AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(9): 17444–17460.
DOI: 10.3934/math.2022961
Received: 03 June 2022
Revised: 18 July 2022
Accepted: 21 July 2022
Published: 28 July 2022

Research article

An improvement in maximum likelihood estimation of the Burr XII distribution parameters

Ali A. Al-Shomrani*

Department of Statistics, Faculty of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

* **Correspondence:** Email: email2aaalshomrani@yahoo.com.

Abstract: In this paper, we discuss the parameters estimation of the Burr XII distribution. We know that the most popular method in the literature for parameter estimation is the maximum likelihood method. However, the maximum likelihood estimators (MLEs) are widely known to be biased for small sample sizes. Therefore, this motivate us to obtain approximately unbiased estimators for this distribution' parameters. Precisely, we focus on two bias-correction techniques (analytical and bootstrap approaches) to reduce the biases of the MLEs to the second order of magnitude. In order to compare the performance of these estimators, Monte Carlo simulations are conducted. Lastly, two real-data examples are provided to show the usefulness of these proposed estimators when sample sizes are small.

Keywords: Burr XII distribution; bias-correction; bootstrap
Mathematics Subject Classification: 60Exx, 62-xx

1. Introduction

[6] listed the Burr system of distributions that include twelve types of cumulative distribution functions which yield a variety of density shapes among one of these distributions is a Burr type XII model. [7] have pointed that choosing the parameters of the Burr distribution appropriately covers a large portion of the curve shape characteristics in the Pearson family. As a result, the Burr type XII distribution has been applied in medical, business, chemical engineering, quality control and reliability studies since its shape parameter produces various shapes to be good fits for different data. For instance, [29] showed that the Burr type XII distribution can be applied roughly to any given set of uni-model data besides its relationship to other different distributions. In addition, [9] used this distribution to fit the uranium survey data. [33–35] used this distribution to fit non-censored and censored lifetimes data. Moreover, [23] estimated the parameters of the Burr type XII distribution

under doubly censored data. [30] obtained point and interval estimates of the parameters of the Burr type XII distribution with multiple-censored using the maximum likelihood method. [22] considered the estimation of the three-parameter Burr XII distribution and its doubly truncated form. [32] provided a technique that takes the use of the relationship between three-parameter Burr XII distribution and the two parameter Weibull distribution, which arises as the distribution's limiting case. Furthermore from Bayesian point of view, [21] studied the Burr type XII distribution as a failure model under various loss functions. In addition, [14] used this distribution to estimate its parameters under random censoring. Furthermore, [2] used this distribution to estimate its parameters under middle-censored data. [1] obtained estimation of stress-strength parameter for Burr type XII distribution based on progressive type-II Censoring. A characterization of the Burr type XII distribution was studied by [3]. Moreover, [5] studied the properties and applications of the Burr XII-Power Cauchy distribution.

It is well-known that the most popular estimation method is the maximum likelihood method. This is because it has attractive mathematical properties such as unbiasedness, efficiency, consistency and asymptotic normality for large sample sizes. These properties, however, may not hold true for small or even moderate sample sizes see [8, 15–17, 19, 26, 31] among others.

As is well known, the most often quoted features of the maximum likelihood technique rely mostly on the large sample size condition. The maximum likelihood estimators, in particular, contain biases of the order $O\left(n^{-1}\right)$, which are typically ignored in practice. As will be shown, the maximum likelihood estimators of the parameters associated with the Burr XII distribution are biased, particularly in small samples, implying that the expected value of estimators varies from the real value of parameters. Given this, it is critical to create virtually unbiased estimators for this distribution. This observation motivates us to use a bias-corrected approach to lower the MLE's bias from order $O\left(n^{-1}\right)$ to order $O\left(n^{-2}\right)$. Furthermore, in terms of bias and standard error, we find superior estimators for this distribution than MLE for practical applications. In the next paragraph, we will use two corrective procedures in addition to MLEs to generate modified MLEs that are unbiased to second order. As demonstrated by our simulation analysis, the suggested two corrected MLEs outperform the uncorrected MLEs in terms of consistency and efficiency. Moreover, the two corrected procedures contain closed-form expressions, making them visually appealing and simple to compute. Additionally, numerical data reveals that bias-corrected estimators are exceptionally accurate even for very small sample sizes and outperform earlier estimators in terms of biases and root mean squared errors.

In this paper, we consider two techniques. The first one is a corrective technique, say "analytical approach" introduced by [11]. This approach corrects the bias of MLEs to the second order of magnitude by subtract it from the MLEs. Some authors such as [18, 20, 28] implemented software programs even though its limited to enable users to calculate the analytic Cox-Snell formula for bias corrections for different pre-specified distributions. Furthermore, the second technique is based on bootstrap re-sampling method that can reduce the bias to the second order, say "bootstrap approach" presented by [12]. We will call these corrected estimators as bias-corrected estimators (or for short, bc estimators) in both techniques. Monte Carlo Simulations and real data applications are considered to illustrate the performances of these estimators.

Because of estimation problem for the three-parameter Burr XII distribution as pointed out by [27], we set without loss of generality the scale parameter to be 1. Hence, if X follows a Burr type XII distribution (denoted by BurrXII(α, β)), then the distribution function (cdf) and the probability density function (pdf) of X are respectively given by (see [6]):

$$F(x;\alpha,\beta) = 1 - \{1 + x^{\alpha}\}^{-\beta}, \tag{1.1}$$

$$f(x;\alpha,\beta) = \alpha\,\beta\,x^{(\alpha-1)}\{1 + x^{\alpha}\}^{-(\beta+1)}, \tag{1.2}$$

where both $\alpha > 0$ and $\beta > 0$ are shape parameters. However, α parameter plays the important role in the shape of this distribution to be decreasing, positively skewed or unimodel as shown in Figure 1. Also, this can be seen from the effectiveness of the shape parameter α in the hazard function ($h(x)$) of this distribution

$$h(x) = \left(\frac{\alpha\,\beta}{x}\right)\{1 + x^{-\alpha}\}^{-1}. \tag{1.3}$$

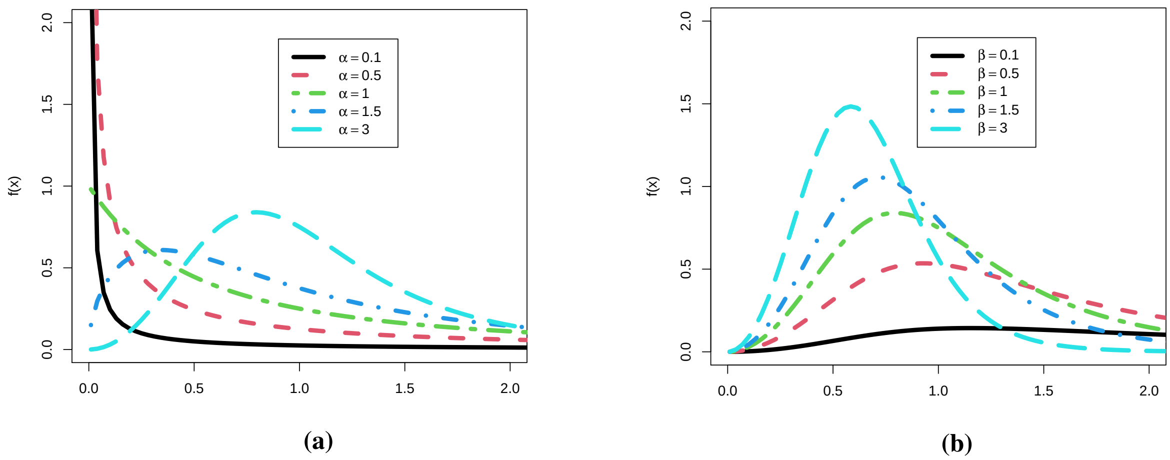

Figure 1. Plots of the pdf of the Burr XII distribution with (a) $\beta = 1$ and different values of α and (b) $\alpha = 1$ and different values of β.

