Generalized Rayleigh Distribution: Different Methods of Estimations

Debasis Kundu[1] Mohammad Z. Raqab[2]

Abstract

Recently Surles and Padgett (2001) introduced two-parameter Burr Type X distribution, which can also be described as generalized Rayleigh distribution. It is observed that this particular skewed distribution can be used quite effectively in analyzing lifetime data. Recently the authors (Raqab and Kundu; 2003) considered this distribution and discussed its different properties. In this paper, we mainly consider different estimators and compare their performance through Monte Carlo simulations.

Keywords: Maximum likelihood estimators; Modified moment estimators; Fisher Information matrix; Asymptotic distribution; Order statistics; Percentile based estimator, Least squares estimators; L-moment estimators.

Corresponding Author:[1] Professor D. Kundu, Department of Mathematics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, INDIA. Phone: 91-512-2597141, Fax: 91-512-2597500, e-mail: kundu@iitk.ac.in

Postal address [2] Department of Mathematics, University of Jordon Amman 11942, JORDON.

1 Introduction

Burr (1942) introduced twelve different forms of cumulative distribution functions for modeling lifetime data. Among those distributions, Burr Type X and Burr Type XII are the most popular ones. Several authors consider different aspects of the Burr Type X and Burr Type XII distributions, see for example Ahmad, Fakhry and Jaheen (1997), Jaheen (1995, 1996), Raqab (1998), Rodriguez (1977), Sartawi and Abu-Salih (1991), Surles and Padgett (1998) and Wongo (1993). For an excellent review for the two distributions the readers are refereed to Johnson, Kotz and Balakrishnan (1995).

Recently, Surles and Padgett (2001) (see also Surles and Padgett; 2004) introduced two-parameter Burr Type X distribution and correctly named as the generalized Rayleigh distribution. Note that the two-parameter generalized Rayleigh distribution is a particular member of the generalized Weibull distribution, originally proposed by Mudholkar and Srivastava (1993), see also Mudholkar, Srivastava and Freimer (1995). In this paper, we also prefer to call the two-parameter Burr Type X distribution as the generalized Rayleigh (GR) distribution. For $\alpha > 0$ and $\lambda > 0$, the two-parameter GR distribution has the distribution function;

$$F(x;\alpha,\lambda) = \left(1 - e^{-(\lambda x)^2}\right)^{\alpha}; \quad x > 0.$$

Therefore, GR distribution has the density function

$$f(x;\alpha,\lambda) = 2\alpha\lambda^2 x e^{-(\lambda x)^2}\left(1 - e^{-(\lambda x)^2}\right)^{\alpha-1}; \quad x > 0,$$

the survival function

$$S(x;\alpha,\lambda) = 1 - \left(1 - e^{-(\lambda x)^2}\right)^{\alpha} \quad x > 0,$$

and the hazard function

$$h(x;\alpha,\lambda) = \frac{2\alpha\lambda^2 x e^{-(\lambda x)^2}\left(1 - e^{-(\lambda x)^2}\right)^{\alpha-1}}{1 - \left(1 - e^{-(\lambda x)^2}\right)^{\alpha}}.$$

Here α and λ are the shape and scale parameters respectively. The two-parameter GR distribution will be denoted by $GR(\alpha,\lambda)$. It is observed by the authors (Raqab and Kundu; 2003) that for $\alpha \le \frac{1}{2}$, the probability density function (PDF) of a GR distribution is a decreasing function and it is a right skewed unimodal function for $\alpha > \frac{1}{2}$. Different forms of the density functions can be found in Raqab and Kundu (2003). It is also observed that the hazard function of a GR distribution can be either bathtub type or an increasing function, depending on the shape parameter α. For $\alpha \le \frac{1}{2}$, the hazard function of $GR(\alpha,\lambda)$ is bathtub type and for $\alpha > \frac{1}{2}$, it has an increasing hazard function. Surles and Padgett (2001) showed that the two-parameter GR distribution can be used quite effectively in modeling strength data and also modeling general lifetime data.

The main aim of this paper is to consider different estimators and study how the estimators of the different unknown parameter/ parameters behave for different sample sizes and for different parameter values. We mainly compare, the maximum likelihood estimators, the modified moment estimators, estimators based on percentiles, least squares estimators, weighted least squares estimators and the modified L-moment estimators by using extensive simulation techniques.

The rest of the paper is organized as follows. In Section 2, we briefly describe the MLE's and their implementations. In Sections 3 to 6, we describe other methods. Simulation results and discussions are provided in Section 7.

2 Maximum Likelihood Estimators:

In this section we consider the maximum likelihood estimators (MLE's) of $GR(\alpha, \lambda)$. First we consider the case when both α and λ are unknown. Let $x_1, \ldots, x_n$ be a random sample of size n from $GR(\alpha, \lambda)$, then the log-likelihood function $L(\alpha, \lambda)$ can be written as;

$$L(\alpha, \lambda) = C + nln\alpha + 2n\ln\lambda + \sum_{i=1}^{n} \ln x_i - \lambda^2 \sum_{i=1}^{n} x_i^2 + (\alpha - 1)\sum_{i=1}^{n} \ln\left(1 - e^{-(\lambda x)^2}\right). \tag{1}$$

The normal equations become;

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \ln\left(1 - e^{-(\lambda x)^2}\right) = 0, \tag{2}$$

$$\frac{\partial L}{\partial \lambda} = \frac{2n}{\lambda} - 2\lambda \sum_{i=1}^{n} x_i^2 + 2\lambda(\alpha - 1)\sum_{i=1}^{n} \frac{x_i^2 e^{-(\lambda x_i)^2}}{1 - e^{-(\lambda x)^2}} = 0. \tag{3}$$

