Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v13n1p229

On the weighted BurrXII distribution: theory and practice

By Shakhatreh, Al-Masri

Published: 02 May 2020

Electronic Journal of Applied Statistical Analysis
Vol. 13, Issue 01, May 2020, 229-255
DOI: 10.1285/i20705948v13n1p229

On the weighted BurrXII distribution: theory and practice

Mohammed K. Shakhatreh*[a] and Abedel-Qader S. Al-Masri[b]

[a] *Department of Mathematics and Statistics, Jordan University of Science and Technology , P.O.Box 3030, Irbid 22110, Jordan*
[b] *Department of Statistics, Yarmouk University, Irbid, Jordan*

Published: 02 May 2020

We take an in-depth look at the weighted Burr-XII distribution. This distribution generalizes Burr-XII, Lomax, and log-logistic distributions. We discuss the distributional characteristics of the probability density function, the failure rate function, and mean residual lifetime of this distribution. Moreover, we obtain various statistical properties of this distribution such as moment generating function, entropies, mean deviations, order statistics and stochastic ordering. The estimation of the distribution parameters via maximum likelihood method and the observed Fisher information matrix are discussed. We further employ a simulation study to investigate the behavior of the maximum likelihood estimates (MLEs). A test concerning the existence of size-bias in the sample is provided. In the end, a real data is presented and it is analyzed using this distribution along with some existing distributions for illustrative purposes.

keywords: Burr-XII distribution, hazard rate function,mean residual lifetime, maximum likelihood estimation, weighted Burr-XII distribution.

1 Introduction

The Burr-XII (BXII) distribution having log-logistic, Lomax, and Weibull distributions as special sub-cases is introduced and studied in Burr (1942). This distribution is commonly used in many real applications such as flood frequency (Shao et al., 2004), reliability (Wingo, 1993; Zimmer et al., 1998), and survival analysis (Shao and Zhou, 2004)

*Corresponding author: mkshakhatreh6@just.edu.jo

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

due to its amenability and versatility. The attractive properties of the BXII distribution have been investigated by many authors including Burr and Cislak (1968); AL-Hussaini (1991); Rodriguez (1977) and Tadikamalla (1980) who gave the details on the connection between the (BXII) distribution and other distributions. According to Soliman (2005), the BXII distribution possessing algebraic tails which are effective for modeling phenomena with less frequency compared to corresponding models that are based on exponential tails. Recently, several generalizations of the BXII distribution have been introduced. For example, the beta BXII , Kumaraswamy BXII, Weibull BXII, Marshall-Olkin exponentiated BXII distributions have been considered and studied by Paranaíba et al. (2011); Paranaíba et al. (2013); Afify et al. (2018); Cordeiro et al. (2017) respectively. Statistical inferences on the parameters of BXII distribution from complete and censored cases have been discussed extensively, see, Wang et al. (1996); Wingo (1993); Shao (2004); Soliman (2005); Wu et al. (2007); Silvaa et al. (2008); Usta (2013).
The idea of weighted (WD) distributions is originally introduced in Fisher (1934). These distributions can be used to model many practical problems in the presence of biased samples that arise naturally in many situations. More precisely, if X represents a random variable of interest with probability density function $f(x)$, then the weighted version of X with respect to the weight function $w(x) > 0$ is a random variable denoted by X_w with probability density function given

$$f_{X_w}(x) = \frac{w(x)f(x)}{\mathbb{E}[w(X)]}, \qquad x > 0, \tag{1}$$

provided that $\mathbb{E}[w(X)] < \infty$. Of particular interest is the generalized length-biased (GLB) distributions of order $\alpha > 0$ which can be obtained as special sub-case from (1) by taking $w(x) = x^\alpha$. The so-called GLB distribution of order α is considered and studied in (Patil and Ord, 1976). In this case, the pdf of the GLB is given by

$$f_w(x) = \frac{x^\alpha f(x)}{\mu_f^\alpha}, \tag{2}$$

where $\mu_f^\alpha = \int_0^\infty x^\alpha f(x)dx$. Notice that if $\alpha = 0$ then the existence of size-bias in the sample does not appear. Real applications of WD distributions are spread out in various fields including environmental sciences (Patil, 2002; Gove, 2003), forestry (Gove and Patil, 1998; Ducey and Gove, 2015), reliability (Gupta and Kirmani, 1990; Sansgiry and Akman, 2006; Jain et al., 2007), water quality(Leiva et al., 2009)
The main objective of this paper is to study the structural properties of the weighted BXII distribution and to provide various statistical measures such as moments, mean residual life, order statistics and entropies. We also focus on the statistical inference on the parameters of WWBXII distribution. In particular, we use the maximum likelihood method to estimate these parameters and we discuss their proprieties. One major interest is to provide some information concerning the existence of size-bias in the sample.

The rest of the paper is organized as follows. The WWBXII distribution along with the shapes of its probability density function and hazard rate function are given in section 2. In Section 3, we provide various statistical measures and properties of the

WWBXII distribution. Estimating the parameters of the WWBXII distribution using maximum likelihood method along with a test concerning the existence of size-bias in the sample are given in Section 4. In Section 5, we conduct simulation study to assess the performance of the MLEs. In Section 6, a real life data analysis is considered and analyzed for illustrative purpose. Concluding remarks are given in Section 7.

2 The weighted BurrXII distribution

Let X be a random variable following the BXII with shape parameters parameters c, $k > 0$. The probability density function (pdf) of X can take the following form

$$f(x) = \frac{k\,c\,x^{c-1}}{(1+x^c)^{k+1}}, \qquad x > 0. \tag{3}$$

The α^{th} moment of X is

$$\mathbb{E}(X^\alpha) = \frac{k\Gamma(k-\frac{\alpha}{c})\Gamma(\frac{\alpha}{c}+1)}{\Gamma(k+1)}, \quad ck > \alpha. \tag{4}$$

Using Equation (2) along with equations (3) and (4), the linked pdf of the weighted BurrXII distribution (WWBXII) given respectively on the support of X is

$$f_w(x) = \frac{c\,x^{\alpha+c-1}}{\mathbb{B}(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c})\,(1+x^c)^{k+1}}, \qquad x > 0, \quad c, k, \alpha > 0, \quad k > \frac{\alpha}{c}, \tag{5}$$

where $\mathbb{B}(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. The cumulative distribution function (cdf) of X is given by

$$F(x) = \mathbb{I}_{\frac{x^c}{1+x^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right) = \frac{1}{\mathbb{B}(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c})}\int_0^{\frac{x^c}{1+x^c}} y^{\frac{\alpha}{c}}(1-y)^{k-\frac{\alpha}{c}-1}dy, \tag{6}$$

where $\mathbb{B}_y(a,b) = \int_0^y w^{a-1}(1-w)^{b-1}dw$ is the incomplete beta function and $\mathbb{I}_y(a,b) = \frac{B_y(a,b)}{B(a,b)}$ is the incomplete beta function ratio. To show Equation (6), we have that

$$F_w(x) = \int_0^x f_w(y)\,dy = \frac{c}{\mathbb{B}(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c})}\int_0^x y^{\alpha+c-1}(1+y^c)^{-(k+1)}dy \tag{7}$$

On substituting $u = \frac{1}{1+y^c}$, Equation (6) reduces to

$$\begin{aligned} F_w(x) &= \frac{1}{\mathbb{B}(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c})}\int_{\frac{1}{1+x^c}}^1 u^{k-\frac{\alpha}{c}-1}(1-u)^{\frac{\alpha}{c}}du = 1 - \frac{\int_0^{\frac{1}{1+x^c}} u^{k-\frac{\alpha}{c}-1}(1-u)^{\frac{\alpha}{c}}du}{\mathbb{B}(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c})} \\ &= 1 - \mathbb{I}_{\frac{1}{1+x^c}}\left(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c}\right) = \mathbb{I}_{\frac{x^c}{1+x^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right), \end{aligned}$$

where the last equality follows from $\mathbb{I}_x(a,b) = 1 - \mathbb{I}_{1-x}(b,a)$. The reliability (rf), hazard (hrf), and reversed hazard(rhf) functions are given respectively by

$$S_w(x) = 1 - F_w(x) = \mathbb{I}_{\frac{1}{1+x^c}}\left(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c}\right), \tag{8}$$

$$h_w(x) = \frac{c\, x^{\alpha+c-1}}{\mathbb{B}_{\frac{1}{1+x^c}}\left(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c}\right)\,(1+x^c)^{k+1}}, \tag{9}$$

$$a_w(x) = \frac{c\, x^{\alpha+c-1}}{\mathbb{B}_{\frac{x^c}{1+x^c}}\left(1 + \frac{\alpha}{c}, k - \frac{\alpha}{c}\right)\,(1+x^c)^{k+1}}, \tag{10}$$

where $k > \alpha/c$. The WWBXII distribution extends naturally the following distribution: When $k = 1$, equation (5) results in the two-parameter weighted log-logistic distribution. When $c = 1$, equation (5) becomes weighted lomax distribution. When $\alpha = 1$ and $\alpha = 2$, then equation (5) results in the length-biased and area-biased WWBXII distribution respectively. Table 1 lists further possible sub-models of the WWBXII distribution. If

