta} + \hat{u}_{\theta}\hat{e}_{\theta}\hat{\sigma}_{\theta\theta} + 0.5(\hat{u}_{\theta}\hat{\sigma}^2{}_{\theta\theta}\hat{\ell}_{\theta\theta\theta})$$

$$\hat{\beta}_L = \hat{\beta} + \hat{u}_{\beta}\hat{e}_{\beta}\hat{\sigma}_{\beta} + 0.5(\hat{u}_{\beta}\hat{\sigma}^2{}_{\beta\beta}\hat{\ell}_{\beta\beta\beta})$$

4.2.Monte Carlo Markov Chain Approximation Method

We use Monte Carlo Markov Chain approximation method (MCMC) to evaluate the Bayes estimates (BEs) for the unknown parameters α , θ and β . The form of the BEs of α , θ and β under square loss function may not be obtained in explicit forms. Consequently, we choose for stochastic simulation procedures to generate samples from the posterior distribution of α , θ and β . Based on the joint posterior density given in Eq. (23), the Monte Carlo (MC) samples of α , θ and β are generated using the following algorithm:

1) Step 1: Generate α_1 from gamma distribution $f_{\alpha}(n + a_1, b_1)$.
2) Step 2: Generate θ_1 from gamma distribution $f_{\theta}(n + a_2, b_2 + \sum_{i=1}^{n} \ln x_i)$.
3) Step 3: Generate β_1 from gamma distribution $f_{\beta}(n\theta_1 + a_3, b_3)$.
4) Step 4: Repeat Steps 1-3 M times to obtain MC samples $(\alpha_i, \theta_i, \beta_i): \ i = 1,...,M$.
5) Step 5: We compute any function of α , θ and β (say $V(\alpha,\theta,\beta)$).

$$E(V(\alpha,\theta,\beta \mid \underline{x})) = \frac{E(V(\alpha,\theta,\beta)e^{-\Delta})}{E(e^{-\Delta})},$$

by averaging the numerator and denominator with respect to these simulations. This gives an estimate of α , θ and β .

5. SIMULATION STUDY AND DATA ANALYSIS

In this section, we divide our work into two subsections. In the first subsection, we perform some numerical computations based on artificial data while in the second subsection we apply these computations on a real life example.

5.1.Simulation Study

In this subsection, we compare the performances of the different estimators proposed in the previous sections using some numerical computations. We perform a simulation study to compare the performances of the different methods in the sense of bias and mean square error (MSE) for different sample sizes and for different parameter values. The generation of the MOP can be easily obtained through the transformation $X = \beta(\frac{\alpha}{1-U} + 1 - \alpha)^{1/\theta}$, where U is a uniform distribution deviateson $(0,1)$. Mathematica 7 and R codes are used for generating the MOP random variables and for solving the non-linear equations as well as for computing the minimization or maximization of the

related functions. The computations of the Bayes estimator of MOP's parameters depends on the second method of Bayesian estimation namely the MCMC method based on the improper prior $(a_1 = a_2 = b_1 = b_2 = c_1 = c_2 = 0)$. We postpone the estimation under many proper priors for the real life example. We do not use the Bayes estimation under Lindley's approximation because it is clearly very difficult and it needs to solve many complicated non linear equations.
Since β is the scale parameter and all the estimators are scale invariant, we take $\beta = 1$ in all cases considered. We consider various choices of the parameters α, θ and sample sizes $n = 10,\ 50$ and 100.
We compute the average relative biases and average relative MSE's over 10,000 runs. This number of runs will give the accuracy in the order $\pm(10,000)^{-5} = \pm 0.01$ (see Karian and Dudewicz (1999)). Therefore, we report all the results up to three decimal places.
First we consider the estimation of α when other parameters are known. If β and θ are known, the MLE's and PCE's of α can be obtained directly from (6) and (17) respectively. The MLE's of β and θ can be obtained directly from equations (7) and (8), respectively. The PCE's of β and θ can be computed directly from Equations (16) and (18), respectively. The MME's of all parameters can be obtained by solving the non-linear Equation (11) when other parameters are known. The LSE's and WLSE's can be obtained by minimizing (21) and (22), respectively, with respect to the needed parameter only. The BEs of the unknown parameters of the MOP distribution can be computed using the algorithm presented in section 4.2 based on Eq.(23). If $\hat{\alpha}$ is an estimate, then we present the average value of $(\hat{\alpha}/\alpha)$ and the average MSE of $(\hat{\alpha}/\alpha)$. The relative average bias and relative average MSE of $(\hat{\gamma}/\gamma)$, where $\hat{\gamma}$ is an estimate of γ, are defined, respectively, as follows:

$$\mathrm{Re}\,lativeBias(\hat{\gamma}/\gamma) = (1/k)\sum_{i=1}^{k}(\hat{\gamma}_i/\gamma)$$

and

$$\mathrm{Re}\,lativeMSE(\hat{\gamma}/\gamma) = (1/k\gamma^2)\sum_{i=1}^{k}(\hat{\gamma}_i - \rho)^2,$$

where $\rho = \gamma\,\mathrm{Re}\,lativeBias(\hat{\gamma}/\gamma)$ and k is the number of iterations. We are calculated the results for $\alpha = 0.2, 0.8$, for $\theta = 0.5, 1.0, 2.0$ and for $n = 10, 50$ and 100. The results are presented in 3 different tables. For each method, the average value of $(\hat{\alpha}/\alpha)$ is given in each box and the corresponding MSE is reported within parenthesis.
It is observed from Table 1 that all of the estimators usually overestimate α for small values of α. For large values of α, most of the estimators tend to be underestimates α for $\theta > 1$. It is also observed that all estimates decrease as the value of θ increase. One can also observe that for each estimation method, the average relative MSE's decreases as the sample size increases and also as the value of α increases.
It is observed from Tables 2 that all of the estimators usually overestimate β for all values of α and when $\theta \le 1$ for small sample size except the PCE's and LME's which are usually underestimate. For large sample size the estimators tends to be underestimate especially for large values of θ. It is also observed that all estimates decrease as the value of θ increase and as the sample size increase except the LME's which increase as the sample size increase. One can also observe that for each estimation method, the average relative MSE's decreases as the sample size increases and also as the values of α and θ increase.
It is observed from Tables 3 that all of the estimators usually underestimate θ for all values of α except MLE's that tend to be overestimate. All estimates tend to be underestimate for large values of θ and large sample sizes. It is also observed that all estimates decrease as the value of θ increases. One can also observe that for each estimation method, the average relative MSE's decreases as the sample size increases and also as the values of α and θ increase.
Among all classical methods of estimation, the MLE's provide the best results for all sample sizes. The WLSE's work better than the LSE's for all sample sizes and all values of α and θ. In the context of computational issues, the

MLE's, MME's and PCE's could not be easily implemented since they involve non-linear equations. The LSE's and WLSE's involve non-linear functions that should be minimized.

According to the bias and MSE's results reported in Tables 1-3, the Bayesian method of estimation provides the best results over all classical methods of estimation. It can be easily noticed that the MSE's results are better than that of all classical methods and also the bias results give a closer value than that of classical methods.

Table 1: Average relative estimators and average relative MSE's of α .