The rest of the paper is organized as follows. In Section 2, we introduce the maximum likelihood estimation (MLE). We provide the bc MLEs for both analytical and bootstrap approachs in Section 3. In Section 4, simulation results to assess the performance of the maximum likelihood estimation method are reported. Furthermore, we demonstrate the performance of these approachs by providing two examples based on real data in Section 5. Finally, we give some concluding remarks in Section 6.

2. Maximum likelihood estimation

Let $X_1, \cdots, X_n$ be a random sample of size n from BurrXII(α,β). The log likelihood function (l) is

$$l(\alpha,\beta) = Log\,L(\alpha,\beta) = n\log(\alpha) + n\log(\beta) + (\alpha-1)\sum_{i=1}^{n}\log(x_i) - (\beta+1)\sum_{i=1}^{n}\log\{1 + (x_i)^{\alpha}\}. \tag{2.1}$$

In order to obtain the MLE's ($\widehat{\alpha}$ *and* $\widehat{\beta}$) of α *and* β respectively, we maximize (2.1) with respect to α *and* β. Thus, we have the following equations:

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log(x_i) - (\beta+1)\sum_{i=1}^{n}\log(x_i)\left[1 + (x_i)^{-\alpha}\right]^{-1}, \tag{2.2}$$

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \log\{1 + (x_i)^{\alpha}\}. \tag{2.3}$$

Even though these non-linear Eqs (2.2) and (2.3) cannot be solved analytically, several programs such as SAS, R, Mathematica, Maple, or Matlab can be used to solve these equations numerically to estimate the MLEs denoted by $\widehat{\eta}_{MLE}$ where $\eta = (\alpha, \beta)^{'}$. As pointed earlier, these MLEs will be biased for small sample sizes. Consequently, the bias may give misleading results which has an effect on interpretations of phenomena in real applications. Therefore, this motivates us to consider approximately unbiased estimators to reduce the bias of these MLEs of Burr XII distribution.

3. Bias-corrected MLEs

Two bias-correction approaches are considered in this section. The first one is the "analytical approach" introduced by [11] which is presented in Section 3.1 and the second one is the "bootstrap approach" introduced by [12] which is presented in Section 3.2.

3.1. Analytical approach

Suppose $l(\eta)$ is the log likelihood function based on a sample of n observations with p-dimenstional parameter vector expressed as $\eta = \left(\eta_1, \cdots, \eta_p\right)^{'}$ and $l(\eta)$ is regular with respect to all derivatives up to the third order, inclusively. Then, the joint cumulants of the derivatives of $l = l(\eta)$ are defined as

$$\varphi_{ij} = E\left[\frac{\partial^2 l}{\partial \eta_i \partial \eta_j}\right], \; i, j = 1, \cdots, p, \tag{3.1}$$

$$\varphi_{ijk} = E\left[\frac{\partial^3 l}{\partial \eta_i \partial \eta_j \partial \eta_k}\right], \; i, j, k = 1, \cdots, p, \tag{3.2}$$

$$\varphi_{ij,k} = E\left[\left(\frac{\partial^2 l}{\partial \eta_i \partial \eta_j}\right)\left(\frac{\partial l}{\partial \eta_k}\right)\right], \; i, j, k = 1, \cdots, p. \tag{3.3}$$

These joint cumulants' derivatives are

$$\varphi_{ij}^{(k)} = \frac{\partial \varphi_{ij}}{\partial \eta_k}, \; i, j, k = 1, \cdots, p. \tag{3.4}$$

Furthermore, we assume all expressions in (3.1) to (3.4) are of order $O(n)$.

Let $\mathrm{I} = \left[-\varphi_{ij}\right]$ be the Fisher information matrix of η, where $i, j = 1, \cdots p$. [11] proved that for independent sample not necessarily identical the bias of the s^{th} element of the MLE of η is

$$Bias\left(\widehat{\eta}_s\right) = \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p} \varphi^{si}\varphi^{jk}\left[\frac{1}{2}\varphi_{ijk} + \varphi_{ij,k}\right] + O\left(n^{-2}\right), \; s = 1, \cdots, p, \tag{3.5}$$

where φ^{ij} is the $(i, j)^{th}$ element of the inverse of the information matrix. Later, [10] showed that Eq (3.5) still hold if the sample data are not identical and non-independent observations provided that all expressions in (3.1) to (3.4) are of order $O(n)$. They expressed Eq (3.5) as

$$Bias\left(\widehat{\eta}_s\right) = \sum_{i=1}^{p} \varphi^{si} \sum_{j=1}^{p} \sum_{k=1}^{p} \varphi^{jk} \left[\varphi_{ij}^{(k)} - \frac{1}{2}\varphi_{ijk}\right] + O\left(n^{-2}\right) \ , \ s = 1, \cdots , p \ . \tag{3.6}$$

This bias expression in Eq (3.6) is generally easier to compute than that of Eq (3.5) because the term of the form in Eq (3.3) was dropped from Eq (3.5).

Define $\omega_{ij}^{(k)} = \varphi_{ij}^{(k)} - \frac{1}{2}\varphi_{ijk}$, $i, j, k = 1, \cdots , p$. Thus, we have these two matrices (Ψ and $\Omega^{(k)}$) as

$$\Psi = \left[\Omega^{(1)} | \ \Omega^{(2)} \ | \ \cdots | \Omega^{(p)}\right] \ where \ \Omega^{(k)} = \left[\omega_{ij}^{(k)}\right] \ , \ i, j, k = 1, \cdots , p. \tag{3.7}$$

Therefore, the bias expression of $\widehat{\eta}$ can be expressed in matrix form as

$$Bias\left(\widehat{\eta}\right) = \mathrm{I}^{-1}\Psi \cdot \mathrm{vec}\left(\mathrm{I}^{-1}\right) + O\left(n^{-2}\right), \tag{3.8}$$

where vec is an operator that stacks column vectors of a matrix below one another. Thus, the bc MLE of η, denoted as $\widehat{\eta}_{BCMLE}$, is given by

$$\widehat{\eta}_{BCMLE} = \widehat{\eta} - \widehat{\mathrm{I}}^{-1}\widehat{\Psi} \cdot \mathrm{vec}\left(\widehat{\mathrm{I}}^{-1}\right), \tag{3.9}$$

where $\widehat{\eta}$ is the MLE of η , $\widehat{\mathrm{I}} = \mathrm{I}|_{\eta=\widehat{\eta}}$ and $\widehat{\Psi} = \Psi|_{\eta=\widehat{\eta}}$. Note that the bias of $\widehat{\eta}_{BCMLE}$ is $O\left(n^{-2}\right)$.

