REVSTAT – Statistical Journal
Volume 19, Number 4, October 2021, 513–530

A TRUNCATED GENERAL-G CLASS OF DISTRIBUTIONS WITH APPLICATION TO TRUNCATED BURR-G FAMILY

Authors: Farrukh Jamal
– Department of Statistics, The Islamia of University Bahawalpur, Bahawalpur, Pakistan
drfarrukh1982@gmail.com

Hassan S. Bakouch
– Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

M. Arslan Nasir
– Department of Statistics, Govt. S.E College, Bahawalpur, Pakistan

Received: May 2019 Revised: August 2019 Accepted: August 2019

Abstract:

- In this paper, we introduce a truncated general-G class of distributions. This class can be viewed as a weighted family of distributions with a general weight function, and also it generalizes the beta generator family proposed by Eugene *et al.* (2002). Some features of the class are stated with a comprehensive study to the truncated Burr-G (TB-G) family as one of the important subclass of the introduced class. The study includes the mixture representation in terms of baseline distribution, moments, moment generating function, stochastic ordering, stress-strength parameter, entropies, estimation by the maximum likelihood. The applicability of some new sub-models of the TB-G family is shown using two practical data sets.

Keywords:

- *family of distributions; Burr distribution; quantile function; simulation; estimation; goodness-of-fit statistics.*

AMS Subject Classification:

- 60E05, 62E15.

1. INTRODUCTION

Over the last two decades, several extensions of the well-known lifetime distributions have been developed for modeling many types of practical data sets. This development is followed by many approaches for generating new families of (probability) distributions which increase chances of modeling data of various random nature. Among those families, we can mention: The beta generator (beta-G) by Eugene *et al.* (2002) [9], the gamma-G (type 1) by Zografos and Balakrishnan (2009) [19], the Kumaraswamy-G (Kw-G) by Cordeiro and de Castro (2011) [7], the gamma-G (type 2) by Ristic and Balakrishnan (2012) [13], the log-gamma-G by Amini *et al.* (2014) [4], beta weighted modified Weibull distribution using the beta generator by Saboor *et al.* (2016) [14], the generalized transmuted family of distributions by Alizadeh *et al.* (2017) [3], the odd-Burr generalized family of distributions by Alizadeh *et al.* (2017) [2], the odd Burr-III family of distributions by Jamal *et al.* (2017) [11], the extended odd family of probability distributions by Bakouch *et al.* (2019) [5] and mid-truncated Burr XII distribution and its applications in order statistics by Saran *et al.* (2019) [15].

In practical life problems, truncation arises in many fields, such as industry, biology, hydrology, reliability theory and medicine. An example of truncation is the progression of a disease which is not an increasing function, but will stabilize after time point. This point is called the truncation for the support of the variable of the interest which may be time, length, height etc. Therefore, many researchers are attracted to analyze such truncated data using truncated versions of the standard statistical distributions. For instance, the truncated Weibull distribution has been applied to analyze the tree diameter and height distributions in forestry, fire size and high-cycle fatigue strength prediction (see Zhang and Xie, 2011 [18]). In Zaninetti and Ferraro (2008) [17], the truncated Pareto distribution is compared to the Pareto distribution using astrophysics data and they concluded, generally, that the truncated Pareto distribution performs better than the Pareto. Burroughs and Tebbens (2002) [6] showed the suitability of truncated power law distributions for data sets of earthquake magnitudes and forest fire areas. Additional applications of the former distributions in hydrology and atmospheric science are given by Aban *et al.* (2006) [1].

Motivated by the importance of general families of distributions and truncation, we introduce a more flexible class of distributions with the cumulative distribution function (cdf)

$$F(x) = \int_0^{G(x,\xi)} r_T(t)\, dt = \int_0^{G(x,\xi)} \frac{r(t)}{R(1)}\, dt = \frac{R\big[G(x,\xi)\big]}{R(1)}, \tag{1.1}$$

where $r_T(t)$ is the probability density function (pdf) of a random variable (rv) with support $[0,1]$, hence it can be any truncated rv T on this support with a cdf, $R(\cdot)$ and $G(x,\xi)$ is the cdf of a real-valued rv X with pdf $g(x,\xi)$, ξ denoting the related parameter vector. Table 1 gives a list of some truncated distribution in the interval [0,1]. The associated pdf of (1.1) is

$$f(x) = \frac{r\big[G(x,\xi)\big]\, g(x,\xi)}{R(1)}, \qquad x \in \mathbb{R}, \tag{1.2}$$

and the survival function based on (1.1) is given as

$$h(x) = \frac{r\big[G(x,\xi)\big]\, g(x,\xi)}{R(1) - R\big[G(x,\xi)\big]}. \tag{1.3}$$

Further, the associated quantile function based on (1.1) having the form

$$Q_x(u) = G^{-1}\Big\{R^{-1}\big[R(1) \times u\big]\Big\}, \tag{1.4}$$

where $u \sim$ uniform$[0, 1]$.

Table 1: List of some truncated distribution in the interval [0,1].

S.r	Distribution	$r(t)$	$r_T(t)$
1.	Uniform	$F(x) = \frac{x}{\theta}$	$F(x) = x$
2.	Exponential	$F(x) = 1 - e^{-\theta x}$	$F(x) = \frac{1-e^{-\theta x}}{1-e^{-\theta}}$
3.	Weibull	$F(x) = 1 - e^{-ax^b}$	$F(x) = \frac{1-e^{-ax^b}}{1-e^{-a}}$
4.	Gamma	$F(x) = \frac{\gamma\left(a, \frac{x}{b}\right)}{\Gamma(a)}$	$F(x) = \frac{\gamma\left(a, \frac{x}{b}\right)}{\gamma\left(a, \frac{1}{b}\right)}$
5.	Lomax	$F(x) = 1 - \left(1 + \frac{x}{a}\right)^{-b}$	$F(x) = \frac{1-\left(1+\frac{x}{a}\right)^{-b}}{1-\left(1+\frac{1}{a}\right)^{-b}}$
6.	log-logistic	$F(x) = 1 - \left(1 + \frac{x^c}{a}\right)^{-1}$	$F(x) = \frac{1-\left(1+\frac{x^c}{a}\right)^{-1}}{1-\left(1+\frac{1}{a}\right)^{-1}}$
7.	Burr XII	$F(x) = 1 - (1 + x^c)^{-k}$	$F(x) = \frac{1-(1+x^c)^{-k}}{1-2^{-k}}$
8.	Burr III	$F(x) = (1 + x^{-c})^{-k}$	$F(x) = \frac{(1+x^{-c})^{-k}}{2^{-k}}$
9.	Frechet	$F(x) = \exp\left[-\left(\frac{a}{x}\right)^b\right]$	$F(x) = \frac{\exp\left[-\left(\frac{a}{x}\right)^b\right]}{\exp\left[-a^b\right]}$
10.	Power function	$F(x) = \left(\frac{x}{\theta}\right)^k$	$F(x) = x^k$
11.	Log normal	$F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$	$F(x) = \frac{\Phi\left(\frac{\ln x-\mu}{\sigma}\right)}{\Phi\left(\frac{-\mu}{\sigma}\right)}$

Some additional motivations of the class defined by (1.2) are as follows. The class (1.2) can be interpreted as weighted family of distributions, for $g(x,\xi)$, with the general weight function $w(X) = r\big(G(x,\xi)\big)$ and normalizing constant $R(1) = E\big\{w(X)\big\}$. Also, the introduced class generalizes the beta generator family (Eugene *et al.*, 2002 [9]) as beta distribution is a sub-model of $r_T(t)$.