From (2), we obtain the MLE of α as a function of λ, say $\hat{\alpha}(\lambda)$, as

$$\hat{\alpha}(\lambda) = -\frac{n}{\sum_{i=1}^{n} \ln\left(1 - e^{-(\lambda x)^2}\right)}. \tag{4}$$

Substituting $\hat{\alpha}(\lambda)$ in (1), we obtain the profile log-likelihood of λ as

$$g(\lambda) = L(\hat{\alpha}(\lambda), \lambda) = C + n\ln\left(-\sum_{i=1}^{n} \ln\left(1 - e^{-(\lambda x)^2}\right)\right) + 2n\ln\lambda$$
$$-\lambda^2 \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} \ln\left(1 - e^{-(\lambda x)^2}\right). \tag{5}$$

Therefore, the MLE of λ, say $\hat{\lambda}_{MLE}$, can be obtained by maximizing (5) with respect to λ. It can be shown that the maximum of (5) can be obtained as a fixed point solution of the following equation;

$$h(\mu) = \mu, \tag{6}$$

where

$$h(\mu) = \left[\frac{\sum_{i=1}^{n} \frac{x_i^2 e^{-\mu x_i^2}}{1-e^{-\mu x_i^2}}}{\sum_{i=1}^{n} \ln\left(1 - e^{-\mu x_i^2}\right)} + \frac{1}{n}\sum_{i=1}^{n} x_i^2 + \frac{1}{n}\sum_{i=1}^{n} \frac{x_i^2 e^{-\mu x_i^2}}{1 - e^{-\mu x_i^2}} \right]^{-1}.$$

If $\hat{\mu}$ is a solution of (6), $\hat{\lambda}_{MLE} = \sqrt{\hat{\mu}}$. Very simple iterative procedure $h(\mu^{(j)}) = \mu^{(j+1)}$, can be used, where $\mu^{(j)}$ is the j^{th} iterate. The iterative procedure works very well. Once $\hat{\lambda}_{MLE}$ is obtained, the MLE of α, say $\hat{\alpha}_{MLE}$, can be obtained from (4) as $\hat{\alpha}_{MLE} = \hat{\alpha}(\hat{\lambda}_{MLE})$. Note that, $\hat{\alpha}_{MLE}$ and $\hat{\lambda}_{MLE}$ are not in explicit form. Further, it is not possible to obtain the variances of $\hat{\alpha}_{MLE}$ and $\hat{\lambda}_{MLE}$. We propose to use the asymptotic normality results of $\hat{\alpha}_{MLE}$ and $\hat{\lambda}_{MLE}$, as derived in Surles and Padgett (2001) or Raqab and Kundu (2003). Using the notation

$$\xi(a, b) = \sum_{i=0}^{\infty} \frac{1}{(i + a)^b},$$

we obtain,

$$V(\hat{\alpha}_{MLE}) \approx \frac{I_{22}}{I_{11}I_{22} - I_{12}^2} \qquad \text{and} \quad V(\hat{\lambda}_{MLE}) \approx \frac{I_{11}}{I_{11}I_{22} - I_{12}^2},$$

where

$$\begin{aligned}
I_{11} &= -\frac{1}{\alpha^2} \\
I_{12} &= \frac{2}{(\alpha-1)\lambda^2}\left(\psi(\alpha) - \psi(1) - \frac{\alpha-1}{\alpha}\right) \quad \text{for} \quad \alpha \neq 1 \\
&= \frac{2}{\lambda^2}\,\xi(2,2) \ \text{ for } \quad \alpha = 1 \\
I_{22} &= -\frac{2}{\lambda^2} - \frac{2}{\lambda^2}\left[\psi(\alpha+1) - \psi(1)\right] - \frac{2\alpha}{\lambda^3}\left(\psi(1) - \psi(\alpha)\right) - 2\frac{\alpha-1}{\lambda^3} \\
&\quad -\frac{4\alpha}{\lambda^3(\alpha-2)}\left((\psi(2) - \psi(\alpha))^2 + \psi'(2) - \psi'(\alpha)\right), \quad \text{for} \quad \alpha \neq 2 \\
&= \frac{-2}{\lambda^2}\left[1 + \psi(\alpha+1) - \psi(1)\right] + \frac{4}{\lambda^3}\left[\xi(2,2) - 4\xi(3,2) - \xi(2,3)\right] \quad \text{for} \quad \alpha = 2.
\end{aligned}$$

Now consider the MLE of α, when the scale parameter λ is known. Without loss of generality, we can assume that $\lambda = 1$. If λ is known, the MLE of α, say $\hat{\alpha}_{MLESCK}$ is

$$\hat{\alpha}_{MLESCK} = -\frac{n}{\sum_{i=1}^{n} \ln(1 - e^{-x_i^2})}. \tag{7}$$

Now note that, if X_i's are independent and identically distributed $GR(\alpha, 1)$, then - $\alpha \sum_{i=1}^{n} \ln(1-e^{-X_i^2})$ follows gamma random variable with shape parameter n and scale parameter 1. Therefore, for $n > 2$,

$$E\left(\hat{\alpha}_{MLESCK}\right) = \frac{n}{n-1}\alpha \qquad \text{and} \quad V\left(\hat{\alpha}_{MLESCK}\right) = \frac{n^2}{(n-1)^2(n-2)}\alpha^2.$$

Using (7), an unbiased estimate of α can be easily obtained as

$$\hat{\alpha}_{USCK} = \frac{n-1}{n}\hat{\alpha}_{MLESCK} = -\frac{n-1}{\sum_{i=1}^{n} \ln(1 - e^{-x_i^2})}, \tag{8}$$

where $V(\hat{\alpha}_{USCK}) = \frac{\alpha^2}{n-2}$.