Since we study the Burr XII distribution, we have $\eta = (\alpha, \beta)^{'}$ and $p = 2$. In order to obtain the bc MLEs, we need to find the higher-order derivatives of the log-likelihood function for the Burr XII distribution taken with respect to α and β as follows:

$$\frac{\partial^2 l}{\partial \alpha^2} = -\frac{n}{\alpha^2} - (\beta + 1) \sum_{i=1}^{n} \{\log (X_i)\}^2 (X_i)^{\alpha} [1 + (X_i)^{\alpha}]^{-2}, \tag{3.10}$$

$$\frac{\partial^2 l}{\partial \beta^2} = -\frac{n}{\beta^2}, \tag{3.11}$$

$$\frac{\partial^2 l}{\partial \alpha \partial \beta} = - \sum_{i=1}^{n} \left\{\log (X_i) (X_i)^{\alpha} [1 + (X_i)^{\alpha}]^{-1}\right\}, \tag{3.12}$$

$$\frac{\partial^3 l}{\partial \alpha^3} = \frac{2n}{\alpha^3} - (\beta + 1) \sum_{i=1}^{n} \left\{[\log (X_i)]^3 (X_i)^{\alpha} [1 + (X_i)^{\alpha}]^{-3} [1 - (X_i)^{\alpha}]\right\}, \tag{3.13}$$

$$\frac{\partial^3 l}{\partial \beta^3} = \frac{2n}{\beta^3}, \tag{3.14}$$

$$\frac{\partial^3 l}{\partial \alpha^2 \partial \beta} = - \sum_{i=1}^{n} \{\log (X_i)\}^2 (X_i)^{\alpha} [1 + (X_i)^{\alpha}]^{-2}, \tag{3.15}$$

$$\frac{\partial^3 l}{\partial \beta^2 \partial \alpha} = 0. \tag{3.16}$$

Observe that if X follows BurrXII(α,β) then $Y = (X)^{\alpha} \sim f(y) = \beta\,[1+y]^{-(\beta+1)}$ $for\ y > 0$. We know that the beta function is $Beta\,(a,b) = \int_0^1 y^{a-1}(1-y)^{b-1}\,dy = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ where $\Gamma(t) = \int_0^\infty y^{t-1}e^{-y}\,dy$ is the gamma function. The following formulas will be needed

$$\int_0^\infty \log(y)\ y^{a-1}[1+y]^{-b}\,dy = Beta\,(a,b-a)\,\{\psi(a) - \psi(b-a)\}\quad for\ a < b, \tag{3.17}$$

$$\int_0^\infty [\log(y)]^2\, y^{a-1}[1+y]^{-b}\,dy = Beta\,(a,b-a)\left\{[\psi(a) - \psi(b-a)]^2 + \psi^{(1)}(a) + \psi^{(1)}(b-a)\right\},\quad for\ a < b, \tag{3.18}$$

and in general

$$\int_0^\infty [\log(y)]^c\, y^{a-1}[1+y]^{-b}\,dy = \frac{\partial^c}{\partial\alpha^c} Beta\,(a,b-a)\quad for\ a < b\ and\ c = 1,2,\cdots, \tag{3.19}$$

see Eqs (2.2.4.24), (2.6.4.6), (2.6.4.7) and (2.6.4.10) in [24], respectively. Therefore, using the previous Eqs (3.17)–(3.19) we get

$$\begin{aligned}
&\mathbb{E}\left\{\log(X)\,(X)^{\alpha}[1+(X)^{\alpha}]^{-1}\right\}\\
&= \left(\frac{1}{\alpha}\right)\mathbb{E}\left\{\log(Y)\ Y\,[1+Y]^{-1}\right\}\\
&= \left(\frac{1}{\alpha}\right)\int_0^\infty \log(y)\ y\,[1+y]^{-1}\,f(y)\,dy\\
&= \left(\frac{\beta}{\alpha}\right)\int_0^\infty \log(y)\ y\,[1+y]^{-(\beta+2)}\,dy\\
&= \left(\frac{\beta}{\alpha}\right) Beta\,(2,\beta)\ \{\psi(2) - \psi(\beta)\}\\
&= \frac{\psi(2) - \psi(\beta)}{\alpha(\beta+1)}.
\end{aligned} \tag{3.20}$$

Let $\psi(t) = \frac{d\log(\Gamma(t))}{dt} = \frac{\Gamma'(t)}{\Gamma(t)}$ be the digamma function. Thus, $\psi^{(m)}(t) = \frac{d^m}{dt^m}\psi(t) = \frac{d^{m+1}}{dt^{m+1}}\log(\Gamma(t))$ be the m^{th} derivative of the digamma function. Hence, $\psi(t) = \psi^{(0)}(t)$. Similarly,

$$\begin{aligned}
&\mathbb{E}\left\{\{\log(X)\}^2(X)^{\alpha}[1+(X)^{\alpha}]^{-2}\right\}\\
&= \mathbb{E}\left\{\left\{\frac{\log(Y)}{\alpha}\right\}^2 Y\,[1+Y]^{-2}\right\}\\
&= \left(\frac{\beta}{\alpha^2(\beta+1)(\beta+2)}\right)\left\{[\psi(2) - \psi(\beta+1)]^2 + \psi^{(1)}(2) + \psi^{(1)}(\beta+1)\right\},
\end{aligned} \tag{3.21}$$

$$\begin{aligned}
&\mathbb{E}\left\{\{\log(X)\}^3(X)^{\alpha}[1+(X)^{\alpha}]^{-3}\right\}\\
&= \mathbb{E}\left\{\left\{\frac{\log(Y)}{\alpha}\right\}^3 Y\,[1+Y]^{-3}\right\}
\end{aligned}$$

$$= \left(\frac{\beta}{\alpha^3(\beta+2)(\beta+3)}\right)\{[\psi(2)-\psi(\beta+2)]\times$$
$$\left([\psi(2)-\psi(\beta+2)]^2+3\left[\psi^{(1)}(2)+\psi^{(1)}(\beta+2)\right]\right)+\psi^{(2)}(2)-\psi^{(2)}(\beta+2)\Big\}, \quad (3.22)$$

$$\mathbb{E}\left\{\{\log(X)\}^3(X)^{2\alpha}[1+(X)^{\alpha}]^{-3}\right\}$$
$$= \mathbb{E}\left\{\left\{\frac{\log(Y)}{\alpha}\right\}^3 Y^2[1+Y]^{-3}\right\}$$
$$= \left(\frac{2\beta}{\alpha^3(\beta+1)(\beta+2)(\beta+3)}\right)\{[\psi(3)-\psi(\beta+1)]\times$$
$$\left([\psi(3)-\psi(\beta+1)]^2+3\left[\psi^{(1)}(3)+\psi^{(1)}(\beta+1)\right]\right)+\psi^{(2)}(3)-\psi^{(2)}(\beta+1)\Big\}. \quad (3.23)$$

For the joint cumulants of the derivatives of the log-likelihood function, see Appendix A. As previously mentioned, we can use any computer program to compute the bc MLE of η ($\widehat{\eta}_{BCMLE}$) given by

$$\widehat{\eta}_{BCMLE} = \widehat{\eta} - \widehat{\mathrm{I}}^{-1}\widehat{\Psi}\cdot vec\left(\widehat{\mathrm{I}}^{-1}\right),$$

where I is the Fisher information matrix of η and $\Psi = \left[\Omega^{(1)}|\,\Omega^{(2)}\right]$ *such that* $\Omega^{(k)} = \left[\omega_{ij}^{(k)}\right]$, $i, j, k = 1, 2$.