Table 1: sub- models of the WWBXII (α, c, k) distribution

Case	α	c	k	Distribution
[1]	0	c	k	BurrXII (BXII)
[2]	1	c	k	Size-biased (SBXII)
[3]	α	1	k	Two-parameter weighted Lomax (TWLM)
[4]	α	c	1	Two-parameter weighted Log-logistic (TWLL)
[5]	0	1	k	Lomax (LM)
[6]	1	1	k	Size-biased Lomax (SLM)
[7]	0	c	1	Log-logistic (LL)
[8]	1	c	1	Size-biased Log-logistic (SLL)

X denotes an WWBXII random variable, we write $X \sim$ WWBXII(α, c, k). The quantile function of WWBXII can be obtained as a solution of $x_q = \inf\{x : F(x) = q\}$, where F is the cdf given in (7) for a given $q \in (0,1)$. In our case F is strictly increasing function, it follows that $x_q = F^{-1}(q)$. The quantile function can be used to generate samples from WWBXII distribution. Here, we describe an algorithm for generating observations from the WWBXII distribution

- Generate independently uniform numbers $u_i \sim U(0,1), \qquad i = 1, \ldots, n$
- Solve

$$\mathbb{I}_{\frac{x_i^c}{1+x_i^c}}\left(1 + \frac{\alpha}{c}, k - \frac{\alpha}{c}\right) - u_i = 0 \tag{11}$$

Figure 1 displays two simulated data sets. These two simulated data sets show clearly that they are consistent with WWBXII distribution

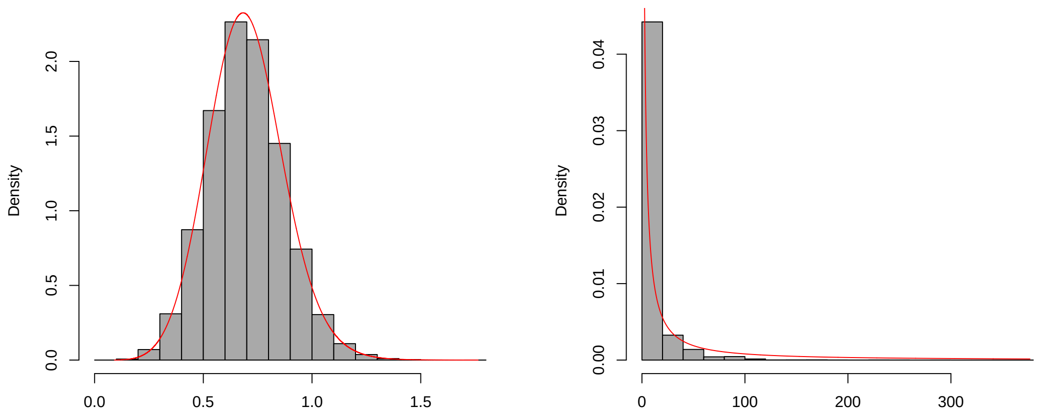

Figure 1: Plots of WWBXII densities for simulated data sets: (left panel) $\alpha = 2$, $c = 4$, $k = 6$ and (right panel) $\alpha = 0.3$, $c = 0.5$, $k = 1.3$

3 Structural properties

In this section, distributional properties of the WWBXII distribution such as shapes of the density function, the shapes of the hazard function, moments, the density of the r^{th} order statistics, and the mean and median deviations are derived and studied in detail.

3.1 Shapes of pdf and hrf

Theorem 3.1. *The pdf of the* WWBXII *distribution is decreasing for* $0 < c + \alpha \leq 1$ *and is unimodal for* $c + \alpha > 1$.

Proof. Set $g(x) = \ln(f_w(x)) = d + (\alpha + c - 1)\ln(x) - (k + 1)\ln(1 + x^c)$, where d is a constant which is free of x. The first derivative of $g(x)$ is $g'(x) = (c + \alpha - 1)x^{-1} - c(k + 1)x^{c-1}(1 + x^c)^{-1}$. If $0 < c + \alpha \leq 1$ then it follows that $g^{'}(x) < 0$, for all $x > 0$. This implies that $f_w(x)$ is a decreasing function. Next, if $c + \alpha > 1$, it then follows that $g'(x) = x^{-1}\omega(x)$, where $\omega(x) = (\alpha + c - 1 - c(k + 1)x^c(1 + x^c)^{-1})$. Now $g'(x) = 0$ if and only if $\omega(x) = 0$. In this case $g^{'}(x)$ has a unique positive solution $x^c = \frac{\alpha+c-1}{(ck-\alpha+1)}$, with $\lim_{x\to 0}\omega(x) = c + \alpha - 1 > 0$ and $\lim_{x\to\infty}\omega(x) = (\alpha - ck - 1) < 0$. Thus $\omega(x)$ is unimodal. Therefore, for $c + \alpha > 1$, there exists $x = x_0$ such that $g(x) > 0$ for $x < x_0$ and $g(x) < 0$ for $x > x_0$. □

Theorem 3.2. *The hazard rate function of the* WWBXII *distribution is decreasing for* $0 < c + \alpha \leq 1$ *and is an upside-down bathtub shaped for* $c + \alpha > 1$*and* $c > 1$.

Proof. Let $\eta_w(x) = -(\partial/\partial x)\ln(f_w(x)) = -(c + \alpha - 1)x^{-1} + c(k + 1)x^{c-1}(1 + x^c)^{-1}$. It follows that $\eta_w^{'}(x) = (c+\alpha-1)x^{-2} + c(c-1)(k+1)x^{c-2}(1+x^c)^{-1} - c^2(k+1)x^{2(c-1)}(1+$

$x^c)^{-2}$. Clearly, $\eta_w^{'}(x) < 0$ for all $x > 0$ whenever $\alpha + c < 1$. So, it follows from a theorem in Glaser (1980) that $h_w(x)$ is a decreasing function. For the second case, rewrite $\eta_w^{'}(x)$ as $\eta_w^{'}(x) = x^{-2}u(x)$, where

$$u(x) = (\alpha + c - 1) + \frac{c(c-1)(k+1)x^c}{(1+x^c)} - \frac{c^2(k+1)x^{2c}}{(1+x^c)^2}.$$

Now suppose that $\alpha + c > 1$ and $c > 1$. The first derivative of $u(x)$ after some algebra lead to

$$u^{'}(x) = \frac{c^2(1+k)x^{c-1}\left[c - 1 - (1+c)x^c\right]}{(1+x^c)^3}.$$

Notice that $u(x)$ is decreasing with a root at that $x_0^c = \left(\frac{c-1}{c+1}\right)$. Therefore $u(x)$ is an upside-down function with $\lim_{x\to 0} u(x) = \alpha + c - 1 > 0$ by the assumption in the the Theorem and $\lim_{x\to\infty} u(x) = \alpha + c - 1 + c(c-1)(k+1) - c^2(k+1) = \alpha - ck - 1 < 0$ by using the fact that $k > \alpha/c$. So $\eta_w(x)$ has a root at $x = t_0$ such that $u(t_0) = 0$, with $\eta_w^{'}(t_0) > 0$ for $x < t_0$ and $\eta_w^{'}(t_0) < 0$ for $x > t_0$. Since $\lim_{x\to 0} f_w(x) = 0$, it follows by Glaser's Theorem (Glaser, 1980) that Equation (9) is upside-down bathtub shaped. □