Let us consider the MLE of λ, say $\hat{\lambda}_{MLESHK}$, when the shape parameter α is known. For known α, $\hat{\lambda}_{MLESHK}$ can be obtained by maximizing

$$u(\lambda) = 2n \ln \lambda - \lambda^2 \sum_{i=1}^{n} x_i^2 + (\alpha - 1) \sum_{i=1}^{n} \ln(1 - e^{-(\lambda x_i)^2}), \tag{9}$$

with respect to λ. Now $\hat{\lambda}_{MLESHK}$, which maximizes $u(\lambda)$ can be obtained as the positive square root of $v(\lambda^2) = \lambda^2$, where

$$v(\lambda^2) = \frac{2}{n} \left[\sum_{i=1}^{n} x_i^2 - (\alpha - 1) \sum_{i=1}^{n} \frac{e^{-(\lambda x_i)^2} x_i^2}{1 - e^{-(\lambda x_i)^2}} \right]^{-1}.$$

From the asymptotic properties of the MLE, it follows that

$$E(\hat{\lambda}_{MLESHK}) \approx \lambda, \qquad \text{and} \qquad Var(\hat{\lambda}_{MLESHK}) \approx \frac{1}{n} I_{22}^{-1}.$$

3 Modified Moment Estimators:

It is observed by Raqab and Kundu (2003), that the moments of a GR distribution can not be expressed in a nice form. Therefore, it may not be possible to compute the moment estimators very easily. But it is observed in Raqab and Kundu (2003) that if X follows $GR(\alpha, \lambda)$, then

$$E(X^2) = \frac{1}{\lambda^2} \left(\psi(\alpha + 1) - \psi(1) \right), \qquad \text{and} \quad E(X^4) - (E(X^2))^2 = \frac{1}{\lambda^4} \left(\psi'(1) - \psi'(\alpha + 1) \right).$$

Here $\psi(.)$ and $\psi'(.)$ denote the digamma and polygamma functions respectively. We propose the following modified moment estimators (MME's). First let us define, U and V as follows;

$$U = \frac{1}{n} \sum_{i=1}^{n} x_i^2, \qquad V = \frac{1}{n} \sum_{i=1}^{n} x_i^4 - U^2.$$

The MME of α can be obtained as the solution of the following non-linear equation;

$$\frac{V}{U^2} = \frac{\psi'(1) - \psi'(\alpha + 1)}{(\psi(\alpha + 1) - \psi(1))^2}. \tag{10}$$

We denote the estimate of α as $\hat{\alpha}_{MME}$. Once $\hat{\alpha}_{MME}$ is obtained, we obtain the MME of λ, say $\hat{\lambda}_{MME}$, as

$$\hat{\lambda}_{MME} = \sqrt{\frac{\psi(\hat{\alpha}_{MME} + 1) - \psi(1)}{U}}. \tag{11}$$

It is not possible to obtain the exact variances of $\hat{\alpha}_{MME}$ or $\hat{\lambda}_{MME}$. The asymptotic variances of $\hat{\alpha}_{MME}$ and $\hat{\lambda}_{MME}$ can be obtained from the following asymptotic property of $\hat{\alpha}_{MME}$ and $\hat{\lambda}_{MME}$:

$$\left[\sqrt{n}(\hat{\alpha}_{MME} - \alpha), \sqrt{n}(\hat{\lambda} - \lambda) \right] \rightarrow N_2[\mathbf{0}, \Sigma],$$

where Σ is a 2×2 matrix, and it can be expressed as follows;

$$\Sigma = \mathbf{D}\mathbf{A}^{-1}\mathbf{C}\mathbf{A}^{-1}\mathbf{D},$$

where

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ & \\ 0 & \lambda \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} -\psi'(\alpha+1) & \psi''(\alpha+1) \\ & \\ \psi(\alpha+1) - \psi(1) & 2(-\psi'(\alpha+1) + \psi'(1)) \end{bmatrix}.$$

Also,

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ & \\ c_{21} & c_{22} \end{bmatrix},$$

where

$$c_{11} = \psi'(1) - \psi'(\alpha+1), \qquad c_{12} = c_{21} = \psi''(\alpha+1) - \psi'(1)$$

and

$$c_{22} = (\psi^{(3)}(1) - \psi^{(3)}(\alpha+1)) - (\psi'(1) - \psi'(\alpha+1))^2.$$

If the scale parameter is known, (without loss of generality, we can assume it to be 1), then the MME of α, say $\hat{\alpha}_{MMESCK}$, can be obtained by solving the non-linear equation

$$U = (\psi(\alpha) - \psi(1)).$$

The asymptotic variance of $\hat{\alpha}_{MMESCK}$ is

$$V(\hat{\alpha}_{MMESCK}) \approx \frac{\psi'(1) - \psi'(\alpha+1)}{n(\psi'(\alpha+1))^2}.$$

If the shape parameter α is known, the MME of λ, say $\hat{\lambda}_{MMESHK}$, is

$$\hat{\lambda}_{MMESHK} = \frac{\sqrt{\psi(\alpha+1) - \psi(1)}}{\sqrt{U}} \qquad \text{and} \qquad V(\hat{\lambda}_{MMESHK}) \approx \frac{1}{4n} \frac{\psi'(1) - \psi'(\alpha+1)}{\sqrt{\psi(\alpha+1) - \psi(1)}}.$$

4 Estimators Based on Percentiles

If the data come from a distribution function which has a closed form, then it is quite natural to estimate the unknown parameters by fitting straight line to the theoretical percentile points obtained from the distribution function and the sample percentile points. This method was originally proposed by Kao (1958, 1959) and it has been used quite successfully for Weibull distribution and for the generalized exponential distribution. In this paper, we apply the same technique for the GR distribution.