As it can be seen from (1.2), we have a truncated general-G class of distributions and the only sub-model we aware of is the truncated Weibull G family proposed by Najarzadegan *et al.* (2017) [12] as a powerful alternative to beta-G family of distributions. Because of having two composite general functions $R(\cdot)$ and $G(\cdot)$, we can not investigate more analytic properties and therefore we aim to study extensively the truncated Burr-G (TB-G) family of distributions by considering $R(\cdot)$ as the cdf of Burr distribution and $G(\cdot)$ is a general cdf. The reason of using Burr is due to its ability of analyzing hydrologic, environmental, survival and reliability data. Another aim is to provide an empirical evidence on the great flexibility of sub-models of the TB-G family to fit practical data from different domains and this is investigated in the application section.

Rest of the paper is outlined as follows. Section 2 concerns with some general mathematical properties of the TB-G family, including mixture representation in terms of baseline distribution, moments, incomplete moments, moment generating function, stochastic ordering of the random variables following such family, stress-strength parameter and entropies (Shannon and Rényi). Also, some new special models of the generated family are considered. In Section 3, estimation of the parameters of the family is implemented through maximum likelihood method with application to two practical data sets. Section 4 gives a simulation study for a sub-model of the family.

2. THE TRUNCATED BURR-G FAMILY: SOME PROPERTIES AND SUB-MODELS

This section gives some general mathematical properties of the TB-G family, including moments, incomplete moments, moment generating function, stochastic ordering, stress-strength parameter and entropies. Further, some new sub-models of the family are obtained.

2.1. The truncated Burr-G family

In this section, we introduce the TB-G family of distributions and give its mixture representation in terms of baseline distribution.

Recall that the Burr distribution has the cdf

$$R(x) = 1-(1+x^c)^{-k}\,, \qquad x>0\,, \tag{2.1}$$

using (1.1), the cdf of the TB-G family is expressed as

$$F(x) \,=\, \frac{1-\big[1+G^c(x,\xi)\big]^{-k}}{1-2^{-k}}\,, \tag{2.2}$$

where c, k are the shape parameters of the family and $G(x,\xi)$ is a baseline cdf, which depends on a parameter vector ξ. Hereafter, for simplicity, we ignore mention of ξ in the functions of interest, e.g., we set $G(x) = G(x,\xi)$, $g(x) = g(x,\xi)$.

The pdf corresponding to (2.2) is given by

$$f(x) \,=\, \frac{c\,k\,g(x)\,G^{c-1}(x)\,\big[1+G^c(x)\big]^{-k-1}}{1-2^{-k}}\,, \qquad x\in\mathbb{R}\,. \tag{2.3}$$

The survival function and hazard rate are, respectively, given by

$$\bar F(x) \,=\, \frac{\big[1+G^c(x)\big]^{-k}-2^{-k}}{1-2^{-k}} \tag{2.4}$$

and

$$\tau(x) \,=\, \frac{c\,k\,g(x)\,G^{c-1}(x)\,\big[1+G^c(x)\big]^{-k-1}}{\big[1+G^c(x)\big]^{-k}-2^{-k}}\,. \tag{2.5}$$

Also, the quantile function of the TB-G family has the form

$$Q_x(u) = G^{-1}\left[\left\{\left[1-(1-2^{-k})\,u\right]^{-\frac{1}{k}} - 1\right\}^{\frac{1}{c}}\right]. \tag{2.6}$$

Further, the shapes of the density and hazard rate functions of the TB-G family can be described analytically using their critical points as follows. The critical points of the TB-G density are the roots of the equation:

$$\frac{g'(x)}{g(x)} + (c-1)\,\frac{g(x)}{G(x)} - c\,(k+1)\,\frac{g(x)\,G^{c-1}(x)}{1-G^c(x)} = 0\,,$$

while the critical points of the hazard rate are the roots of the equation:

$$\frac{g'(x)}{g(x)} + (c-1)\,\frac{g(x)}{G(x)} - c\,(k+1)\,\frac{g(x)\,G^{c-1}(x)}{1-G^c(x)} + k\,c\,\frac{g(x)\,G^{c-1}(x)\left[1+G^c(x)\right]^{-k-1}}{\left[1+G^c(x)\right]^{-k} - 2^{-k}} = 0\,.$$

Note that the equation above may have more than one root.

Now, we close this subsection by obtaining the mixture representation of the TB-G in terms of baseline distribution as follows.

Consider the series expansion, for $|z|<1$,

$$(1-z)^{-b} = \sum_{i=0}^{\infty}\binom{b+i-1}{i} z^i\,, \tag{2.7}$$

the cdf in equation (2.2) can be written as

$$F(x) = \frac{1}{1-2^{-k}}\left[1-\sum_{i=0}^{\infty}\binom{k+i-1}{i}(-1)^i\,G^{ic}(x)\right]. \tag{2.8}$$

Also, it can be rewritten in the form

$$F(x) = \sum_{l=0}^{\infty} b_l\,H_l(x)\,, \tag{2.9}$$

where $b_l = \frac{1}{1-2^{-k}}\sum_{i=1}^{\infty}\sum_{j=l}^{\infty}\binom{k+i-1}{i}\binom{c\,i}{j}\binom{j}{l}(-1)^{i+j+l+1}$ and $H_l(x) = G^l(x)$ is the exp-G distribution function with power parameter l.

Similarly, simple derivation of the previous equation gives the pdf

$$f(x) = \sum_{l=0}^{\infty} b_l\,h_{l-1}(x)\,, \tag{2.10}$$

where $h_{l-1}(x) = l\times g(x)\,G^{l-1}(x)$ is the exp-G density function with power parameter $l-1$. Thus, some mathematical properties of the proposed family can be derived from (2.10) and those of exp-G properties. For example, the ordinary and incomplete moments and moment generating function (mgf) of X can be obtained from those exp-G quantities, see the next subsection.

2.2. Moments and moment generating function

In this subsection, we will discuss the r^{th} moments, m^{th} incomplete moments and moment generating function of the TB-G family.

The moments of the TB-G family of distributions can be obtained by using the infinite mixture representation

$$E(X^r) = \sum_{l=0}^{\infty} b_l \int_{-\infty}^{\infty} x^r h_{l-1}(x)\, dx\,, \tag{2.11}$$

where b_l and $h_{q-1}(x)$ are defined in (2.10).

The s^{th} incomplete moment of the TB-G family can be obtained as

$$T_s'(x) = \sum_{l=0}^{\infty} b_l \int_{-\infty}^{x} x^s h_{l-1}(x)\, dx\,. \tag{2.12}$$

The moment generating function of the TB-G family of distributions is

$$M_X(t) = \sum_{l=0}^{\infty} b_l \int_{-\infty}^{\infty} e^{tx} h_{l-1}(x)\, dx\,.$$

Bonferroni and Lorenz curves, defined for a given probability, π, by $B(\pi) = T_1'(q)/(\pi\mu_1')$ and $L(\pi) = T_1'(q)/\mu_1'$, respectively, where $\mu_1' = E(X)$, $T_1'(x) = \sum_{l=0}^{\infty} b_l \int_{-\infty}^{x} x\, h_{l-1}(x)\, dx$ and $q = Q(\pi)$ is the quantile function of X at π. These curves for the Truncated Burr log logistic (TBLL) distribution (see definition of TBLL in the next subsection) as functions of π, are plotted for some parameter values in Figure 1. These curves are very useful in economics, reliability, demography, insurance and medicine. The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships form equation (2.11).