		(α, θ)					
n	*method*	(0:2, 0:5)	(0:2, 1:0)	(0:2, 2:0)	(0:8, 0:5)	(0:8, 1:0)	(0:8, 2:0)
10	MLE	1.201(0.319)	1.161(0.289)	0.978(0.212)	1.961(0.119)	1.115(0.088)	0.983(0.055)
	MME	1.780(0.598)	1.238(0.390)	0.882(0.284)	1.880(0.494)	1.793(0.224)	0.719(0.106)
	PCE	1.653(1.454)	1.297(0.674)	1.219(0.422)	1.205(0.478)	0.901(0.362)	0.866(0.253)
	LSE	2.461(0.499)	1.031(0.235)	0.913(0.218)	1.420(0.413)	1.341(0.304)	0.954(0.201)
	WLSE	2.343(0.361)	1.018(0.224)	0.821(0.215)	1.433(0.411)	1.352(0.303)	0.849(0.201)
	LME	1.366(1.686)	0.914(0.078)	0.915(0.077)	0.665(0.145)	0.669(0.143)	0.631(0.140)
	BE	1.113(0.301)	1.018(0.064)	0.921(0.061)	1.033(0.111)	1.042(0.033)	0.949(0.021)
50	MLE	1.093(0.174)	1.063(0.171)	0.943(0.149)	1.070(0.112)	0.975(0.074)	0.896(0.051)
	MME	1.629(0.257)	1.148(0.159)	0.938(0.143)	1.958(0.350)	1.906(0.208)	1.840(0.076)
	PCE	1.497(0.573)	1.205(0.385)	1.176(0.249)	1.097(0.371)	1.029(0.272)	0.923(0.148)
	LSE	2.419(0.348)	1.022(0.213)	0.911(0.207)	1.424(0.305)	1.326(0.302)	0.846(0.201)
	WLSE	2.255(0.319)	1.013(0.206)	0.731(0.204)	1.405(0.393)	1.349(0.282)	0.841(0.201)
	LME	1.363(1.212)	0.933(0.081)	0.936(0.074)	0.783(0.127)	0.758(0.123)	0.777(0.125)
	BE	1.103(0.261)	1.020(0.064)	0.927(0.057)	1.053(0.091)	1.032(0.031)	0.959(0.020)
100	MLE	1.042(0.108)	1.019(0.141)	0.917(0.132)	1.041(0.087)	1.011(0.068)	0.826(0.044)
	MME	1.543(0.213)	1.055(0.148)	0.917(0.135)	0.908(0.311)	1.897(0.185)	0.812(0.058)
	PCE	1.231(0.467)	1.250(0.291)	1.104(0.197)	1.069(0.309)	1.036(0.269)	0.929(0.142)
	LSE	2.404(0.331)	1.016(0.208)	0.810(0.204)	1.412(0.303)	1.322(0.302)	0.739(0.159)
	WLSE	2.217(0.310)	1.009(0.203)	0.678(0.202)	1.345(0.302)	1.253(0.201)	0.739(0.201)
	LME	0.946(0.076)	0.945(0.076)	0.943(0.072)	0.931(0.075)	0.941(0.078)	0.938(0.074)
	BE	1.024(0.095)	1.013(0.060)	0.929(0.042)	1.023(0.064)	1.012(0.028)	0.979(0.017)

Table 2: Average relative estimators and average relative MSE's of β .

(α, θ)