3.2. Bootstrap approach

[12] introduced bootstrap resampling method to generate pseudo-samples from the original sample. In order to obtain the bc MLEs, we subtract the estimated bias obtained from these samples from the original MLEs as follows:

Let $\mathbf{x} = (x_1, \cdots, x_n)'$ be a random sample of size n from F, the distribution function (cdf). Let $\widehat{\tau}$ be the estimator of τ where the parameter $\tau = t(F)$ be some function of F. We resample $\mathbf{x}^* = \left(x_1^*, \cdots, x_n^*\right)'$ as pseudo-samples of size n from the original sample $\mathbf{x}$ by generating observations with replacement. From these pseudo-samples, we obtain the bootstrap replicates of $\widehat{\tau}$, denoted by $\widehat{\tau}^* = g(\mathbf{x}^*)$. The empirical cdf (ecdf) of $\widehat{\tau}^*$ can be used to estimate the cdf of $\widehat{\tau}$, $F_{\widehat{\tau}}$. The bootstrap bias of the estimator $\widehat{\tau} = g(\mathbf{x})$ can be estimated by

$$B_F(\widehat{\tau}, \tau) = \mathbb{E}_F[\widehat{\tau}] - \tau(F). \quad (3.24)$$

By replacing F by $F_{\widehat{\tau}}$ in Eq (3.24) since it is consistent estimator, we obtain the estimated bootstrap bias as

$$\hat{B}_{F_{\widehat{\tau}}}(\widehat{\tau}, \tau) = \mathbb{E}_{F_{\widehat{\tau}}}[\widehat{\tau}^*] - \widehat{\tau}. \quad (3.25)$$

If we have bootstrap estimates of size N, say $\left(\widehat{\tau}^{*(1)}, \cdots, \widehat{\tau}^{*(N)}\right)$, and N is sufficiently large then the estimator $\mathbb{E}_{F_{\widehat{\tau}}}[\widehat{\tau}^*]$ in Eq (3.25) can be approximated by

$$\widehat{\tau}^{*(\cdot)} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\tau}^{*(i)}.$$

Therefore, the estimated bootstrap bias becomes

$$\hat{B}_{F_{\widehat{\tau}}}(\widehat{\tau}, \tau) = \widehat{\tau}^{*(\cdot)} - \widehat{\tau}. \tag{3.26}$$

Hence, the bc estimator (τ^B) according to Efron's bootstrap resampling approach is

$$\tau^B = \widehat{\tau} - \hat{B}_{F_{\widehat{\tau}}}(\widehat{\tau}, \tau) = 2\,\widehat{\tau} - \widehat{\tau}^{*(\cdot)}. \tag{3.27}$$

Therefore, in our case we let $\widehat{\eta}_{BCBOOT} = \tau^B$.

4. A simulation study

In this section, we perform Monte Carlo simulations to assess the effectiveness of the several considered estimators of the Burr XII distribution with cdf and pdf presented respectively in Eqs (1.1) and (1.2). We choose random samples of size n = 5, 10, 20, 35, 50, 75, 100, 200 , 500 with α = 0.1, 0.5, 1, 1.5, 3 and β = 0.1, 0.5, 1, 1.5, 3. We used $M = 5000$ Monte Carlo replications with $B = 1000$ bootstrap replications in our simulations for each combination of (n, α, β).

For an estimator $\hat{\eta}$ of the parameter $\eta = (\alpha, \beta)'$, we calculate the average bias and the RMSEs, which are denoted by Bias($\widehat{\eta}$)=$\frac{1}{M}\sum_{i=1}^{M}(\widehat{\eta}_i - \eta)$, and RMSE($\widehat{\eta}$)= $\sqrt{\frac{1}{M}\sum_{i=1}^{M}(\widehat{\eta}_i - \eta)^2}$, respectively.

The average biases and the RMSEs of the estimates of α and β across sample sizes were depicted in Figures 2 and 3. The following are some conclusions that can be derived.

(1) For each simulation under consideration, the MLE estimators of α seem to have a positive bias. This shows that they overestimate the real value of the parameter α on average, in particular when the sample size is small. Furthermore, when the real value of the parameter β is less than one, the MLE estimators frequently appear to have a negative bias, i.e., they underestimate the real value of the parameter β for different sample sizes. On the other hand, when the real value of the parameter β is equal to or larger than one, the MLE estimators often tend to be positively biased. This means that they, on average, overestimate the real value of the parameter β for different sample sizes.

(2) In each simulation for varied sample sizes, the BCMLEs of α and β beat the MLE estimators in terms of bias and RMSE. As a result, if bias is an issue, the BCMLEs would be excellent choices for estimating α and β. In addition, in terms of bias and RMSE for most simulations, the BCBOOTs of α and β beat the MLE estimators for various sample sizes. As a consequence, the BCBOOTs would be preferable choices for estimating α and β.

(3) As sample size n becomes large, the biases and RMSEs of all investigated estimators will decrease as expected. This is mostly due to the fact that in statistical theory, most estimators perform better as the sample size n increases. As expressed above, the reductions in bias and RMSE are very significant for the bc estimators for small sample sizes. Tables 1 and 2 give point estimates of the parameters α and β of Burr XII distribution with their corresponding biases and RMSEs by using the three different estimation methods (MLE, BCMLE and BCBOOT) across samples sizes, respectively. For example, where n = 10, α = 1.5, and β =1, we have Bias($\hat{\alpha}$ MLE)=0.2641, Bias($\hat{\alpha}$ BCMLE)=-0.0099, Bias($\hat{\alpha}$ BCBOOT)= 0.0864, Bias($\hat{\beta}$ MLE) = 0.0786, Bias($\hat{\beta}$ BCMLE) = 0.0026, Bias($\hat{\beta}$ BCBOOT) =0.0018, RMSE($\hat{\alpha}$ MLE) = 0.7336, RMSE($\hat{\alpha}$ BCMLE) = 0.5055, RMSE($\hat{\alpha}$ BCBOOT)= 0.6304, RMSE($\hat{\beta}$ MLE) = 0.4299, RMSE($\hat{\beta}$ BCMLE) = 0.3621, RMSE($\hat{\beta}$ BCBOOT) = 0.3952.