First let us consider the case when both parameters are unknown. Since

$$F(x;\alpha,\lambda) = \left(1 - e^{-(\lambda x)^2}\right)^{\alpha};$$

therefore

$$-\frac{1}{\lambda^2}\ln\left[1 - (F(x;\alpha,\lambda))^{\frac{1}{\alpha}}\right] = x^2.$$

Let us denote $X_{(i)}$ as the i-th order statistic, *i.e*, $X_{(1)} < \ldots < X_{(n)}$. If p_i denotes some estimate of $F(x_{(i)};\alpha,\lambda)$, then the estimate of α and λ can be obtained by minimizing

$$\sum_{i=1}^{n}\left[x_{(i)}^2 + \frac{1}{\lambda^2}\ln\left(1 - p_i^{\frac{1}{\alpha}}\right)\right]^2, \tag{12}$$

with respect to α and λ. Note that (12) is a non-linear function and it has to be minimized using some non-linear optimization technique. We call the corresponding estimators as the percentile estimators or PCE's. Several estimators of p_i can be used here, see for example Mann, Schafer and Singpurwalla (1974). In this paper, we mainly consider $p_i = \frac{i}{n+1}$, which is the expected value of $F(X_{(i)})$.

When α is known, then the PCE of λ, say $\hat{\lambda}_{PCESHK}$, becomes

$$\hat{\lambda}_{PCESHK} = \sqrt{-\frac{\sum_{i=1}^{n}\left(\ln(1-p_i^{\frac{1}{\alpha}})\right)^2}{\sum_{i=1}^{n} x_{(i)}^2 \ln(1-p_i^{\frac{1}{\alpha}})}}. \tag{13}$$

Interestingly, $\hat{\lambda}_{PCESHK}$ has an explicit form unlike the corresponding MLE.

Now let us consider the case when the scale parameter λ is known. Without loss of generality, we can assume $\lambda = 1$. If we denote $F(x;\alpha) = F(x;\alpha,1)$ then

$$\ln F(x;\alpha) = \alpha\ln(1 - e^{-x^2}).$$

Therefore, the PCE of α, say $\hat{\alpha}_{PCESCK}$, can be obtained by minimizing

$$\sum_{i=1}^{n}\left(\ln p_i - \alpha\ln(1 - e^{-x_{(i)}^2})\right)^2,$$

with respect to α and it is

$$\hat{\alpha}_{PCESCK} = \frac{\sum_{i=1}^{n}\ln p_i \ln(1 - e^{-x_{(i)}^2})}{\sum_{i=1}^{n}\left[\ln(1 - e^{-x_{(i)}^2})\right]^2}. \tag{14}$$

See for, example Mann, Schafer and Singpurwalla (1974) for a detailed discussion on this.

5 Least Squares and Weighted Least Squares Estimators:

The least squares estimators and weighted least squares estimators were originally proposed by Swain, Venkataraman and Wilson (1988) to estimate the parameters of Beta distributions. The method can be described as follows: Suppose $Y_1, \ldots, Y_n$ is a random sample of size n from a distribution function $G(.)$ and $Y_{(1)} < \ldots, Y_{(n)}$ denote the order statistics of the observed sample. It is well known that

$$E(G(Y_{(j)})) = \frac{j}{n+1} \quad \text{and} \quad V(G(Y_{(j)})) = \frac{j(n-j+1)}{(n+1)^2(n+2)}. \tag{15}$$

The proposed method use (15) and obtain estimates of the unknown parameters.

The least squares estimators of the unknown parameters can be obtained by minimizing

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

with respect to the unknown parameters. Therefore, in this case, the least squares estimators of α and λ, say $\hat{\alpha}_{LSE}$ and $\hat{\lambda}_{LSE}$ respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} \left(\left[1 - e^{-(\lambda x_{(j)})^2}\right]^{\alpha} - \frac{j}{n+1} \right)^2, \tag{16}$$

with respect to α and λ.

The weighted least squares estimators of the unknown parameters can be obtained by minimizing

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

with respect to the unknown parameters. Therefore, in this case, the weighted least squares estimators of α and λ, say $\hat{\alpha}_{WLSE}$ and $\hat{\lambda}_{WLSE}$ respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} \frac{(n+1)^2(n+2)}{j(n-j+1)} \left(\left[1 - e^{-(\lambda x_{(j)})^2}\right]^{\alpha} - \frac{j}{n+1} \right)^2, \tag{17}$$

with respect to α and λ.

6 Modified L-Moment Estimators

In this section, we propose a method of estimating the unknown parameters based on the linear combination of order statistics. This method is popularly known as the L-moment estimators or L-estimators. The readers are refereed to David (1981) or Hosking (1990). It is observed (Hosking; 1990) that the L-moment estimators have certain advantages over the conventional moment estimators.