Figure 2 displays some shapes for the pdf of WWBXII distribution

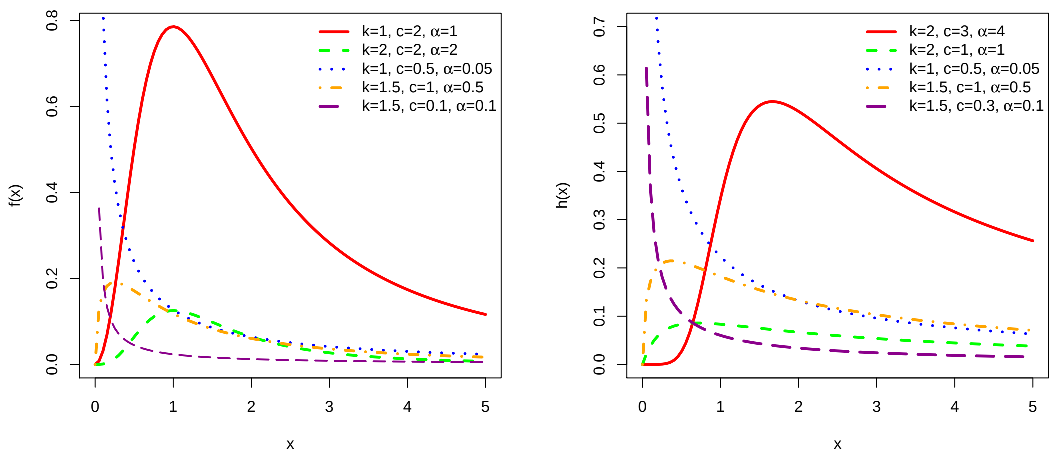

Figure 2: plots of probability density function (left panel) and hazard rate function (right panel) for some selected parameter values.

3.2 Moment generating function

The moment generating function plays an essential role in determining the probability distribution for a given random X. The moments of X can be used to describe most important features and characteristics of a distribution. Let $X \sim \mathsf{WWBXII}(\alpha, c, k)$. the MGF of X is

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_{r=0}^{\infty} (\frac{t^r}{r!}) E(X^r) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\mathbb{B}(k - \frac{r+\alpha}{c}, 1 + \frac{r+\alpha}{c})}{\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c})}. \quad (12)$$

The r^{th} moment can be readily obtained from Equation (12)

$$\mathbb{E}(X^r) = \frac{\mathbb{B}(k - \frac{r+\alpha}{c}, 1 + \frac{r+\alpha}{c})}{\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c})}, \quad k > \frac{r + \alpha}{c}.$$

The first four moments are given below provided that $k > c^{-1}(\alpha + 4)$.

$$\mu = \mathbb{E}(X) = \frac{(\alpha+1)\,\Gamma(k - \frac{\alpha+1}{c})\Gamma(\frac{\alpha+1}{c})}{\alpha\,\Gamma(k - \frac{\alpha}{c})\Gamma(\frac{\alpha}{c})}, \quad \mu_2 = \mathbb{E}(X^2) = \frac{(\alpha+2)\Gamma(k - \frac{\alpha+2}{c})\Gamma(\frac{\alpha+2}{c})}{\alpha\,\Gamma(k - \frac{\alpha}{c})\Gamma(\frac{\alpha}{c})},$$
$$\mu_3 = \mathbb{E}(X^3) = \frac{(\alpha+3)\,\Gamma(k - \frac{\alpha+3}{c})\Gamma(\frac{\alpha+3}{c})}{\alpha\,\Gamma(k - \frac{\alpha}{c})\Gamma(\frac{\alpha}{c})}, \quad \mu_4 = \mathbb{E}(X^4) = \frac{(\alpha+4)\Gamma(k - \frac{\alpha+4}{c})\Gamma(\frac{\alpha+4}{c})}{\alpha\,\Gamma(k - \frac{\alpha}{c})\Gamma(\frac{\alpha}{c})}.$$

By using these moments, the variance, skewness and kurtosis are obtained from well-known relations among these moments,

$$\sigma^2 = \mu_2 - \mu^2, \quad \gamma_3 = \frac{\mu_3 - 3\mu\sigma^2 - \mu^3}{\sigma^2}, \quad \gamma_4 = \frac{\mu_4 - 4\mu\mu_3 + 6\mu^2\mu_2 - 3\mu^4}{\sigma^4}.$$

Figure 3 shows the skewness and kurtosis measures for $\mathsf{WWBXII}(\alpha, c, k)$. Clearly, the two measures namely, the skewness and kurtosis decrease as c and k increase and then start increasing slightly.

3.3 Mean Deviations

The mean deviations about the mean $\mu = \mathbb{E}(X)$ and about the median M are defined respectively by

$$\delta_1(X) = \int_0^{\infty} |x - \mu| f_w(x) dx \qquad \text{and} \qquad \delta_2(X) = \int_0^{\infty} |x - M| f_w(x) dx.$$

We obtain μ from Equation (12) and M is obtained as the solution of the non-linear solution $\mathbb{I}_{\frac{x^c}{1+x^c}}\left(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c}\right) = 0.5$. These measures can be further simplified as $\delta_1(X) = 2\mu F_w(\mu) - 2\mathbb{L}_w(\mu)$ and $\delta_2(X) = \mu - 2\mathbb{L}_w(M)$, where $\mathbb{L}_w(t) = \int_0^t x f_w(x) dx$ and $F_w(x)$ is given in. To compute $\mathbb{L}_w(t)$, let $u = (1 + x^c)^{-1}$ implies $du = -cx^{c-1}(1 + x^c)^{-2}$. Simple algebra leads to

$$\begin{aligned} \int_0^t x f_w(x) dx &= \frac{1}{\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c})} \int_{\frac{1}{1+t^c}}^{1} u^{k - \frac{\alpha+1}{c} - 1} (1 - u)^{\frac{\alpha+1}{c}} du \\ &= 1 - \mathbb{I}_{\frac{1}{1+t^c}}(k - \frac{\alpha+1}{c}, 1 + \frac{\alpha+1}{c}) = \mathbb{I}_{\frac{t^c}{1+t^c}}(1 + \frac{\alpha+1}{c}, k - \frac{\alpha+1}{c}). \end{aligned}$$

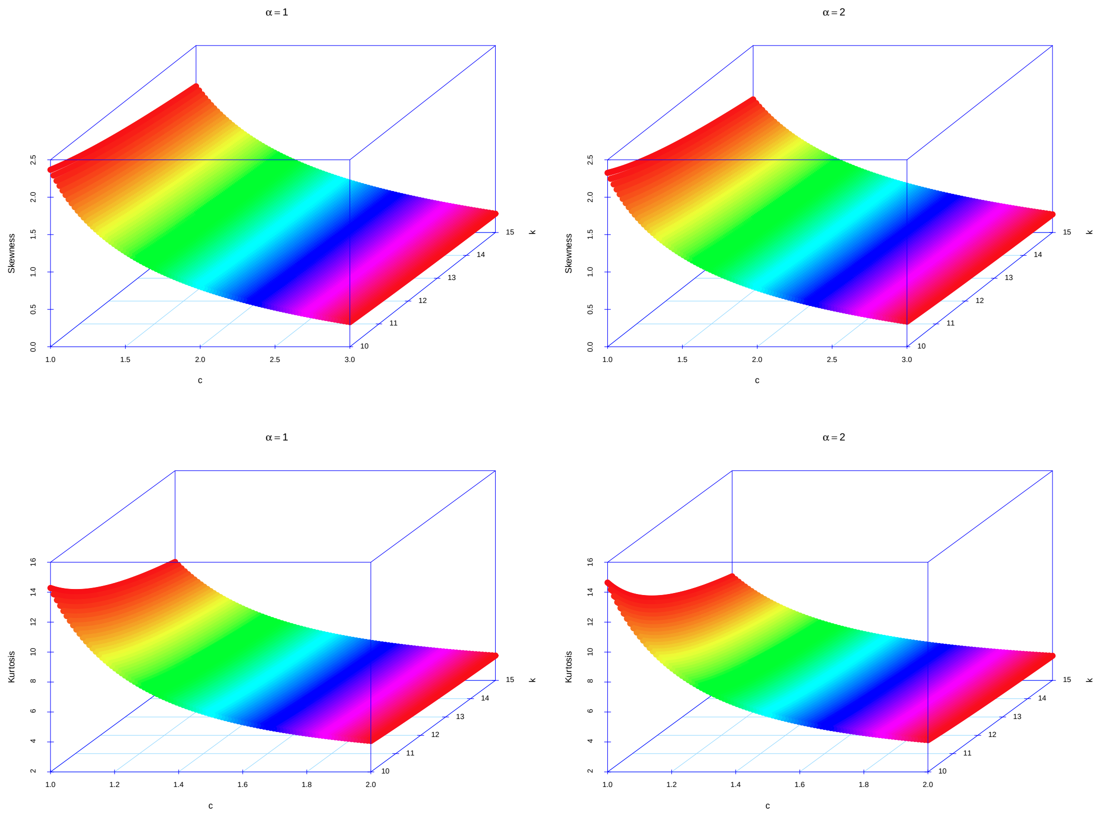

Figure 3: The plots of skewness and kurtosis as a function of c and k when $\alpha = 1, 2$.