n	*method*	(0:2; 0:5)	(0:2; 1:0)	(0:2; 2:0)	(0:8; 0:5)	(0:8; 1:0)	(0:8; 2:0)
10	MLE	1.186(0.135)	1.096(0.118)	1.083(0.105)	1.175(0.080)	1.164(0.077)	1.065(0.071)
	MME	1.098(0.137)	1.076(0.109)	1.085(0.107)	1.060(0.081)	1.062(0.077)	1.058(0.072)
	PCE	0.982(0.223)	0.975(0.134)	0.958(0.082)	0.974(0.072)	0.972(0.069)	0.973(0.068)
	LSE	2.521(0.579)	1.075(0.172)	1.066(0.133)	1.364(0.114)	1.202(0.108)	1.110(0.105)
	WLSE	2.758(0.609)	1.052(0.161)	0.892(0.135)	1.430(0.118)	1.305(0.111)	1.206(0.108)
	LME	0.964(0.072)	0.963(0.071)	0.955(0.071)	0.968(0.070)	0.944(0.069)	0.947(0.068)
	BE	1.143(0.122)	1.018(0.112)	0.956(0.097)	1.123(0.067)	1.112(0.062)	0.987(0.060)
50	MLE	1.122(0.132)	1.116(0.120)	1.110(0.101)	1.114(0.075)	1.098(0.073)	1.012(0.071)
	MME	1.024(0.133)	1.019(0.121)	1.018(0.101)	1.013(0.081)	1.016(0.073)	1.010(0.071)
	PCE	0.960(0.146)	0.970(0.133)	0.978(0.084)	0.982(0.080)	0.985(0.078)	0.985(0.069)
	LSE	2.220(0.235)	1.011(0.128)	0.956(0.107)	1.380(0.102)	1.205(0.102)	1.101(0.102)
	WLSE	2.556(0.208)	1.010(0.125)	0.615(0.106)	1.428(0.104)	1.240(0.103)	0.935(0.102)
	LME	1.072(0.169)	0.983(0.060)	0.976(0.061)	0.967(0.064)	0.951(0.063)	0.950(0.061)
	BE	1.120(0.101)	1.064(0.094)	0.939(0.084)	1.103(0.057)	1.027(0.048)	0.980(0.041)
100	MLE	1.111(0.113)	1.092(0.110)	1.067(0.108)	1.104(0.107)	1.101(0.106)	1.001(0.106)
	MME	1.013(0.114)	1.013(0.110)	1.006(0.108)	1.004(0.077)	1.006(0.076)	1.005(0.066)
	PCE	0.966(0.116)	0.977(0.111)	0.977(0.078)	0.978(0.070)	0.983(0.066)	0.983(0.066)
	LSE	2.175(0.147)	1.015(0.114)	1.073(0.103)	1.303(0.101)	1.193(0.101)	1.108(0.101)
	WLSE	2.073(0.115)	1.014(0.113)	0.603(0.103)	1.409(0.102)	1.204(0.101)	0.678(0.101)
	LME	1.750(0.276)	0.994(0.056)	0.983(0.047)	0.988(0.017)	0.972(0.014)	0.959(0.010)
	BE	1.107(0.096)	1.024(0.085)	0.921(0.081)	1.003(0.051)	0.998(0.031)	0.973(0.029)

Table 3: Average relative estimators and average relative MSE's of θ.

		$(\alpha,\ \theta)$					
n	*method*	(0:2; 0:5)	(0:2; 1:0)	(0:2; 2:0)	(0:8; 0:5)	(0:8; 1:0)	(0:8; 2:0)
10	MLE	1.218(1.944)	1.061(0.075)	1.041(0.071)	1.043(0.070)	1.042(0.069)	1.044(0.069)
	MME	1.340(0.050)	1.293(0.106)	0.852(0.037)	0.656(0.026)	0.692(0.027)	0.791(0.011)
	PCE	1.386(1.658)	0.946(0.080)	0.972(0.080)	0.954(0.074)	0.534(0.196)	0.534(0.189)
	LSE	2.136(0.072)	1.029(0.018)	0.728(0.017)	0.543(0.021)	0.377(0.015)	0.311(0.013)
	WLSE	2.097(0.054)	1.016(0.014)	0.719(0.006)	0.534(0.041)	0.376(0.024)	0.310(0.013)
	LME	1.762(0.195)	0.987(0.039)	0.855(0.031)	0.634(0.010)	0.594(0.006)	0.584(0.005)
	BE	1.178(0.051)	1.059(0.022)	0.974(0.015)	1.033(0.011)	1.002(0.011)	0.992(0.009)
50	MLE	1.204(0.468)	1.039(0.061)	1.042(0.064)	1.043(0.060)	1.045(0.060)	1.043(0.061)
	MME	1.249(0.335)	1.194(0.111)	0.908(0.077)	0.666(0.046)	0.900(0.015)	0.753(0.010)
	PCE	0.923(0.075)	0.914(0.074)	0.904(0.072)	0.909(0.071)	0.921(0.073)	0.916(0.069)
	LSE	1.277(0.064)	1.008(0.052)	0.655(0.021)	0.514(0.021)	0.376(0.015)	0.310(0.012)
	WLSE	1.264(0.032)	1.004(0.031)	0.659(0.031)	0.507(0.041)	0.386(0.021)	0.319(0.011)
	LME	1.250(0.027)	0.999(0.031)	0.810(0.011)	0.651(0.008)	0.616(0.006)	0.588(0.004)
	BE	1.152(0.036)	1.023(0.019)	0.921(0.012)	1.019(0.008)	1.002(0.008)	0.949(0.007)
100	MLE	1.142(0.186)	1.042(0.054)	1.041(0.050)	0.998(0.059)	1.001(0.060)	1.002(0.060)
	MME	1.222(0.463)	1.160(0.113)	0.948(0.076)	0.706(0.049)	0.859(0.003)	0.754(0.011)
	PCE	0.993(0.243)	0.925(0.069)	0.924(0.066)	0.922(0.065)	0.865(0.087)	0.876(0.084)
	LSE	1.083(0.054)	1.004(0.042)	0.651(0.021)	0.508(0.021)	0.376(0.011)	0.308(0.011)
	WLSE	1.035(0.032)	1.002(0.032)	0.659(0.031)	0.504(0.041)	0.388(0.011)	0.321(0.011)
	LME	1.114(0.020)	0.997(0.019)	0.810(0.011)	0.652(0.005)	0.604(0.003)	0.580(0.002)
	BE	1.087(0.019)	1.009(0.014)	0.911(0.010)	1.008(0.008)	0.994(0.007)	0.915(0.005)

