STATISTICS IN TRANSITION new series, December 2021
Vol. 22, No. 4 pp. 171–189, DOI 10.21307/stattrans-2021-044
Received – 30.05.2020; accepted – 26.03.2021

Estimation of the density and cumulative distribution functions of the exponentiated Burr XII distribution

Amal S. Hassan[1], Salwa M. Assar[2], Kareem A. Ali[3], Heba F. Nagy[4,]

ABSTRACT

The exponentiated Burr Type XII (EBXII) distribution has wide applications in reliability and economic studies. In this article, the estimation of the probability density function and the cumulative distribution function of EBXII distribution is considered. We examine the maximum likelihood estimator, the uniformly minimum variance unbiased estimator, the least squares estimator, the weighted least squares estimator, the maximum product spacing estimator, the Cramér–von-Mises estimator, and the Anderson–Darling estimator. We derive analytical forms for the bias and mean square error. A simulation study is performed to investigate the consistency of the suggested methods of estimation. Data relating to the wind speed and service times of aircraft windshields are used with the studied methods. The simulation studies and real data applications have revealed that the maximum likelihood estimator performs more efficiently than its remaining counterparts.

Key words: exponentiated Burr Type XII model, least squares estimator, maximum likelihood estimator, uniform minimum variance unbiased estimator, weighted least squares estimator.

Mathematical Subject Classification: 62F10.

1. Introduction

The Burr Type XII (BXII) distribution has gained special attention in physics, actuarial studies, reliability and applied statistics. Characteristics of the BXII distribution are near to several distributions like exponential, normal, lognormal, etc. Extra properties about the BXII distribution can be found in Headrick et al. (2010).

[1] Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research, Egypt. E-mail: dr.amalelmoslamy@gmail.com. ORCID: https://orcid.org/0000-0003-4442-8458.

[2] Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research, Egypt. E-mail: salwaassar@yahoo.com. ORCID: https://orcid.org/0000-0001-7450-7486.

[3] Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research, Egypt. E-mail: ali.kareem64@yahoo.com. ORCID: https://orcid.org/0000-0002-2522-7733.

[4] Corresponding Author, Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research. Egypt, E-mail: heba_nagy_84@yahoo.com. ORCID: https://orcid.org/0000-0003-0262-205X.

The exponentiated Burr Type XII distribution is a generalization to the BXII distribution through adding a new shape parameter. The cumulative distribution function (CDF) of the EBXII distribution is defined as follows:

$$G(x)=\left(1-\left(1+x^{c}\right)^{-k}\right)^{\beta};\ x,k,c\ ,\beta>0, \tag{1}$$

where, k, c and β are shape parameters. The probability density function (PDF) of the EBXII corresponding to (1) is specified by

$$g\left(x\right)=ck\,\beta x^{c-1}\left(1+x^{c}\right)^{-k-1}\left(1-\left(1+x^{c}\right)^{-k}\right)^{\beta-1};\ x,k,c\ ,\beta>0. \tag{2}$$

Statistical developments on the EBXII model have been studied by several authors. Among them, AL-Hussaini and Hussein (2011) studied maximum likelihood (ML) and Bayesian estimation to the parameters of the EBXII distribution under Type II censored data. Kumar et al. (2017) established several explicit expressions and recurrence relations for single and product moments of r-th lower record values from the EBXII distribution.

Statistical inference is one of the most popular topics in research and scientific studies whether from the theoretical or applied aspects. Most traditional studies have been focused on inferring the parameter(s) involved in the distribution. The importance of statistical distributions is not limited to the characterization of statistical phenomena, but rather to the calculation of many population metrics such as moments, probability weighted moments, failure rate function, etc. However, it would be more useful to study the efficient estimation of the PDF and CDF. The estimation of the PDF and the CDF is important for many reasons. For instance, the best estimators for the PDF can be used to estimate functionals of the PDF such as estimation of Kullback-Leibler divergence, as provided by Hurvich et al. (1990), estimation of Fisher information (see Mielniczuk and Wojtyś (2010), the estimation of the differential entropy (see Nilsson and Kleijn (2007), and estimation of the Rényi entropy. Similarly, the best estimators for the CDF can be used to estimate functionals of the CDF like estimation of quantiles (see Saleh et al. (1988) and estimation of the Lorenz curve (see Woo and Yoon (2001)).

Some studies on the estimation of PDF and CDF have appeared in recent literature for some continuous distributions, for instance, Pareto distribution by Asrabadi (1990) and (Dixit and Nooghabi (2010), Dixit and Nooghabi (2011)), exponentiated Pareto distribution by Jabbari (2010), generalized Rayleigh distribution by Alizadeh et al. (2013), generalized exponential Poisson distribution by Bagheri et al. (2014), exponentiated Weibull by Alizadeh et al. (2015a), generalized exponential distribution by Alizadeh et al. (2015b), exponentiated Gumbel distribution by Bagheri et al. (2016a), Weibull extension distribution by Bagheri et al. (2016b), Lindley distribution by Maiti and Mukherjee (2018), generalized logistic distribution by Tripathi et al. (2017), Frechet distribution by Maleki and Deiri (2017) and Topp-Leone distribution by Benkhelifa (2017), exponentiated gamma distribution (Rasekhi (2018)), and Gompertz distribution (Dey et al. (2018)).

Our objective here is to investigate the efficient estimation of the PDF and the CDF of the EBXII model due to its wide statistical applications and developments. Different parametric methods of estimation, namely ML, uniformly minimum variance unbiased (UMVU), least squares (LS), weighted least squares (WLS), Cramér-von-Mises (CvM), Anderson–Darling (AD) and maximum product spacing (MPS) are considered. This paper is organized as follows. Sections (2) and (3) provide ML and UMVU estimators of the PDF and CDF with their mean square errors (MSEs). Section (4) includes other parametric methods of estimation. Section (5) comprises a simulation study in order to compare different suggested estimators. A real data set is analyzed for illustrative purpose in Section (6). The article ends with concluding remarks in Section (7).