Consequently,

$$\delta_1(x) = 2\,\mu \mathbb{I}_{\frac{\mu^c}{1+\mu^c}}(1 + \frac{\alpha}{c}, k - \frac{\alpha}{c}) - 2\mathbb{I}_{\frac{\mu^c}{1+\mu^c}}(1 + \frac{\alpha+1}{c}, k - \frac{\alpha+1}{c}),$$

and

$$\delta_2(x) = \mu - 2\mathbb{I}_{\frac{M^c}{1+M^c}}(1 + \frac{\alpha+1}{c}, k - \frac{\alpha+1}{c}).$$

3.4 Mean Life Residual

The mean residual lifetime is considered as one of the most important measures of the lifetime distribution that is used extensively in reliability, risk, and survival analysis. It is described by the conditional mean of $X - t|X > t$, $t > 0$ and is denoted by $\mu_w(t)$. Let X be a WWBXII random variable, then the mean residual life of X is given by

$$\mu_w(t) = \mathbb{E}(X - t|X > t) = \frac{\mathbb{I}_w(t)}{S_w(t)} - t, \tag{13}$$

where $\mathbb{I}_w(t) = \int_t^\infty x f_w(x)dx$. The computation of $\mathbb{I}_w(t)$ is similar to the computation of $\mathbb{L}_w(t)$. In this case, we have that

$$\mathbb{I}_w(t) = \int_t^\infty x f_w(x)dx = \mathbb{1}_{\frac{1}{1+t^c}}(k - \frac{\alpha+1}{c}, 1 + \frac{\alpha+1}{c}).$$

By substituting the value of the above integral into Equation (13), the mean residual life is obtained and is equal to

$$\mu_w(t) = \frac{\mathbb{1}_{\frac{1}{1+t^c}}(k - \frac{\alpha+1}{c}, 1 + \frac{\alpha+1}{c})}{\mathbb{1}_{\frac{1}{1+t^c}}(k - \frac{\alpha+1}{c}, 1 + \frac{\alpha}{c})} - t$$

By considering the relation between the mean residual life and hazard rate function, we have the following theorem which describes its shape.

Theorem 3.3. *Let $h_w(x)$ be the hazard rate function corresponding to the* WWBXII *distribution, then $\mu_w(x)$ is bathtub shaped with a unique change point $x \in (0, \widehat{t}]$, where $\widehat{t}$ is the change point of the hazard rate function.*

Proof. We have that $\mu_w(0) = \left[\alpha\Gamma(k - \frac{\alpha}{c})\Gamma(\frac{\alpha}{c})\right]^{-1}(\alpha+1)\Gamma(k-\frac{\alpha+1}{c})\Gamma(\frac{\alpha+1}{c}) < \infty$ provided that $k > (1+\alpha)/c$. Since $h_w(0) = 0$ this implies that $\mu_w(0)h_w(0) = 0 < 1$. So it follows by a theorem from Guess et al. (1998) that $\mu_w(x)$ is bathtub shaped with change point $0 < x < \widehat{t}$. □

3.5 Order Statistics

Let $X_1, X_2, \ldots, X_n$ be independent random variables having WWBXII distribution with order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, where $X_{(j)}$ denotes the j^{th} order statistics for $j = 1, \ldots, n$. The (pdf) of $X_{(j)}$ is given as follows .

$$f_{x_{(j)}}(x; c, k, \alpha) = \frac{n!}{(j-1)!(n-j)!} f_w(x)(F_w(x))^{j-1}(s_w(x))^{n-j},$$

By substituting the values of $f_w(x)$, $F_w(x)$, and $S_w(x)$ from Equations (5), (6) and (8) respectively we have that

$$f_{w_{j:n}}(x) = j\binom{n}{j}\frac{c\,x^{\alpha+c-1}\,(1+x^c)^{-(k+1)}}{\mathbb{B}(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c})}\left(\mathbb{1}_{\frac{x^c}{1+x^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right)\right)^{j-1}$$
$$\times\left(\mathbb{1}_{\frac{1}{1+x^c}}\left(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c}\right)\right)^{n-j}.$$

Extreme order statistics namely, the minimum $X_{(1)} = \min(X_1, \ldots, X_n)$ and the maximum $X_{(n)} = \max(X_1, \ldots, X_{(n)})$ have important applications in modeling life of series and parallel system. The asymptotic distributions of these extreme order statistics are established in the following theorem.

Theorem 3.4. *Let $X_1, \dots, X_n$ be a random sample from* WWBXII *distribution. Let $X_{(1)} = \min(X_1, \dots, X_n)$ and $X_{(n)} = \max(X_1, \dots, X_n)$ be the smallest and largest order statistics from this sample. Then*

(i) $\lim_{n\to\infty} P\left(X_{(1)} \le b_n x + a_n\right) = 1 - \exp(-(x^{(\alpha+c)}))$, *where $x > 0$, $a_n = 0$, and b_n can be obtained as a solution of* $\mathbb{I}_{\frac{b_n^c}{1+b_n^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right) = n^{-1}$.

(ii) $\lim_{n\to\infty} P\left(X_{(n)} \le c_n x + d_n\right) = \exp(-x^{-(ck-\alpha)})$ *where $x > 0$, $d_n = 0$ and c_n can be obtained as a solution of* $\mathbb{I}_{\frac{1}{1+b_n^c}}\left(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c}\right) = n^{-1}$.

Proof. The asymptotical result for $X_{(1)}$ reported in Arnold et al. (1992) indicates that a necessary and sufficient conditions of the limiting distribution of $X_{(1)}$, say $F(.)$ to belong to the min domain of attraction of the Weibull type, i.e., $\lim_{n\to\infty} P(X_{(1)} \le b_n x + a_n) = 1 - \exp(-x^{\beta}), \quad x > 0, \beta > 0$ where $b_n = F^{-1}(\frac{1}{n}) - F^{-1} - F^{-1}(0)$ if and only if $F^{-1}(0) < \infty$ and for every $x > 0$ there exists $\beta > 0$ such that

$$\lim_{\epsilon\to 0} \frac{F(F^{-1}(0) + x\epsilon)}{F(F^{-1}(0) + \epsilon)} = x^{\beta}.$$

For our case, we have that $F_w^{-1}(0) = 0 < \infty$ and moreover

$$\lim_{\epsilon\to 0} \frac{F(x\epsilon)}{F(\epsilon)} = \lim_{\epsilon\to 0} \frac{\mathbb{I}_{\frac{(\epsilon x)^c}{1+(\epsilon x)^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right)}{\mathbb{I}_{\frac{\epsilon^c}{1+\epsilon^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right)} \doteq x^{\alpha+c} \lim_{\epsilon\to 0} \left(\frac{1+\epsilon^c}{1+(\epsilon x)^c}\right)^{k+1} = x^{\alpha+c},$$

where $\doteq$ means that L' Hopitals rule is used. It then follows from Theorem (8.3.6) of Arnold et al. (1992) that there exists normalizing constants $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ and $(b_n)_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ such that

$$\lim_{n\to\infty} P(X_{(1)} \le b_n x + a_n) = 1 - \exp(-x^{(\alpha+c)}),$$

where these normalized constants can be determined by choosing $a_n = 0$ and $b_n = F^{-1}(\frac{1}{n}) - F^{-1} - F^{-1}(0)$. So b_n can be obtained as a solution to

$$\mathbb{I}_{\frac{b_n^c}{1+b_n^c}}\left(1+\frac{\alpha}{c}, k-\frac{\alpha}{c}\right) = n^{-1}.$$