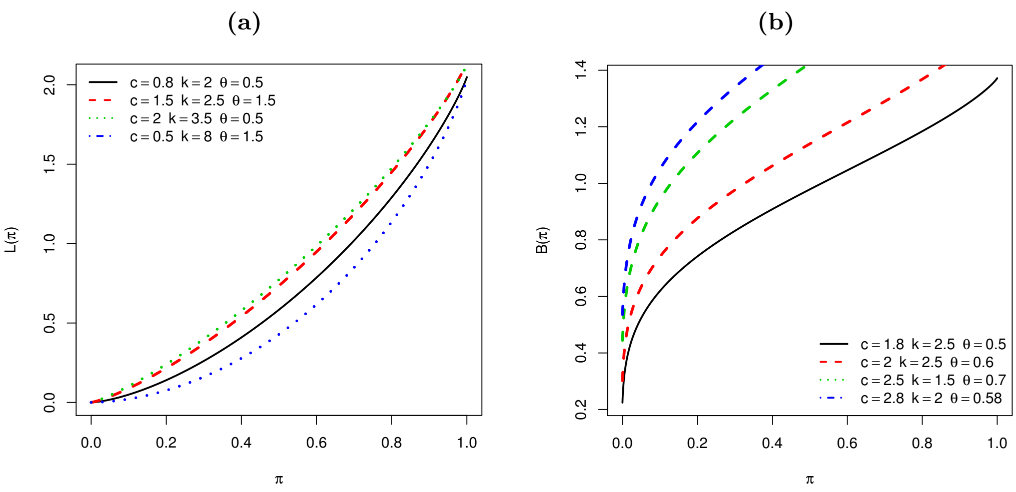

Figure 1: Plots of $B(\pi)$ and $L(\pi)$ versus π for the TB-LL distribution.

Plots of skewness and kurtosis of the TBLL distribution for $\theta = 1.5$ are displayed in Figure 2. Based on these plots, we conclude that, if c and k increase, the skewness and kurtosis decrease.

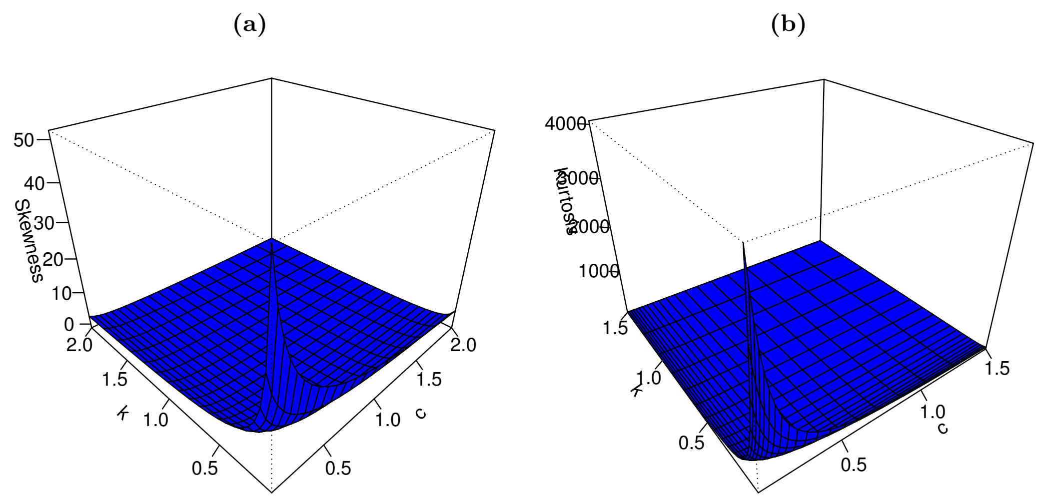

Figure 2: Plots for skewness and kurtosis of the TB-LL distribution.

2.3. Stochastic ordering and reliability parameter

Comparative behavior of random variables can be measured by stochastic ordering concept (Shaked and Shanthikumar, 1994 [16]) that is summarized in the next proposition.

Proposition 2.1. *Let* $X_1 \sim \text{TB-G}(c, k_1, \xi)$ *and* $X_2 \sim \text{TB-G}(c, k_2, \xi)$, *then the likelihood ratio* $\frac{f(x)}{g(x)}$ *is*

$$\frac{f(x)}{g(x)} = \frac{k_1}{k_2}\left[1 + G^c(x)\right]^{k_2 - k_1} \frac{1 - 2^{-k_2}}{1 - 2^{-k_1}}.$$

Taking derivative with respect to x, we have

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{k_1}{k_2}\frac{1 - 2^{-k_2}}{1 - 2^{-k_1}}\left[1 + G^c(x)\right]^{k_2 - k_1 - 1}(k_2 - k_1)\, c\, g(x)\, G^{c-1}(x),$$

then $\frac{d}{dx}\frac{f(x)}{g(x)} < 0$ for $k_2 < k_1$. So, the likelihood ratio exists and this implies that the random variable X_1 is a likelihood ratio order than X_2, that is $X_1 \leq_{\text{lr}} X_2$. Other stochastic ordering behaviors follow using $X_1 \leq_{\text{lr}} X_2$, such as hazard rate order ($X_1 \leq_{\text{hr}} X_2$), mean residual life order ($X_1 \leq_{\text{mrl}} X_2$) and stochastically greater ($X_1 \leq_{\text{st}} X_2$).

The stress strength model is a common approach used in various applications of engineering and physics. Let X_1 and X_2 be two independent random variables with TB-G(c, k_1, ξ) and TB-G(c, k_2, ξ) distributions. Then the stress strength model is given by

$$R = \int_{-\infty}^{\infty} f_1(x)\, F_2(x)\, dx.$$

Now, by using the mixture representation given in (2.10) and (2.9), we have

$$R = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty} b_l\, b_m \int_{-\infty}^{\infty} h_{l-1}(x)\, H_m(x)\, dx\,,$$

where $h_{l-1}(x)$ and $H_m(x)$ are already defined by equations (2.9) and (2.10).

2.4. Entropies

The entropy of a random variable X with density function $f(x)$ is a measure of variation of the uncertainty of physical systems. Two popular entropy measures are due to Shannon entropy and Rényi entropy. A large value of the entropy may indicate the greater uncertainty in the data; conversely, a small entropy means less uncertainty. The Rényi entropy is defined by

$$I_\delta = \frac{1}{1-\delta}\,\log\left(\int_{-\infty}^{\infty} f^\delta(x)\, dx\right), \qquad \delta > 0 \quad \text{and} \quad \delta \neq 1\,. \tag{2.13}$$

Let $f(x)$ follow the TB-G family, then we have

$$f^\delta(x) = \frac{(ck)^\delta\, g^\delta(x)\, G^{\delta(c-1)}(x)\, \left[1 + G^c(x,\xi)\right]^{-\delta(k+1)}}{(1-2^{-k})^\delta}\,.$$

After some algebra, we get

$$f^\delta(x) = \left(\frac{ck}{1-2^{-k}}\right)^\delta \sum_{j=0}^{\infty} \binom{\delta(k+1)+j-1}{j} (-1)^j\, g^\delta(x)\, G^{c(j+\delta)-\delta}(x)\,.$$

Rewriting the above expression as

$$f^\delta(x) = \sum_{j=0}^{\infty} w_j(\delta)\, g\big(x; \delta, c(j+\delta)\big)\,,$$

where $w_j(\delta) = \left(\frac{ck}{1-2^{-k}}\right)^\delta \binom{\delta(k+1)+j-1}{j} (-1)^j$ and $g\big(x;\delta,c(j+\delta)\big) = g^\delta(x)\, G^{c(j+\delta)-\delta}(x)$.