2. Maximum likelihood estimators

In this section we obtain the ML estimators of the PDF and the CDF of the EBXII distribution. Let $X_1, X_2, \ldots, X_n$ be a random sample with size n from the EBXII distribution with known parameters k and c. The log likelihood function of the EBXII distribution is given by

$$L\left(\beta|\underline{x},c,k\right) = nln\left(\beta\right) + nln\left(k\right) + nln\left(c\right) + \left(c-1\right)\sum_{i=1}^{n}\ln\left(x_i\right) - \left(k+1\right)\sum_{i=1}^{n}\ln\left(1+x_i^c\right) + \left(\beta-1\right)\sum_{i=1}^{n}\ln\left(1-\left(1+x_i^c\right)^{-k}\right). \tag{3}$$

The ML estimator of β, say $\hat{\beta}$, is given as

$$\hat{\beta} = \frac{n}{-\sum_{i=1}^{n}\ln\left(1-\left(1+x_i^c\right)^{-k}\right)} = \frac{n}{T}, \qquad T = -\sum_{i=1}^{n}\ln\left(1-\left(1+x_i^c\right)^{-k}\right).$$

We can rewrite CDF(1) as follows:

$$-\ln G(x) = \beta V, \quad V = -\ln(1-(1+x^c)^{-k}).$$

It can be seen that V has an exponential distribution with scale parameter β. Then T has a gamma (n, β), random variable with density function given by

$$f\left(t\right) = \frac{\beta^n}{\Gamma(n)} t^{n-1} e^{-\beta t}, \ t > 0. \tag{4}$$

Therefore, $\hat{\beta}=\frac{n}{T}=S$, has an inverse gamma $(n,n\beta)$ distribution with PDF given by

$$f(s)=\frac{(n\beta)^{n}}{\Gamma(n)}s^{-n-1}e^{-\frac{n\beta}{s}},\quad s>0\,. \tag{5}$$

Applying the invariance property of the ML method, the required PDF and CDF estimators are obtained as follows:

$$\hat{g}(x)=ck\,\hat{\beta}x^{c-1}\left(1+x^{c}\right)^{-k-1}\left(1-\left(1+x^{c}\right)^{-k}\right)^{\hat{\beta}-1},$$

and $$\hat{G}(x)=\left(1-\left(1+x^{c}\right)^{-k}\right)^{\hat{\beta}}. \tag{6}$$

Now, we show that $\hat{g}(x)$ and $\hat{G}(x)$ are biased estimators of $g(x)$ and $G(x)$ respectively. Further, the MSEs are obtained. Theorem (1) calculates $E(\hat{g}(x)^{r})$ and $E(\hat{G}(x)^{r})$.

Theorem 1: We have

$$E(\hat{g}(x)^{r})=\frac{2c^{r}k^{r}b^{r}d^{r\left(\frac{1}{k}+1\right)}(1-d)^{-r}(n\beta)^{n}}{\Gamma(n)}\left(\frac{n\beta}{-rln(1-d)}\right)^{\frac{r-n}{2}}\mathrm{K}_{r-n}(2\sqrt{-n\beta rln(1-d)}),$$

and

$$E(\hat{G}(x)^{r})=2\frac{(n\beta)^{n}}{\Gamma(n)}\left(\frac{n\beta}{-rln(1-d)}\right)^{\frac{-n}{2}}\mathrm{K}_{-n}(2\sqrt{-n\beta rln(1-d)}),$$

where $b=x^{c-1}, d=(1+x^{c})^{-k}$.

Proof: First by using $b=x^{c-1}, d=(1+x^{c})^{-k}$, then $\hat{g}(x)$ can be rewritten as follows: $\hat{g}(x)=cksbd^{\frac{1}{k}+1}(1-d)^{s-1}$, and $\hat{G}(x)=(1-d)^{s}$.

Thus,

$$E\left(\hat{g}(x)^{r}\right)=\int_{0}^{\infty}c^{r}k^{r}s^{r}b^{r}d^{r\left(\frac{1}{k}+1\right)}(1-d)^{rs-r}\frac{(n\beta)^{n}}{\Gamma(n)}s^{-n-1}e^{\frac{-n\beta}{s}}ds$$

$$=c^{r}k^{r}b^{r}d^{r\left(\frac{1}{k}+1\right)}(1-d)^{-r}\frac{(n\beta)^{n}}{\Gamma(n)}\int_{0}^{\infty}s^{r-n-1}e^{-s(-rln(1-d))}e^{\frac{-n\beta}{s}}ds$$

$$=\frac{2c^r k^r b^r d^{r\left(\frac{1}{k}+1\right)}(1-d)^{-r}(n\beta)^n}{\Gamma(n)}\left(\frac{n\beta}{-rln(1-d)}\right)^{\frac{r-n}{2}}\mathrm{K}_{r-n}(2\sqrt{-n\beta rln(1-d)}.$$

Here, $\mathrm{K}_\nu(.)$ denotes the modified Bessel's function of the second kind of order ν (see equation (3.471.9) in (Gradshteyn and Ryzhik (2000)). Similarly, $E(\hat{G}(x)^r)$ takes the following form:

$$E(\hat{G}(x)^r)=2\frac{(n\beta)^n}{\Gamma(n)}\left(\frac{n\beta}{-rln(1-d)}\right)^{\frac{-n}{2}}\mathrm{K}_{-n}(2\sqrt{-n\beta rln(1-d)}.$$

Theorem 2: The MSEs for $\hat{g}(x)$ and $\hat{G}(x)$ respectively are given by

$$MSE(\hat{g}(x))=c^2k^2b^2d^{2\left(\frac{1}{k}+1\right)}(1-d)^{-2}(2\frac{(n\beta)^n}{\Gamma(n)}\left(\frac{n\beta}{-2ln(1-d)}\right)^{\frac{2-n}{2}}\mathrm{K}_{2-n}(2\sqrt{-2n\beta ln(1-d)}$$

$$-4\beta(1-d)^{-\beta}\frac{(n\beta)^n}{\Gamma(n)}\left(\frac{n\beta}{-ln(1-d)}\right)^{\frac{1-n}{2}}\mathrm{K}_{1-n}(2\sqrt{-n\beta ln(1-d)})+\beta^2(1-d)^{-2\beta}).$$

and

$$MSE(\hat{G}(x))=2\frac{(n\beta)^n}{\Gamma(n)}\left(\frac{n\beta}{-2ln(1-d)}\right)^{\frac{-n}{2}}\mathrm{K}_{-n}(2\sqrt{-2n\beta ln(1-d)}-4(1-d)^{\beta}\frac{(n\beta)^n}{\Gamma(n)}\left(\frac{n\beta}{-ln(1-d)}\right)^{\frac{-n}{2}}\mathrm{K}_{-n}(2\sqrt{-n\beta ln(1-d)}+(1-d)^{2\beta}.$$

Proof: Since
$MSE(\hat{g}(x))=E\left(\hat{g}(x)^2\right)-2g(x)E(\hat{g}(x))+g(x)^2$.

Hence, $E\left(\hat{g}(x)^2\right)$ and $E(\hat{g}(x))$ can be obtained by setting r =1 and r =2 in Theorem (1), hence the $MSE(\hat{g}(x))$ is easily calculated. The proof of $MSE(\hat{G}(x))$ is similar.

3. Uniformly minimum variance unbiased estimators

In this section, UMVU estimators of the PDF and CDF of the EBXII distribution are considered. In addition, the rth moment and the MSE of these estimators are derived.

Let $X_1,\ldots,X_n$ be a random sample of size n from the EBXII distribution. Then, $T=-\sum_{i=1}^{n}\ln\left(1-\left(1+x_i^c\right)^{-k}\right)$ is complete sufficient statistic for the parameter β

(assumed k and c are known parameters). Recall that T has a gamma (n, β) distribution with density function (4). According to the Lehmann-Scheffe theorem, if $g(x_1|t) = g^*(t)$ is the conditional PDF of $X_1|T$, we have

$$E\left(g^*(T)\right) = \int g(x_1|t) f(t) dt = \int g(x_1, t) dt = g(x_1),$$

where $g(x_1, t)$ is the joint PDF of X_1 and T. Therefore, $g^*(t)$ is the UMVU estimator of $g(x)$.