Similarly, the asymptotical distribution for the largest order statistics belongs to the max domain of attraction of Fréchet type, i.e., $\lim_{n\to\infty} P(X_{(n)} \le c_n x + d_n) = 1 - \exp(-t^{\delta}), \quad x > 0, \delta > 0$ where $d_n = 0$ and $c_n = F^{-1}(1 - \frac{1}{n})$ if and only if $F^{-1}(1) = \infty$ and for every $t > 0$ there exits $\delta > 0$,

$$\lim_{t\to\infty} \frac{1 - F(xt)}{1 - F(t)} = x^{-\delta}.$$

For the second part (ii),

$$\lim_{t\to\infty} \frac{S_w(xt)}{S_w(t)} = \lim_{t\to\infty} \frac{\mathbb{I}_{\frac{1}{1+(xt)^c}}\left(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c}\right)}{\mathbb{I}_{\frac{1}{1+t^c}}\left(k-\frac{\alpha}{c}, 1+\frac{\alpha}{c}\right)} \doteq x^{c} \lim_{t\to\infty} \left(\frac{1+t^c}{1+(xt)^c}\right)^{k-\frac{\alpha}{c}+2} = x^{-(ck-\alpha)-c},$$

where $\doteq$ means that L' Hopitals rule is gain used. So it follows by Theorem (8.3.5) of Arnold et al. (1992) that there exists deterministic constants $(d_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ and $(c_n)_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ such that

$$\lim_{n\to\infty} P(X_{(n)} \le c_n x + d_n) = \exp(-x^{-(ck-\alpha)-c}),$$

where these normalized constants are given by $d_n = F_w^{-1}(0) = 0$ and c_n can be obtained as a solution to $\mathbb{I}_{\frac{1}{1+b_n^c}}\left(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c}\right) = n^{-1}$. □

3.6 Entropies

The variation of uncertainty of a random variable X can be measured by the notion of an entropy. Two well known entropy measures are: Renyi entropy and Shannon entropy. Suppose that $X \sim \mathsf{WWBXII}(\alpha, c, k)$, the Renyi entropy is defined as

$$I_{R(r)}(x) = \frac{1}{1-r} \ln\left(\int_0^\infty f_w^r(x)dx\right), \tag{14}$$

where $r > 0$ and $r \neq 1$.

Theorem 3.5. *Let X be an* WWBXII *random variable. The Renyi Entropy is*

$$I_{R(r)}(x) = -\log(c) + \frac{r}{1-r}\log[\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c})] + \frac{1}{1-r}\log[\mathbb{B}(rk - \frac{r\alpha - r + 1}{c}, r + \frac{r\alpha - r + 1}{c})]$$

Proof. We have that

$$\int_0^\infty f_w^r(x)dx = \frac{c^r}{[\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c})]^r} \int_0^\infty x^{r\alpha + cr - r}\left(1 + x^c\right)^{-r(k+1)} dx.$$

On substituting $u = \frac{1}{1+x^c}$, it then follows

$$\begin{aligned}\int_0^\infty [f_w(x)]^r dx &= \frac{c^{r-1}}{(\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c}))^r} \int_0^1 (1-u)^{\frac{r\alpha - r + 1}{c} + r - 1} u^{rk - 1 - \frac{r\alpha - r + 1}{c}} du \\ &= \frac{c^{r-1}\mathbb{B}(rk - \frac{r\alpha - r + 1}{c}, r + \frac{r\alpha - r + 1}{c})}{(\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c}))^r}.\end{aligned}$$

On substituting the above quantity into Equation (14), we get the desired result. □

The Shannon entropy can is a special case of Renyi entropy and is defined as $I_S(x) = \lim_{r\to 1} I_{R(r)}(x) = \mathbb{E}\left[-\ln(f(X))\right]$. The value of the Shannon entropy is given in the following theorem

Theorem 3.6. *Let X be an* WWBXII *random variable, the Shannon Entropy is*

$$\begin{aligned}I_S(x) = &\ln(\mathbb{B}(k - \frac{\alpha}{c}, 1 + \frac{\alpha}{c})) - \ln(c) - \frac{(\alpha + c - 1)}{c}\left[\psi(1 + \frac{\alpha}{c}) - \psi(k - \frac{\alpha}{c})\right] \\ &+ (k+1)\left[\psi(k+1) - \psi(k - \frac{\alpha}{c})\right].\end{aligned}$$

Proof. We have that

$$\mathbb{E}[-\ln f_w(X)] = \mathbb{E}\left[-\ln\left(\frac{c}{\beta(k-\frac{\alpha}{c},1+\frac{\alpha}{c})}\,x^{\alpha+c-1}\,(1+x^c)^{-(k+1)}\right)\right]$$
$$= \ln\left(\frac{\mathbb{B}(k-\frac{\alpha}{c},1+\frac{\alpha}{c})}{c}\right) - (\alpha+c-1)\,\mathbb{E}[\ln(X)] + (k+1)\mathbb{E}[\ln(1+X^c)]. \tag{15}$$

Now consider the computation of $\mathbb{E}[\ln(X)]$. Observe that

$$E[\ln(X)] = \frac{c}{\mathbb{B}(k-\frac{\alpha}{c},1+\frac{\alpha}{c})}\int_0^\infty \ln(x)\,x^{\alpha+c-1}\,(1+x^c)^{-(k+1)}\,dx.$$

Set $u=(1+x^c)^{-1}$ implies $du = -c\,x^{c-1}(1+x^c)^{-2}dx$. It then follows that

$$\begin{aligned}
E[\ln(X)] &= \frac{1}{c\,\mathbb{B}(k-\frac{\alpha}{c},1+\frac{\alpha}{c})}\int_0^1 \ln\left(\frac{1-u}{u}\right)u^{k-1-\frac{\alpha}{c}}\,(1-u)^{\frac{\alpha}{c}}\,du\\
&= \frac{1}{c\,\mathbb{B}(k-\frac{\alpha}{c},1+\frac{\alpha}{c})}\left[\int_0^1 \ln(1-u)\,u^{k-1-\frac{\alpha}{c}}\,(1-u)^{\frac{\alpha}{c}}\right.\\
&\qquad \left.-\int_0^1 \ln(u)\,u^{k-1-\frac{\alpha}{c}}\,(1-u)^{\frac{\alpha}{c}}\,du\right]\\
&= \frac{1}{c}[\psi(1+\frac{\alpha}{c}) - \psi(k+1) - \psi(k-\frac{\alpha}{c}) + \psi(k+1)].
\end{aligned}$$

It can be seen that the integrals in the above equation can be computed by differentiating the beta function. Specifically, the beta function $\mathbb{B}(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}dt$. The first partial derivative of $\mathbb{B}(a,b)$ with respect to a and b respectively are :

$$(\partial/\partial a)\mathbb{B}(a,b) = \int_0^1 \ln(t)t^{a-1}(1-t)^{b-1}dt = \mathbb{B}(a,b)[\psi(a)-\psi(a+b)], \tag{16}$$

and

$$(\partial/\partial b)\mathbb{B}(a,b) = \int_0^1 \ln(1-t)t^{a-1}(1-t)^{b-1}dt = \mathbb{B}(a,b)[\psi(b)-\psi(a+b)], \tag{17}$$

where $\psi(a)=(\partial/\partial a)\ln(\Gamma(a))$. By using equations (16) and (17) we have

$$\mathbb{E}[\ln(X)] = \frac{\psi(1+\frac{\alpha}{c}) - \psi(k-\frac{\alpha}{c})}{c}. \tag{18}$$

The computation of $\mathbb{E}[\ln(1+x^c)]$ can be treated similarly and hence

$$\mathbb{E}[\ln(1+x^c)] = \psi(k+1) - \psi(k-\frac{\alpha}{c}). \tag{19}$$

By substituting (18) and (19) into Equation in (15), the result follows. □

3.7 Stochastic ordering

Stochastic ordering of distribution functions plays an essential role in the statistical methodology. The likelihood ratio ordering is considered as one of the most important ordering among various stochastic orderings. Let X and Y be two random variables with distributions F and G and corresponding probability density functions f and g. We say that X is smaller than Y in (1) likelihood ratio order ($X \prec_{lr} Y$) if $g(x)/f(x)$ is increasing in x;(2) hazard rate order ($X \prec_{hr} Y$) if $(1-G(x))/(1-F(x))$ is increasing in x; (3) standard stochastic order ($X \prec_{st} Y$) if $G(x) > F(x)$; (4) mean residual life order ($X \prec_{mrl} Y$) if $\mathbb{E}(x-t|x>t) \leq \mathbb{E}(Y-t|x>t.)$

The relation among these stochastic orders implies that $X \prec_{lr} Y \Rightarrow X \prec_{hs} Y \Rightarrow X \prec_{st} Y$. The reader is referred to Shakeda and Shanthikumar (2007) for further information and results. In this section, we are interested in comparing $\mathsf{BXII}(\mathsf{c},\mathsf{k})$ and $\mathsf{WWBXII}(\alpha,\mathsf{c},\mathsf{k})$

Theorem 3.7. *Let* $X \sim \mathsf{BXII}(\mathsf{c},\mathsf{k})$, $Y \sim \mathsf{WWBXII}(\alpha_1,\mathsf{c},\mathsf{k})$, *and* $Z \sim \mathsf{WWBXII}(\alpha_2,\mathsf{c},\mathsf{k})$, *then it follows that (1)* $X \prec_{lr} Y$. *and (2)* $Y \prec_{lr} Z$. *provided that* $\alpha_1 < \alpha_2$.