Now equation (2.13) becomes

$$I_\delta = \frac{1}{1-\delta}\,\log\left[\sum_{j=0}^{\infty} w_j(\delta) \int_{-\infty}^{\infty} g\big(x;\delta,c(j+\delta)\big)\, dx\right].$$

The above expression depends only for any choice of baseline distribution.

On the other side, the Shannon entropy of the TB-G family can be obtained using its definition as

$$\eta = -E\big[\log f(X)\big]\,. \tag{2.14}$$

Using the pdf of the TB-G family, we have

$$-E\big[\log f(X)\big] = \log[1-2^{-k}] - \log(ck) - E\big[\log g(X)\big] - (c-1)\,E\big[\log G(X)\big] + (k+1)\,E\Big[\log\big\{1+G^c(X)\big\}\Big]. \tag{2.15}$$

Making use of the expansions, for $|x|<1$,

$$\log(1+x) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}\, x^i\,,$$

$$\log x = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}\, (x-1)^i\,,$$

we obtain

$$E\Big[\log\big\{1+G^c(X)\big\}\Big] = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}\, E\big[G^{c\,i}(X)\big]\,,$$

$$E\big[\log G(X)\big] = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} \sum_{j=0}^{i} \binom{i}{j} (-1)^j\, E\big(G^{i-j}(X)\big)\,.$$

Hence, equation (2.15) becomes

$$\begin{aligned} -E\big[\log f(X)\big] &= \log[1-2^{-k}] - \log(ck) - E\big[\log g(X)\big] \\ &\quad - (c-1) \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} \sum_{j=0}^{i} \binom{i}{j} (-1)^j\, E\big(G^{i-j}(X)\big) \\ &\quad + (k+1) \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}\, E\big[G^{c\,i}(X)\big]\,. \end{aligned}$$

The expression above depends only on an arbitrary choice of the baseline distribution.

2.5. Some sub-models

In this subsection, we present four sub-models of the TB-G family by selecting some baseline distributions and the plots of their density and hazard rate functions. The plots indicate various shapes for both functions which proves the flexibility of the family. This flexibility is also confirmed by comparing those sub-models with other competing distributions for some practical data in Section 3.

Truncated Burr Uniform (TBU) distribution

Consider the uniform distribution on $(0,\theta)$ as the baseline distribution with the pdf and cdf, $g(x,\theta)=\frac{1}{\theta}$ and $G(x,\theta)=\frac{x}{\theta}$, respectively. Then the pdf and cdf of the TBU distribution are given by

$$f(x;c,k,\theta) = \frac{ck}{\theta}\,\frac{\left(\frac{x}{\theta}\right)^{c-1}}{1-2^{-k}} \left[1+\left(\frac{x}{\theta}\right)^c\right]^{-k-1}$$

and

$$F(x;c,k,\theta) = \frac{1-\left[1+\left(\frac{x}{\theta}\right)^c\right]^{-k}}{1-2^{-k}}\,, \qquad 0<x<\theta\,.$$

Figure 3 gives the plots of density and hrf of the TBU distribution.

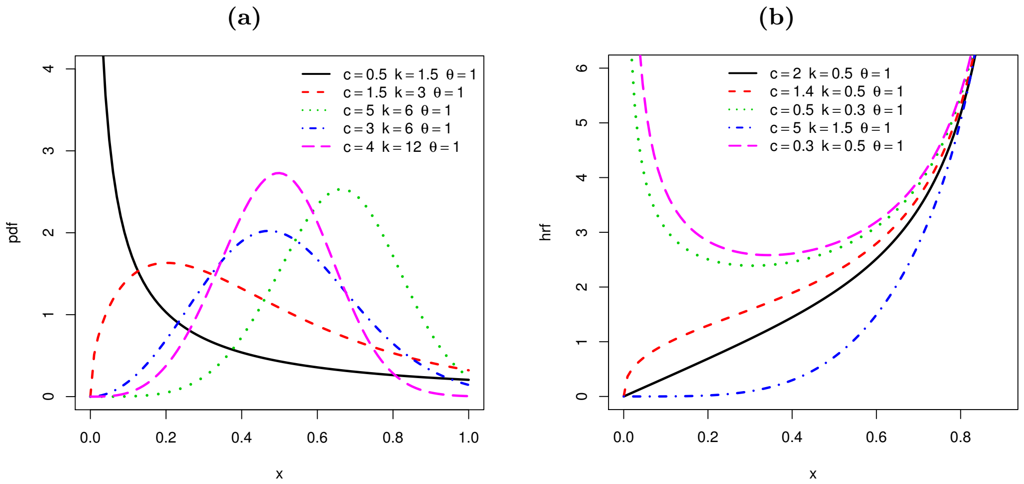

Figure 3: Plots for density and hrf of the TBU.

Truncated Burr Weibull (TBW) distribution

Let the Weibull distribution be the baseline one with the associated pdf and cdf, $g(x,a,b) = a\,b\,x^{b-1}e^{-ax^b}$ and $G(x,a,b) = 1 - e^{-ax^b}$, respectively. Then the pdf and cdf of the TBW distribution are given by

$$f(x;c,k,a,b) \;=\; \frac{c\,k\,a\,b\,x^{b-1}e^{-ax^b}}{1-2^{-k}}\,\frac{\left[1-e^{-ax^b}\right]^{c-1}}{\left[1+\{1-e^{-ax^b}\}^c\right]^{k+1}}$$

and

$$F(x;c,k,a,b) \;=\; \frac{1-\left[1+\{1-e^{-ax^b}\}^c\right]^{-k}}{1-2^{-k}}, \qquad 0<x<\infty\,.$$

Figure 4 displays the plots of density and hrf of the TBW distribution.

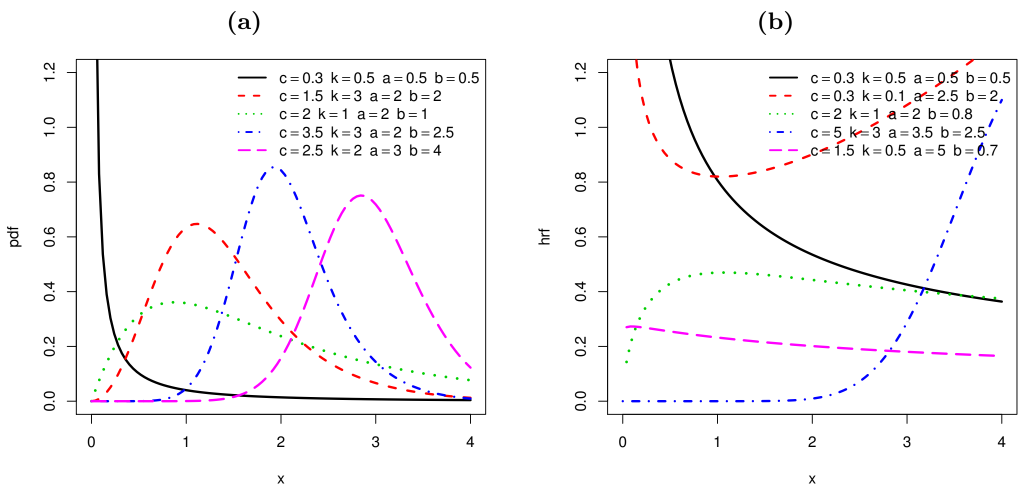

Figure 4: Plots for density and hrf of the TBW.