Lemma 1: The conditional distribution of V given $T = t$ is obtained as

$$g_{V|T}(v|t) = \frac{(n-1)(t-v)^{n-2}}{t^{n-1}}, v < t < \infty, \qquad V = -\ln\left(1-\left(1+x_1^c\right)^{-k}\right).$$

Proof: We have

$$g_{V|T}(v|t) = \frac{g(v, t-v)}{f(t)} = \frac{\beta^n (t-v)^{n-2} e^{-\beta t} \Gamma(n)}{\Gamma(n-1)\beta^n t^{n-1} e^{-\beta t}} = \frac{(n-1)(t-v)^{n-2}}{t^{n-1}}, v < t < \infty.$$

In the following theorem the UMVU estimators for $g(x)$ and $G(x)$ are obtained.

Theorem 3: The uniformly minimum variance unbiased estimators for $g(x)$ and $G(x)$ are given by

$$\tilde{g}(x) = g^*(t) = \frac{(n-1)\left(t+\ln\left(1-\left(1+x^c\right)^{-k}\right)\right)^{n-2}}{t^{n-1}} \times \frac{kcx^{c-1}\left(1+x^c\right)^{-k-1}}{1-\left(1+x^c\right)^{-k}}, \quad -\ln\left(1-\left(1+x^c\right)^{-k}\right) < t < \infty,$$

and

$$\tilde{G}(x) = \left\{\frac{t+\ln\left(1-\left(1+x^c\right)^{-k}\right)}{t}\right\}^{n-1}.$$

Proof: The estimator $\tilde{g}(x)$ is the UMVU estimator for $g(x)$ can be proved by the Lehmann-Scheffe theorem and Lemma (1). In addition, $\tilde{G}(x)$ is the UMVU estimator of $G(x)$ from the fact that

$$\frac{d}{dx}\tilde{G}(x) = \frac{d}{dx}\left(\frac{t+\ln(1-(1+x^c)^{-k})}{t}\right)^{n-1} = \tilde{g}(x).$$

Further, we compute the MSEs for the two UMVU estimators of $\tilde{g}(x)$ and $\tilde{G}(x)$, suppose that

$$M = \frac{(n-1)kcx^{c-1}\left(1+x^c\right)^{-k-1}}{1-\left(1+x^c\right)^{-k}} \text{ and } p(x) = -\ln\left(1-\left(1+x^c\right)^{-k}\right).$$

Then, we obtain the following expectation:

$$E\left(\tilde{g}(x)^r\right) = \int_{p(x)}^{\infty} \tilde{g}(x)^r f(t)dt = M^r \frac{\beta^n}{\Gamma(n)} \int_{p(x)}^{\infty} \frac{(t-p(x))^{nr-2r}}{t^{nr-r}} t^{n-1} e^{-\beta t} dt.$$

After some simplification, we obtain

$$E\left(\tilde{g}(x)^r\right) = \frac{M^r}{\Gamma(n)} \sum_{i=0}^{nr-2r} \binom{nr-2r}{i} (-1)^i p(x)^i \beta^{i+r} \int_{p(x)\beta}^{\infty} u^{n-i-r-1} e^{-u} du,$$

which $\int_{p(x)\beta}^{\infty} u^{n-i-r-1} e^{-u} du$ is the upper incomplete gamma function, so $E\left(\tilde{g}(x)^r\right)$ can be formulated as follows:

$$E\left(\tilde{g}(x)^r\right) = \sum_{i=0}^{nr-2r} \frac{\beta^{i+r}}{\Gamma(n)} \binom{nr-2r}{i} (-1)^i M^r p(x)^i \Gamma(n-i-r, p(x)\beta). \tag{7}$$

Similarly, we can prove that

$$E\left(\tilde{G}(x)^r\right) = \sum_{i=0}^{nr-r} \frac{\beta^i}{\Gamma(n)} \binom{nr-r}{i} (-1)^i p(x)^i \Gamma(n-i, p(x)\beta). \tag{8}$$

Theorem 4. The mean square errors for $\tilde{g}(x)$ and $\tilde{G}(x)$, respectively, are given by

$$MSE\left(\tilde{g}(x)\right) = \sum_{i=0}^{2n-4} \frac{M^2}{\Gamma(n)} \binom{2n-4}{i} (-1)^i \beta^{i+2} p(x)^i \Gamma(n-i-2, p(x)\beta) - (\mathrm{g}(x))^2,$$

and

$$MSE\left(\tilde{G}(x)\right) = \sum_{i=0}^{2n-2} \frac{\beta^i}{\Gamma(n)} \binom{2n-2}{i} (-1)^i p(x)^i \Gamma(n-i, p(x)\beta) - (\mathrm{G}(x))^2.$$

Proof: Since

$$MSE\left(\tilde{g}(x)\right) = E\left(\tilde{g}(x)^2\right) - g(x)^2,$$

where $E\left(\tilde{g}(x)^2\right)$ can be obtained by setting r =2 in (7), hence we can calculate $MSE\left(\tilde{g}(x)\right)$. The proof of $MSE\left(\tilde{G}(x)\right)$ is similar.

4. Other parametric methods of estimation

In this section, several methods of estimation such as LS, WLS, MPS, CvM and AD are considered. All these methods are based on the CDF. Let $X_{i:n}, i=1,\ldots,n,$ be the order statistics of a random sample from the EBXII distribution and assumed k and c are known parameters. Then, the LS, WLS, MPS, CvM and AD estimators of the PDF and the CDF of the EBXII distribution are derived in the following subsections.

4.1. Least squares and weighted least squares estimators

The ordinary least squares and the weighted least squares (Swain et al. (1988)) are well-known methods used for estimating the unknown parameters. The LS estimator of β, say, β' and the WLS estimator of β, say, $\bar{\beta}$, are given by minimizing the following quantities with respect to β

$$\sum_{i=1}^{n}\left(G(x_{i:n})-\frac{i}{n+1}\right)^2, \tag{9}$$

$$\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i\,(n-i+1)}\left(G(x_{i:n})-\frac{i}{n+1}\right)^2. \tag{10}$$

There is no closed form solution for β, in minimizing Equations (9) and (10), so the numerical technique is applied to find β' and $\bar{\beta}$. Hence, the LS and WLS estimators of the CDF and PDF for the EBXII distribution are obtained, respectively, as follows:

$$G'(x)=\left(1-\left(1+x^c\right)^{-k}\right)^{\beta'},\quad g'(x)=ck\,\beta' x^{c-1}\left(1+x^c\right)^{-k-1}\left(1-\left(1+x^c\right)^{-k}\right)^{\beta'-1},$$

and

$$\bar{G}(x)=\left(1-\left(1+x^c\right)^{-k}\right)^{\bar{\beta}},\ \bar{g}(x)=ck\,\bar{\beta}x^{c-1}\left(1+x^c\right)^{-k-1}\left(1-\left(1+x^c\right)^{-k}\right)^{\bar{\beta}-1}.$$