The proof of the theorem is straightforward. This theorem tells us that WWBXII distribution has thinner tail than BXII distribution and if $\alpha_1 < \alpha_2$ then $\mathsf{WWBXII}(\alpha_2,\mathsf{c},\mathsf{k})$ has thinner tail than $\mathsf{WWBXII}(\alpha_1,\mathsf{c},\mathsf{k})$.As it has the likelihood ratio ordering, this implies there exists a uniformly power test (UMP) for one-sided or two-sided hypothesis on the shape parameter α, when the other parameters are known.

4 Model Estimation

In this section, we discuss the estimation of the model parameters using the maximum likelihood method.

4.1 Maximum Likelihood Estimate

Let $X_1 \ldots, X_n$ be independent and identically random sample from WWBXII distribution with the vector parameters $\boldsymbol{\theta} = (\alpha, c, k)^T$, where $\boldsymbol{\Omega} = \{(\alpha, c, k)^T : \alpha, c, k > 0, \quad k > \frac{\alpha}{c}\}$ is the parameter space. The log-likelihood function is given by

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}) &= n\ln(k) + 2n\ln(c) + n\ln(\Gamma(k)) - n\ln(\alpha) - n\ln(\Gamma(k-\frac{\alpha}{c})) \\
&\quad -n\ln(\Gamma(\frac{\alpha}{c})) + (\alpha+c-1)\sum_{i=1}^{n}\ln(x_i) - (k+1)\sum_{i=1}^{n}\ln(1+x_i^c).
\end{aligned}
$$

Then we take the partial derivative of the log-likelihood function with respect to the parameters give the following non-linear equations.

$$\frac{\partial \mathcal{L}}{\partial \alpha} = \frac{-n}{\alpha} + \frac{n}{c}\,\psi(k - \frac{\alpha}{c}) - \frac{n}{c}\,\psi(\frac{\alpha}{c}) + \sum_{i=1}^{n} \ln(x_i) \tag{20}$$

$$\frac{\partial \mathcal{L}}{\partial k} = \frac{n}{k} + n\,\psi(k) - n\,\psi(k - \frac{\alpha}{c}) - \sum_{i=1}^{n} \ln(1 + x_i^c) \tag{21}$$

$$\frac{\partial \mathcal{L}}{\partial c} = \frac{2n}{c} - \frac{\alpha n}{c^2}\,\psi(k - \frac{\alpha}{c}) + \frac{\alpha n}{c^2}\,\psi(\frac{\alpha}{c}) + \sum_{i=1}^{n} \ln(x_i) - (k+1)\sum_{i=1}^{n} x_i^c\,(1 + x_i^c)^{-1}\,\ln(x_i), \tag{22}$$

where $\psi(a) = (\partial/\partial a)\ln(\Gamma(a))$. The maximum likelihood estimates of the parameters c, k, α is the solution to the above equations. Due to non explicit solution of these estimating equations (20), (21) and (22), the MLEs can be obtained numerically by using for example Newton-Raphson procedure in `R` software. However, the MLEs can be obtained efficiently by implementing the functions `optim` or `nlm` in the statistical software `R` (Team, 2013).

4.2 Existence and Uniqueness of MLEs

Theorem 4.1. *Let $h(\alpha; c, k, x^n)$ denote the function on the right-hand-side (RHS) of the equation* (20), *where $\alpha \in (0, ck)$.Then there exists a unique solution for $h(k; c, \alpha, x^n)$ provided that $k > 1$ and $\sum_{i=1}^{n} \ln(x_i) > 0$.*

Proof. Since $-\gamma = \lim_{x\to 0}(\psi(x) + x^{-1})$ and $\psi(x) = -\gamma + \sum_{\ell=0}^{\infty} \frac{x-1}{(\ell+1)(\ell+x)}$, where γ is the Euler constant, it follows that

$$\begin{aligned}
\lim_{x\to 0} h(\alpha; c, k, x^n) &= \lim_{x\to 0}\left(\frac{-n}{\alpha} + \frac{n}{c}\,\psi(k - \frac{\alpha}{c}) - \frac{n}{c}\,\psi(\frac{\alpha}{c}) + \sum_{i=1}^{n} \ln(x_i)\right) \\
&= \frac{n}{c}\lim_{x\to 0}\left(\psi(\frac{\alpha}{c}) + \frac{c}{\alpha}\right) + \frac{n}{c}\psi(k) + \sum_{i=1}^{n} \ln(x_i) \\
&= \frac{n\gamma}{c} - \frac{n\gamma}{c} + \sum_{\ell=0}^{\infty} \frac{k-1}{(\ell+1)(\ell+x)} + \sum_{i=1}^{n} \ln(x_i).
\end{aligned}$$

So, it follows that $\lim_{x\to 0} h(\alpha; c, k, x^n) > 0$ if and only if $k > 1$ and $\sum_{i=1}^{n} \ln(x_i) > 0$. Similarly, $\lim_{\alpha\to ck} h(\alpha; c, k, x^n) = \lim_{\alpha\to ck}\left(\frac{-n}{\alpha} + \frac{n}{c}\psi(k - \frac{\alpha}{c}) - \frac{n}{c}\psi(\frac{\alpha}{c}) + \sum_{i=1}^{n} \ln(x_i)\right) = -\infty$. Hence, there exits a root for $h(\alpha; c, k, x^n)$ in the interval $\alpha \in (0, ck)$. □

Theorem 4.2. *Let $g(k; c, \alpha, x^n)$ denote the function on the right-hand-side (RHS) of the equation* (21), *where $k \in (\frac{\alpha}{c}, \infty)$. Then there exists a root for $g(k; c, \alpha, x^n)$. The root is unique if and only if $k > 2\frac{\alpha}{c}$.*

Proof. Since $\psi(y) \approx \ln(y) - (2y)^{-1}$ for large y then equation (21), and after some manipulation, reduces

$$g(c; k, \alpha, x^n) = \frac{n}{2k} + \frac{nc}{2(ck-\alpha)} + n\ln\left(1 + \frac{\alpha}{(ck-\alpha)}\right) - \sum_{i=1}^{n} \ln(1 + x_i^c).$$

Now, $\lim_{k\to\infty} g(c; k, \alpha, x^n) = -\sum_{i=1}^{n} \ln(1 + x_i^c) < 0$. On the other hand and since that

$$\lim_{k\to\frac{\alpha}{c}} \left[\frac{nc}{2(ck-\alpha)} + n\ln\left(1 + \frac{\alpha}{ck-\alpha}\right)\right] = \infty,$$

and hence $\lim_{c\to\frac{\alpha}{c}} g(k; c, \alpha, x^n) = \infty$. This implies that there exits a root for $g(k; c, \alpha, x^n)$ in $k \in (\frac{\alpha}{c}, \infty)$. To prove the uniqueness of the root, consider the first derivative of $g(k; c, \alpha, x^n)$,

$$g^{'}(k; c, \alpha, x^n) = -\frac{n}{2k^2} - \frac{nc^2}{(ck-\alpha)^2} - \frac{nk\alpha}{(ck-\alpha)(ck-2\alpha)} < 0,$$

if and only if $k > 2\frac{\alpha}{c}$. □

Theorem 4.3. *Let $w(c; k, \alpha, x^n)$ denote the function on the right-hand-side (RHS) of the equation (22), where $c \in (\frac{\alpha}{k}, \infty)$. Then there exists a root for $w(c; k, \alpha, x^n)$ provided that the sample data $x^n = (x_1, \ldots, x_n)^T$ contains at least one measurement less than unity, i.e., $(x_i < 1)$ for some $i \in \{1, \ldots, n\}$.*

