Properties of the Transmuted Burr XII Distribution, Regression and its Applications

Ahmed Z. Afify[1], Gauss M. Cordeiro[2], Marcelo Bourguignon[3]
, and Edwin M. M. Ortega[4]

[1]*Department of Statistics, Mathematics and Insurance,Benha University, Egypt* [2]*Departamento de Estatıstica, Universidade Federal de Pernambuco, 50740-540, Recife, PE, Brazil* [3]*Departamento de Estatıstica,Universidade Federal de Pernambuco, 50740-540 Recife, PE, Brazil*[4]*Departamento de Ciˆencias Exatas, Universidade de S˜ao Paulo,50740-540 Piracicaba, SP, Brazil*

Abstract: In this paper, we define and study a four-parameter model called the transmuted Burr XII distribution. We obtain some of its mathematical properties including explicit expressions for the ordinary and incomplete moments, generating function, order statistics, probability weighted moments and entropies. We formulate and develop a log-linear model using the new distribution so-called the log-transmuted Burr XII distribution for modeling data with a unimodal failure rate function, as an alternative to the log-McDonald Burr XII, log-beta Burr XII, log-Kumaraswamy Burr XII, log-Burr XII and logistic regression models. The flexibility of the proposed models is illustrated by means of three applications to real data sets.

Keywords: Entropy; Maximum likelihood; Moment; Order statistic; Transmuted family.

1. Introduction

Recently there has been an increased interest among statisticians in proposing new methods to generate univariate continuous distributions, which have been extensively used for modeling data in several applied sciences such as engineering, lifetime analysis, economics, medical and actuarial studies, demography, finance and insurance. So, there are many well-known generators in the statistical literature. For example, Marshall and Olkin (1997) pioneered the Marshall- Olkin-G (MO-G), Eugene et al. (2002) proposed the beta-G (B-G), Zografos and Balakrishnan (2009) defined the gamma-G (G-G), Cordeiro and de Castro (2011) studied the Kumaraswamy- G (Kw-G), Alexander et al. (2012) defined the McDonald-G (Mc-G), Alzaatreh et al. (2013) proposed the transformer (T-X), Torabi and Montazari (2014) introduced the logistic-G (Lo-G), Bourguignon et al. (2014) investigated the Weibull-G (W-G), Nofal et al. (2017) proposed the generalized transmuted-G (GT-G), Afify et al. (2017b) introduced the beta transmuted-H (BT- H), Afify et al. (2016a) defined the transmuted

geometric-G (TG-G) and Afify et al. (2017a) studied the complementary geometric transmuted-G families. However, there is a clear need for generating new classes to provide more flexibility in modeling and analyzing lifetime data.

Consider a baseline cumulative distribution function (cdf) tt(x; φ) and probability density function (pdf) g(x; φ) depending on a parameter vector $\varphi = (\varphi_1, \dots, \varphi_k)$. The cdf and pdf of the transmuted-G (TG) family of distributions are defined by

$$F_1(x; \lambda, \varphi) = G(x; \varphi)[1 + \lambda - \lambda G(x; \varphi)] \tag{1}$$

and

$$f_1(x; \lambda, \varphi) = g(x; \varphi)[1 + \lambda - 2\lambda G(x; \varphi)] \tag{2}$$

respectively, where $|\lambda| \leq 1$.

The TG density function is a mixture of the baseline density and the exponentiated-G (Exp-G) density with power parameter two. The baseline cdf tt(x; φ) is clearly a special case of (1) when $\lambda = 0$. Further details of the TG family were explored by Shaw and Buckley (2007).

The pdf and cdf of the TG class can be expressed, respectively, as

$$f_1(x; \lambda, \varphi) = g(x; \varphi)[1 + \lambda - 2\lambda \overline{G}(x; \varphi)] \tag{3}$$

and

$$F(x; \lambda, \varphi) = 1 - \overline{G}(x; \varphi)[1 - \lambda + \lambda \overline{G}(x; \varphi)] \tag{4}$$

The reliability function (rf), hazard rate function (hrf) and cumulative hazard rate function (chrf) corresponding to (3) are given by

$$R(x; \lambda, \varphi) = \overline{G}(x; \varphi)[1 - \lambda + \lambda \overline{G}(x; \varphi)$$

$$h(x; \lambda, \varphi) = \tau(x; \varphi) \frac{[1 - \lambda + 2\lambda \overline{G}(x; \varphi)]}{[1 - \lambda + \lambda \overline{G}(x; \varphi)]}$$

and

$$H(x; \lambda, \varphi) = -\log\{\overline{G}(x; \varphi)[1 - \lambda + \lambda \overline{G}(x; \varphi)]\}$$

respectively, where $\overline{G}(x; \varphi) = 1 - G(x; \varphi)$ and $\tau(x; \varphi)$ is the the baseline hrf.

A simple motivation for the TG family follows by taking two iid random variables, say T1 and T2, with cdf $G(x; \varphi) = 1 - \overline{G}(x; \varphi)$. Let $T_{1:2} = \min(T_1, T_2)$ and $T_{2:2} = \max(T_1, T_2)$. Next, consider the random variable X defined by

$$X = \begin{cases} T_{1:2}, & \text{with probability } \dfrac{\lambda + 1}{2}; \\ T_{2:2}, & \text{with probability } \dfrac{1 - \lambda}{2}. \end{cases}$$

Then, the cdf of X is given by (4).

Recently, some authors proposed generalizations of the BXII distribution. For example, the beta BXII (Paranaíba et al., 2011), Kumaraswamy BXII (Paranaíba et al., 2013), McDonald BXII (Gomes et al., 2015), beta exponentiated BXII (Mead, 2014), Marshall-Olkin extended BXII (Al-Saiarie et al., 2014), Kumaraswamy exponentiated BXII (Mead and Afify, 2017) and Weibull BXII (Afify et al., 2016b), generalized Burr-normal (Nasiret al., 2017) distributions.

Bourguignon et al. (2016) and Al-Khazaleh (2016) introduced the transmuted Burr XII (TBXII)

distribution. However, they do not investigated its regression, applications and several properties of the TBXII distribution. Therefore, the main objective of this paper is to study the TBXII distribution defined from the TG family and give a comprehensive account of some of its mathematical properties. Further, we prove empirically that the TBXII model provides better fits than at least three other competitive models in two applications, each model having more parameters.

This paper is organized as follows. In Section 2, we define the TBXII model and provide some plots of its pdf and hrf. In Section 3, we obtain some of its structural properties. The performance of the maximum likelihood estimators (MLEs) is investigated by means of a simu- lation study in Section 4. In Section 5, we propose a log-transmuted Burr XII regression model of location-scale form, in addition to the MLEs. In Section 6, we give three applications to real data to illustrate the flexibility of the TBXII model. Finally, some concluding remarks are offered in Section 7.

2. The TBXII Distribution

In this section, we define the TBXII distribution and provide plots of its pdf and hrf. The cdf and pdf (for x > 0) of the BXII distribution with positive parameters in $\varphi = (\alpha, \beta, \theta)^T$ are $G(x; \alpha, \beta, \theta) = 1 - [1 + (x/\theta)^\alpha]^{-\beta}$ and $g(x; \alpha, \beta, \theta) = \alpha\beta x^{\alpha-1}[1 + (x/\theta)^\alpha]^{-\beta-1}/\theta^\alpha$, respectively.

For ease of notation, we omit sometimes the parameters in the pdf, cdf, hrf, etc. The cdf and pdf of the TBXII model (for x > 0) are, respectively, given by

$$F(x) = 1 - \left[1 + \left(\frac{x}{\theta}\right)^\alpha\right]^{-\beta} \left\{1 - \lambda + \lambda\left[1 + \left(\frac{x}{\theta}\right)^\alpha\right]^{-\beta}\right\} \tag{5}$$

and

$$f(x) = \frac{\alpha\beta}{\theta^\alpha} x^{\alpha-1}\left[1 + \left(\frac{x}{\theta}\right)^\alpha\right]^{-\beta-1}\left\{1 - \lambda + 2\lambda\left[1 + \left(\frac{x}{\theta}\right)^\alpha\right]^{-\beta}\right\}, \tag{6}$$

where α and β are positive shape parameters, $\theta > 0$ is a scale parameter and $\lambda \in [-1, 1]$ is a shape parameter. Henceforth, we denote a random variable X having pdf (6) by X ~TBXII($\alpha, \beta, \theta, \lambda$).

The quantile function (qf) of X is obtained by inverting the TBXII cdf (5). We obtain

$$Q(u; \alpha, \beta, \theta, \lambda) = \begin{cases} Q_G\left(\dfrac{1 + \lambda - \sqrt{(1+\lambda)^2 - 4\lambda u}}{2\lambda}; \alpha, \beta, \theta, \lambda\right), if\ \lambda \neq 0; \\ Q_G(u; \alpha, \beta, \theta, \lambda) \qquad\qquad , if\ \lambda = 0, \end{cases} \tag{7}$$

where $Q_G(u; \alpha, \beta, \theta, \lambda) = G^{-1}(u; \alpha, \beta, \theta)\theta[(1 - u)^{-\frac{1}{\beta}} - 1]^{1/\alpha}$.

The survival function (sf), hrf and cumulative hazard rate function (chrf) of X are, respec- tively, given by

$$S(x) = [1 + (\frac{x}{\theta})^\alpha]^{-\beta}\left\{1 - \lambda + \lambda\left[1 + \left(\frac{x}{\theta}\right)^\alpha\right]^{-\beta}\right\},$$

$$h(x) = \frac{\alpha\beta x^{\alpha-1}}{[\theta^{\alpha} + x^{\alpha}]} \frac{\{1 - \lambda + 2\lambda\left[1 + \left(\frac{x}{\theta}\right)^{\alpha}\right]^{-\beta}\}}{\{1 - \lambda + \lambda\left[1 + \left(\frac{x}{\theta}\right)^{\alpha}\right]^{-\beta}\}}$$

and

$$H(x) = -\log\left(\left[1 + \left(\frac{x}{\theta}\right)^{\alpha}\right]^{-\beta}\left\{1 - \lambda + \lambda\left[1 + \left(\frac{x}{\theta}\right)^{\alpha}\right]^{-\beta}\right\}\right).$$

The TBXII distribution is a very flexible model that leads to different distributions when its parameters change. It contains the following special models:

• The transmuted log-logistic (TLL) distribution when $\beta = 1$ and $\theta = m-1$.