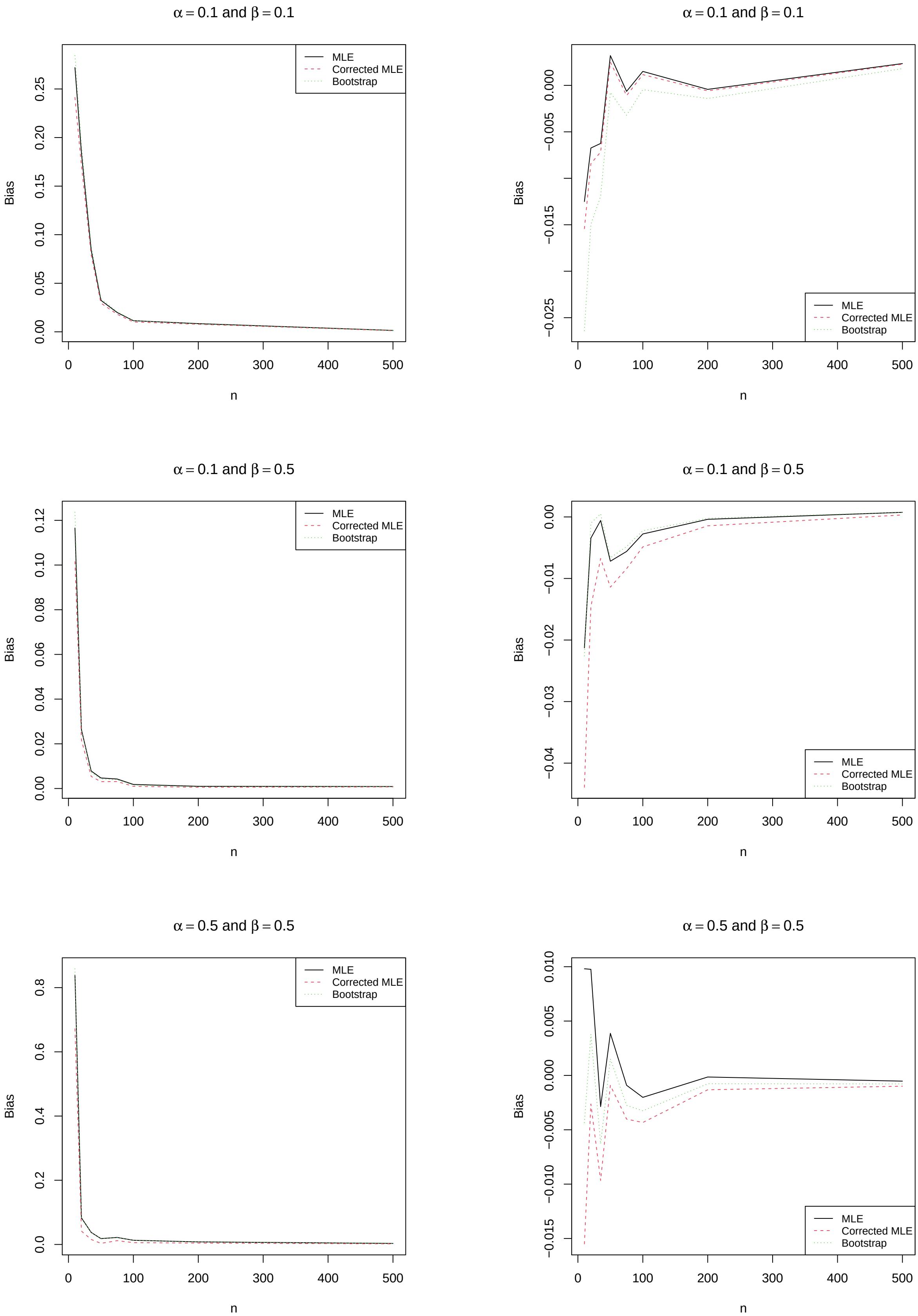

Figure 2. Plots of the average biases comparisons of the three different estimation methods for α (left side) and β (right side).

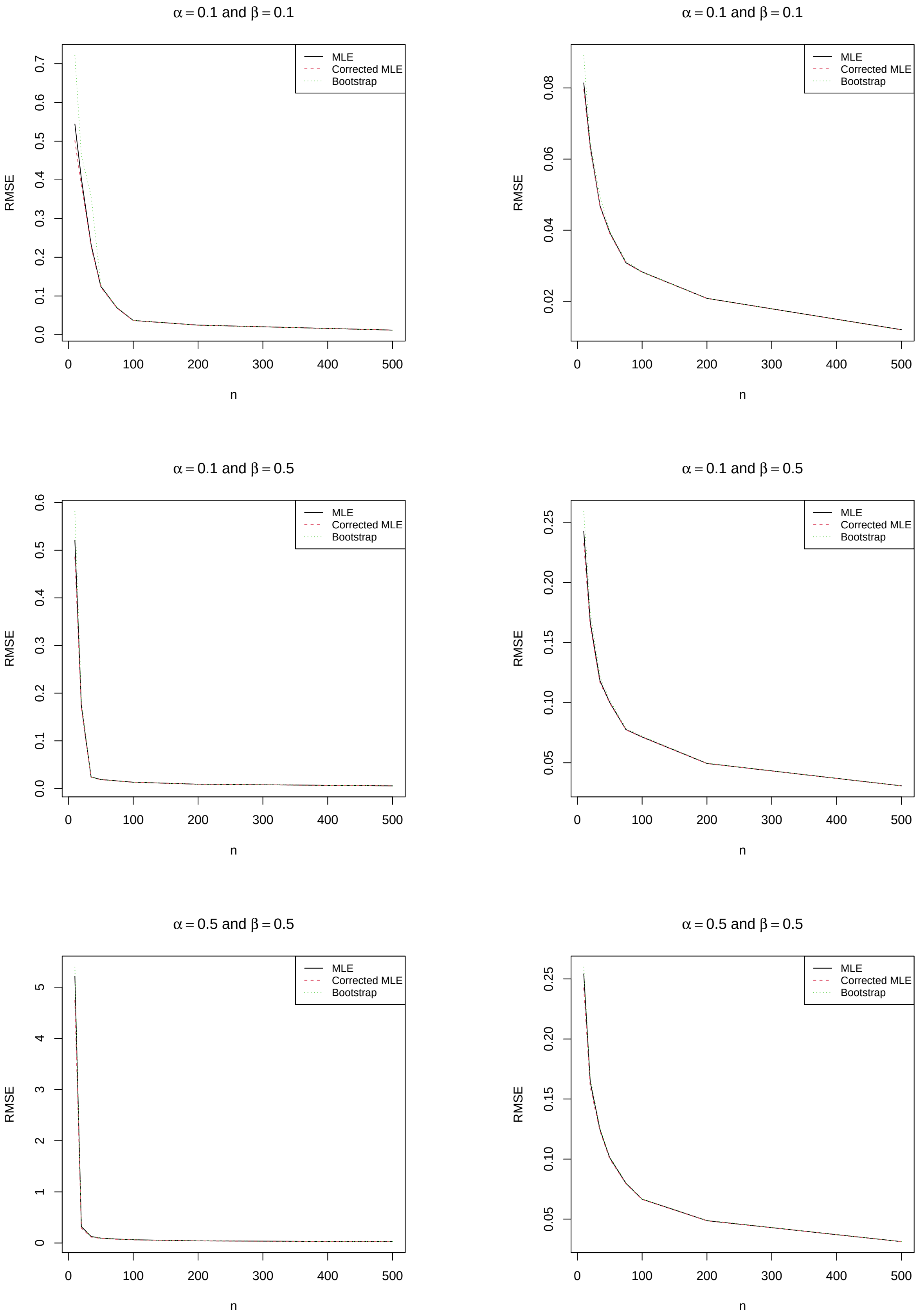

Figure 3. Plots of the RMSEs Comparisons of the three different estimation methods for α (left side) and β (right side).

Table 1. Point estimates of the parameter α of Burr XII distribution with their corresponding biases and RMSEs by using the three different estimation methods (MLE, BCMLE and BCBOOT) across sample sizes.

		$\widehat{\alpha}$			Bias			RMSE		
n	α	MLE	BCMLE	BCBOOT	MLE	BCMLE	BCBOOT	MLE	BCMLE	BCBOOT
10	1.5	1.7641	1.4901	1.5864	0.2641	-0.0099	0.0864	0.7336	0.5055	0.6304
20	1.5	1.6585	1.5336	1.551	0.1585	0.0336	0.051	0.4478	0.3788	0.3918
35	1.5	1.5779	1.5114	1.517	0.0779	0.0114	0.017	0.2869	0.2623	0.2638
50	1.5	1.5562	1.51	1.5124	0.0562	0.01	0.0124	0.2209	0.2057	0.2074
75	1.5	1.5293	1.4992	1.5	0.0293	-0.0008	0	0.1686	0.1619	0.1616
100	1.5	1.5269	1.5044	1.5045	0.0269	0.0044	0.0045	0.1431	0.138	0.139
200	1.5	1.508	1.4969	1.4977	0.008	-0.0031	-0.0023	0.1002	0.099	0.0997
500	1.5	1.5049	1.5005	1.5007	0.0049	0.0005	0.0007	0.063	0.0626	0.0627

Table 2. Point estimates of the parameter β of Burr XII distribution with their corresponding biases and RMSEs by using the three different estimation methods (MLE, BCMLE and BCBOOT) across sample sizes.