4.2. Maximum product of spacing estimators

The MPS has been proposed by Cheng and Amin (1979) as an alternative method for the ML for the estimation parameters of continuous univariate distribution. Let a sample of size n be available from EBXII, we define the corresponding uniform spacings as follows:

$D_i=G(x_{i:n})-G(x_{i-1:n}),\qquad i=1,2,\ldots,n$ where $G(x_{0:n})=0,\ G(x_{n+1:n})=1,$

$\sum_{i=1}^{n+1}D_i=1.$

The MPS estimator of β, say $\breve{\beta}$, can be obtained by maximizing the geometric mean of the spacings,

$$D(\beta)=\left(\prod_{i=1}^{n+1} D_i\right)^{\frac{1}{n+1}}=\prod_{i=1}^{n+1}\left(\mathrm{G}(x_{i:n})-G(x_{i-1:n})\right)^{\frac{1}{n+1}},$$

with respect to β. Equivalently, the following expression

$$D(\beta)=\frac{1}{n+1}\sum_{i=1}^{n+1}\ln\left(G(x_{i:n})-G(x_{i-1:n})\right), \tag{11}$$

can be maximized to obtain the desired estimator of β. It can be shown that $\breve{\beta}$ satisfies

$$\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{\left(G(x_{i:n})-G(x_{i-1,n})\right)}\left[D_0(x_{i:n})-D_0(x_{i-1:n})\right]=0, \tag{12}$$

$D_0(x_{i:n})=\left(1-\left(1+x_{i:n}{}^{c}\right)^{-k}\right)^{\beta}\ln\left(1-\left(1+x_{i:n}{}^{c}\right)^{-k}\right)$. There is no closed form solution for β in (12), so the numerical technique is applied to find $\breve{\beta}$. Now, the MPS estimators of the CDF and PDF are obtained as follows:

$$\breve{G}(x)=\left(1-\left(1+x^{c}\right)^{-k}\right)^{\breve{\beta}},\quad \breve{g}(x)=ck\breve{\beta}x^{c-1}\left(1+x^{c}\right)^{-k-1}\left(1-\left(1+x^{c}\right)^{-k}\right)^{\breve{\beta}-1}.$$

4.3. Cramér-von-Mises and Anderson Darling estimators

Cramér-von-Mises is a type of minimum distance estimators (also called maximum goodness of fit estimators), which is based on the difference between the estimate of the CDF and its empirical distribution function. The CvM estimator of β, say $\ddot{\beta}$, is obtained by minimizing the following function with respect to β

$$C(\beta)=\frac{1}{12n}+\sum_{i=1}^{n}\left(G(x_{i:n})-\frac{2i-1}{2n}\right)^{2}. \tag{13}$$

Equivalently, we solve the equation

$$\sum_{i=1}^{n}\left(G(x_{i:n})-\frac{2i-1}{2n}\right)D_0(x_{i:n})=0,$$

to obtain β in the previous equation. So, the CvM estimators of the CDF and PDF are obtained as follows:

$$\ddot{G}(x)=\left(1-\left(1+x^{c}\right)^{-k}\right)^{\ddot{\beta}},\quad \ddot{g}(x)=ck\ddot{\beta}x^{c-1}\left(1+x^{c}\right)^{-k-1}\left(1-\left(1+x^{c}\right)^{-k}\right)^{\ddot{\beta}-1}.$$

Additionally, the Anderson-Darling method was initially discussed by Anderson and Darling (1952) as another type of minimum distance estimators. The AD estimator of β , say β°, is obtained by minimizing, with respect to β the function

$$A(\beta) = -n + \sum_{i=1}^{n} (2i-1)\left[\log(G(x_{i:n}) + \log(1-G(x_{n+1-i:n})\right]. \tag{14}$$

Analogously, these estimators can also be obtained by solving the following non-linear equation

$$\sum_{i=1}^{n} (2i-1)\left[\frac{D_0(x_{i:n})}{G(x_{i:n})} + \frac{D_0(x_{n+1-i:n})}{1-G(x_{n+1-i:n})}\right] = 0,$$

with respect to β . Therefore, the AD estimators of the CDF and PDF for the EBXII distribution are obtained as follows:

$$G^{\circ}(x) = \left(1-\left(1+x^{c}\right)^{-k}\right)^{\beta^{\circ}}, \quad g^{\circ}(x) = ck\,\beta^{\circ} x^{c-1}\left(1+x^{c}\right)^{-k-1}\left(1-\left(1+x^{c}\right)^{-k}\right)^{\beta^{\circ}-1}.$$

5. Simulation study

A simulation study is carried out in this section to determine the efficient estimate of the PDF and CDF between the following estimates: ML, UMVU, LS, WLS, MPS, CvM and AD. The MSE is used to compare between these estimators. One thousand random samples with sizes n=10,25,50,75 and 100 are generated from the EBXII distribution with different sets of parameters (β, k, c) = (0.5,2,2.5), (1,2,2.5), (2,2,2.5), (1.5,2,1.5), (2,2,1.5) and (4,2,1.5). The MSEs of each estimate are displayed in Tables 1 and 2. It can be detected from these tables that ML and UMVU estimates are more efficient than the other corresponding estimates.

In addition, Figures 1 to 6 display the MSEs values for the PDF and CDF of the EBXII distribution by proposed methods. The left-hand side graph in each figure is related to the PDF estimates and the corresponding right-hand side graph is the CDF estimates. Generally, the efficiency of all estimates improves as sample size increases.

Figure 7 and Figure 8 show the performance of a different set of the parameters. We detect from these figures that the ML and UMVU estimates of the set of parameters (β, k, c)= (4, 2, 1.5) have better statistical properties than the other corresponding studied sets.