• For $\alpha = 1$, we obtain the transmuted Pareto type II (TPII) distribution.

• Setting $\lambda = 0$, we have the BXII distribution.

• The case $\beta = 1$, $\theta = m-1$ and $\lambda = 0$ refers to the log-logistic (LL) distribution.

• The Pareto type II (PII) distribution follows when $\alpha = 1$ and $\lambda = 0$.

Figure 1 displays plots of the TBXII density for some parameter values. The hrf plots are given in Figure 2.

3. The TBXII Properties

Henceforth, let T be a random variable having the BXII density $g(x; \alpha, \beta, \theta) = \alpha\beta x^{\alpha-1}[1 + (\frac{x}{\theta})^{\alpha}]^{-\beta-1}/\theta^{\alpha}$, where α, β and θ are positive parameters. The TBXII density (6) can be expressed as

$$f(x) = (1-\lambda)\frac{\alpha\beta}{\theta^{\alpha}}x^{\alpha-1}\left[1 + \left(\frac{x}{\theta}\right)^{\alpha}\right]^{-\beta-1} + \lambda\frac{\alpha(2\beta)}{\theta^{\alpha}}x^{\alpha-1}\left[1 + \left(\frac{x}{\theta}\right)^{\alpha}\right]^{-2\beta-1} \tag{8}$$

$$= (1-\lambda)g(x; \alpha, \beta, \theta) + \lambda g(x; \alpha, 2\beta, \theta).$$

Equation (8) reveals that the TBXII density function is a linear mixture of two BXII densities. So, several structural properties of the TBXII model can be determined from those properties of the BXII distribution.
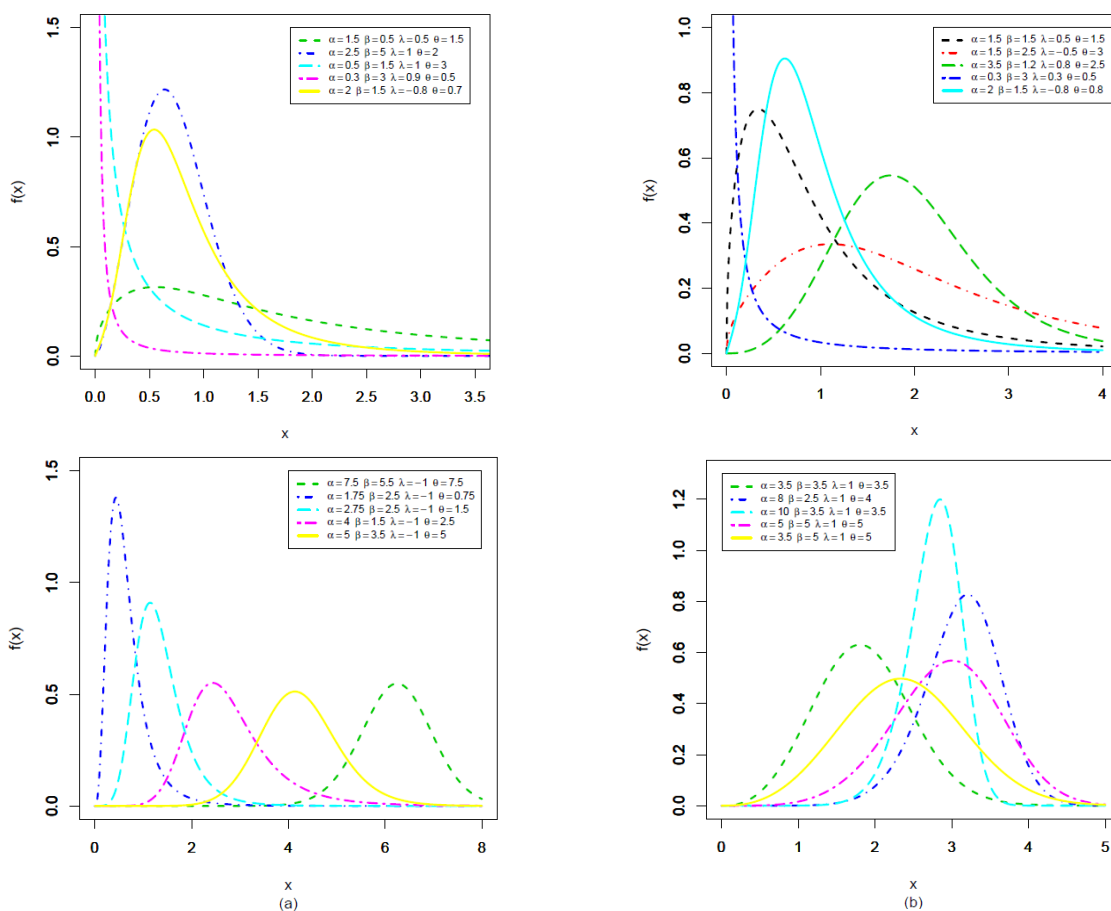

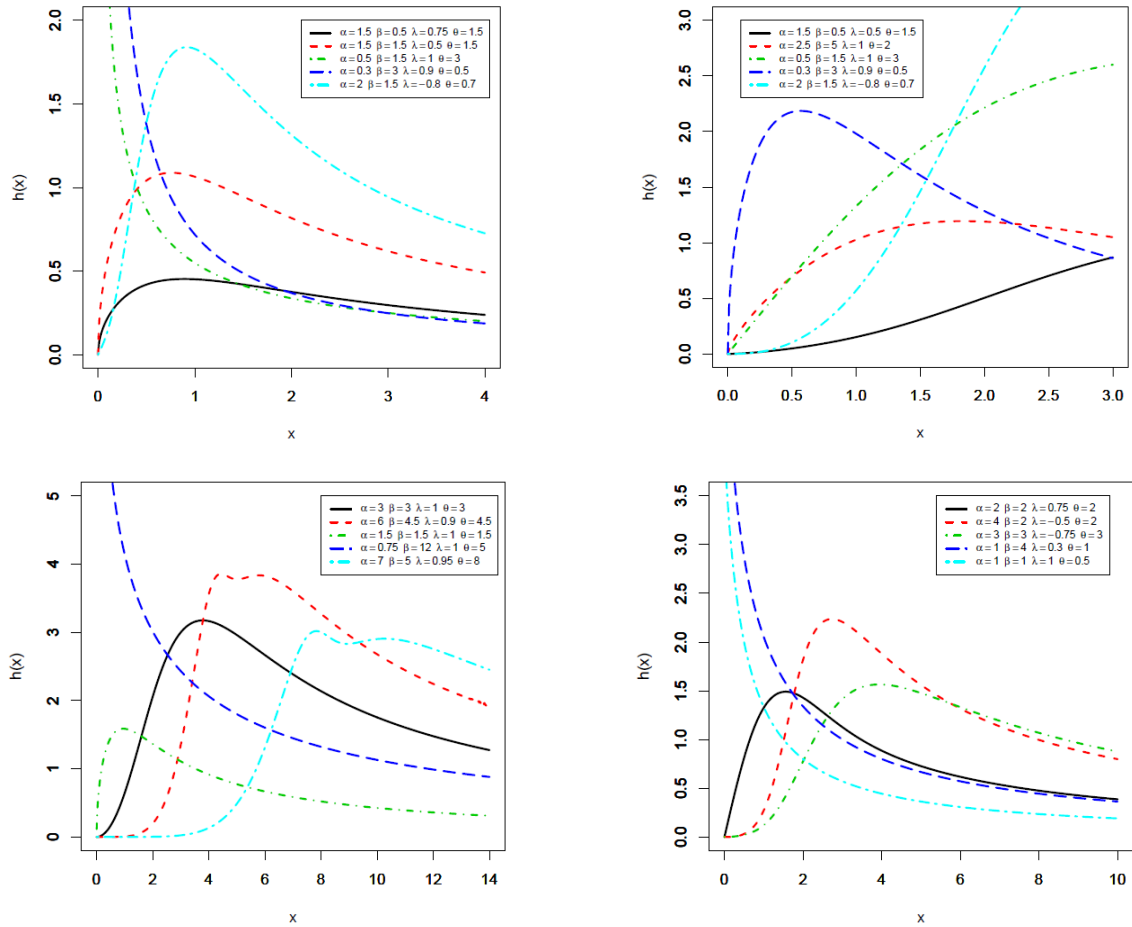Figure 1: Plots of the TBXII density.

Figure 2: Plots of the TBXII hrf.

3.1 Moments

For $r < \alpha\beta$, the rth ordinary and lower incomplete moments of T are, respectively, given by

$$\mu_r{}' = \beta\theta^\gamma\, B(\beta - \tfrac{r}{\alpha}, \tfrac{r}{\alpha} + 1)\ \ \text{and} \varphi_r(z) = \beta\theta^\gamma\ B\left(z^\alpha; \beta - \tfrac{r}{\alpha}, \tfrac{r}{\alpha} + 1\right),$$

where $B(a, b) = \int_0^\infty t^{\alpha-1}(1+t)^{-(a+b)}dt$ and $B(z\,; a, b) = \int_0^z t^{\alpha-1}(1+t)^{-(a+b)}dt$ are the beta and incomplete beta functions of the second type, respectively.

Then, the nth ordinary moment of X is given by

$$\mu_n{}' = E(X^n) = (1 - \lambda)\int_0^\infty x^n\, g(x; \alpha, \beta, \theta)\, dx + \lambda\int_0^\infty x^n\, g(x; \alpha, 2\beta, \theta)\, dx.$$

For $n < \alpha\beta$, we obtain

$$\mu_n{}' = E(X^n) = (1 - \lambda)\beta\theta^n\, B\left(\beta - \tfrac{n}{\alpha}, \tfrac{n}{\alpha} + 1\right) + 2\lambda\beta\theta^n B\left(2\beta - \tfrac{n}{\alpha}, \tfrac{n}{\alpha} + 1\right). \tag{9}$$

Setting n = 1 in (9), follows the mean of X.

the central moments (μs) and cumulants (κs) of X can be obtained from

$$\mu_s = E(X - \mu_1{}')^s = \sum_{i=0}^s (-1)^i \binom{s}{i} (\mu_n{}')^s \mu_{s-i}{}'$$

and

$$K_s = \mu_s' - \sum_{i=0}^{s-1} \binom{s-1}{i-1} K_r \mu_{s-r}',$$

where $K_1 = \mu_1'$. The skewness and kurtosis measures of follow from the third and fourth standardized cumulants using well-known relationships.