The standard method to compute the L-moment estimators is to equate the sample L-moments with the population L-moments. When both parameters are unknown we need to equate the first two sample L-moments with the population L-moments. It is possible to express the L-moments in terms of the ordinary moments, see for example Hosking (1990). Unfortunately, for GR distribution all the moments can not be expressed in a compact form. Since the even moments can be expressed in terms of the $\psi(.)$ function, we propose the following modified L-moment estimators (MLME's), similar to the MME's as follows. Transform the data $y_i = x_i^2$, for $i = 1, \ldots, n$. If $y_{(1)} < \ldots < y_{(n)}$ denote the ordered transformed sample, then using the same notation as Hosking (1990), we obtain the first and second sample L-moments as

$$l_1 = \frac{1}{n}\sum_{i=1}^{n} y_{(i)} \qquad \text{and} \qquad l_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n}(i-1)y_{(i)} - l_1. \tag{18}$$

Similarly, the first two population L-moments of the transformed random variable $Y = X^2$ are

$$\lambda_1 = \frac{1}{\lambda}\left(\psi(\alpha+1) - \psi(1)\right) \qquad \text{and} \qquad \lambda_2 = \frac{1}{\lambda}\left(\psi(2\alpha+1) - \psi(\alpha+1)\right), \tag{19}$$

respectively. Note that Y follows GE(α, λ), therefore, (19) follows from the properties of the GE distribution (see Gupta and Kundu; 2001). Therefore, MLME's can be obtained by solving the following two equations;

$$l_1 = \frac{1}{\lambda}\left(\psi(\alpha+1) - \psi(1)\right) \qquad \text{and} \qquad l_2 = \frac{1}{\lambda}\left(\psi(2\alpha+1) - \psi(\alpha+1)\right).$$

First we obtain the MLME of α, say $\hat{\alpha}_{MLME}$, as the solution of the following non-linear equation;

$$\frac{\psi(2\alpha+1) - \psi(\alpha+1)}{\psi(\alpha+1) - \psi(1)} = \frac{l_2}{l_1}. \tag{20}$$

Once $\hat{\alpha}_{MLME}$ is obtained from the non-linear equation (20), we compute the MLME of λ, say $\hat{\lambda}_{MLME}$, as

$$\hat{\lambda}_{MLME} = \frac{\psi(\hat{\alpha}_{MLME}+1) - \psi(1)}{l_1}.$$

Note that if $\alpha(\lambda)$ is known, the MLME of λ (α) is same as the corresponding MME, obtained in Section 3.

7 Numerical Experiments and Discussions

In this section we present results of some numerical experiments to compare the performance of the different estimators proposed in the previous sections. We perform extensive Monte Carlo simulations to compare the performance of the different estimators, mainly with respect to their biases and mean squared errors (MSE's) for different sample sizes and for different parameter values. Note that, λ is the scale parameter and all the estimators are scale invariant. Therefore, we take $\lambda = 1$ in all cases considered. We consider α = 0.25, 0.5, 1.0, 2.0, 2.5 and n = 10, 20, 30, 50 and 100. We compute the average relative estimates and average relative MSE's over 10,000 replications. Note that 10,000 replications will give the accuracy in the order $\pm$ (10,000)$^{-5}$ = $\pm$ 0.01 (Karian and Dudewicz; 1999).

For scarcity of space we report the results only for $n = 20$ and 50. The other results can be obtained from the corresponding author on request.

First consider the estimation of α when λ is known. When λ is known the MLE, unbiased estimator (UBE) and PCE of α can be obtained from (7), (8) and (14) respectively. The modified moment estimator of α can be obtained by solving the non-linear equation (10). Similarly the least squares and weighted least squares estimators of α can be obtained by minimizing (16) and (17) respectively with respect to α only. The results are reported in Table 1.

It is observed in Table 1 that most of the estimators usually overestimate α, except PCE, which underestimates all the times. As far as biases are concerned, the UBE's are more or less unbiased as expected and considering the minimum MSE's, the PCE's produce the best results for all sample sizes and for different values of α considered here. The MSE's of the UBE's are also quite close to the PCE's.

In the context of computational complexities, UBE, MLE and PCE are easiest to compute. They do not involve any non-linear equation solving, whereas the MME, LSE and WLSE involve solving non-linear equations and they need to be calculated by some iterative processes. Comparing all the methods, we conclude that for known scale parameter, the UBE should be used for estimating α.

Now consider the estimation of λ when α is known. In this case the MLE of λ can be obtained by maximizing (9) with respect to λ. The MME and the PCE of λ can be obtained from (11) and (13) respectively. Finally the LSE and the WLSE can be obtained by minimizing (16) and (17) with respect to λ only for fixed α. The results are reported in Table 2.

In this case also it is observed that except PCE, other estimators overestimate λ. The PCE under estimates λ in most considered cases, except when the sample size is very small and $\alpha \leq 0.5$. Comparing the biases of all the estimators, it is observed that the PCE performs the best for small sample sizes ($n \leq 10$) and for moderate or large sample sizes it works the best when $\alpha \leq 0.5$. For moderate or large sample sizes WLSE has the minimum biases when $\alpha > 0.5$. The performance of the MLE's are quite close to the WLSE for large sample sizes. As far as MSE's are concerned, MME outperforms others for all cases considered here. Comparing the computational complexities of the different estimators, it is observed that when the shape parameter is known, PCE and MME can be computed directly where as some iterative techniques are needed to compute MLE, LSE and WLSE. Summing up, we recommend to use the MME for estimating λ when the shape parameter is known.