Table 1. The MSEs of the PDF and CDF estimates of the EBXII distribution for (β, k, c)= (0.5,2,2.5), (1,2,2.5) and (2,2,2.5)

Sample size	Methods	$\beta = 0.5$		$\beta = 1$		$\beta = 2$	
		PDF	CDF	PDF	CDF	PDF	CDF
n=10	ML	0.11	7.917e-3	0.026	8.763e-3	0.021	7.979e-3
	UMVU	0.115	8.007e-3	0.031	8.753e-3	0.026	7.957e-3
	MPS	0.201	8.227e-3	0.03	8.849e-3	0.023	8.047e-3
	CvM	0.107	0.012	0.04	0.012	0.037	0.013
	AD	0.111	8.739e-3	0.028	9.135e-3	0.024	8.788e-3
	LS	0.111	0.013	0.042	0.013	0.039	0.014
	WLS	0.116	0.012	0.039	0.012	0.037	0.013
n=25	ML	0.039	2.952e-3	9.415e-3	2.884e-3	8.208e-3	2.962e-3
	UMVU	0.037	2.915e-3	0.01	2.871e-3	8.865e-3	2.951e-3
	MPS	0.056	2.965e-3	0.01	2.971e-3	8.741e-3	3.058e-3
	CvM	0.032	3.902e-3	0.013	3.824e-3	0.011	3.917e-3
	AD	0.045	3.257e-3	0.011	3.333e-3	9.48e-3	3.368e-3
	LS	0.033	3.906e-3	0.013	4.017e-3	0.011	3.939e-3
	WLS	0.032	3.626e-3	0.012	3.482e-3	9.925e-3	3.495e-3
n=50	ML	0.012	1.483e-3	4.635e-3	1.393e-3	4.181e-3	1.46e-3
	UMVU	0.012	1.482e-3	4.844e-3	1.398e-3	4.365e-3	1.464e-3
	MPS	0.016	1.535e-3	5.113e-3	1.47e-3	4.501e-3	1.534e-3
	CvM	0.015	1.969e-3	6.104e-3	1.832e-3	5.476e-3	1.914e-3
	AD	0.014	1.752e-3	5.435e-3	1.629e-3	4.875e-3	1.709e-3
	LS	0.015	1.972e-3	6.126e-3	1.836e-3	5.485e-3	1.917e-3
	WLS	0.014	1.791e-3	5.635e-3	1.686e-3	5.048e-3	1.765e-3
n=75	ML	8.442e-3	9.864e-4	3.014e-3	9.171e-4	2.685e-3	9.409e-4
	UMVU	8.49e-3	9.893e-4	3.07e-3	9.121e-4	2.729e-3	9.351e-4
	MPS	0.01	1.032e-3	3.142e-3	9.3e-4	2.74e-3	9.485e-4
	CvM	0.011	1.319e-3	4.13e-3	1.235e-3	3.733e-3	1.291e-3
	AD	9.722e-3	1.175e-3	3.616e-3	1.086e-3	3.247e-3	1.13e-3
	LS	0.011	1.323e-3	4.134e-3	1.236e-3	3.732e-3	1.291e-3
	WLS	9.86e-3	1.186e-3	3.716e-3	1.111e-3	3.319e-3	1.151e-3
n=100	ML	5.985e-3	7.63e-4	2.508e-3	7.553e-4	2.135e-3	7.462e-4
	UMVU	6.005e-3	7.619e-4	2.562e-3	7.567e-4	2.183e-3	7.481e-4
	MPS	6.877e-3	7.802e-4	2.672e-3	7.855e-4	2.257e-3	7.785e-4
	CvM	7.305e-3	9.491e-4	3.329e-3	1.005e-3	2.808e-3	9.835e-4
	AD	6.731e-3	8.642e-4	2.959e-3	8.928e-4	2.5e-3	8.764e-4
	LS	7.351e-3	9.495e-4	3.339e-3	1.007e-3	2.814e-3	9.858e-4
	WLS	6.819e-3	8.725e-4	2.978e-3	8.966e-4	2.515e-3	8.792e-4

Table 2. The MSEs of the PDF and CDF estimates of the EBXII distribution for the set of parameters (β, k, c) = (1.5,2,1.5), (2,2,1.5) and (4,2,1.5)

Sample size	Methods	$\beta = 1.5$		$\beta = 2$		$\beta = 4$	
		PDF	CDF	PDF	CDF	PDF	CDF
n=10	ML	0.022	7.947e-3	0.013	7.405e-3	7.499e-3	7.8e-3
	UMVU	0.028	7.907e-3	0.017	7.287e-3	9.843e-3	7.826e-3
	MPS	0.029	7.971e-3	0.017	7.291e-3	9.129e-3	7.948e-3
	CvM	0.03	0.012	0.019	0.011	0.012	0.011
		0.026	8.779e-3	0.016	8.207e-3	9.04e-3	8.78e-3
	AD	0.032	0.014	0.02	0.011	0.013	0.011
	LS	0.03	0.012	0.019	0.01	0.012	0.011
	WLS						
n=25	ML	8.38e-3	3.083e-3	6.48e-3	3.386e-3	3.285e-3	3.042e-3
	UMVU	9.201e-3	3.076e-3	7.156e-3	3.398e-3	3.574e-3	2.997e-3
	MPS	0.01	3.187e-3	7.699e-3	3.544e-3	3.601e-3	3.008e-3
	CvM	0.011	3.973e-3	8.25e-3	4.274e-3	4.578e-3	4.11e-3
	AD	9.843e-3	3.544e-3	7.366e-3	3.782e-3	3.99e-3	3.596e-3
	LS	0.011	3.985e-3	8.338e-3	4.285e-3	4.616e-3	4.119e-3
	WLS	0.01	3.709e-3	7.932e-3	4.043e-3	4.29e-3	3.823e-3
n=50	ML	3.934e-3	1.504e-3	3.015e-3	1.522e-3	1.631e-3	1.377e-3
	UMVU	4.112e-3	1.499e-3	3.178e-3	1.529e-3	1.714e-3	1.377e-3
	MPS	4.453e-3	1.544e-3	3.43e-3	1.614e-3	1.801e-3	1.437e-3
	CvM	5.304e-3	1.961e-3	3.972e-3	1.935e-3	2.246e-3	1.884e-3
	AD	4.687e-3	1.732e-3	3.507e-3	1.716e-3	1.979e-3	1.659e-3
	LS	5.34e-3	1.964e-3	4.004e-3	1.944e-3	2.258e-3	1.889e-3
	WLS	4.858e-3	1.788e-3	3.634e-3	1.761e-3	2.052e-3	1.712e-3
n=75	ML	2.699e-3	1.01e-3	1.972e-3	9.846e-4	1.121e-3	9.636e-4
	UMVU	2.768e-3	1.004e-3	2.048e-3	9.896e-4	1.153e-3	9.608e-4
	MPS	2.934e-3	1.023e-3	2.203e-3	1.043e-3	1.201e-3	9.888e-4
	CvM	3.451e-3	1.252e-3	2.615e-3	1.299e-3	1.435e-3	1.219e-3
	AD	3.109e-3	1.129e-3	2.339e-3	1.158e-3	1.277e-3	1.086e-3
	LS	3.465e-3	1.253e-3	2.63e-3	1.303e-3	1.439e-3	1.22e-3
	WLS	3.169e-3	1.148e-3	2.361e-3	1.165e-3	1.297e-3	1.102e-3
n=100	ML	1.951e-3	7.399e-4	1.354e-3	7.04e-4	8.727e-4	7.468e-4
	UMVU	1.987e-3	7.365e-4	1.38e-3	7.005e-4	8.922e-4	7.453e-4
	MPS	2.087e-3	7.477e-4	1.434e-3	7.1e-4	9.219e-4	7.614e-4
	CvM	2.619e-3	9.819e-4	1.815e-3	9.273e-4	1.189e-3	1.007e-3
	AD	2.318e-3	8.687e-4	1.61e-3	8.245e-4	1.052e-3	8.908e-4
	LS	2.625e-3	9.817e-4	1.819e-3	9.276e-4	1.192e-3	1.007e-3
	WLS	2.361e-3	8.829e-4	1.634e-3	8.345e-4	1.065e-3	9.022e-4

Note: in the tables, e denoted to base 10.