The sth lower incomplete moment, say $\varphi_s(t)$, is given by $\varphi_s(t) = \int_0^t x^s f(x)\, dx$.

. We can write from equation (8)

$$\varphi_s(t) = (1-\lambda) \int_0^t x^s g(x; \alpha, \beta, \theta)\, dx + \lambda \int_0^t x^s g(x; \alpha, 2\beta, \theta)\, dx,$$

and using the incomplete moment of T given before, we obtain (for $s < \alpha\beta$)

$$\varphi_s(t) = (1-\lambda)\beta\theta^s\, B\left(t^\alpha; \beta - \frac{s}{\alpha}, \frac{s}{\alpha} + 1\right) + 2\lambda\beta\theta^s B\left(t^\alpha; 2\beta - \frac{s}{\alpha}, \frac{s}{\alpha} + 1\right). \qquad (10)$$

The first incomplete moment $(\varphi_1(t))$ is related to the Bonferroni and Lorenz curves, the mean residual and mean waiting times. It follows from the above equation by setting $s = 1$. In a different way, it can be expressed from equation (8) as

$$\varphi_1(t) = (1-\lambda)\alpha\beta\theta \int_{t/\theta}^{\infty} x^\alpha [1 + \left(\frac{x}{\theta}\right)^\alpha]^{-\beta-1}\, dx + 2\lambda\beta\theta \int_{t/\theta}^{\infty} x^\alpha [1 + \left(\frac{x}{\theta}\right)^\alpha]^{-2\beta-1}\, dx.$$

Setting $y = x^\alpha$, we obtain

$$\varphi_1(t) = (1-\lambda)\beta\theta \int_{\left(\frac{t}{\theta}\right)^\alpha}^{\infty} y^{\frac{1}{\alpha}} (1+y)^{-\beta-1}\, dy + 2\lambda\beta\theta \int_{\left(\frac{t}{\theta}\right)^\alpha}^{\infty} y^{\frac{1}{\alpha}} (1+y)^{-2\beta-1}\, dy.$$

The following integral (for $t > 0$ and $s, \beta > 0$) is obtained using Maple as

$$J(t; s, \beta) = \int_t^{\infty} y^s (1+y)^{-\beta}\, dy$$

$$= -\left[\frac{2F_1[(\beta, s+1); (2+s); -t]t^{s+1}}{(s+1)} + \frac{\Gamma(\beta - s - 1)\pi \csc(\pi s)}{\Gamma(\beta)\Gamma(-s)}\right],$$

where

$$2F_1(a, b, c; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!},$$

is the hypergeometric function and $(a)_k = a(a+1)\ldots(a+k-1)$ denotes the ascending factorial.

Thus, we can write

$$\varphi_1(t) = (1-\lambda)\beta\theta J\left(\left(\frac{t}{\theta}\right)^\alpha; \frac{1}{\alpha}, \beta + 1\right) + 2\lambda\beta J\left(\left(\frac{t}{\theta}\right)^\alpha; \frac{1}{\alpha}, 2\beta + 1\right). \qquad (11)$$

For a given probability $\pi > 0$, the Lorenz and Bonferroni curves are given by

$L(\pi) = \varphi_1(q)/\mu_1'$ and $B(\pi) = \varphi_1(q)/(\pi\mu_1')$, respectively, where $q = F^{-1}(\pi) = Q(\pi)$

is given by (7).In economics, if π is the proportion of units whose income is lower than or equal to q,

$L(\pi)$ gives the proportion of total income volume accumulated by the set of units with an income lower than or equal to q. The index $L(\pi)$ yields fractions of the total income, wheres $B(\pi)$ refers to relative income levels. These curves are important in economics and also in reliability,demography, insurance and medicine. For the proposed model, they can be determined from (11) as

$$L(\pi) = \frac{(1-\lambda)J\left(\left(\frac{t}{\theta}\right)^{\alpha};\frac{1}{\alpha},\beta+1\right) + 2\lambda J\left(\left(\frac{t}{\theta}\right)^{\alpha};\frac{1}{\alpha},2\beta+1\right)}{\theta^{r-1}[(1-\lambda)B\left(\beta-\frac{\gamma}{\alpha},\frac{\gamma}{\alpha}+1\right) + 2\lambda B\left(2\beta-\frac{\gamma}{\alpha},\frac{\gamma}{\alpha}+1\right)}$$

and

$$B(\pi) = \frac{(1-\lambda)J\left(\left(\frac{t}{\theta}\right)^{\alpha};\frac{1}{\alpha},\beta+1\right) + 2\lambda J\left(\left(\frac{t}{\theta}\right)^{\alpha};\frac{1}{\alpha},2\beta+1\right)}{\pi\theta^{r-1}[(1-\lambda)B\left(\beta-\frac{\gamma}{\alpha},\frac{\gamma}{\alpha}+1\right) + 2\lambda B\left(2\beta-\frac{\gamma}{\alpha},\frac{\gamma}{\alpha}+1\right)},$$

respectively.

The nth moment of the residual life of X uniquely determines F (x) and it is defined by (for t > 0 and n = 1, 2, . . .)

$$m_n(t) = \frac{1}{S(t)}\int_t^\infty (x-t)^n dF(x).$$

We can write based on equations (8), (9) and (10)

$$m_n(t) = \frac{1}{S(t)}\sum_{r=0}^n \binom{n}{r}(-t)^{n-r}[\mu_r' - \varphi_r(t)].$$

The mean residual life (MRL) of X represents the expected additional life length for a unit which is alive at age t and follows by setting n = 1 in the last equation.

The nth moment of the reversed residual life of X uniquely determines F (x) and it is defined by (for t > 0 and n = 1, 2, . . .)

$$M_n(t) = \frac{1}{F(t)}\int_t^\infty (x-t)^n dF(x).$$

Then,$M_n(t)$ can be expressed as

$$M_n(t) = \frac{1}{F(t)}\sum_{r=0}^n \binom{n}{r}(-1)^r t^{n-r}\varphi_r(t).$$

The mean inactivity time (MIT) given by $M_1(t)$ represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in (0, t).

3.2 Means Deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. In this section, we determine the mean deviations, which depend on $\varphi_1(t)$. The mean deviations about the mean $[\delta_1 = E(|X - \mu_1'|)]$ and about the median $M[\delta_2 = E(|X - M|)]$of X are given, respectively, by

$$\delta_1 = 2\mu_1' F(\mu_1') - 2\varphi_1(\mu_1') \quad \text{and} \quad \delta_2 = \mu_1' - 2\varphi_1(M),$$

where $\mu_1' = E(X)$, $M = Q(1/2)$, $F(M)$ and $F(\mu_1')$ are easily obtained from (5) and $\phi1(t)$ follows from (10) with s = 1 or (11).

3.3 Generating Function

Parana´ıba et al. (2011) derived the the moment generating function (mgf) of the BXII(α, δ, θ) distribution denoted here by Mδ(t). The mgf M (t) of X can be expressed from equation (8) as

$$M(t) = (1 - \lambda)M_\beta(t) + \lambda M_{2\beta}(t). \tag{12}$$

We can write $M_\delta(t)$ (for t < 0) (see Parana´ıba et al., 2011)

$$M_\delta(t) = \alpha\delta \int_0^\infty e^{t\theta y} y^{\alpha-1} (1 + y^\alpha)^{-(\delta+1)} \, dy.$$

This integral can be determined only in special cases. Now, we assume that α = m/r, where m and r are positive integers. This condition is not restrictive since every positive real number can be approximated by a rational number. We can obtain, using a result of Prudnikov et al. (1992, p.21), a simple representation for Mδ(t) (for t < 0) given by

$$M_\delta(t) = \frac{m\delta}{r} \int_0^\infty e^{\theta y t} y^{\frac{m}{r}-1} (1 + y^{\frac{m}{r}})^{-(\delta+1)} dy = \frac{m\delta}{r} I\left(-\theta t, \frac{m}{r} + 1, \frac{m}{r}, -\delta - 1\right).$$

The function I($\cdot$) can be expressed in terms of the Meijer G-function (Gradshteyn and Ryzhik, 2000) (and evaluated in Mathematica software) defined by

$$G_{p,q}^{m,n}\left(x \bigg| \begin{matrix} a_1 & ,\dots, & a_p \\ b_1 & ,\dots, & b_p \end{matrix}\right) = \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j + t) \sum_{j=1}^n \Gamma(1 - a_j - t)}{\prod_{j=n+1}^p \Gamma(a_j + t) \prod_{j=m+1}^p \Gamma(1 - b_j - t)} x^{-t} dt,$$

where $i = \sqrt{-1}$ is the complex unit and L denotes an integration path (Gradshteyn and Ryzhik, 2000, Section 9.3). The Meijer G-function contains as particular cases many integrals with elementary and special functions.

For m and k positive integers, $\mu > -1$ and p > 0, the following result holds

$$I\left(p, \mu, \frac{m}{r}, v\right) = \int_0^\infty \exp(-px) \, x^\mu (1 + x^{\frac{m}{r}})^v \, dx$$

$$= \frac{r^{-v} m^{\mu+\frac{1}{2}}}{(2\pi)^{\frac{(m-1)}{2}} \Gamma(-v) p^{\mu+1}} \times G_{k+m,k}^{k,k+m}\left(\frac{m^m}{p^m} \bigg| \begin{matrix} \Delta(m, -\mu) & ,\Delta(r, v+1) \\ \Delta(r, 0) \end{matrix}\right), \tag{13}$$

where $\Delta(r, a) = \frac{a}{r}, \frac{a+1}{r}, \dots, \frac{a+r}{r}$. Equation (13) allows to determine Mδ(t) when α = m/r. Other special cases for Mδ(t) can be found in Parana´ıba et al. (2013).

3.4 Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Let

X1, . . . , Xn be a random sample from the TG family. The pdf of the ith order statistic, say Xi:n, is given by

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i}(-1)^j \binom{n-j}{j} F^{j+i-1}(x). \tag{14}$$

Using (4) and after applying the generalized binomial expansion, we obtain

$$F^{j+i-1}(x) = \sum_{l=0}^{\infty}(-1)^l \binom{j+i-1}{l}(1-\lambda)^l \bar{G}(x)^l[1+s\bar{G}(x)]^l,$$

where $s = \lambda/(1-\lambda)$. By expanding zβ in Taylor series, we have

$$Z^\beta = \sum_{k=}^{\infty}\frac{(\beta)_k}{k!}(z-1)^k, \tag{15}$$

where $(\beta)_k = \beta(\beta-1)\dots(\beta-k+1)$is the descending factorial. Equation (15) converges for every z > 0.