Now we consider the estimation of α and λ when both of them are unknown. We consider six different estimators, namely MLE, MME, PCE, LSE, WLSE and MLME. The results for α and λ are reported in Tables 3 and 4 respectively. For better understanding, the average relative biases and the relative MSE's of the different estimators of λ is presented in Figures 1 - 2 when sample size is 30.

The performance of most of the methods are quite bad when the sample size is very small ($n = 10$). In particular, the estimation of α becomes very difficult for small sizes. It may be mentioned that for all the methods the biases are quite severe particularly for small or moderate sample sizes ($n \leq 20$). Some bias correction technique like jackknifing may be used for practical purposes, although it is not pursued here.

Comparing the performance of all the estimators, it is observed that as far as biases are concerned, the LSE performs quite well for $n \leq 20$. For $n \geq 30$, the WLSE outperforms LSE marginally. Considering the MSE's, the WLSE performs better than the rest in most cases considered. Interestingly, while estimating λ, the biases and MSE's of the WLSE are lower than the other estimators most of the times. Computationally, the LSE and WLSE involve two dimensional optimization whereas rest of the estimators involve only one dimensional optimization. Even though, the WLSE can be obtained by performing two dimensional optimization, we recommend to use it for estimating α and λ when both are unknown.

ACKNOWLEDGMENTS: The authors would like to thank the Co-Editor, Professor Erricos John Kontoghiorghes, the Associate Editor, and two referees for their helpful comments.

References

[1] Ahmad, K.E., Fakhry, M.E. and Jaheen, Z.F. (1997), "Empirical Bayes estimation of $P(Y < X)$ and characterization of Burr-type X model", *Journal of Statistical Planning and Inference*, vol. 64, 297-308.

[2] Burr, I.W. (1942), "Cumulative frequency distribution", *Annals of Mathematical Statistics*, vol. 13, 215-232.

[3] David, H.A. (1981), *Order Statistics*, 2nd Edition, Wiley, New York.

[4] Gupta, R.D. and Kundu, D. (2002), "Generalized exponential distribution: different methods of estimation", *Journal of Statistical Computation and Simulation*, vol. 59, 315-337.

[5] Hosking, J.R.M. (1990), "L-Moment: Analysis and estimation of distributions using linear combinations of order statistics", *Journal of Royal Statistical Society*, Ser. B, vol. 52, 105-124.

[6] Jaheen. Z.F. (1995), "Bayesian approach to prediction with outliers from the Burr type X model", *Microelectron. Rel.*, vol. 35, 45-47.

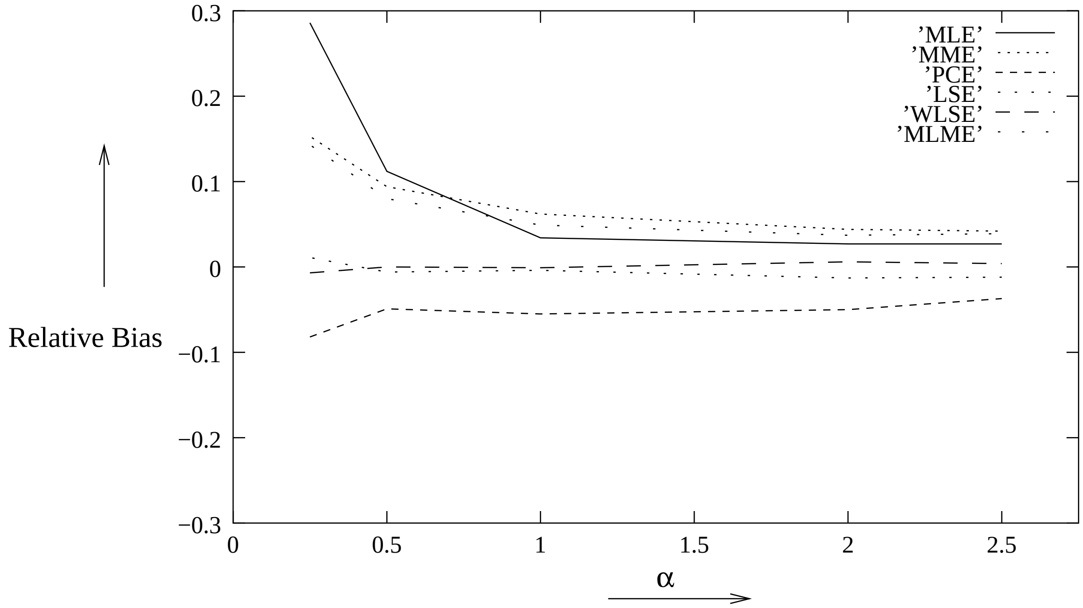

Figure 1: Average relative biases of the different estimators of λ for sample size 30