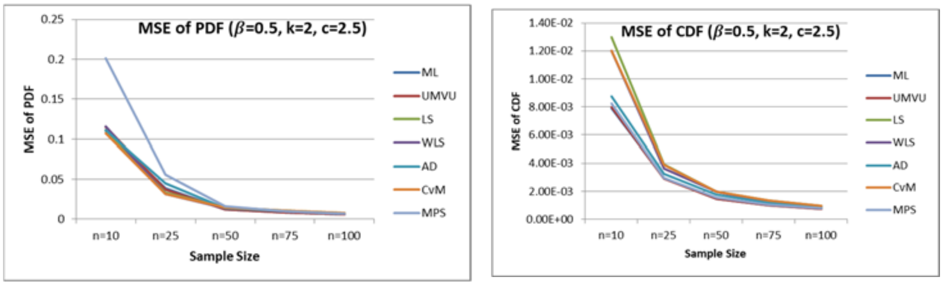

Figure1. The MSEs of the PDF and CDF for the parameter set (β, k, c) = (0.5,2,2.5)

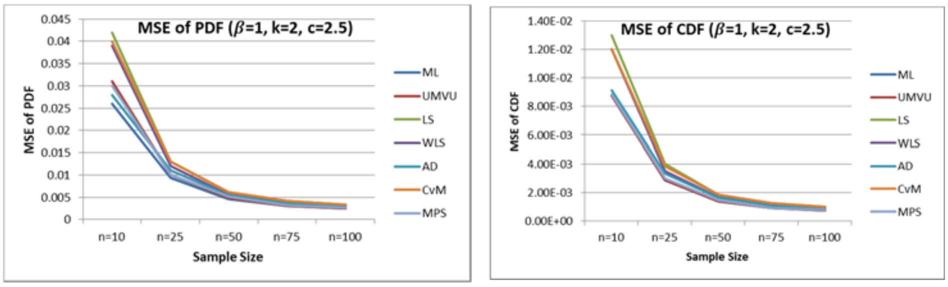

Figure 2. The MSEs of the PDF and CDF for the parameter set (β, k, c) = (1,2,2.5)

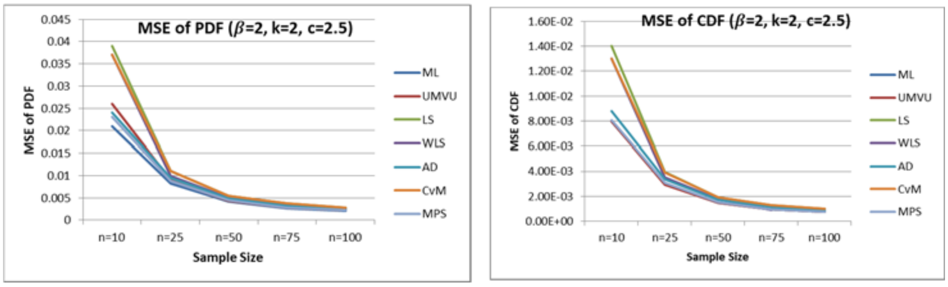

Figure 3. The MSEs of the PDF and CDF for the parameter set (β, k, c) = (2,2,2.5)

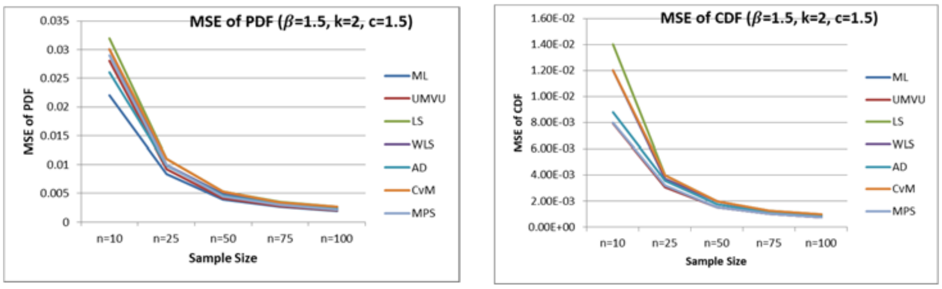

Figure 4. The MSEs of the PDF and CDF for the parameter set (β, k, c) = (1.5,2,1.5)

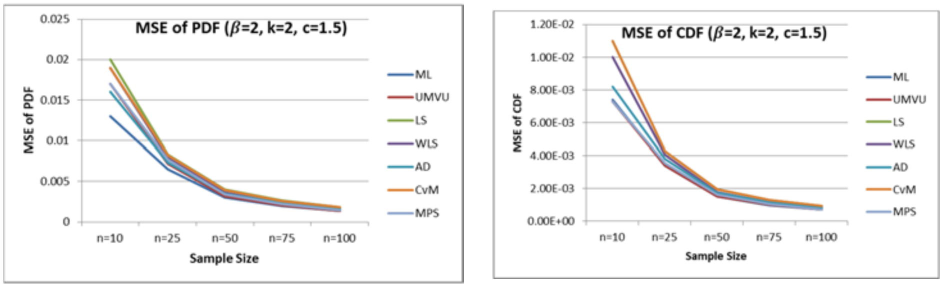

Figure 5. The MSEs of the PDF and CDF for the parameter set (β, k, c) = (2,2,1.5)

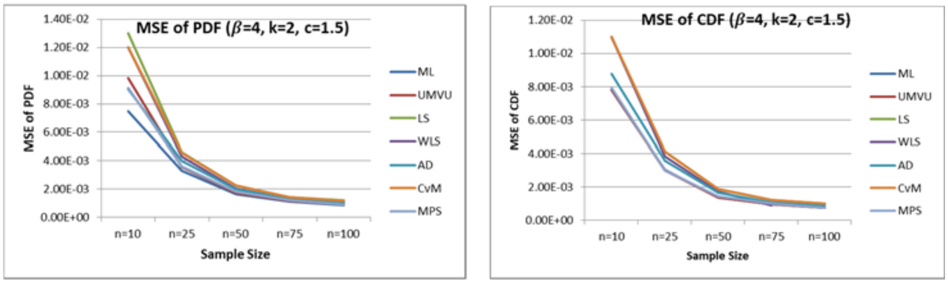

Figure 6. The MSEs of the PDF and CDF for the parameter set (β, k, c) = (4,2,1.5)