By using the expansion in (15) and the pdf (6), we have

$$f(x)F^{j+i+1}(x) = \sum_{l=1,r=0}^{\infty}\frac{(-1)^l(1-\lambda)s^r(l)_r}{r!}\binom{j+i-1}{l}[1+(\tfrac{x}{\theta})^\alpha]^{-(l+\gamma+1)\beta-1} \times \frac{\alpha\beta}{\theta^\alpha}x^{\alpha-1}\{1-\lambda$$

$$+ 2\lambda[1+(\tfrac{x}{\theta})^\alpha]^{-\beta}\}.$$

Then, we can write

$$f(x)F^{j+i+1}(x) = \sum_{l=1,r=0}^{\infty}\frac{(-1)^l(1-\lambda)s^r(l)_r}{r!}\binom{j+i-1}{l}$$

$$\times \left[\frac{(1-\lambda)}{(l+r+1)}g_{(l+r+1)\beta}(x) + \frac{2\lambda}{(l+r+2)}g_{(l+r+2)\beta}(x)\right], \tag{16}$$

where $g_\delta(x)$is the BXII pdf with parameters α, δ, θ.

By inserting (16) in equation (14), the pdf of Xi:n can be reduced to

$$f_{i:n}(x) = \omega_{l,r}\left[\frac{(1-\lambda)}{(l+r+1)}g_{(l+r+1)\beta}(x) + \frac{2\lambda}{(l+r+2)}g_{(l+r+2)\beta}(x)\right], \tag{17}$$

where

$$\omega_{l,r} = \frac{1}{B(i, n-i+1)}\sum_{j=0}^{n-i}\frac{(-1)^{l+j}(1-\lambda)s^r(l)_r}{r!}\binom{n-i}{j}\binom{j+i-1}{l}.$$

Thus, the density function of the TBXII order statistics is an infinite linear combination of BXII densities. Based on equation (17), we can obtain some structural properties of Xi:n from those of the BXII model.

For example, the sth ordinary moment of Xi:n is given by

$$E(X_{i:n}^s) = \omega_{l,r}[(1-\lambda)\beta\theta^s B\left([l+r+1]\beta - \frac{s}{\alpha}, \frac{s}{\alpha}+1\right)$$

$$+ 2\lambda\beta\theta^s B\left([l+r+2]\beta - \frac{s}{\alpha}, \frac{s}{\alpha}+1\right)$$

(18)

3.5 Probability Weighted Moments

The (s, h)th probability weighted moment (PWM) of X, say $\rho_{s,h}$, is defined by

$$\rho_{s,h} = E[X^s F(X)^h] = \int_{-\infty}^{\infty} x^s F(x)^h f(x)\, dx.$$

From equation (16), we can write

$$f(x)F^h(x) = \sum_{l=0,r=0}^{\infty} d_{l,r}\left[\frac{(1-\lambda)}{(l+r+1)}g_{(l+r+1)\beta}(x) + \frac{2\lambda}{(l+r+2)}g_{(l+r+2)\beta}(x)\right],$$

where $d_{l,r} = \frac{(-1)^l(1-\lambda)s^r(l)_r}{r!}\binom{h}{l}$.

Then, the quantity $\rho_{s,r}$ can be expressed as

$$\rho_{s,h} = \sum_{l=1,r=0}^{\infty} d_{l,r}\left[\frac{(1-\lambda)}{(l+r+1)}\int_0^{\infty} x^s g_{(l+r+1)\beta}(x)\, dx + \frac{2\lambda}{(l+r+1)}x^s g_{(l+r+2)\beta}(x)\, dx\right].$$

We can obtain $\rho_{s,h}$ based on the ordinary moments of the BXII distribution given at the beginning of Section 3.1

3.6 Entropies

The Rényi entropy of a random variable X represents a measure of variation of the uncertainty and it is defined by

$$I_\delta(X) = \frac{1}{(1-\delta)}\log\left(\int_{-\infty}^{\infty} f(x)^\delta\, dx\right), \delta > 0 \text{ and } \delta \neq 1.$$

By using the pdf (3), we can write

$$f(x)^\delta = (1-\lambda)^\delta g(x)^\delta \frac{\{1 + d\bar{G}(x)\}^\delta}{A}.$$

(19)

Applying (15) to the quantity A gives

$$A = \sum_{k=0}^{\infty} \frac{(\delta)_k d^k}{k!}\bar{G}(x)^k.$$

(20)

By inserting (20) in equation (19), we obtain

$$f(x)^\delta = (1-\lambda)^\delta \sum_{k=0}^{\infty} \frac{(\delta)_k d^k}{k!} g(x)^\delta \bar{G}(x)^k.$$

Thus,

$$f(x)^\delta = (\frac{\alpha\beta}{\theta^\alpha})^\delta (1-\lambda)^\delta \sum_{k=0}^{\infty} \frac{(\delta)_k d^k}{k!} x^{\delta(\alpha-1)} [1 + (\frac{x}{\theta})^\alpha]^{-\beta(k+\delta)-\delta}.$$

Then, the R´enyi entropy of X is given by

$$I_\delta(X) = \frac{1}{(1-\delta)} \log\{(\frac{\alpha\beta}{\theta^\alpha})^\delta \left(1 - \lambda)^\delta \sum_{k=0}^{\infty} \frac{(\delta)_k d^k}{k!} \int_0^\infty x^{\delta(\alpha-1)} [1 + (\frac{x}{\theta})^\alpha]^{-\beta(k+\delta)-\delta} dx\right\},$$

The last integral can be determined from the syth ordinary moment of the BXII

$g(x; \alpha, \beta^*, \theta)$density, where $s^* = (\alpha-1)(1-\delta)$ and $\beta^* = \beta(k+\delta)+\delta-1$. Then,

$$I_\delta(X) = \frac{1}{(1-\delta)} \log\{(\frac{\alpha\beta}{\theta^\alpha})^\delta \left(1 - \lambda)^\delta \sum_{k=0}^{\infty} \frac{(\delta)_k d^k \theta^{s^*+\alpha}}{\alpha k!} B\left(\beta(k+\delta)+\delta-1-\frac{s^*}{\alpha}, \frac{s^*}{\alpha}+1\right)\right\}.$$

The Shannon entropy, say S, of X is defined by

$$S = E\{-\log[f(X)]\}.$$

It is a special case of the R´enyi entropy when $\delta \uparrow 1$. Thus, based on equation (3), we can write

$$S = E\{-\log[q(X)]\} + E\{-\log[1 - \lambda + 2\lambda\bar{G}(X)]\}.$$

The two expectations in S can be evaluated numerically for any baseline G model.

3.7 Simulation

Simulating the TBXII random variable is straightforward. Let U be a uniform variate on the unit interval (0, 1). Using the qf (7), the random variable

$$x = \begin{cases} \theta\{[1 - \frac{1 + \lambda - \sqrt{(1+\lambda)^2 - 4\lambda\mu}}{2\lambda}]^{-\frac{1}{\beta}} - 1\}^{\frac{1}{\alpha}}, if\ \lambda \neq 0; \\ \theta[(1-\mu)^{-\frac{1}{\beta}} - 1]^{\frac{1}{\alpha}}, if\ \lambda = 0, \end{cases}$$

follows (6), i.e., X ~TBXII(α, β, θ, λ).

Some plots comparing the exact TBXII densities and histograms from two simulated data sets for some parameter values are displayed in Figure 3. These plots indicate that the simulated values are consistent with the TBXII theoretical density function.
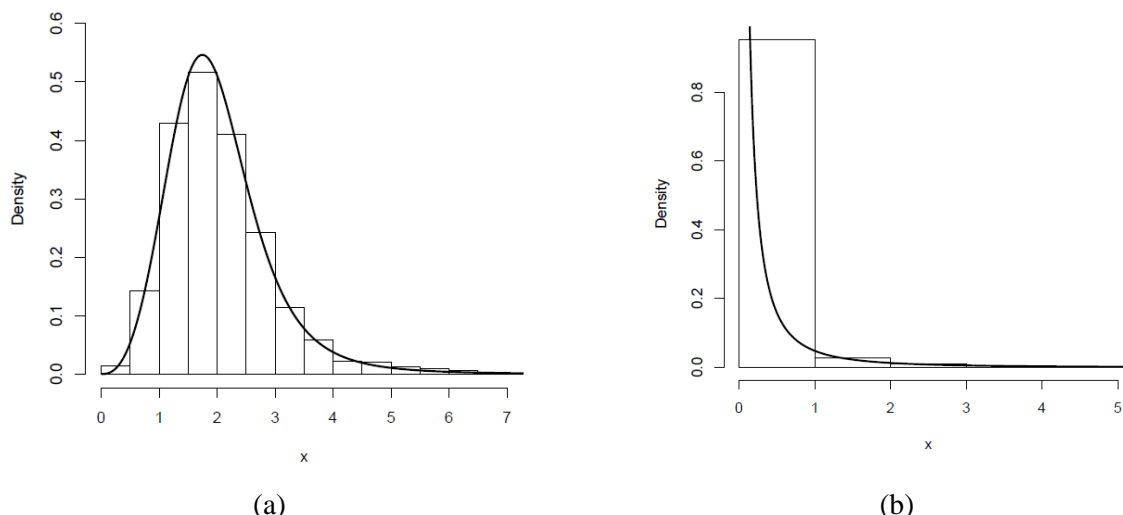
(a)   (b)

Figure 3: Plots of the TBXII densities for simulated data sets: (a) α = 3.5, β = 1.2, θ = 2.5, λ = 0.8. (b) α = 0.5, β = 4.5, θ = 2.0, λ = 0.5.