5.2.Data Analysis

In this section we have taken one real-life data set from Lawless (1982), and will use the MOP model to analyze the data. Many authors used this data to study some models related to Pareto model, see for example Sankaran and Kundu (2014). The data set consists of failure times or censoring times for 36 appliances subjected to an automated life test. Failures are mainly classified into 18 different modes, though among 33 observed failures only 7 modes are present and only model 6 and 9 appear more than once. We are mainly interested about the failure mode 9. The data are given below:
Data Set: 1167, 1925, 1990, 2223, 2400, 2471, 2551, 2568, 2694, 3034, 3112, 3214, 3478, 3504, 4329,

6976, 7846.
The MLEs of α , β and θ are computed numerically using the function optim in R statistical package. These MLEs, Kolmogorov-Smirnov (K-S) distance between the fitted and the empirical distribution functions and the corresponding p-values (between parentheses) are respectively: $\hat{\alpha} = 33.2920$, $\hat{\beta} = 1167$, $\hat{\theta} = 3.9630$ and $K - S = 0.1157(0.9786)$. When we use Pareto distribution with MLEs $\hat{\alpha} = 0.1252$, $\hat{\theta} = 1167$ to fit these data, we find that $K - S = 0.7881(9.5 \times 10^{-12})$. Therefore, this indicates that the MOP distribution fits the data set well and better than using Pareto distribution. This is also an example that proves the needs of new distributions in managing some sets of data.
Based on the results of the simulation study presented in the previous subsection that shows that the MLE is the best classical estimator for the three parameters for all cases. In Table 4, we present the average relative estimators and the average relative MSE's for the MLEs and BEs for the parameters α and θ. From Table 4, we can clearly notice that the parameters' estimations based on Bayesian approach are much better than that are based on the classical approaches.

Table 4: Average relative estimators and average relative MSE's of MLEs and BEs for α and θ.

method	α	β	θ
MLE	1.3218(1.944)	1.1713(0.079)	1.4915(1.031)
BE	1.1781(0.051)	1.1009(0.034)	0.9774(0.017)

To study the sensitivity of the variation in the specification of prior parameters on our Bayesian analysis, further MCMC simulations were undertaken using proper and improper priors. The proper priors on α , β and θ are chosen such that the prior mean of α is equal to 4 and its standard deviation is equal to 2, for β the prior mean is 2 with a standard deviation1,and that for θ, the prior mean is 1 with a standard deviation0.125. This corresponds to $a_1 = 4, a_2 = 8, b_1 = 4, b_2 = 2, c_1 = 1$ and $c_2 = 8$ (call it π_1). For the second proper prior (π_2) on the parameters, we assume that only small amount of prior information is available and assign small integer values to a_i and b_i. That is, we assume that $a_1 = a_2 = 2$, $b_1 = b_2 = 1$ and $c_1 = c_2 = 1$. The third prior (π_3) is limiting improper priors by setting $a_1 = a_2 = 0.5, b_1 = b_2 = 1, c_1 = c_2 = 0$. Table 5 presents the MCMC based posterior means and standard deviations for the model parameters. It can be seen from Table 5 that our results are sensitive to the assumed values of the prior parameters.