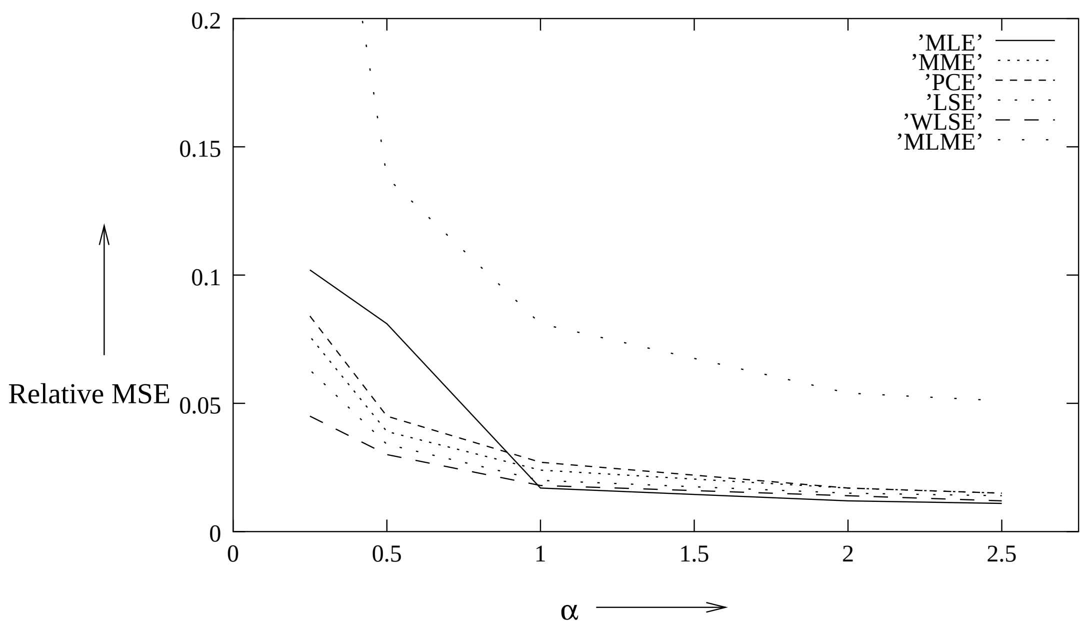

Figure 2: Average relative MSE's of λ when the sample size is 30

[7] Jaheen. Z.F. (1996), "Empirical Bayes estimation of the reliability and failure rate functions of the Burr type X failure model", *Journal of Applied Statistical Science*, vol. 3, 281-288.

[8] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distribution Vol. 1*, 2nd Ed., New York, Wiley.

[9] Kao, J.H.K. (1958), "Computer methods for estimating Weibull parameters in reliability studies", *Transactions of IRE-Reliability and Quality Control*, vol. 13, 15-22.

[10] Kao, J.H.K. (1959), "A graphical estimation of mixed Weibull parameters in life testing electron tube", *Technometrics*, vol. 1, 389-407.

[11] Karian, Z.A. and Dudewicz, E.J. (1999), *Modern Statistical Systems and GPSS Simulations*, 2nd Edition, CRC Press, Florida.

[12] Mann, N.R., Schafer, R. E. and Singpurwalla, N.D. (1974), *Methods for Statistical Analysis of Reliability and Life Data*, Wiley, New York.

[13] Mudholkar, G.S. and Srivastava, D.K. (1993), " Exponentiated Weibull family for analyzing bathtub failure data", *IEEE Transactions on Reliability*, vol. 42, 299-302.

[14] Mudholkar, G.S., Srivastava, D.K. and Freimer, M. (1995), " The exponentiated Weibull family; a re-analysis of the bus motor failure data", *Technometrics*, vol. 37, 436-445.

[15] Raqab, M.Z. (1998), "Order statistics from the Burr type X model", *Computers Mathematics and Applications*, vol. 36, 111-120.

[16] Raqab, M.Z. and Kundu, D. (2003), "Burr Type X distribution; revisited", submitted for publication.

[17] Rodriguez, R.N. (1977), "A guide to Burr Type XII distributions", *Biometrika*, vol. 64, 129-134.

[18] Sartawi, H.A. and Abu-Salih, M.S. (1991), "Bayes prediction bounds for the Burr type X model", *Communications in Statistics Theory and Methods*, vol. 20, 2307-2330.

[19] Surles, J.G. and Padgett, W.J. (1998), "Inference for $P(Y < X)$ in the Burr Type X model", *Journal of Applied Statistical Science*, 225-238.

[20] Surles, J.G. and Padgett, W.J. (2001), "Inference for reliability and stress-strength for a scaled Burr Type X distribution", *Lifetime Data Analysis*, vol. 7, 187-200.

[21] Surles, J.G. and Padgett, W.J. (2004), "Some properties of a scaled Burr type X distribution", to appear in the *Journal of Statistical Planning and Inference*.

[22] 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*, vol. 29, 271-297.

[23] Wongo, D.R. (1993), "Maximum likelihood methods for fitting the Burr Type XII distribution to multiply (progressively) censored life test data", *Metrika*, vol. 40, 203-210.

Table 1: Average relative estimates and average relative MSE's of α are presented, when λ is known. The average relative MSE's are reported within brackets against each average relative estimates.

n	Method	$\alpha = 0.25$	$\alpha = 0.50$	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 2.5$
20	MLE	1.053(0.071)	1.052(0.069)	1.053(0.064)	1.053(0.063)	1.051 (0.063)
	MME	1.043(0.293)	1.039(0.193)	1.042(0.147)	1.044(0.122)	1.039(0.116)
	PCE	0.926(0.055)	0.926(0.055)	0.928(0.058)	0.928(0.058)	0.926(0.055)
	LSE	1.062(0.085)	1.034(0.079)	1.047(0.081)	1.039(0.080)	1.045(0.090)
	WLSE	1.065(0.115)	1.043(0.076)	1.030(0.073)	1.015(0.069)	1.046(0.082)
	UBE	1.001(0.061)	1.000(0.060)	1.000(0.055)	1.000(0.055)	1.000 (0.054)
50	MLE	1.022(0.024)	1.021(0.025)	1.020(0.022)	1.019(0.022)	1.022 (0.022)
	MME	1.007(0.136)	1.023(0.089)	1.017(0.064)	1.021(0.055)	1.017(0.053)
	PCE	0.948(0.022)	0.947(0.022)	0.946(0.023)	0.946(0.023)	0.942(0.022)
	LSE	1.015(0.022)	1.014(0.029)	1.021(0.032)	1.020(0.030)	1.009(0.028)
	WLSE	1.013(0.020)	1.016(0.026)	1.018(0.028)	1.011(0.027)	1.009(0.025)
	UBE	0.999(0.026)	0.995(0.029)	1.000(0.021)	0.998(0.021)	1.002 (0.021)