4. Maximum Likelihood Estimation

Several approaches for parameter estimation were proposed in the statistical literature but the maximum likelihood method is the most commonly employed. The MLEs enjoy desirable properties for constructing confidence intervals. In this section, we discuss the estimation and inference for the TBXII distribution by maximum likelihood for complete data sets. Let $x_1, , , x_n$ be a random sample from X ~TBXII(α, β, θ, λ) and let $\eta = (\alpha, \beta, \theta, \lambda)^T$ be the vector of the model parameters. The log-likelihood function for η reduces to

$$l(\eta) = n\log(\alpha\beta) - n\alpha\log(\theta) + (\alpha - 1)\sum_{i=1}^{n}\log(x_i) - (\beta + 1)\sum_{i=1}^{n}\log[1 + (\frac{x_i}{\theta})^\alpha]$$

$$+ \sum_{i=1}^{n}\log\{1 - \lambda + 2\lambda[1 + (\frac{x_i}{\theta})^\alpha]^{-\beta}\}. \qquad (21)$$

The components of the score vector $U(\eta) = (\frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \beta}, \frac{\partial l}{\partial \theta}, \frac{\partial l}{\partial \lambda})^T$ are given by

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

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

$$\frac{\partial l}{\partial \theta} = -\frac{n\alpha}{\theta} + (\beta + 1)\sum_{i=1}^{n}\frac{\alpha x_i{}^\alpha\theta^{-(\alpha+1)}}{[1 + (\frac{x_i}{\theta})^\alpha]} - \sum_{i=1}^{n}\frac{2\alpha\beta\lambda x^\alpha\theta^{-(\alpha+1)}(1 + x^\alpha\theta^{-\alpha})^{-(\beta+1)}}{1 - \lambda + 2\lambda[1 + (\frac{x_i}{\theta})^\alpha]^{-\beta}},$$

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} \frac{2[1 + (\frac{x_i}{\theta})^\alpha]^{-\beta} - 1}{1 - \lambda + 2\lambda[1 + (\frac{x_i}{\theta})^\alpha]^{-\beta}}.$$

The complicated system of nonlinear equations obtained from U (η) = 0 can be solved using numerical Newton-Raphson type methods. However, it is much simpler to maximize (21) directly by means of the R (optim function), SAS (PROC NLMIXED), Mathcad or Ox program (sub-routine MaxBFGS).

Next, we conduct a small Monte Carlo simulation study to assess on the finite sample behavior of the MLEs of α, β, θ and λ. All results are obtained from 1, 000 Monte Carlo replications and the simulations are carried out using the statistical software R. In each replication, a random sample of size n is drawn from the TBXII(α, β, θ, λ) distribution using the inversion method. Table 1 presents the means of the MLEs of the parameters of the TBXII distribution and the corresponding root of mean squared errors (RMSEs) for sample sizes n = 200 and n = 400.

λ	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\theta}$	$\widehat{\lambda}$
		$n = 200$		
−0.8	1.6143 (0.2898)	1.6226 (0.7507)	1.2390 (0.6299)	−0.5812 (0.4147)
−0.5	1.5205 (0.2256)	1.8428 (1.0852)	1.4027 (0.9649)	−0.3095 (0.4290)
−0.2	1.4776 (0.2032)	1.8293 (1.1622)	1.3396 (1.0753)	−0.1039 (0.3943)
0.2	1.4701 (0.1815)	2.0628 (6.4035)	1.3266 (3.0914)	0.1731 (0.3770)
0.5	1.5126 (0.1584)	1.7203 (1.3484)	1.0738 (0.7509)	0.3941 (0.3513)
0.8	1.5703 (0.1669)	1.5956 (0.5422)	0.9314 (0.3582)	0.5694 (0.3916)
		$n = 400$		
−0.8	1.5585 (0.2516)	1.6024 (0.5359)	1.1520 (0.3787)	−0.6785 (0.3285)
−0.5	1.4995 (0.1958)	1.7003 (0.6248)	1.2381 (0.5162)	−0.3685 (0.3817)
−0.2	1.4574 (0.1725)	1.7346 (0.7123)	1.2287 (0.5835)	−0.1365 (0.3840)
0.2	1.4805 (0.1224)	1.5806 (0.5895)	1.0795 (0.3853)	0.2246 (0.3323)
0.5	1.5121 (0.1265)	1.5062 (0.5436)	0.9592 (0.3010)	0.4166 (0.3476)
0.8	1.5393 (0.1089)	1.5935 (0.4336)	0.9611 (0.2707)	0.6456 (0.3074)

Table 1: Empirical means and RMSEs (in parentheses); α = 1.5, β = 1.5 and θ = 1.

5. The log-transmuted Burr XII regression model with censored data

In this section, we extend the log-Burr XII model (Silva et al., 2008) and log-beta Burr XII model (Cordeiro et al., 2016) by replacing the BXII distribution by the TBXII distribution. From the random variable X having the TBXII density function (6), we define the random variable Y = log(X) having the log-transmuted Burr XII (LTBXII) distribution. The density function of Y , parameterized in terms of $\alpha = \sigma^{-1}$ and θ = exp(μ), becomes (for y ∈ R)

$$f(y) = \frac{\beta}{\sigma} \exp\left(\frac{y - \mu}{\sigma}\right) [1 + \exp\left(\frac{y - \mu}{\sigma}\right)]^{-\beta-1} \{1 - \lambda + 2\lambda \left[1 + \exp\left(\frac{y - \mu}{\sigma}\right)\right]^{-\beta}\}, \qquad (22)$$

where λ, β and σ are positive parameters and μ ∈ R.

We refer to equation (28) as the new LTBXII distribution, say Y ∼ LTBXIIλ, β, σ, μ), where μ is the location parameter, σ > 0 is the scale parameter and λ and β are positive shape parameters. Thus,

if $X \sim \text{TBXII}(\alpha, \beta, \theta, \lambda)$ then $Y = \log(X) \sim \text{LTBXII}(\lambda, \beta, \sigma, \mu)$.

It simplifies to the log-Burr XII (LBXII) distribution when $\lambda = 0$. The corresponding survival function is given by

$$S(y) = [1 + \exp\left(\frac{y - \mu}{\sigma}\right)]^{-\beta}\{1 - \lambda + \lambda\left[1 + \exp\left(\frac{y - \mu}{\sigma}\right)\right]^{-\beta}\}. \tag{23}$$

We define the standardized random variable $Z = (Y - \mu)/\sigma$ having density function

$$\pi(z; a, b, c, k) = \beta\exp(z)[1 + \exp(z)]^{-\beta-1}\{1 - \lambda + 2\lambda[1 + \exp(z)]^{-\beta}\}, y\epsilon R. \tag{24}$$

In many practical applications, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric models to estimate univa- riate survival functions and for censored data regression problems are widely used. A parametric model that provides a good fit to lifetime data tends to yield more precise estimates of the quantities of interest. Based on the LTBXII density, we propose a linear location-scale regression model or log-linear regression model linking the response variable yi and the explanatory variable vector $v_i^T = \left(v_{i1,..,}v_{ip}\right)$ as follows

$$y_i = v_i^T\tau + \sigma z_i, \text{i} = 1,,,\text{n} \tag{25}$$

where the random error zi has density function (24), $\tau = (\tau_1, ..., \tau_p)^T$, σ, λ and β are unknown parameters. The parameter $\mu_i = V_i^T\tau$ is the location of yi. The location parameter vector $\mathbf{\mu} = (\mu_1, ..., \mu_n)^T$ is represented by a linear model $\mathbf{\mu} = \mathbf{V\tau}$, where $\mathbf{V} = (V_1, ..., V_n)^T$ is a known model matrix. Equation (25) is referred to as the LTBXII regression model for censored data. It is an extension of an accelerated failure time model using the TBXII distribution for censored data and it opens new possibilities for fitting many different types of data.

Consider a sample (y1, v1), . . . , (yn, vn) of n independent observations, where each random response is defined by $y_i = \min\{\log(X_i), \log(D_i)\}$. We assume non-informative censoring such that the observed lifetimes and censoring times are independent. Let F and D be the sets of individuals for which yi is the log-lifetime or log-censoring, respectively. Conventional like- lihood estimation techniques can be applied here. The log-likelihood function for the vector of parameters $\mathbf{\eta} = (\lambda, \beta, \sigma, \tau^T)^T$ from model (25) has the form $l(\eta) = \sum_{i\in F} l_i(\eta) + \sum_{i\in D} l_i^{(d)}(\eta)$, where $l_i(\eta) = \log[f(y_i|v_i)], l_i^{(d)}(\eta) = \log[S(y_i|v_i)]$, $f(y_i|v_i)$ is the density (28) and S(yi|vi) is the survival function (23) of Yi. The total log-likelihood function for η reduces to

$$l(\hat{\eta}) = \text{rlog}\left(\frac{\beta}{\sigma}\right) + \sum_{i\in F}\log[1 + \exp(z_i)] + \sum_{i\in F}\log\{1 - \lambda + 2\lambda[1 + \exp(z_i)]^\beta\}$$
$$- \beta\sum_{i\in C}\log[1 + \exp(z_i)] + \sum_{i\in C}\log\{1 - \lambda + \lambda[1 + \exp(z_i)]^\beta\}, \tag{26}$$

where $\hat{z}_i = (y_i - v_i^T\hat{\tau}/\hat{\sigma})$, and r is the number of uncensored observations (failures). The MLE $\hat{\eta}$ of the vector of unknown parameters can be determined by maximizing the log-likelihood (26). We use the subroutine NLMixed in SAS to calculate $\hat{\eta}$.

The NLMixed procedure of SAS has been exhaustively used to estimate the model parameters for several distributions published in recent years.

Initial values for τ and σ are taken from the fit of the LBXII regression model with $\lambda = 0$ (see, Silva et al., 2008). The fitted LTBXII model yields the estimated survival function for

$yi(\hat{z}_i = \left(y_i - \frac{v_i^T \hat{\tau}}{\hat{\sigma}}\right))$ given by

$$S\left(y_{i}; \hat{\lambda}, \hat{\beta}, \hat{\sigma}, \hat{\tau}^T\right) = [1 + \exp(\frac{y_i - v_i^T \hat{\tau}}{\hat{\sigma}})]^{-\hat{\beta}}\{1 - \hat{\lambda} + \hat{\lambda}\left[1 + \exp\left(\frac{y_i - v_i^T \hat{\tau}}{\hat{\sigma}}\right)\right]^{-\hat{\beta}}\}. \tag{27}$$

The asymptotic covariance matrix $K(\eta)^{-1}$ of $\hat{\eta}$ can be approximated by the inverse of the $(p + 3)$ $\times (p + 3)$ observed information matrix $-\ddot{L}(\eta)$. The elements of the observed information matrix $-\ddot{L}(\eta)$, namely $-L_{\lambda\lambda}, -L_{\lambda\beta}, -L_{\lambda\sigma}, -L_{\lambda\tau_j}, L_{\beta\beta}, -L_{\beta\sigma}, -L_{\beta\tau_j}, -L_{\sigma\sigma}, -L_{\sigma\tau_j}, and -L_{\tau_j\tau_k}$ for $j, k = 1, , , p$, can be calculated numerically.   The multivariate normal distribution $N_{p+3}(0, -\ddot{L}(\eta)^{-1})$ for $\hat{\eta}$ can be used in the classical way to construct   approximate confidence regions for some parameters in η.