Table 5: Bayesian analysis using different priors.

Prior	π_1	π_2	π_3
$E(\alpha \mid data)$	33.7982	34.7326	35.2302
$Std(\alpha \mid data)$	1.5131	1.0453	1.6346
$E(\beta \mid data)$	1165.77	1166.78	1162.73
$Std(\beta \mid data)$	1.1043	1.2049	1.7059

$E(\theta \mid data)$	3.8751	3.7278	3.7398
$Std(\theta \mid data)$	2.1043	1.9048	1.5993

RECEIVED: DECEMBER, 2017
REVISED: JUNE, 2018

REFERENCES

[1] ABDUL HAQ, ALOMARI, A. (2016). Bayes Estimation and Prediction of a Three Component Mixture of Rayleih Distribution Under Type I Censoring, **Revista Investigation Operatcional** 37, 22-37.
[2] AHSANULLAH, M. (1992). Inference and prediction problems of the generalized Pareto distribution based on record values, In: P.K. Sen and LA. Salama (Eds.), **Order Statistics and Nonparametrics: Theory and Applications,**Elsevier Science Publishers, Amsterdam, 47-57.
[3] Alice, T. and Jose, K. K. (2004). Marshall-Olkin Pareto distributions and its reliability applications, IAPQR Trans. 29, 1-9.
[4] BDAIR, O. (2012). Different Methods of Estimation for Marshall Olkin Exponential Distribution, **Journal of Applied Statistical Science** 19, 13-29.
[5] CASTILLO, E., HADI, A.S., BALAKRISHNAN, N., and SARABIA, J.M. (2004). **Extreme Value And Related Models With Applications In Engineering and Science**, Wiley, New Jersey.
[6] DAVID, H. A. and NAGARAJA, H. N. (2003). **Order Statistics**. John Wiley and Sons, Inc., New York.
[7] GHITANY, M. E.(2005). Marshall-Olkin extended Pareto Distribution and its Application, **International Journal of Applied Mathematics** 18, 17-31.
[8] GHITANY, M. E., AL-HUSSAINI, E. K. and AL-JARALLAH, R. A.(2005). Marshall-Olkin Extended Weibull Distribution and its Application to Censored Data, **Journal of Applied Statistics** 32, 1025-1034.
[9] GHITANY, M. and E., KOTZ, S. (2007). Reliability Properties of Extended Linear Failure-Rate Distributions, **Probability in the Engineering and informational Sciences** 21, 441-450.
[10] GUPTA, R. and KUNDU, D. (2001). Generalized Exponential Distribution: Different Method of Estimations, **Journal of Statistical Computation and Simulation** 69, 315-337.
[11] GUPTA, R. C., GUPTA, R. D. and GUPTA, P.L. (1998). Modeling Failure Time Data by Lehman Alternatives, **Communications in Statistics: Theory and Methods** 27, 887-904.
[12] HOGG, R. V., MCKEAN, J. W. and CRAIG, A. T. (2005). **Introduction to Mathematical Statistics,** 6th ed. Pearson Prentice-Hall, New Jersey.
[13] HOSKING, J. R. M., (1990). L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, **Journal of the Royal Statistical Society**. Series B 52, 105-124.
[14] JOHNSON, N. L., KOTZ, S. and BALAKRISHNAN, N. (1995). **Continuous Univariate Distribution,** Vol. 1, 2nd Edition. Wiley, New York.
[15] JOSE, K. K. (2011). Marshall-Olkin Family of Distributions and their Applications in Reliability Theory, Time Series Modeling and Stress-Strength Analysis, **International Statistical Institute Proceeding, 58th World Statistical Congress,** Dublin. (Session CPS 005).
[16] JOSE, K. K. and ALICE, T. (2001). Marshall-Olkin Generalized Weibull Distributions and Applications, STARS, **International Journal** 2, 1-8.
[17] JOSE, K. K. and ALICE, T. (2005). Marshall-Olkin Family of Distributions. **Applications in Time Series Modeling and Reliability,**J.C Publications, Palakkad.
[18] JOSE, K. And K., UMA, P.(2009). On Marshall-Olkin Mittag-Leffler Distributions and Processes, **Far East Journal of Theoretical Statistics** 28,189-199.
[19] KAO, J. H. K. (1958). Computer Methods for Estimating Weibull Parameters in Reliability Studies, **Transactions of IRE-Reliability and Quality Control** 13, 15-22.
[20] KAO, J. H. K. (1959). A Graphical Estimation of Mixed Weibull Parameters in Life Testing Electron Tube, **Technometrics** 1, 389-407.
[21] KARIAN, Z. A. and DUDEWICZ, E. J. (1999). **Modern Statistical Systems and GPSS**