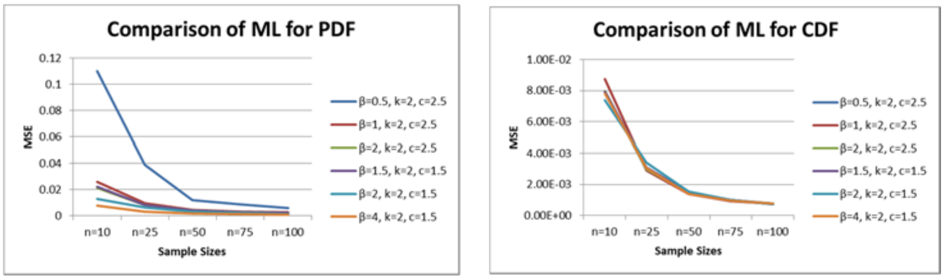

Figure 7. The MSEs of the PDF and CDF of the ML estimators for all the parameter sets

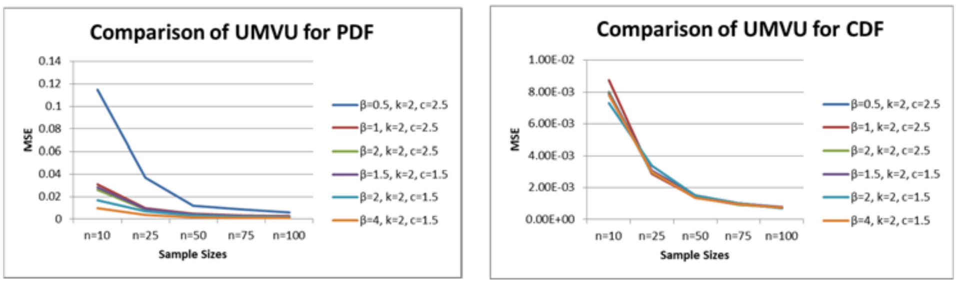

Figure 8. The MSEs of the PDF and CDF of the UMVU estimators for all parameter sets

6. Application to real data

Real data sets are considered to compare between ML, LS, WLS, CvM, AD and MPS methods. The first data consist of 31 observations that represent the Average Monthly Wind Speed (m/s) at Kolkata (from 1st March, 2009 to 31st March, 2009); these data were introduced by Bhattacharya and Bhattacharjee (2010). The second data set represents the data on service times of 63 aircraft windshield given by Murthy et al. (2004). For both data sets, all the three parameters are considered as unknown parameters. The parameters are estimated by ML, MPS, LS, WLS, CvM and AD methods. ML, MPS estimators are obtained by maximizing Equations (3) and (11), respectively, with respect to β, k and c. LS, WLS, CvM and AD estimators can be obtained by minimizing Equations (9), (10), (13) and (14), respectively, with respect to β , k and c. We compared the estimation methods by means of model selection criteria. The criteria like Akaike information criterion (AIC), Bayesian information criterion (BIC), and corrected Akaike information criterion (AICc) are considered. The model with the minimum AIC, BIC and AICc is chosen as the best model to fit the data. In addition, the PDF plot (estimated PDFs versus the empirical histogram for the data) and the CDF plot (estimated CDFs versus the empirical CDF for the data) are used in the model selection. Tables 3 and 4 give the parameter estimates and the values of the model selection for different methods.

Table 3. Estimates of the parameters and the corresponding AIC, BIC and AICc for first data

Methods	c Estimate	k Estimate	β Estimate	AIC	BIC	AICc
ML	2.139	1.631	1.333	56.304	60.605	56.946
LS	3.295	0.686	0.599	57.841	62.143	58.484
WLS	3.113	0.856	0.701	56.921	61.223	57.564
CvM	3.278	0.757	0.647	57.237	61.539	57.88
AD	3.461	0.701	0.585	57.317	61.619	57.96
MPS	1.198	1.040	0.989	79.264	83.566	79.687

Table 4. Estimates of the parameters and the corresponding AIC, BIC and AICc for second data

Methods	c Estimate	k Estimate	β Estimate	AIC	BIC	AICc
ML	1.378	1.206	1.988	235.331	241.76	235.527
LS	1.148	2.103	6.789	270.693	277.123	270.89
WLS	1.174	2.302	8.321	291.356	297.786	291.553
CvM	1.154	2.142	7.156	275.685	282.115	275.882
AD	1.381	1.242	2.456	237.674	244.103	237.87
MPS	1.402	1.1	1.734	235.739	242.168	235.939

As seen from Tables 3 and 4, the ML estimates give the smallest values compared with the other estimates. Figures 9 and 10 represent plots of the CDFs and PDFs of the EBXII distribution based on the fitted ML, LS, WLS, CvM, AD and MPS methods to the data, the figures indicate the superiority of the ML method over the other methods.

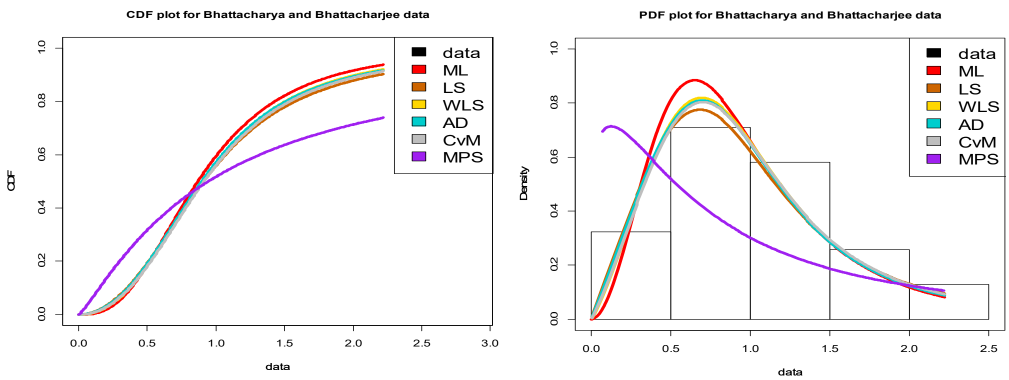

Figure 9. CDF and PDF plots for Wind Speed (m/s) data fitted by different methods of estimation

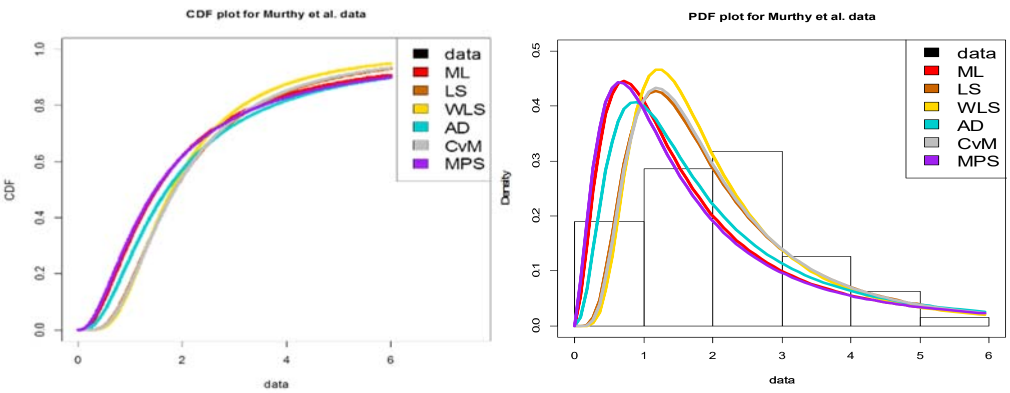

Figure 10. CDF and PDF plots for service times of 63 aircraft windshield fitted by different methods of estimation