		$\widehat{\beta}$			Bias			RMSE		
n	β	MLE	BCMLE	BCBOOT	MLE	BCMLE	BCBOOT	MLE	BCMLE	BCBOOT
10	1	1.0786	1.0026	1.0018	0.0786	0.0026	0.0018	0.4299	0.3621	0.3952
20	1	1.0273	0.996	0.9923	0.0273	-0.004	-0.0077	0.271	0.2526	0.2588
35	1	1.0191	1.0022	1.0013	0.0191	0.0022	0.0013	0.1917	0.1841	0.1874
50	1	0.998	0.987	0.9867	-0.002	-0.013	-0.0133	0.1574	0.154	0.1564
75	1	1.0095	1.002	1.0014	0.0095	0.002	0.0014	0.1342	0.1316	0.1318
100	1	1.0014	0.9959	0.9962	0.0014	-0.0041	-0.0038	0.113	0.1117	0.1127
200	1	1.006	1.0033	1.0034	0.006	0.0033	0.0034	0.0786	0.078	0.0787
500	1	0.9963	0.9953	0.995	-0.0037	-0.0047	-0.005	0.0474	0.0474	0.0478

5. Illustrative applications

We consider two real data sets to compare the performance of these estimators (the maximum likelihood estimator denoted by $\widehat{\eta}_{\text{MLE}}$, the bc maximum likelihood estimator using analytical approach denoted by $\widehat{\eta}_{\text{BCMLE}}$ and the bc estimator using bootstrap approach denoted by $\widehat{\eta}_{\text{BCBOOT}}$).

The first set of data indicates the survival rate (in percentage) of the batches of rats exposed to varied amounts of radiation. This data set was introduced by [13]. It is presented as follows: 44.0, 55.0, 16.0, 13.0, 4.0, 1.96, 6.12, 0.5, 0.32, 0.110, 0.015, 0.019, 0.7 and 0.006.

The estimated values of the parameters of the Burr XII distribution are listed in Table 3. Table 3 shows that the MLE estimates are larger than the bc MLE and bootstrap estimates of α and β, implying that the MLE technique overestimates both parameters. Figure 4 illustrates the evaluation of the pdf and cdf of the Burr XII distribution with respect to the estimates of α and β in Table 3. The density shape based on the MLE method may be misleading as shown in this Figure, and thus we propose the use of bc MLEs for this data set.

Table 3. Point estimates of the parameters (α and β) of Burr XII distribution for the first data.

Estimate	α	β
MLE	0.5677	1.0382
BCMLE	0.5276	0.9956
BCBOOTP	0.5387	1.0324

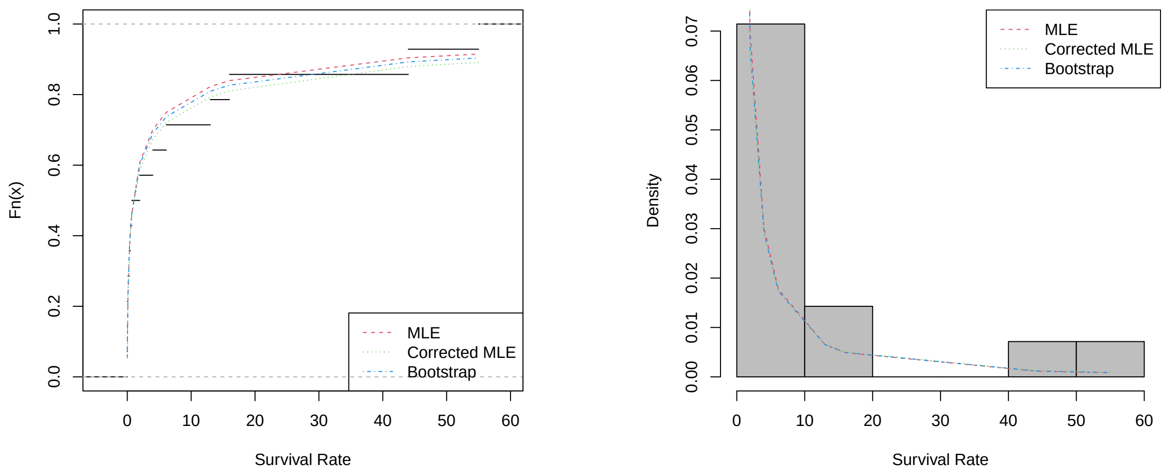

Figure 4. The cdf and pdf of the Burr XII distribution fitted to the survival rate data using the various estimates of α and β.

The second data set represents the time to failure (in months) of 20 electronic components on test. This data set was taken from [33]. Furthermore, it was analyzed by [25]. For convenience, we present it as follows: 0.1, 0.1, 0.2, 0.3, 0.4, 0.5, 0.5, 0.6, 0.7, 0.8, 0.9, 0.9, 1.2, 1.6, 1.8, 2.3, 2.5, 2.6, 2.9 and 3.1.

Similarly, The estimated values of the parameters of the Burr XII distribution are listed in Table 4. Table 4 shows that the MLE estimates are larger than the bc MLE and bootstrap estimates of α and β, implying that the MLE technique overestimates both parameters, particularly α. Figure 5 illustrates the evaluation of the pdf and cdf of the Burr XII distribution with respect to the estimates of α and β in Table 4. The density shape based on the MLE method may be misleading as shown in this Figure, and thus we propose the use of bc MLEs for this data set.

Table 4. Point estimates of the parameters (α and β) of Burr XII distribution for the second data.

Estimate	α	β
MLE	1.5606	1.2883
BCMLE	1.4574	1.2415
BCBOOTP	1.4712	1.2509

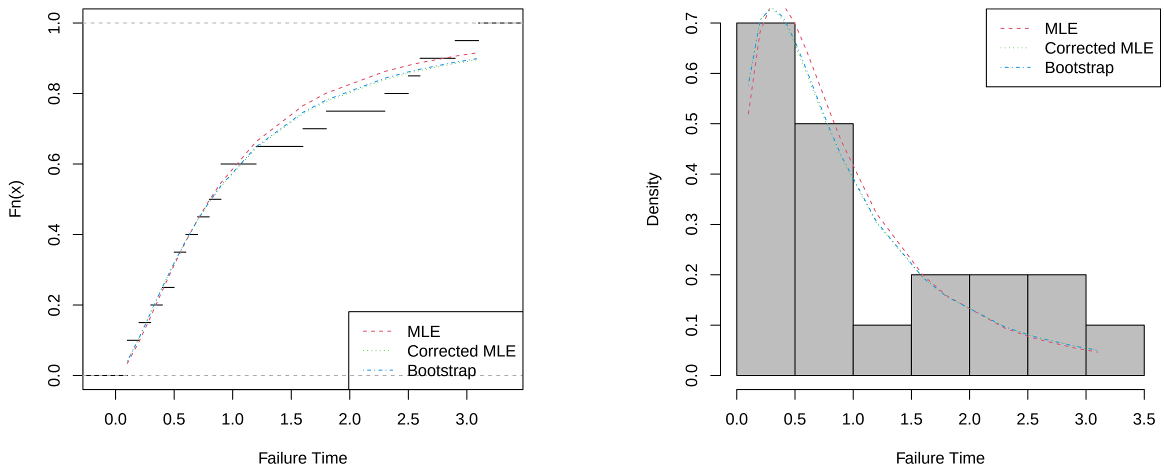

Figure 5. The cdf and pdf of the Burr XII distribution fitted to the failure time data using the various estimates of α and β.