Simulations, 2nd edition. CRC Press, Florida.
[22] KARANDIKAR, R.L. (2006). On Markov Chain Monte Carlo (MCMC) Method, **Sadhana** 31, 81-104.
[23] KOTZ, S., and NADARAJAH, S.(2000). **Extreme value distributions: theory and applications**, Imperial College Press, London.
[24] KUMARASWAMY, P. (1980). A Generalized Probability Density Function for Double-Bounded Random Processes, **Journal of Hydrology** 46,79-88.
[25] KUNDU, D. and RAQAB, M. (2005). Generalized Rayleigh Distribution: Different Methods of Estimations, **Computational Statistics & Data Analysis** 49, 187-200.
[26] LAWLESS, J. F. (1982). **Statistical Models and Methods For Lifetime Data**. John Wiley & Sons, New York.
[27] LAI, C. D. (2013). Constructions and applications of lifetime distributions, **Appl. Stochastic Models Bus. Ind.** 29, 127-140.
[28] LINDLEY, D.V. (1980). Approximate Bayes Method, **Trabajos de Estadistica** 31, 223-237.
[29] MANN, N. R., SCHAFER, R. E. and SINGPURWALLA, N. D. (1974). **Methods for Statistical Analysis of Reliability and Life Data.** New York, Wiley.
[30] MARSHALL, A. W. and OLKIN, I. (1997). A New Method for Adding a Parameter to a Family of Distributions with Application to the Exponential and Weibull Families, **Biometrika** 84, 641-652.
[31] PICKANDS, J.(1975). Statistical Inference Using Extreme Order Statistics, **Annals of Statistics** 3, 119-131.
[32] PREDA, V., PANAITESCU, E., and CONSTANTINESCU, A. (2010). Bayes Estimators of Modified--Weibull Distribution Parameters Using Lindley's Approximation, **Wseas Transactions on Mathematics** 9.
[33] SANKARAN, P. G. and JAYAKUMAR, K.(2006). On Proportional Odds Model, **Statistical Papers** 49, 779-789.
[34] SANKARAN, P. G,. and KUNDU, D.(2014). On a Bivariate Pareto Model, **Statistics** 48, 241-255.
[35] SINGH, R., SINGH, S. K., SINGH, U., and SINGH, G. P. (2008). Bayes Estimator of the Generalized Exponential Parameters under Linex Loss Function Using Lindley's Approximation, **Data Science Journal** 7.
[36] SINGH, S. K., SINGH, U. andYADAV, A. S. (2014). Bayesian Estimation of Marshall-Olkin Extended Exponential Parameters Under Various Approximation Techniques, **Hacettepe Journal of Mathematics and Statistics** 43, 341-354.
[37] SWAIN, J., VENKATRAMAN, S. and WILSON, J. (1988). Least Squares Estimation of Distribution Function in Johnson's Translation System, **Journal of Statistical Computation and Simulation** 29, 271-297.