Truncated Burr Logistic (TBL) distribution

Consider the Logistic as the baseline distribution with associated pdf and cdf, $g(x,\theta) = \left\{1 - e^{-\theta x}\right\}^{-1}$ and $G(x,\theta) = \theta\, e^{-\theta x} \left\{1 - e^{-\theta x}\right\}^{-2}$, respectively. Then the pdf and cdf of the TBL distribution are given by

$$f(x;c,k,\theta) = \frac{c\,k\,\theta\,e^{-\theta\,x}}{\left[1-2^{-k}\right]\left\{1-e^{-\theta x}\right\}^{2}} \left[1-e^{-\theta x}\right]^{1-c} \left[1+\left\{1-e^{-\theta x}\right\}^{-c}\right]^{-k-1}$$

and

$$F(x;c,k,\theta) = \frac{1-\left[1+\left\{\left[1-e^{-\theta x}\right]^{-c}\right\}^{-k}\right]}{1-2^{-k}}, \qquad 0 < x < \infty .$$

In Figure 5 we give the plots of density and hrf of the TBL distribution.

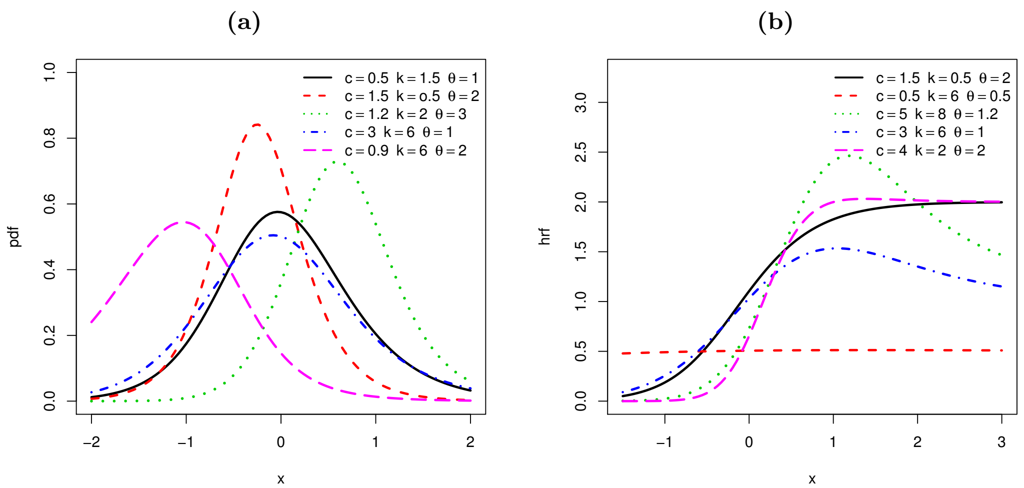

Figure 5: Plots for density and hrf of the TBL.

Truncated Burr log logistic (TBLL) distribution

Let log logistic be the baseline distribution with the associated pdf and cdf, $g(x,\theta) = \frac{\theta\,x^{\theta}}{\left(1+x^{\theta}\right)^{2}}$ and $G(x,\theta) = \frac{x^{\theta}}{1+x^{\theta}}$, respectively. Then the pdf and cdf of the TBLL distribution are given by

$$f(x;c,k,\theta) = \frac{c\,k\,\theta\,x^{\theta}}{\left[1-2^{-k}\right]\left(1+x^{\theta}\right)^{2}} \left[\frac{x^{\theta}}{1+x^{\theta}}\right]^{c-1} \left[1+\left\{\frac{x^{\theta}}{1+x^{\theta}}\right\}^{c}\right]^{-k-1}$$

and

$$F(x;c,k,\theta) = \frac{1-\left[1+\left\{\frac{x^{\theta}}{1+x^{\theta}}\right\}^{c}\right]^{-k}}{1-2^{-k}} .$$

Figure 6 portrays the plots of density and hrf of the TBLL distribution.

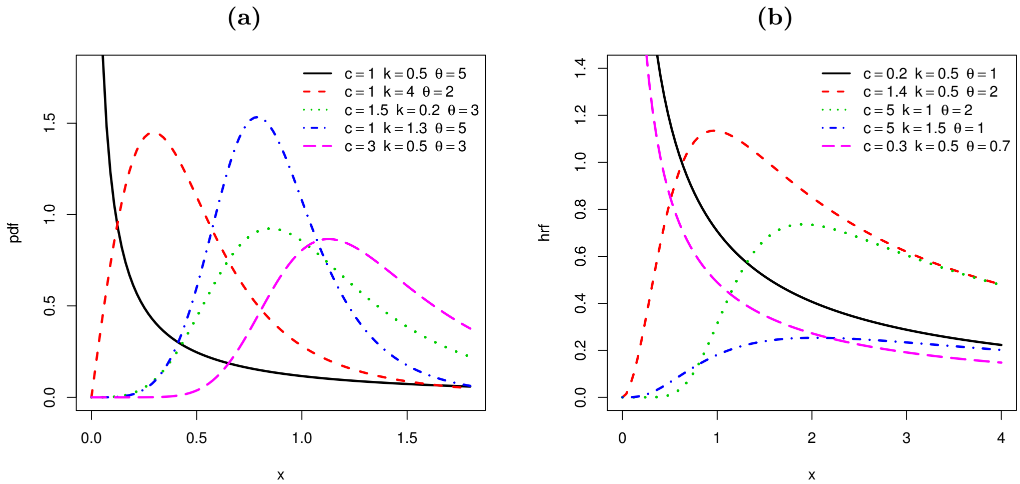

Figure 6: Plots for density and hrf of the TBLL.

3. ESTIMATION OF PARAMETERS WITH APPLICATIONS

In this section, we give the maximum likelihood estimators (MLEs) of the unknown parameters of the TB-G family for complete samples. Using those estimators we check the capability of some sub-models of this family for fitting some practical data sets. Let $x_1, x_2, ..., x_n$ be the observed values of a random sample of size n from the TB-G family given in equation (2.3). The log-likelihood function for the vector parameter $\Theta = [c, k, \xi]^{\mathsf{T}}$ can be expressed as

$$\begin{aligned} \ell(\Theta) &= -n\log(1-2^{-k}) + n\log(ck) + \sum_{i=1}^{n}\log g(x_i) + (c-1)\sum_{i=1}^{n}\log G(x_i) \\ &- (k+1)\sum_{i=1}^{n}\log\{1+G^c(x_i)\}\,. \end{aligned}$$

The components of score vector $U = (U_k, U_c, U_\xi)^{\mathsf{T}}$ are given by

$$\begin{aligned} U_k &= -n\,\frac{2^{-k}\log 2}{1-2^{-k}} + \frac{n}{k} - \sum_{i=1}^{n}\log\{1+G^c(x_i)\}\,, \\ U_c &= \frac{n}{c} + \sum_{i=1}^{n}\log G(x_i) - (k+1)\sum_{i=1}^{n}\left[\frac{c\,g(x_i)\,G^{c-1}(x_i)}{1+G^c(x_i)}\right], \\ U_\xi &= \sum_{i=1}^{n}\left[\frac{g^{\xi}(x_i)}{g(x_i)}\right] + (c-1)\sum_{i=1}^{n}\left[\frac{G^{\xi}(x_i)}{G(x_i)}\right] - (k+1)\sum_{i=1}^{n}\left[\frac{c\,G^{\xi}(x_i)\,G^{c-1}(x_i)}{1+G^c(x_i)}\right]. \end{aligned}$$

The equations above are non-linear and hence can not be solved analytically, but can be solved numerically using software like R language. The rest of this section provides two applications of four sub-models of the TB-G family, namely, the TBW, TBLL, TBU and TBL distributions given in Subsection 2.5. Truncated Weibull-BXII (TW-BXII) and Truncated Weibull-Weibull (TW-W) introduced by Najarzadegan *et al.* (2017) [12] are used as competitive models for those sub-models. For comparison purposes, we consider two practical data sets, one is taken from El-deeb (2015) [8] and another from Hinkley (1977) [10]. Description of both data sets is as follows.