We can use likelihood ratio (LR) statistics for comparing some special models with the LTBXII model. We consider the partition $\eta = (\eta_1^T, \eta_2^T)^T$, where η_1 is a subset of parameters of interest and η_2 is a subset of nuisance parameters. The LR statistic for testing the null hypothesis $H_0: \eta_1 = \eta_1^{(0)}$ versus the alternative hypothesis $H_0: \eta_1 \neq \eta_1^{(0)}$ is given by $w = 2\{l(\hat{\eta}) - l(\tilde{\eta})\}$, where $\tilde{\eta}$ and $\hat{\eta}$ are the estimates under the null and alternative hypotheses, respectively. The statistic w is asymptotically (as $n \to \infty$) distributed as X_v^2, where v is the dimension of the subset η_1 of parameters of interest.

6   Applications

In this section, we prove empirically the flexibility of the TBXII model by means of two applications to real data. The MLEs of the model parameters are evaluated numerically and the fit of this model is compared with those of other competitive models. These results are obtained using the Mathcad program. We also provide an application for the LTBXII regression model.

For the two data sets, we shall compare the fits of the TBXII distribution with those of other competitive models, namely: the Kumaraswamy Weibull (KwW) (Cordeiro et al., 2010), gamma Weibull (GW) (Provost et al., 2011), McDonald Weibull (McW) (Cordeiro et al., 2014), generalized transmuted Burr X (GTBX) (Nofal et al., 2017), modified beta Weibull (MBW) (Khan, 2015) and exponentiated transmuted generalized Rayleigh (ETGR) (Afify et al., 2015) distributions with corresponding densities (for x > 0):

- KwW:

$$f(x) = \frac{ab\beta\alpha^\beta x^{\beta-1}}{\exp[(\alpha x)^\beta]}\{1 - \exp[-(\alpha x)^\beta]\}^{a-1}\{1 - \exp[-(\alpha x)^\beta\}^{b-1};$$

- GW:

$$f(x) = \frac{1}{\Gamma(1+\gamma/\beta)} \beta \alpha^{\frac{\gamma}{\beta}+1} x^{\beta+\gamma-1} \exp(-\alpha x^\beta);$$

- McW:

$$f(x) = \frac{\beta c \alpha^\beta}{B(\frac{a}{c},b)\exp[(\alpha x)^\beta]} x^{\beta-1}\{1 - \exp[-(\alpha x)^\beta]\}^{a-1}\{1 - \{1 - \exp[-(\alpha x)^\beta]\}^c\}^{b-1};$$

- GTBX:

$$f(x) = \frac{2\alpha\beta^2 x}{\exp[(\beta x)^2]}\{1 - \exp[-(\beta x)^2]\}^{\alpha a-1}\{a(1+\lambda) - \lambda(a+b)\{1 - \exp[-(\beta x)^2]\}^{\alpha b}\};$$

- MBW:

$$f(x) = \frac{\beta \alpha^{-\beta} c^a x^{\beta-1}}{B(a,b)\exp[b(\frac{x}{\alpha})^\beta]}\{1 - \exp[-(\frac{x}{\alpha})^\beta]\}^{a-1}\{1 - (1-c)[1 - \exp[-(\frac{x}{\alpha})^\beta]]\}^{-a-b};$$

- ETGR:

$$f(x) = \frac{2\alpha\gamma\beta^2 x}{\exp[(\beta x)^2]}\frac{\{1 + \lambda - 2\lambda[1 - \exp[-(\beta x)^2]]^\alpha\}}{\{1 - \exp[-(\beta x)^2]\}^{-\alpha\gamma+1}}\{1 + \lambda - \lambda\{1 - \exp[-(\beta x)^2]\}^\alpha\}^{\gamma-1}.$$

The parameters of the above densities are all positive real numbers except for the ETGR and GTBX distributions for which $|\lambda| \le 1$.

In order to compare the fitted models, we use the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Further, we adopt the Anderson-Darling (A^*) and Cram´er-von Mises (W^*) statistics to compare the fits of the TBXII model with other non-nested models. These statistics, widely used to determine how closely a specific cdf fits the empirical distribution to a given data set, are given by

$$A^* = \left(\frac{9}{4n^2} + \frac{3}{4n} + 1\right)\left\{n + \frac{1}{n}\sum_{j=1}^{n}(2j-1)\log[z_i(1 - z_{n-j+1})]\right\}$$

and

$$w^* = \left(\frac{1}{2n} + 1\right)\left\{\sum_{j=1}^{n}(z_i - \frac{2j-1}{2n})^2 + \frac{1}{12n}\right\},$$

respectively, where zi = F (yj) and the yj's are the ordered observations. Generally, the smaller these statistics are, the better the fit is.

6.1 Application 1: The Gauge Lengths Data

The first data set on gauge lengths of 20 mm (Kundu and Raqab, 2009) consists of n = 74 observations. Table 2 provides the MLEs of the model parameters, their corresponding standard errors (SEs) and the values of the AIC, BIC, W^* and A^* statistics. The plots of the fitted TBXII pdf and other fitted pdfs (defined before). to the current data, are displayed in Figure4. They reveal that the TBXII distribution yields the best fit and it can be considered a very competitive model to other distributions with positive support.

6.2 Application 2: The Nicotine Data

The second data set refers to nicotine measurements, made from several brands of cigarettes in 1998, collected by the an independent agency "Federal Trade Commission" of the US govern- ment. The report entitled tar, nicotine, and carbon monoxide of the smoke of 1206 varieties of domestic cigarettes for the year of 1998 consists of the data sets and some information about the source of the data, smokers behavior and beliefs about nicotine, tar and carbon monoxide contents in cigarettes.

Table 3 lists the MLEs of the model parameters, their SEs and the values of the AIC, BIC, W^* and A^* statistics. Some plots of the fitted TBXII pdf and other fitted pdfs are displayed in Figure 5. They also indicate that the TBXII model yields the best fit to the nicotine data compared to the other models.

Table 2: MLEs, their SEs (in parentheses) and the W ∗, A∗, AIC and BIC statistics for the gauge lengths.

Model	Estimates					W^*	A^*	AIC	BIC
TBXII$(\alpha, \beta, \theta, \lambda)$	6.5235	6.3067	3.5559	0.211		0.026	0.195	110.3	119.5
	(0.964)	(6.009)	(0.372)	(1.077)					
KwW(α, β, a, b)	0.304	2.720	2.670	9.116		0.026	0.209	110.3	119.5
	(0.151)	(4.825)	(5.760)	(40.899)					
GW(α, β, γ)	5.465	1.229	19.387			0.082	0.543	111.7	118.7
	(9.227)	(0.671)	(10.959)						
GTBX$(\alpha, \beta, \lambda, a, b)$	3.490	0.661	0.002	2.519	0.016	0.104	0.688	118.1	129.6
	(2.084)	(0.120)	(0.048)	(1.503)	(0.428)				
McW(α, β, a, b, c)	1.438	0.583	83.720	14.428	3.460	0.119	0.779	118.8	130.3
	(1.447)	(0.211)	(78.890)	(15.870)	(9.663)				
MBW(α, β, a, b, c)	1.765	1.426	36.336	3.361	3.096	0.124	0.811	119.1	130.7
	(1.097)	(1.488)	(4.439)	(6.695)	(4.714)				
ETGR$(\alpha, \beta, \lambda, \gamma)$	2.121	0.698	0.320	7.790		0.207	1.340	121.4	121.9
	(0.315)	(0.040)	(0.228)	(1.727)					

Figure 6 displays the estimated cdf of the TBXII distribution. The QQ-plots of the fitted TBXII model are shown in Figure 7 for both data.

In summary, we compare the fits of the TBXII model with the KwW, McW, GTBX, MBW, GW and ETGR models for the gauge lengths and nicotine data in 2 and 3. We can conclude that the TBXII model has the lowest values for the AIC, BIC, W ∗ and A∗ statistics among the fitted models. Thus, the TBXII model could be chosen as the best model to both data sets. It is quite clear from the figures in these Tables 2 and 3 that this model provides the best fits to these two data sets. So, we prove empirically that it can be better model than other competitive models with positive support.

Table 3: MLEs, their SEs (given in parentheses) and the W ∗, A∗, AIC and BIC statistics for the nicotine data.

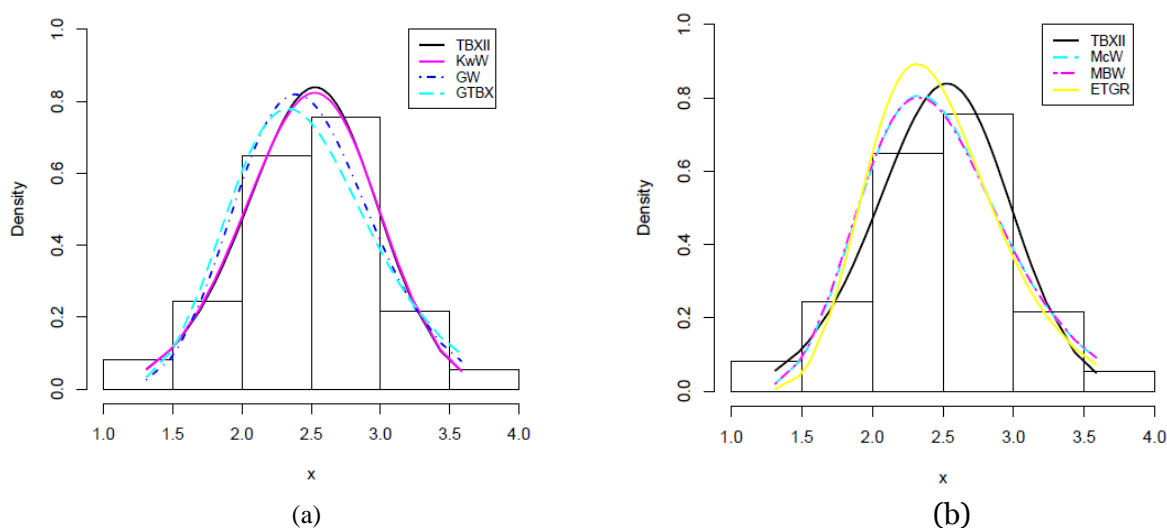| Model	Estimates					W^*	A^*	AIC	BIC
TBXII$(\alpha, \beta, \theta, \lambda)$	2.9018	72.249	4.5353	0.3887		0.619	3.535	236.8	252.2
	(0.118)	(151.565)	(3.4)	(0.25)					
KwW(α, β, a, b)	2.507	0.483	11.814	18.795		0.997	5.825	263.0	278.4
	(1.191)	(0.076)	(3.707)	(6.212)					
GW(α, β, γ)	25.276	0.441	9.708			1.643	9.753	310.5	322.0
	(6.912)	(0.066)	(1.337)						
GTBX$(\alpha, \beta, \lambda, a, b)$	1.365	1.252	0.002	1.158	0.007	0.692	4.023	247.9	267.2
	(8.100)	(0.039)	(0.588)	(6.872)	(0.101)				
McW(α, β, a, b, c)	1.307	0.531	16.858	10.104	1.164	1.617	9.594	312.7	331.9
	(0.596)	(0.079)	(4.219)	(3.995)	(0.793)				
MBW(α, β, a, b, c)	2.509	0.626	20.633	9.408	3.644	1.933	11.525	340.8	360.0
	(0.838)	(0.062)	(4.202)	(3.163)	(0.400)				
ETGR$(\alpha, \beta, \lambda, \gamma)$	0.293	0.949	0.824	8.599		1.160	6.845	281.9	297.2
	(0.073)	(0.040)	(0.064)	(3.112)					