6. Conclusions

The second-order bc MLEs of the Burr XII distribution have been derived based on the "analytical approach" developed by [11]. bc MLEs not only have clear expressions, but they also minimize the bias and root mean square errors (RMSEs) of the parameters of the Burr XII distribution at the same time. For parameter estimation, we also evaluated the "bootstrap approach" resampling method described by [12]. The numerical findings of both simulation studies and real-data applications clearly imply that bc MLEs should be recommended for use in practical applications, particularly when the sample size is small or moderate. Further work could be bias reduction of MLE's for Burr type XII distribution (see [4]).

Acknowledgments

The author would like to thank the editor and the anonymous referees for their insightful comments that helped to enhance the paper.

Conflict of interest

The author declares that he has no conflict of interest.

References

1. A. Abd-Elfattah, M. Abu-Moussa, Estimation of stress-strength parameter for Burr type XII distribution based on progressive type-II censoring, *ANGLISTICUM J. Assoc.-Inst. English Lang. Am. Stud.*, **3** (2013), 1857–8187. http://dx.doi.org/10.0001/(aj).v3i0.700
2. A. H. Abuzaid, The estimation of the Burr-XII parameters with middle-censored data, *SpringerPlus*, **4** (2015), 101. https://doi.org/10.1186/s40064-015-0856-3

3. E. K. Al-Hussaini, A characterization of the Burr type XII distribution, *Appl. Math. Lett.*, **4** (1991), 59–61. https://doi.org/10.1016/0893-9659(91)90123-D
4. J. Arrué, R. B. Arellano-Valle, H. W. Gómez, Bias reduction of maximum likelihood estimates for a modified skew-normal distribution, *J. Stat. Comput. Simul.*, **86** (2016), 2967–2984. https://doi.org/10.1080/00949655.2016.1143471
5. F. A. Bhatti, G. G. Hamedani, M. M. A. Sobhi, M. Korkmaz, On the Burr XII-power Cauchy distribution: Properties and applications, *AIMS Math.*, **6** (2021), 7070–7092. https://doi.org/10.3934/math.2021415
6. I. W. Burr, Cumulative frequency functions, *Ann. Math. Stat.*, **13** (1942), 215–232. https://doi.org/10.1214/aoms/1177731607
7. I. W. Burr, P. J. Cislak, On a general system of distributions: I. Its curve-shape characteristics; II. The sample median, *J. Am. Stat. Assoc.*, **63** (1968), 627–635. https://doi.org/10.2307/2284033
8. J. E. Contreras-Reyes, F. Kahrari, D. D. Cortés, On the modified skew-normal-cauchy distribution: properties, inference and applications, *Commun. Stat.-Theor. M.*, **50** (2021), 3615–3631. https://doi.org/10.1080/03610926.2019.1708942
9. R. D. Cook, M. E. Johnson, Generalized Burr-Pareto-logistic distributions with applications to a uranium exploration data set, *Technometrics*, **28** (1986), 123–131. https://doi.org/10.1080/00401706.1986.10488113
10. G. M. Cordeiro, R. Klein, Bias correction in arma models, *Stat. Probabil. Lett.*, **19** (1994), 169–176. https://doi.org/10.1016/0167-7152(94)90100-7
11. D. R. Cox, E. J. Snell, A general definition of residuals, *J. Roy. Stat. Soc. B*, **30** (1968), 248–265. https://doi.org/10.1111/j.2517-6161.1968.tb00724.x
12. B. Efron, Bootstrap methods: Another look at the jackknife, *Ann. Stat.*, **7** (1979), 1–26. https://doi.org/10.1007/978-1-4612-4380-9_41
13. B. Efron, Computer-intensive methods in statistical regression, *SIAM Rev.*, **30** (1988), 421–449. https://doi.org/10.1137/1030093
14. M. E. Ghitany, S. Al-Awadhi, Maximum likelihood estimation of Burr XII distribution parameters under random censoring, *J. Appl. Stat.*, **29** (2002), 955–965. https://doi.org/10.1080/0266476022000006667
15. D. E. Giles, Bias reduction for the maximum likelihood estimators of the parameters in the half-logistic distribution, *Commun. Stat.-Theor. M.*, **41** (2012), 212–222. https://doi.org/10.1080/03610926.2010.521278
16. D. E. Giles, H. Feng, R. T. Godwin, On the bias of the maximum likelihood estimator for the two-parameter lomax distribution, *Commun. Stat.-Theor. M.*, **42** (2013), 1934–1950. https://doi.org/10.1080/03610926.2011.600506
17. D. E. Giles, H. Feng, R. T. Godwin, Bias-corrected maximum likelihood estimation of the parameters of the generalized pareto distribution, *Commun. Stat.-Theor. M.*, **45** (2016), 2465–2483. https://doi.org/10.1080/03610926.2014.887104
18. P. H. Johnson, Y. Qi, Y. C. Chueh, *Csck mle bias calculation*, Working papers, University of Illinois, Urbana-Champaign, Champaign, IL, 2012.

19. X. Ling, D. E. Giles, Bias reduction for the maximum likelihood estimator of the parameters of the generalized rayleigh family of distributions, *Commun. Stat.-Theor. M.*, **43** (2014), 1778–1792. https://doi.org/10.1080/03610926.2012.675114

20. J. Mazucheli, A. F. B. Menezes, S. Nadarajah, mle.tools: An R package for maximum likelihood bias correction, *R J.*, **9** (2017), 268–290. https://doi.org/10.32614/RJ-2017-055

21. D. Moore, A. S. Papadopoulos, The Burr type XII distribution as a failure model under various loss functions, *Microelectron. Reliab.*, **40** (2000), 2117–2122. https://doi.org/10.1016/S0026-2714(00)00031-7

22. M. K. Okasha, M. Y. Matter, On the three-parameter Burr type XII distribution and its application to heavy tailed lifetime data, *J. Adv. Math.*, **10** (2015), 3429–3442.