Data set 1: This data set is given by El-deeb (2015) [8] and consists of failure times of (67) truncated Aircraft windshield. The windshield on an aircraft is a complex piece of equipment, comprised basically of several layers of material, all laminated under high temperature and pressure. Failures of these items are not structural failures. Instead, they typically involve damage or delimitation of the nonstructural outer ply or failure of the heating system. These failures do not result in damage to the aircraft, but do result in replacement of the windshield. The values of this data set are: 1.866, 2.385, 3.443, 1.876, 2.481, 3.467, 1.899, 2.610, 3.478, 1.911, 2.625, 3.578, 1.912, 2.632, 3.595, 1.070, 1.914, 2.646, 3.699, 1.124, 1.981, 2.661, 3.779,1.248, 2.010, 2.688, 3.924, 1.281, 2.038, 2.82,3, 3.000, 1.281, 2.085, 2.890, 1.303, 2.089, 2.902, 1.432, 2.097, 2.934, 1.480, 2.135, 2.962, 1.505, 2.154, 2.964, 1.506, 2.190, 3.000, 1.568, 2.194, 3.103, 1.615, 2.223, 3.114, 1.619, 2.224, 3.117, 1.652, 2.229, 3.166, 1.652, 2.300, 3.344, 1.757, 2.324, 3.376.

Data set 2: This data set is given by Hinkley (1977) [10] and consists of thirty successive values of March precipitation (in inches) in Minneapolis/St. Paul. In meteorology, precipitation is most commonly rainfall, but also includes hail, snow and other forms of liquid and frozen water falling to the ground and it is measured by inches in some time period. The data values are 0.77, 1.74, 0.81, 1.2, 1.95, 1.2, 0.47, 1.43, 3.37, 2.2, 3, 3.09, 1.51, 2.1, 0.52, 1.62, 1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.9, 2.05.

For each distribution, the MLEs are computed using Quasi-Newton code for Bound Constrained Optimization (L-BFGS-B) and the log-likelihood function is evaluated. Consequently, the goodness-of-fit measures: Anderson–Darling (A^*), Cramer–von Mises (W^*), Akaike information criterion (AIC) and Bayesian information criterion (BIC) are computed. Lower values of those measures indicate better fit. The value for the Kolmogorov–Smirnov (KS) statistic and its p-value are also provided. The required computations are carried out using the R software.

The obtained results are presented in Tables 2–5. As we can see from Tables 2 and 4, the four sub-models of the TB-G family are strong competitor to the compared models.

Table 2: MLEs and their standard errors (in parentheses) for data set 1.

Distribution	c	k	θ	a	b
TBW	0.4564 (1.9144)	86.9870 (45.4333)	— —	9.1067 (2.1784)	7.9149 (3.2404)
TBLL	13.6258 (2.3252)	193.8078 (34.7291)	0.7890 (0.2350)	— —	— —
TBU	3.5954 (0.3412)	498.2935 (15.2232)	14.9104 (12.1123)	— —	— —
TBL	23.3433 (7.0993)	0.0024 (0.0018)	1.6699 (0.1944)	— —	— —
TW-BXII	1.2904 (0.3253)	11.4013 (13.4118)	32.4704 (35.6313)	37.8343 (40.8586)	3.4896 (2.4676)
TB-W	2.8676 (2.7877)	0.8444 (0.6816)	— —	31.2399 (2.1419)	6.7846 (8.0910)

Moreover, among all compared models, the TBLL distribution has the smallest values of the AIC, BIC, A^{*}, W^{*}, and KS, and the largest value of p-value. Thus, we can conclude that the TBLL distribution is the best fit among those models. Figures 7 and 8 display the plots of the fitted pdfs and cdfs of the compared distributions for visual comparison with the histogram and empirical cdf for both data sets. Those figures show the best fit of TBLL distribution.

Table 3: The Value, *AIC*, *BIC*, *A**, *W**, *KS*, P-Value values for data set 1.

Distribution	ℓ	*AIC*	*BIC*	*A**	*W**	*KS*	P-Value
TBW	75.1080	158.2162	167.0942	0.5552	0.0951	0.0992	0.5147
TBLL	74.8708	155.7418	162.4003	0.4637	0.0740	0.0808	0.7379
TBU	75.0909	156.1819	162.8404	0.5564	0.0954	0.0997	0.5080
TBL	76.2189	158.4378	165.0963	0.5855	0.0859	0.0927	0.6016
TW-BXII	75.0635	160.1271	171.2246	0.5051	0.0841	0.0893	0.6487
TW-W	75.0454	158.0909	166.9690	0.4889	0.0798	0.0835	0.7299

Table 4: MLEs and their standard errors (in parentheses) for data set 2.

Distribution	c	k	θ	a	b
TBW	0.3446 (2.8251)	30.8825 (17.3728)	- -	11.9180 (10.6096)	5.3663 (4.4130)
TBLL	8.6122 (6.0513)	123.2974 (12.2964)	0.4892 (0.4066)	- -	- -
TBU	1.8150 (0.2482)	259.5434 (12.1122)	40.3962 (33.2333)	- -	- -
TBL	7.7107 (2.1529)	0.5621 (3.0901)	1.3198 (0.3681)	- -	- -
TW-BXII	1.0579 (1.1048)	86.6647 (71.9193)	60.8969 (69.5585)	0.0024 (4.5165)	3.0599 (6.3469)
TB-W	9.7190 (12.7756)	6.2763 (9.6175)	- -	19.3190 (46.5365)	0.2883 (0.4437)

Table 5: The Value, *AIC*, *BIC*, *A**, *W**, *KS*, P-Value values for data set 2.

Distribution	ℓ	*AIC*	*BIC*	*A**	*W**	*KS*	P-Value
TBW	38.5661	85.1322	90.7370	0.1571	0.0203	0.0648	0.9996
TBLL	38.0934	82.1868	86.3904	0.1019	0.0137	0.0576	1
TBU	38.6334	83.2668	87.4701	0.1680	0.0217	0.0683	0.9990
TBL	38.9520	83.9040	88.1076	0.1466	0.0185	0.0692	0.9988
TW-BXII	38.0919	86.1839	93.1899	0.1037	0.0141	0.0605	0.9999
TW-W	38.6431	85.2862	90.8910	0.1690	0.0219	0.0688	0.9989

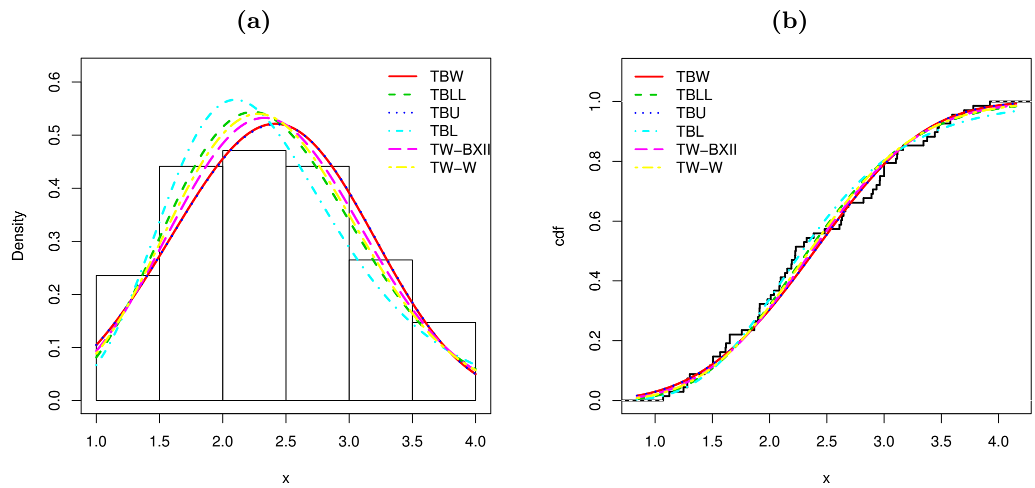

Figure 7: Estimated pdfs and cdfs for data set 1.