Proof. Similar to the proof of theorem 4.2 and by using $\psi(y) \approx \ln(y) - (2y)^{-1}$ for large y then equation (22) after some manipulation reduces to

$$w(c; k, \alpha, x^n) = \frac{2n}{c} - \frac{n\alpha}{c^2}\ln\left(\frac{ck-\alpha}{\alpha}\right) - \frac{n}{2}\frac{ck-2\alpha}{c(ck-\alpha)} + \sum_{i=1}^{n}\ln(x_i) - (k+1)\sum_{i=1}^{n} x_i^c\,(1+x_i^c)^{-1}\ln(x_i).$$

Now, let $v(c; k, \alpha, x^n) = \frac{2n}{c} - \frac{n\alpha}{c^2}\ln\left(\frac{ck-\alpha}{\alpha}\right) - \frac{n}{2}\frac{ck-2\alpha}{c(ck-\alpha)}$. Since $\lim_{c\to\infty} v(c; k, \alpha, x^n) = 0$, it follows that

$$\begin{aligned}
\lim_{c\to\infty} w(c; k, \alpha, x^n) &= \lim_{c\to\infty} v(c; k, \alpha, x^n) + \sum_{i=1}^{n}\ln(x_i) - (k+1)\lim_{c\to\infty}\sum_{i=1}^{n}\frac{x_i^c \ln(x_i)}{(1+x_i^c)} \\
&= \sum_{i=1}^{n}\ln(x_i) - (k+1)\lim_{c\to\infty}\sum_{i=1}^{n}\frac{x_i^c \ln(x_i)}{(1+x_i^c)} \\
&= \sum_{i=1}^{n}\ln(x_i) - (k+1)\lim_{c\to\infty}\sum_{\{i:\, x_i<1\}}\frac{x_i^c \ln(x_i)}{(1+x_i^c)} - (k+1)\lim_{c\to\infty}\sum_{\{i:\, x_i\geq 1\}}\frac{x_i^c \ln(x_i)}{(1+x_i^c)} \\
&= \sum_{i=1}^{n}\ln(x_i) - (k+1)\sum_{\{i:\, x_i\geq 1\}}^{n}\ln(x_i) = \sum_{\{k:x_k<1\}}\ln(x_k) - k\sum_{\{k:\, x_k>1\}}\ln(x_k) < 0.
\end{aligned}$$

On the other hand and because

$$-\lim_{c\to\frac{\alpha}{k}}\left[\frac{n\alpha}{c^2}\ln\left(\frac{ck-\alpha}{\alpha}\right)-\frac{n}{2}\frac{ck-2\alpha}{c(ck-\alpha)}\right]=\infty,$$

we have that $\lim_{c\to\frac{\alpha}{k}} w(c;k,\alpha,x^n)=\infty$. This implies that there exits a root for $w(c;k,\alpha,x^n)$ in $c\in(\frac{\alpha}{k},\infty)$. □

The asymptotic distribution of the MLEs is approximately multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and approximate variance-covariance matrix $\mathbf{J^{-1}}(\widehat{\boldsymbol{\theta}})$, where $\mathbf{J}(\widehat{\boldsymbol{\theta}})=-\frac{\partial^2\mathcal{L}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$ is the observed information matrix. Its elements are given in Appendix. In this case, we can obtain an approximate confidence interval to the parameters of our model. An $100\%(1-\delta)$ asymptotic confidence interval for each parameter $\boldsymbol{\theta}_\ell$ is given by

$$\widehat{\alpha}\pm z_{\delta/2}\sqrt{\mathbf{J}^{(\widehat{\alpha},\widehat{\alpha})}},\qquad \widehat{c}\pm z_{\delta/2}\sqrt{\mathbf{J}^{(\widehat{c},\widehat{c})}}\qquad \widehat{k}\pm z_{\delta/2}\sqrt{\mathbf{J}^{(\widehat{k},\widehat{k})}}$$

where $\mathbf{J}^{(\ell,\ell)}$ is the (ℓ,ℓ) diagonal element of $\mathbf{J^{-1}}(\widehat{\boldsymbol{\theta}})$, for $\ell=\widehat{\alpha},\widehat{c},\widehat{k}$ and $z_{\delta/2}$ is the quantile $1-\delta/2$ of the standard normal distribution.

4.3 Testing of existence of size-bias in the sample

Of major importance is to check the existence of size-bias in the sample. This is equivalent to test the null hypothesis $\mathsf{H}_0:\alpha=0$ against the alternative hypothesis $\mathsf{H}_a:\alpha>0$. In this case we can compute the maximized restricted log-likelihood $\mathcal{L}_r(0,c,k)$ and the full model (unrestricted log-likelihood) $\mathcal{L}_f(\alpha,c,k)$ in order to construct the likelihood ratio test statistics $\Lambda(x)=-2\left(\mathcal{L}_f(\alpha,c,k)-\mathcal{L}_r(0,c,k)\right)$. In this context, the standard asymptotic distribution of $\Lambda(x)$ does not work, i.e., $\Lambda(x)$ does not converge to χ^2_1 as $n\to\infty$ because the value of the parameter α under H_0 lies in the boundary of the parameter space of $\boldsymbol{\alpha}=(0,\infty)$, i.e., $0\in\partial\boldsymbol{\alpha}$. In this regard, we use Theorem 3 of Self and Liang (1987) which implies that $\Lambda(x)$ converges in distribution to

$$\frac{1}{2}+\frac{1}{2}\chi^2_1.$$

5 Simulation and an application

We employ a simulation study to assess the performance of the precision of the point estimates to the parameters of $\mathsf{WBXII}(\alpha,c,k)$ and the coverage probability of the confidence intervals for these parameters as well. The assessment procedure is based on the root mean squared error (MSEs) for each parameter and the average length of confidence intervals (mLCs). The simulation study was carried out using (1000) samples of size 50, 80, 100, 150, and 250 from $\mathsf{WWBXII}(\alpha,\ c\,,k)$ distribution with different choices of the parameter $(\alpha,c,\ k)$. For each generated data, the maximum likelihood is used to provide point estimates of the parameters and the mean of the MLEs (mMLEs) along with the root mean squared errors are calculated and reported in Table 2. Besides to the point estimates, approximate 95% confidence intervals of these parameters are calculated using observed Fisher information matrix and the mean length of confidence intervals (mLCs) along with their corresponding coverage probability (see,Table 2). Simulation outcomes indicate that the absolute differences between the average point estimates of the parameters and the true values decay to 0 as the sample size increases. The root mean squared errors decrease as the sample size increases which confirm that the MLEs are behaved consistency. As expected, the average lengths of the mLCs decrease as the sample sizes get grow. Additionally, the coverage probabilities for these confidence intervals are close to 95%, as the sample sizes increase. Generally speaking, the average length of the 95% confidence interval for the parameter c is shorter than that for the rest of the parameters, but the coverage probabilities of the parameters are similar.

Table 2: Simulation outcomes: The mean of MLEs estimates(mMLEs) and the root mean squared errors(msd)(within parentheses) of the WWBXII distribution along with the mean length of 95% confidence intervals (mLCs) and their corresponding coverage probability (within parentheses)

True values	n	mMLEs (msd)			mLCs (95% C.I.)		
		$\widehat{\alpha}$	$\hat{c}$	$\hat{k}$	C.I($\widehat{\alpha}$)	C.I.($\hat{c}$)	C.I($\hat{k}$)
(2,3,3)	50	2.2439	3.0736	3.1889	1.867	1.028	1.505
		(0.4763)	(0.2622)	(0.3840)	(0.927)	(0.942)	(0.945)
	80	2.2027	3.0376	3.1596	1.692	0.761	1.241
		(0.4316)	(0.1942)	(0.3166)	(0.949)	(0.949)	(0.941)
	100	2.2147	3.0207	3.1724	1.646	0.666	1.245
		(0.4195)	(0.1699)	(0.3176)	(0.937)	(0.952)	(0.951)
	150	2.1604	3.0013	3.1339	1.461	0.516	1.105
		(0.3727)	(0.1317)	(0.2820)	(0.946)	(0.951)	(0.942)
	250	2.1194	2.9964	3.1036	1.227	0.452	0.905
		(0.3131)	90.1153)	(0.2308)	(0.951)	(0.957)	(0.945)
(1,3,3)	50	0.9836	3.0972	3.0494	1.346	1.583	1.610
		(0.3433)	(0.4040)	(0.4107)	(0.931)	(0.936)	(0.945)
	80	0.9634	3.0959	2.9927	1.318	1.297	1.392
		(0.3362)	(0.3308)	(0.3552)	(0.932)	(0.935)	(0.955)
	100	0.9567	3.0532	3.0106	1.296	1.237	1.312
		(0.3306)	(0.3157)	(0.3347)	(0.936)	(0.951)	(0.952)
	150	0.9649	3.0429	3.0051	1.221	1.068	1.217
		(0.3114)	(0.2723)	(0.3104)	(0.941)	(0.945)	(0.954)
	250	0.9510	3.0481	2.9758	1.121	0.877	1.095
		(0.2861)	(0.2238)	(0.2794)	(0.951)	(0.950)	(0.951)