23. H. Panahi, Estimation for the parameters of the Burr type XII distribution under doubly censored sample with application to microfluidics data, *Int. J. Syst. Assur. Eng.*, **10** (2019), 510–518. https://doi.org/10.1007/s13198-018-0735-8

24. A. Prudnikov, *Integrals and series*, Taylor & Francis, 1986.

25. M. K. Rastogi, Y. M. Tripathi, Estimating a parameter of Burr type XII distribution using hybrid censored observations, *Int. J. Qual. Reliab. Ma.*, **28** (2011), 885–893. https://doi.org/10.1108/02656711111162532

26. J. Reath, J. Dong, M. Wang, Improved parameter estimation of the log-logistic distribution with applications, *Comput. Stat.*, **33** (2018), 339–356. https://doi.org/10.1007/s00180-017-0738-y

27. Q. Shao, Notes on maximum likelihood estimation for the three-parameter Burr XII distribution, *Comput. Stat. Data An.*, **45** (2004), 675–687. https://doi.org/10.1016/S0167-9473(02)00367-5

28. B. D. Stošić, G. M. Cordeiro, Using maple and mathematica to derive bias corrections for two parameter distributions, *J. Stat. Comput. Simul.*, **79** (2009), 751–767. https://doi.org/10.1080/00949650801911047

29. P. R. Tadikamalla, A look at the Burr and related distributions, *Int. Stat. Rev./Revue Int. Stat.*, **48** (1980), 337–344. https://doi.org/10.2307/1402945

30. F. Wang, J. Keats, W. J. Zimmer, Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data, *Microelectron. Reliab.*, **36** (1996), 359–362. https://doi.org/10.1016/0026-2714(95)00077-1

31. M. Wang, W. Wang, Bias-corrected maximum likelihood estimation of the parameters of the weighted lindley distribution, *Commun. Stat.-Simul. Comput.*, **46** (2017), 530–545. https://doi.org/10.1080/03610918.2014.970696

32. A. J. Watkins, An algorithm for maximum likelihood estimation in the three parameter Burr XII distribution, *Comput. Stat. Data An.*, **32** (1999), 19–27. https://doi.org/10.1016/S0167-9473(99)00024-9

33. D. R. Wingo, Maximum likelihood estimation of Burr type XII distribution parameters under type II censoring, *Microelectron. Reliab.*, **33** (1993), 1251–1257. https://doi.org/10.1016/0026-2714(93)90126-J

34. D. R. Wingo, Maximum likelihood methods for fitting the Burr type XII distribution to multiply (progressively) censored life test data, *Metrika*, **40** (1993), 203–210. https://doi.org/10.1007/BF02613681

35. D. Wingo, Maximum likelihood methods for fitting the Burr type XII distribution to life test data, *Biometrical J.*, **25** (1983), 77–84. https://doi.org/10.1002/bimj.19830250109

Appendices

A. The joint cumulants of the derivatives of the log-likelihood function:

$$\varphi_{11}=\mathbb{E}\left\{\frac{\partial^2 l}{\partial\alpha^2}\right\}=-\frac{n}{\alpha^2}-\left(\frac{n\beta}{\alpha^2(\beta+2)}\right)\left\{[\psi(2)-\psi(\beta+1)]^2+\psi^{(1)}(2)+\psi^{(1)}(\beta+1)\right\},$$

$$\varphi_{22}=\mathbb{E}\left\{\frac{\partial^2 l}{\partial\beta^2}\right\}=-\frac{n}{\beta^2},\qquad \varphi_{12}=\mathbb{E}\left\{\frac{\partial^2 l}{\partial\alpha\partial\beta}\right\}=\frac{n\left[\psi(\beta)-\psi(2)\right]}{\alpha(\beta+1)},$$

$$\begin{aligned}\varphi_{111}=\mathbb{E}\left\{\frac{\partial^3 l}{\partial\alpha^3}\right\}=\frac{2n}{\alpha^3}-\frac{n\beta}{\alpha^3(\beta+2)(\beta+3)}&\left\{(\beta+1)\left\{[\psi(2)-\psi(\beta+2)]\left([\psi(2)-\psi(\beta+2)]^2+\right.\right.\right.\\&3\left[\psi^{(1)}(2)+\psi^{(1)}(\beta+2)\right]\Big)+\psi^{(2)}(2)-\psi^{(2)}(\beta+2)\Big\}-2\ \{[\psi(3)-\\&\psi(\beta+1)]\left([\psi(3)-\psi(\beta+1)]^2+3\left[\psi^{(1)}(3)+\psi^{(1)}(\beta+1)\right]\right)+\psi^{(2)}(3)-\psi^{(2)}(\beta+1)\Big\}\Big\},\end{aligned}$$

$$\varphi_{222}=\mathbb{E}\left\{\frac{\partial^3 l}{\partial\beta^3}\right\}=\frac{2n}{\beta^3},\qquad \varphi_{122}=\mathbb{E}\left\{\frac{\partial^3 l}{\partial\beta^2\partial\alpha}\right\}=0,$$

$$\varphi_{112}=\mathbb{E}\left\{\frac{\partial^3 l}{\partial\alpha^2\partial\beta}\right\}=-\left(\frac{n\,\beta}{\alpha^2(\beta+1)(\beta+2)}\right)\left\{[\psi(2)-\psi(\beta+1)]^2+\psi^{(1)}(2)+\psi^{(1)}(\beta+1)\right\},$$

$$\varphi_{11}^{(1)}=\frac{\partial\varphi_{11}}{\partial\alpha}=\frac{2\,n}{\alpha^3}+\left(\frac{2\,n\,\beta}{\alpha^3(\beta+2)}\right)\left\{[\psi(2)-\psi(\beta+1)]^2+\psi^{(1)}(2)+\psi^{(1)}(\beta+1)\right\},$$

$$\begin{aligned}\varphi_{11}^{(2)}=\frac{\partial\varphi_{11}}{\partial\beta}=-\left(\frac{2\,n}{\alpha^2(\beta+2)^2}\right)&\left\{[\psi(2)-\psi(\beta+1)]^2+\psi^{(1)}(2)+\psi^{(1)}(\beta+1)-\right.\\&\left.\beta(\beta+2)\psi^{(1)}(\beta+1)[\psi(2)-\psi(\beta+1)]+\left(\frac{\beta\ (\beta+2)}{2}\right)\beta\ (\beta+2)\psi^{(2)}(\beta+1)\right\},\end{aligned}$$

$$\varphi_{22}^{(1)}=\frac{\partial\varphi_{22}}{\partial\alpha}=0,\qquad \varphi_{22}^{(2)}=\frac{\partial\varphi_{22}}{\partial\beta}=\frac{2\,n}{\beta^3},$$

$$\varphi_{12}^{(1)}=\frac{\partial\varphi_{12}}{\partial\alpha}=-\frac{n\left[\psi(\beta)-\psi(2)\right]}{(\beta+1)\,\alpha^2},\varphi_{12}^{(2)}=\frac{\partial\varphi_{12}}{\partial\beta}=\frac{n\left[(\beta+1)\,\psi^{(1)}(\beta)\ -\psi(\beta)+\psi(2)\right]}{\alpha\,(\beta+1)^2}.$$

Thus, $\omega_{ij}^{(k)} = \varphi_{ij}^{(k)} - \frac{1}{2}\varphi_{ijk}$

$$\omega_{11}^{(1)} = \varphi_{11}^{(1)} - \frac{1}{2}\varphi_{111} = \frac{n}{\alpha^3} + \left(\frac{n\beta}{\alpha^3(\beta+2)(\beta+3)}\right)\Big\{2\ (\beta+3)\ [\psi(2)-\psi(\beta+1)]^2 + 2\ (\beta+3)\ \psi^{(1)}(2) + 2\ (\beta+3)\ \psi^{(1)}(\beta+1) - (\beta+1)\ \Big\{[\psi(2)-\psi(\beta+2)]\Big([\psi(2)-\psi(\beta+2)]^2 + 3\Big[\psi^{(1)}(2)+\psi^{(1)}(\beta+2)\Big]\Big) + \psi^{(2)}(2) - \psi^{(2)}(\beta+2)\Big\} + 2\ [\psi(3)-\psi(\beta+1)]\times \Big([\psi(3)-\psi(\beta+1)]^2 + 3\ \Big[\psi^{(1)}(3)+\psi^{(1)}(\beta+1)\Big]\Big) + 2\,\psi^{(2)}(3) - 2\,\psi^{(2)}(\beta+1)\Big\}.$$