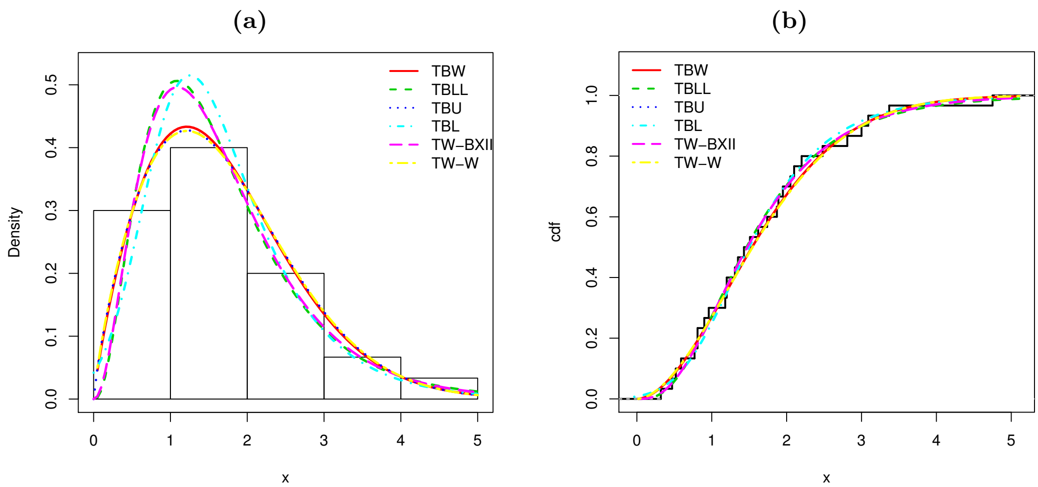

Figure 8: Estimated pdfs and cdfs for data set 2.

4. SIMULATION STUDY

In this section, the performance of the MLEs of the TBLL distribution parameters is discussed by means of Monte-Carlo simulation study. The following measures are used to evaluate the simulation results: Estimated bias, Root mean square error (RMSE) and coverage probability (CP). The simulation experiment was repeated $N = 1{,}000$ times each with sample sizes $n = 20$, 50, 100, 200, 300 and 500, where the samples are generated from the TBLL distribution, with $\theta = 4.5$, $c = 2.8$, $k = 0.8$, by using the inverse transform method. The MLEs of the parameters of TBLL distribution are obtained for each generated sample, $(\hat{\theta}, \hat{c}, \hat{k})$. The formulas for biases, RMSEs and CPs are given as follows.

Estimated bias of MLE $\hat{\Theta}$ of the parameter $\Theta = (\theta, c, k)$ is

$$\frac{1}{N} \sum_{i=1}^{N} \left(\hat{\Theta} - \Theta\right).$$

Root mean squared error (RMSE) of the MLE $\hat{\Theta}$ of the parameter $\Theta = (\theta, c, k)$ is

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\Theta}-\Theta)^2}.$$

Coverage probability (CP) of 95% confidence intervals of the parameter $\Theta = (\theta, c, k)$ is the percentage of intervals that contain the true value of parameter Θ.

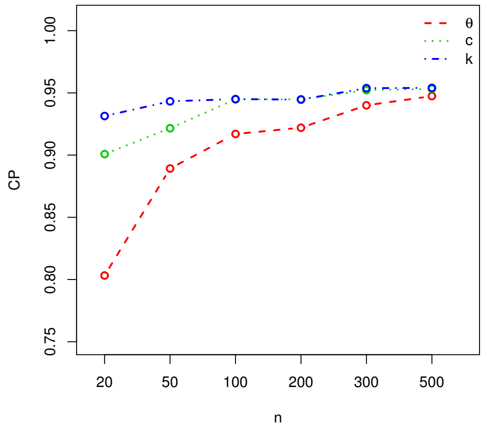

Figure 9: Estimated CPs for the selected parameters.

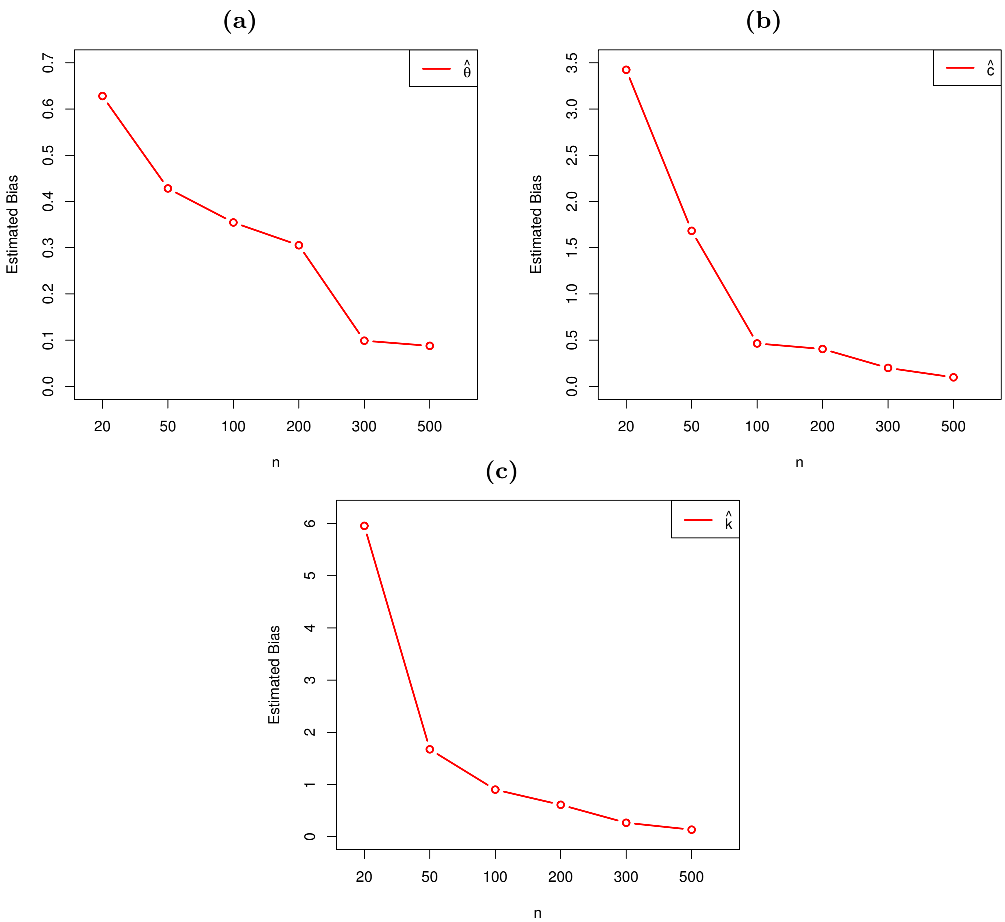

Figure 10: Estimated CPs for the selected parameters.