Table 2: Average relative estimates and average relative MSE's of λ are presented, when α is known. The average relative MSE's are reported within brackets against each average relative estimates.

n	Method	$\alpha = 0.25$	$\alpha = 0.50$	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 2.5$
20	MLE	1.116(0.115)	1.078(0.068)	1.018(0.014)	1.010(0.007)	1.008 (0.006)
	MME	1.074(0.068)	1.038(0.029)	1.019(0.014)	1.011(0.007)	1.008(0.006)
	PCE	0.981(0.310)	0.959(0.118)	0.959(0.057)	0.966(0.030)	0.966(0.025)
	LSE	1.087(0.158)	1.022(0.042)	1.007(0.017)	1.001(0.009)	1.002(0.006)
	WLSE	1.050(0.147)	1.033(0.048)	1.009(0.018)	1.005(0.008)	1.007(0.006)
50	MLE	1.104(0.079)	1.008(0.005)	1.004(0.003)	1.003(0.002)	0.000 (0.000)
	MME	1.030(0.021)	1.013(0.010)	1.008(0.005)	1.004(0.003)	1.003(0.002)
	PCE	0.952(0.084)	0.955(0.041)	0.968(0.022)	0.974(0.012)	0.975(0.010)
	LSE	1.040(0.039)	1.004(0.016)	1.007(0.007)	1.003(0.003)	1.001(0.003)
	WLSE	1.013(0.037)	1.009(0.013)	1.005(0.006)	1.002(0.003)	1.001(0.003)

Table 3: Average relative estimates and average relative MSE's of α are presented when λ is unknown. The average relative MSE's are reported within brackets against each average relative estimates.

n	Method	$\alpha = 0.25$	$\alpha = 0.50$	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 2.5$
20	MLE	1.111(0.354)	1.151(0.147)	1.167(0.210)	1.218(0.341)	1.244 (0.435)
	MME	1.484(0.440)	1.388(0.405)	1.352(0.450)	1.410(0.630)	1.443(1.062)
	PCE	1.124(0.426)	1.023(0.322)	0.968(0.255)	0.999(0.423)	0.982(0.312)
	LSE	1.021(0.090)	1.035(0.151)	1.035(0.218)	1.061(0.361)	1.086(0.555)
	WLSE	1.043(0.090)	1.056(0.134)	1.060(0.272)	1.081(0.276)	1.121(0.553)
	MLME	1.048(0.215)	1.070(0.190)	1.103(0.225)	1.153(0.345)	1.181 (0.467)
50	MLE	0.990(0.199)	1.106(0.070)	1.063(0.049)	1.080(0.071)	1.084 (0.080)
	MME	1.203(0.150)	1.157(0.118)	1.136(0.107)	1.142(0.125)	1.152(0.141)
	PCE	0.958(0.163)	0.996(0.140)	0.951(0.106)	0.927(0.086)	0.924(0.095)
	LSE	1.025(0.034)	1.003(0.038)	1.014(0.055)	1.014(0.068)	1.028(0.081)
	WLSE	1.027(0.030)	1.028(0.034)	1.030(0.048)	1.035(0.062)	1.031(0.072)
	MLME	1.019(0.071)	1.019(0.058)	1.031(0.058)	1.047(0.074)	1.057 (0.082)

Table 4: Average relative estimates and average relative MSE's of λ are presented when α is unknown. The average relative MSE's are reported within brackets against each average relative estimates.

n	Method	$\alpha = 0.25$	$\alpha = 0.50$	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 2.5$
20	MLE	1.307(0.278)	0.951(0.107)	0.964(0.044)	0.967(0.021)	0.962 (0.027)
	MME	1.224(0.126)	1.137(0.062)	1.087(0.038)	1.067(0.025)	1.064(0.023)
	PCE	0.947(0.123)	1.117(0.084)	1.052(0.028)	1.042(0.020)	1.041(0.018)
	LSE	1.048(0.172)	0.991(0.058)	0.978(0.032)	0.977(0.024)	0.984(0.021)
	WLSE	0.993(0.113)	1.017(0.054)	1.003(0.034)	0.990(0.021)	0.998(0.021)
	MLME	1.241(0.732)	1.134(0.263)	1.087(0.146)	1.068(0.096)	1.064 (0.089)
50	MLE	1.160(0.059)	1.107(0.068)	1.023(0.010)	1.018(0.007)	1.017 (0.007)
	MME	1.092(0.042)	1.056(0.023)	1.038(0.015)	1.026(0.010)	1.026(0.009)
	PCE	1.061(0.067)	0.960(0.027)	0.952(0.016)	0.957(0.010)	0.957(0.009)
	LSE	1.002(0.033)	0.995(0.019)	0.994(0.011)	0.994(0.009)	0.997(0.008)
	WLSE	0.987(0.042)	1.009(0.017)	1.004(0.010)	1.002(0.007)	0.998(0.007)
	MLME	1.085(0.149)	1.039(0.074)	1.029(0.045)	1.021(0.030)	1.026 (0.029)