7. Conclusion

In this paper, we consider seven different estimators of the PDF and CDF of the EBXII distribution when the shape parameters k and c are assumed to be known. Maximum likelihood estimator, uniformly minimum variance unbiased estimator, least squares estimator, weighted least squares estimator, maximum product spacing estimator, Cramér-von-Mises estimator and Anderson-Darling estimator are obtained. The MSEs of the maximum likelihood and uniformly minimum variance unbiased estimators are given in explicit forms. A simulation study is performed to compare the behaviours of the proposed estimates. A real data set is considered for illustrative purposes. The results show that the maximum likelihood and uniformly minimum variance unbiased estimates perform better than the other estimators.

References

al-Hussaini, E. K., Hussein, M., (2011). Estimation using censored data from exponentiated Burr type XII population. *American Open Journal of Statistics,* 1, pp. 33–45.

Alizadeh, M., Bagheri, F., M Khaleghy Moghaddam, M., (2013). Efficient estimation of the density and cumulative distribution function of the generalized Rayleigh distribution. *Journal of Statistical Research of Iran JSRI,* 10(1), pp. 1–22.

Alizadeh, M., Bagheri, S., Jamkhaneh, E. B., Nadarajah, S., (2015a). Estimates of the PDF and the CDF of the exponentiated Weibull distribution. *Brazilian Journal of Probability and Statistics,* 29(3), pp. 695–716.

Alizadeh, M., Rezaei, S., Bagheri, S., Nadarajah, S., (2015b). Efficient estimation for the generalized exponential distribution. *Statistical Papers,* 56(4), pp. 1015–1031.

Anderson, T. W., Darling, D. A., (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, pp. 193–212.

Asrabadi, B. R., (1990). Estimation in the Pareto distribution. *Metrika,* 37(1), pp. 199–205.

Bagheri, S., Alizadeh, M., Baloui Jamkhaneh, E., Nadarajah, S., (2014). Evaluation and comparison of estimations in the generalized exponential-Poisson distribution. *Journal of Statistical Computation and Simulation,* 84(11), pp. 2345–2360.

Bagheri, S., Alizadeh, M., Nadarajah, S., (2016a). Efficient estimation of the PDF and the CDF of the exponentiated Gumbel distribution. *Communications in Statistics-Simulation and Computation,* 45(1), pp. 339–361.

Bagheri, S., Alizadeh, M., Nadarajah, S., Deiri, E., (2016b). Efficient estimation of the PDF and the CDF of the Weibull extension model. *Communications in Statistics-Simulation and Computation,* 45(6), pp. 2191–2207.

Benkhelifa, L., (2017). Efficient estimation in the Topp-Leone distribution. *arXiv preprint arXiv:1701.03822*.

Bhattacharya, P., Bhattacharjee, R., (2010). A study on Weibull distribution for estimating the parameters. *Journal of Applied Quantitative Methods,* 5(2), pp. 234–241.

Cheng, R. C. H. , Amin, N. A. K., (1979). Maximum product of spacings estimation with application to the lognormal distribution. *Cardiff,Math.* University of Wales Institute of Science and Technology.

Dey, S., Kayal, T., Tripathi, Y. M., (2018). Evaluation and comparison of estimators in The Gompertz Distribution. *Annals Of Data Science,* 5(2), pp. 235–258.

Dixit, U., Nooghabi, M. J., (2010). Efficient estimation in the Pareto distribution. *Statistical Methodology,* 7(6), pp. 687–691.

Dixit, U., Nooghabi, M. J., (2011). Efficient estimation in the Pareto distribution with The Presence Of Outliers. *Statistical Methodology,* 8(4), pp. 340–355.

Gradshteyn, I., Ryzhik, I., (2000). *Table of Integrals, Series, and Products, Academic Press, San Diego, CA*.

Headrick, T., Pant, M., Sheng, Y., (2010). On simulating univariate and multivariate Burr Type III and Type XII distributions. *Applied Mathematical Sciences,* 4, pp. 2207–2240.

Hurvich, C. M., Shumway, R., Tsai, C.-L., (1990). Improved estimators of Kullback–Leibler information for autoregressive model selection in small samples. *Biometrika,* 77(4), pp. 709–719.

Jabbari, N. H., (2010). Efficient estimation of PDF, CDF and rth moment for the exponentiated Pareto distribution in the presence of outliers. *Statistics: A Journal of Theoretical and Applied Statistics,* 44(4), pp. 1–20.

Kumar, D., Saran, J., Jain, N., (2017). The exponentiated burr xii distribution: moments and estimation based on lower record values. *Sri Lankan Journal of Applied Statistics,* 18(1), pp. 1–18.

Maiti, S. S., Mukherjee, I., (2018). On estimation of the PDF and CDF of the Lindley distribution. *Communications in Statistics-Simulation and Computation,* 47(5), pp. 1370–1381.

Maleki, F., Deiri, E., (2017). Efficient Estimation of the PDF and the CDF of the Frechet Distribution. *Annals of Data Science,* 4(2), pp. 211–225.

Mielniczuk, J., Wojtyś, M., (2010). Estimation of Fisher information using model selection. *Metrika,* 72(2), pp. 163–187.

Murthy, D. P., Xie, M., Jiang, R., (2004). *Weibull Models,* Haboken, NJ, USA, John Wiley & Sons.

Nilsson, M., Kleijn, W. B., (2007). On the estimation of differential entropy from data located on embedded manifolds. *IEEE Transactions on Information Theory, 53*(7), pp. 2330–2341.

Rasekhi, M., (2018). A study on methods for estimating the PDF and the CDF in the exponentiated gamma distribution. *Communications in Statistics-Simulation and Computation*, pp. 1–15.

Saleh, A. M. E., Hassanein, K., Ali, M. M., (1988). Estimation and testing of hypotheses about the quantile function of the normal distribution. *Journal of Information and Optimization Sciences,* 9(1), pp. 85–98.

Swain, J. J., Venkatraman, S., Wilson, J. R., (1988). Least-squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation,* 29(4), pp. 271–297.

Tripathi, Y. M., Mahto, A. K., Dey, S., (2017). Efficient Estimation of the PDF and the CDF of a Generalized Logistic Distribution. *Annals of Data Science,* 4(1), pp. 63–81.

Woo, J. S., Yoon, G. E., (2001). Estimations of Lorenz curve and Gini index in a Pareto distribution. *Communications for Statistical Applications and Methods,* 8(1), pp. 249–256.