From Figures 9–11 we conclude that the estimated biases are positive for all parameters. The estimated biases decrease as the sample size n increases. Further, the estimated RMSEs are so closed to zero for large sample sizes. This result reveals the consistency property of the MLEs. The CP approaches to the nominal value (0.95) when the sample size increases.

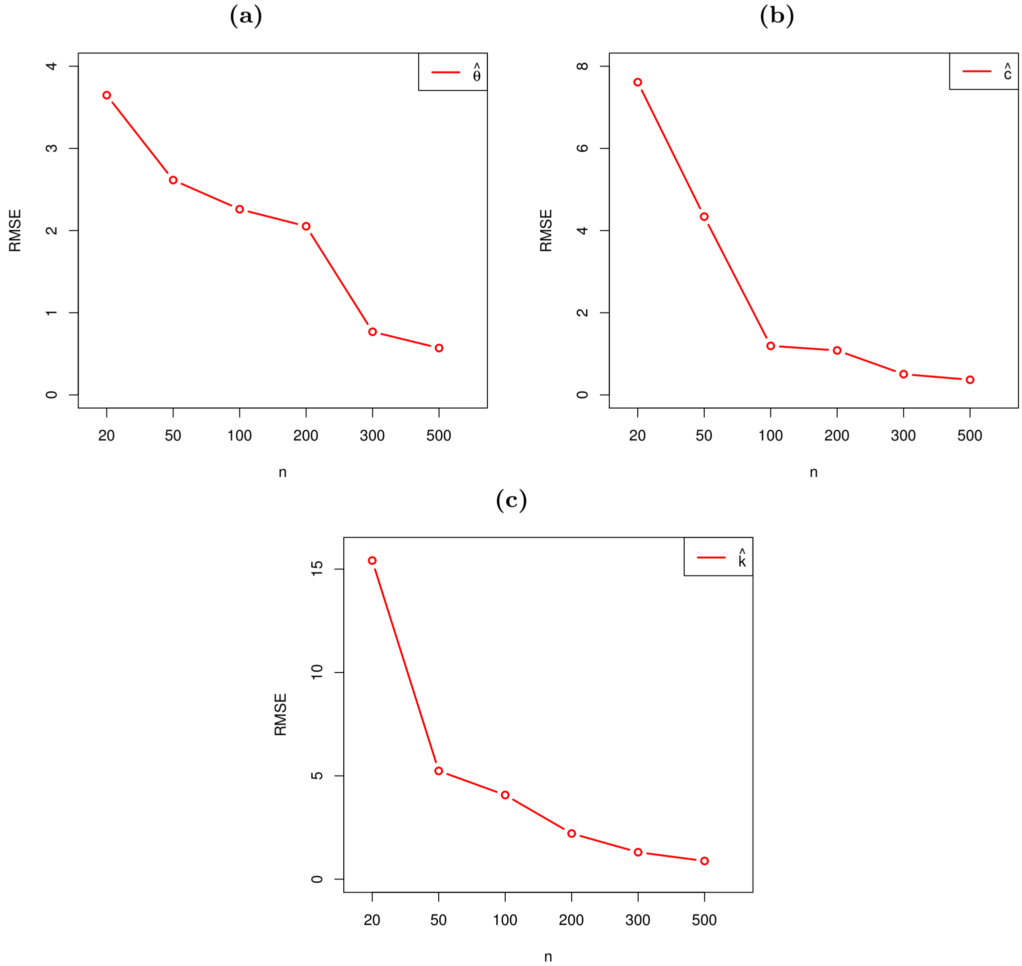

Figure 11: Estimated CPs for the selected parameters.

ACKNOWLEDGMENTS

The authors express their sincere thanks to the managing editor and the referee for careful reading of the article and providing valuable suggestions.

REFERENCES

[1] Aban, I.B.; Meerschaert, M.M. and Panorska, A.K. (2006). Parameter estimation for the truncated Pareto distribution, *Journal of the American Statistical Association*, **101**, 270–277.

[2] Alizadeh, M.; Cordeiro, Gauss M.; Nascimento, Abraão D.C.; Lima, Maria do Carmo S. and Ortega, Edwin M.M. (2017). Odd-Burr generalized family of distributions with some applications, *Journal of Statistical Computation and Simulation*, **87**, 367–389.

[3] Alizadeh, M.; Merovci, F. and Hamedani, G.G. (2017). Generalized transmuted family of distributions: properties and applications, *Hacettepa Journal of Mathematics and Statistics*, **46**, 645–667.

[4] Amini, M.; MirMostafaee, S.M.T.K. and Ahmadi, J. (2014). Log-gamma-generated families of distributions, *Statistics*, **48**, 913–932.

[5] Bakouch, H.S.; Chesneau, C. and Khan, M.N. (2019). The extended odd family of probability distributions with practice to a submodel, *FILOMAT*, **33**, 3855–3867.

[6] Burroughs, S.M. and Tebbens, S.F. (2002). The upper-truncated power law applied to earthquake cumulative frequency-magnitude distributions, *Bulletin of the Seismological Society of America*, **92**, 2983–2993.

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

[8] El-deeb, A.M.H. (2015). *Weibull–Lomax distribution and its properties and applications*, A Thesis, Al-Azhar University – Gaza Deanship of Postgraduate Studies Faculty of Economics and Administrative Sciences Department of Statistics, page 113, unpublished thesis.

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

[10] Hinkley, D. (1977). On quick choice of power transformations, *Journal of the Royal Statistical, Series C*, **26**(1), 67–69.

[11] Jamal, F.; Nasir, M.A.; Tahir, M.H. and Montazeri, N.H. (2017). The odd Burr-III family of distributions, *Journal of Statistics Applications and Probability*, **6**(1), 105–122.

[12] Najarzadegan, H.; Alamatsaz, M.H. and Hayati, S. (2017). Truncated Weibull-G more flexible and more reliable than beta-G distribution, *International Journal of Statistics and Probability*, **6**(5), 1–17.

[13] Ristic, M.M. and Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution, *Journal of Statistical Computation and Simulation*, **82**, 1191–1206.

[14] Saboor, A.; Bakouch, H.S. and Khan, M.N. (2016). Beta Sarhan–Zaindin modified Weibull distribution, *Applied Mathematical Modelling*, **40**, 6604–6621.

[15] Saran, J.; Pushkarna, N. and Sehgal, S. (2019). Mid-truncated Burr XII distribution and its applications in order Statistics, *Statistics, Optimization and Information Computing*, **7**, 171–191.

[16] Shaked, M. and Shanthikumar, J.G. (1994). *Stochastic Orders and Their Applications*, Academic Press, New York.

[17] Zaninetti, L. and Ferraro, M. (2008). On the truncated Pareto distribution with application, *Central European Journal of Physics*, **6**, 1–6.

[18] Zhang, T. and Xie, M. (2011). On the upper truncated Weibull distribution and its reliability implications, *Reliability Engineering and System Safety*, **96**, 194–200.

[19] Zografos, K. and Balakrishnan, N. (2009). On families of beta and generalized gamma-generated distributions and associated inference, *Statistical Methodology*, **6**, 342–362.