Figure 4: The fitted densities for the: (a) TBXII, KwW, GW and GTBX models;
(b) TBXII, McW, MBW and ETGR models (gauge lengths).

6.3 Application 3: Vitamin A data regression model

In this subsection, we consider a data set provided by the Instituto de Saúde Coletiva Universidade Federal da Bahia. This data set was designed to evaluate the effect of vitamin A supplementation on recurrent diarrheal episodes in small children (see Barreto et al., 1994). Censoring times are random, and we aimed at modeling the treatment effect in time until the first occurrence of diarrheal episodes, this can be done by means of an appropriate regression model with censored data. The data from a randomized community trial that was designed to evaluated the effect of vitamin A supplementation on diarrheal episodes in 1,207 pre-school children, aged 6-48 months, who were assigned to receive either placebo or vitamin A in a small city in the Northeast of Brazil from December 1990 to December 1991. The vitamin A dosage was 100,000 IU for children younger than 12 months and 200,000 IU for older children, which
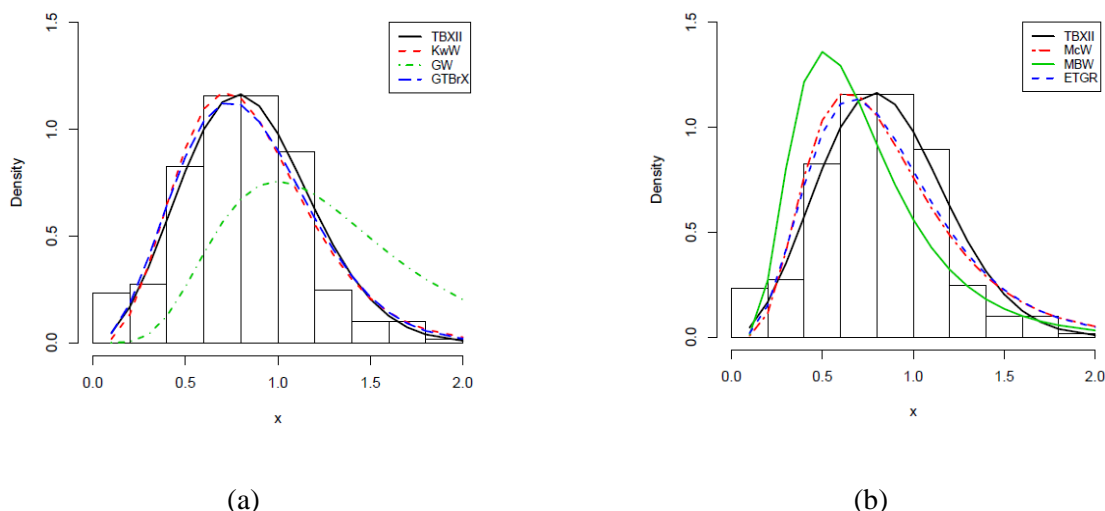
Figure 5: The fitted densities for the: (a) TBXII, KwW, GW and GTBX models.

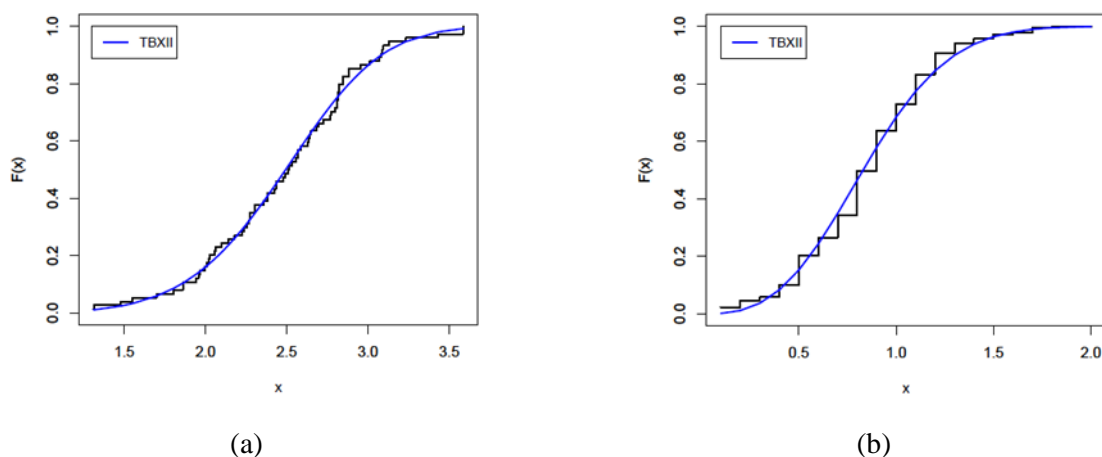(b) TBXII, McW, MBW and ETGR models (nicotine).

Figure 6: The estimated cdf of the TBXII model for (a) the gauge lengths data (b) nicotine data.

is the highest dosage guideline established by the World Health Organization (WHO) for the prevention of vitamin A deficiency.

The total time was defined as the time from the first dose of vitamin A until the occurrence of an episode of diarrhea. An episode of diarrhea was defined as a sequence of days with diarrhea and a day with diarrhea was defined when 3 or more liquid or semi-liquid motions were reported in a 24-hour period. The information on the occurrence of diarrhea collected at each visit corresponds to a recall period of 48-72 hours. The number of liquid and semi-liquid motions per 24 hours was recorded.

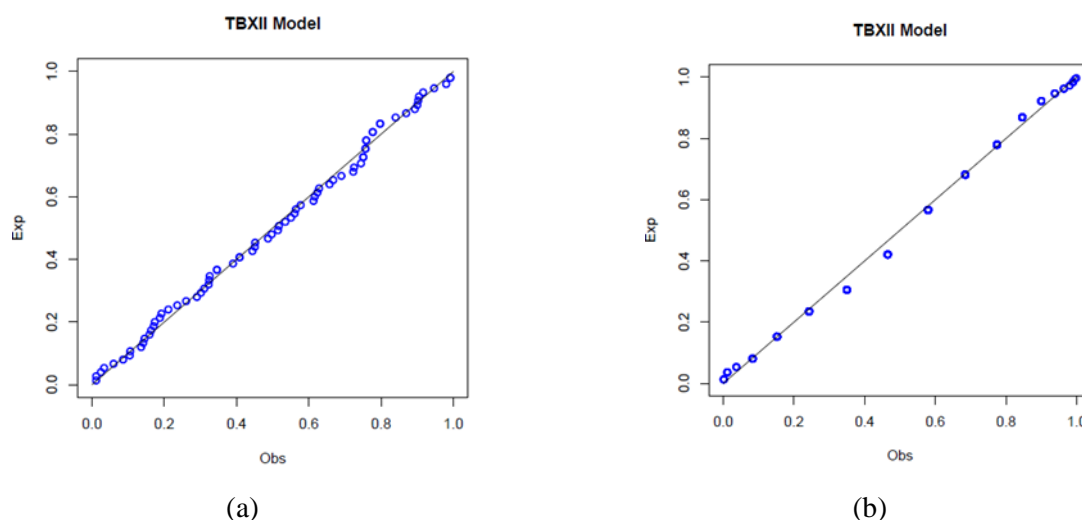The covariates considered in the regression models are:

Figure 7: QQ-plots of the TBXII model for the (a) gauge lengths data. (b) nicotine data.

- x_{i1}: age at baseline (in months);
- x_{i2}: treatment (0 = placebo, 1 = vitamin A);
- x_{i3}: gender (0 = girl, 1 = boy).

We now present results on fitting the model

$$y_i = \beta_0 + \tau_1 x_{i1} + \tau_2 x_{i2} + \tau_3 x_{i3} + \sigma z_i$$

where the variable Yi follows the LBBXII distribution given in (28), i = 1, 2, . . . , 1207.

Recently, Cordeiro et al.(2016) define the log-McDonald Burr XII (LMcBXII) distribution.The density function of Y , is given by (for y ∈ R)

$$f(y) = \frac{ck}{\sigma B(a,b)} \exp\left(\frac{y-\mu}{\sigma}\right) [1 + \exp\left(\frac{y-\mu}{\sigma}\right)]^{-k-1} \{1 - [1 + exp\left(\frac{y-\mu}{\sigma}\right)]^{-k}\}^{ac-1} \times [1$$

$$- \{1 - [1 + exp\left(\frac{y-\mu}{\sigma}\right)]^{-k}\}^c]^{b-1},$$

(28)

where a, b, c, σ and k are positive parameters and μ ∈ R. The LMcBXII distribution contains as special models some well-known distributions. If c =   1,   it reduces to the log-beta Burr   XII (LBBXII) distribution. If a = 1, it yields the log-Kumaraswami Burr XII (LKwBXII) distribution. It simplifies to the log-Burr XII (LBXII) distribution when a = b = c = 1 (Silva   it at el., 2008). Further, if a = b = c = 1, in addition to k = 1, it becomes the logistic distribution (Lawless, 2008). Cordeiro et al. (2016) based on the LMcBXII density, proposed a log-linear regression model linking the response variable yi and the explanatory variable vector $V_i^T = (v_{i1}, ..., v_{ip})$.

We now compare the LTBXII regression model with the LMcBXII, LBBXII, LKwBXII, LBXII and logistic regression models. The MLEs of the model parameters are calculated using the NLMixed procedure in SAS. Iterative maximization of the the log likelihood function (26) starts with initial values for β and σ which are taken from the fit of the LBXII regression model with a = b = c = 1.

A summary of the values of the AIC, BIC and CAIC to compare the LTBXII, LMcBXII, LBBXII, LKwBXII, LBXII regression models is given in Table 4. These results indicate that the LTBXII regression model has the lowest AIC, BIC and CAIC values among those values of the fitted models, and therefore it could be chosen as the best model.

Table 4: AIC, BIC and CAIC statistics for comparing the LTBXII, LMcBXII, LBBXII, LKw- BXII, LBXII and logistic regression models.

Regression model	AIC	CAIC	BIC
LTBXII	3217.3	3217.4	3253.0
LMcBXII	3221.4	3221.5	3267.2
LBBXII	3224.5	3224.7	3265.3
LKwBXII	3223.5	3223.6	3264.2
LBXII	3235.3	3235.4	3265.9
Logistic	3288.3	3288.4	3313.8

The MLEs and their SEs of the parameters for the LTXBII regression model are given in Table 5.

Table 5: MLEs for the parameters from the LTBXII regression model on the complete vitamin A data set

Parameter	Estimate	SE	p-value
λ	3.2163	0.8757	-
β	0.0321	0.0119	-
σ	0.2840	0.0286	-
τ_0	1.9316	0.1001	< 0.001
τ_1	0.0218	0.0026	< 0.001
τ_2	0.0859	0.0573	0.1341
τ_3	0.0397	0.0574	0.4890

We note from the fitted LBBXII regression model that x1 is significant at 1% and that there is a significant difference between the age for the survival times. The MLEs of the parameters for the final model are given in Table 6. We can interpret the estimated coefficients of the final model

Table 6: MLEs for the parameters from the LTBXII regression model on the complete vitamin A data set - final model

Parameter	Estimate	SE	p-value
λ	3.2400	0.8919	-
β	0.0316	0.0118	-
σ	0.2827	0.0281	-
β_0	1.9959	0.0908	< 0.01
β_1	0.0216	0.0026	< 0.01

as the follows The median survival time should increase approximately 2.18% ($e0.0216 \times 100\%$) when age increases one unit.

7. Conclusions

Statistical distributions are very useful in describing and predicting real world phenomena. Numerous generalized distributions have been extensively used over the last decades for modeling data in several areas. Recent developments focus on defining new families to extend well-known distributions and at the same time providing great flexibility in modeling data in practice. So, several classes to generate new distributions by adding one or more parameters have been proposed in the existing literature. We define the transmuted Burr XII (TBXII) model and obtain some of its mathematical properties including explicit expressions for the ordinary and incomplete moments, moments of the residual and reversed residual lifetimes, mean deviations, generating function, order statistics, probability weighted moments and R´enyi and Shannon entropies. Also, a log-transmuted Burr XII regression model with the presence of censored data is proposed as an alternative to model lifetime when the failure rate function presents unimodal shape. We show that the TBXII model can provide better fits than other models generated from well-known families by means of two real data sets. Further, we give an application for the new regression model

References

[1] Afify A.Z., Alizadeh, M., Yousof, H.M., Aryal, G. and Ahmad, M. (2016a). The transmuted geometric-G family of distributions: theory and applications. Pakistan Journal Statistics, 32, 139–160.

[2] Afify, A.Z., Cordeiro, G.M., Ortega, E.M.M., Yousof, H.M. and Butt, N.S. (2016b). The four-parameter Burr XII distribution: properties, regression model and applications. Communications in Statistics-Theory and Methods, DOI: 10.1080/03610926.2016.1231821.

[3] Afify, A.Z., Nofal, Z.M. and Ebraheim, A.N. (2015). Exponentiated transmuted generalized Rayleigh distribution: a new four parameter Rayleigh distribution. Pakistan Journal of Statistics and Operations Research, 11, 115–134.

[4] Afify, A.Z., Cordeiro, G.M., Nadarajah, S., Yousof, H.M.,O¨ zel, G., Nofal, Z.M. and Altun, E. (2017a). The complementary geometric transmuted-G family of distributions: Model, properties and application. Hacettepe University Bulletin of Natural Sciences and Engi- neering Series B: Mathematics and Statistics, 47.

[5] Afify A. Z., Yousof, H. M. and Nadarajah, S. (2017b). The beta transmuted-H family for lifetime data. Statistics and Its Interface, 10, 505-520.

[6] Alexander, C., Cordeiro, G.M., Ortega, E.M.M. and Sarabia, J.M. (2012). Generalized beta-generated distributions. Computational Statistics and Data Analysis, 56, 1880–1897.

[7]   Al-Khazaleh, A.M.H. (2016). Transmuted Burr type XII distribution: A generalization of the Burr type XII distribution. International Mathematical Forum, 11, 547-556.

[8]   Al-Saiari, A.Y., Baharith, L.A. and Mousa, S.A. (2014). Marshall-Olkin extended Burr type XII distribution. International Journal of Statistics and Probability, 3, 78–84.

[9]   Alzaatreh, A., Famoye, F. and Lee, C. (2013). A new method for generating families of continuous distributions. Metron, 71, 63–79.

[10] Barreto, M.L., Santos, L.M.P., Assis, A.M.O., Araujo, M.P.N., Farenzena, G.G., Santos, P.A.B. and Fiaccone, R.L. (1994). Effect of vitamin A supplementation on diarrohea and acute lower-respiratory-tract infections in young children in Brazil. Lancet, 344, 228-231.

[11] Bourguignon, M., Silva, R.B. and Cordeiro, G.M. (2014). The Weibull-G family of proba- bility distributions. Journal of Data Science, 12, 53–68.

[12] Bourguignon, M., Ghosh, I., and Cordeiro, G.M. (2016). General Results for the Transmuted Family of Distributions and New Models. Journal of Probability and Statistics, 1-12.

[13] Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions. Journal of Statistical Computation and Simulation, 81, 883–893.

[14] Cordeiro, G.M., Hashimoto, E.M. and Ortega, E.M.M. (2014). The McDonald Weibull model. A Journal of Theoretical and Applied Statistics, 48, 256–278.

[15] Cordeiro, G.M., Ortega, E.M.M. and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. Journal of the Franklin Institute, 347, 1399– 1429.

[16] Cordeiro, G.M., Ortega, E.M.M., Hamedani, G.G.and Garcia,   D.A. (2016). The McBurr XII and log-McBurr XII models with applications to lifetime data. International Journal of Statistics and Probability, 5, 1-18.

[17] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. Communications in Statistics-Theory and Methods, 31, 497–512.

[18] Gradshteyn, I. S. and Ryzhik, I. M. (2000). Table of Integrals, Series and Products (sixth edition). San Diego: Academic Press.

[19] Gomes, E.E., da-Silva, C.Q. and Cordeiro, G.M. (2015). Two extended Burr models: Theory and practice. Communication in Statistics Theory-Methods, 44, 1706–1734.

[20] Khan, M.N. (2015). The modified beta Weibull distribution.   Hacettepe Journal of Mathematics and Statistics, 44, 1553–1568.

[21] Kundu, D. and Raqab, M.Z. (2009). Estimation of   R = P (Y < X) for three-parameter Weibull distribution. Statistics and Probability Letters, 79, 1839–1846.

[22] Marshall, A.N. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. Biometrika, 84, 641–652.

[23] Mead, M.E. (2014). The beta exponentiated Burr XII distribution. Journal of Statistics: Advances in Theory and Applications, 12, 53–73.

[24] Mead, M.E. and Afify, A.Z. (2017). On five parameter Burr XII distribution: properties and applications. South African Statistical Journal, forthcomig.

[25] Nasir, M.A., Tahir, M.H., Jamal, F. and O¨ zel, G. (2017). A New Generalized Burr Familyof Distributions for the Lifetime Data. J. Stat. Appl., 6, 1-17.

[26] Nofal, Z.M., Afify, A.Z., Yousof, H.M. and Cordeiro, G.M. (2017). The generalized transmuted-G family of distributions. Communications in Statistics-Theory and Methods, 46, 4119-4136.

[27] Parana´ıba, P.F., Ortega, E.M.M., Cordeiro, G.M. and Pescim, R.R. (2011). The beta Burr XII distribution with application to lifetime data. Computation Statistics and Data Anal- ysis, 55, 1118–1136.

[28] Parana´ıba, P.F., Ortega, E.M.M., Cordeiro, G.M. and de Pascoa, M. (2013). The Ku- maraswamy Burr XII distribution: theory and practice. Journal of Statistical Computation and Simulation, 83, 2117–2143.

[29] Provost, S.B., Saboor, A. and Ahmad, M. (2011). The gamma Weibull distribution. Pak- istan Journal Statistics, 27, 111–131.

[30] Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1992). Integrals and Series, 4.Gordon and Breach Science Publishers, Amsterdam.

[31] Shaw, W.T. and Buckley, I.R.C. (2007). The alchemy of probability distributions: beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmu- tation map. arXiv preprint arXiv:0901.0434.

[32] Silva, G.O., Ortega, E.M.M., Cancho, V.G. and Barreto, M.L. (2008). Log-Burr XII regres- sion models with censored data. Computational Statistics and Data Analysis, 52, 3820-3842.

[33] Torabi, H. and Montazari, N.H. (2014). The logistic-uniform distribution and its applica- tion. Communications in Statistics-Simulation and Computation, 43, 2551–2569.

[34] Zografos, K. and Balakrishnan, N. (2009). On families of beta- and generalized gammagen- erated distributions and associated inference. Statistical Methodology, 6, 344–362

[1]Ahmed Z.Afify

Department of Statistics, Mathematics and Insurance,

Benha University, Egypt

e-mail:Ahmed.Afify@fcom.bu.edu.eg

[2]Gauss M. Cordeiro

Departamento de Estat´ıstica,

Universidade Federal de Pernambuco, 50740-540, Recife, PE, Brazil

e-mail:gauss@de.ufpe.br

[3]Marcelo Bourguignon

Departamento de Estat´ıstica,

Universidade Federal de Pernambuco, 50740-540 Recife, PE, Brazil

e-mail:m.p.bourguignon@gmail.com