Intial values	n	mMLEs (msd)			mLCs (95% C.I.)		
		$\widehat{\alpha}$	$\widehat{c}$	$\widehat{k}$	C.I($\widehat{\alpha}$)	C.I.($\widehat{\alpha}$)	C.I($\widehat{\alpha}$)
	50	2.0832	4.0047	5.0751	1.826	1.601	1.580
		(0.4657)	(0.2622)	(0.3840)	(0.934)	(0.964)	(0.955)
	80	2.0850	4.0086	5.1021	1.708	1.431	1.46
		(0.4356)	(0.3651)	(0.3723)	(0.934)	(0.932)	(0.951)
(2,4,5)	100	1.9994	4.0411	5.0127	1.357	1.232	1.358
		(0.3953)	(0.3146)	(0.3463)	(0.954)	(0.951)	(0.947)
	150	1.9876	4.0354	4.9910	1.501	1.129	1.324
		(0.3829)	(0.2881)	(0.3378)	(0.958)	(0.952)	(0.949)
	250	1.9839	4.0386	5.0038	1.364	0.993	1.153
		(0.3479)	(0.2534)	(0.2940)	(0.963)	(0.947)	(0.967)
	50	5.0890	5.085800	10.2177	2.320	2.516	3.086
		(0.5918)	(0.6420)	(0.7874)	(0.951)	(0.946)	(0.954)
	80	5.1173	5.0449	10.2338	1.7077	1.4313	1.4593
		(0.5798)	(0.5268)	(0.7253)	(0.951)	(0.954)	(0.949)
(5,5,10)	100	5.1350	5.0025	10.2004	2.265	1.910	2.804
		(0.5777)	(0.4872)	(0.7152)	(0.948)	(0.951)	(0.953)
	150	5.1318	5.0014	10.2101	2.194	1.645	2.526
		(0.5597)	(0.4198)	(0.6444)	(0.952)	(0.951)	(0.950)
	250	5.1727	4.9730	10.1893	2.158	1.328	2.255
		(0.5506)	(0.3389)	(0.5751)	(0.960)	(0.956)	(0.957)

6 Data analysis

In this section, we illustrate the adequacy and flexibility of the WWBXII(α, c, k) distribution in fitting real data set. We consider a data set from a study carried out by Bjerkedal (1960). The data set comprise survival times in days of 72 Guinea pigs injected with various amount of tubercle. From the study, we consider only the data in which pigs in a single cage are under the same regimen. The data are given below:
12, 15, 22, 24, 24, 32, 32, 33, 34, 38 ,38, 43, 44, 48, 52, 53, 54, 54, 55, 56, 57, 58, 58, 59, 60, 60, 60,60, 61, 62, 63,65, 65, 67, 68, 70, 70, 72, 73, 75, 76, 76, 81, 83, 84, 85, 87, 91, 95, 96, 98, 99, 109, 110, 121, 127, 129, 131, 143, 146,146, 175 ,175 ,211 ,233 ,258 ,258 ,263 ,297 ,341, 341, 376. Table 3 gives some descriptive summary statistics to these data which have positive skewness and kurtosis but have large standard deviation. This indicates that the data are positively skewed, and hence the considered distribution could be used to model this data.

In many applications, empirical information about the hazard rate shape can be useful in selecting a particular distribution. In this regard, a tool called the total time on test (TTT) plot can be used in identifying the suitable model, see for detail Aarset (1987). The TTT-plot for WWBXII for our data is displayed in Figure 4 which indicates an upside-down bathtub hazard rate function (unimodal). This indicates the appropriateness of the WWBXII distribution to fit the current data.

Table 3: Descriptive Statistics of the Guinea pigs data

Mean	Standard deviation	Skewness	Kurtosis	Minimum	Maximum
99.82	81.12	1.76	2.46	12	376

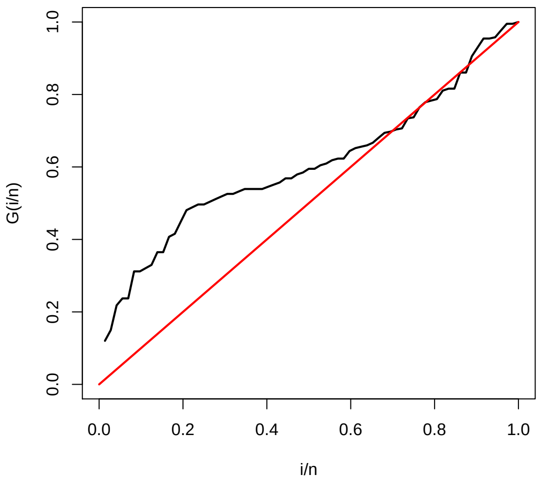

Figure 4: TTT-plot the Guinea pigs data

Next, we fit the WWBXII model and eight models as special sub-cases using MLEs (standard errors) along with the computation of the maximum of the log-likelihood function($\mathcal{L}$), the values of Akaike information criterion (AIC), and the Bayesian information criterion (BIC) statistics. Table 4 lists the MLEs of the parameters of WWBXII along with the estimates of its sub-distributions and compare among them using the statistics $\mathcal{L}$, AIC, and BIC. The numerical values of the these statistics correspond to the WWBXII are the lowest among those fitted sub-models and therefore our model is the best model.

Table 4: MLEs of the model parameters along with their corresponding Standard error (given in parentheses) for Guinea pigs

Model	MLEs of the parameters			Measures		
	$\widehat{\alpha}$	$\widehat{c}$	$\widehat{k}$	$-\mathcal{L}$	AIC($\widehat{\boldsymbol{\theta}}$)	BIC($\widehat{\boldsymbol{\theta}}$)
WWBXII(α, c, k)	31.1736	0.3532	107.8556	390.25	786.5	793.33
	(6.4808)	(0.0846)	(26.7517)			
TWLM(α, k)	118.0781	–	120.0918	392.60	789.2	793.7533
	(20.8566)	(–)	(21.1319)			
TWLL(α, c)	0.7607	1.0183	–	479.97	963.94	968.4933
	(0.2019)	(0.2001)	(–)			
BXII(c, k)	–	3.0486	0.0753	490.55	985.1	989.65
	(–)	(4.1154)	(0.1017)			
SBXII(c, k)	–	0.4205	3.34601	467.71	939.42	943.97
	(–)	(0.0056)	(0.4833)			
LL(c)	–	0.3531	–	526.97	1055.94	1058.22
	(–)	(0.0322)	(–)			
SBLL(c)	–	1.2541	–	480.45	962.9	965.18
	(–)	(0.0285)	(–)			
LM(k)	–	–	0.2294	492	986	988.28
	(–)	(–)	(0.0270)			
SLM(k)	–	–	1.2795	476.95	955.9	958.18
	(–)	(–)	(0.0322)			

We also apply formal non-parametric goodness-of-fit tests such as Kolomogorov-Sminrov (K-S), Cramer-von-Mises (W_n^*), and Anderson-Darling (A_n^*) statistics in order check which distribution fits better these data. For further information about W_n^* and A_n^* statistics, see Chen and Balakrishnan (1995). Usually, the smaller the values of these statistics, the better the fit to the data. These statistics are computed for the current data and given in Table 5. Based on the values of these non-parametric goodness-of-fit tests reported in this table, the WWBXII distribution outperforms its special sub-distributions and therefore fits much better these data. In particular, the K-S distance between the empirical and fitted WWBXII distribution functions is 0.101 with the corresponding p-value is 0.4581 which indicates that the data follow the WWBXII distribution is strongly accepted whereas the p-values correspond to the K-S distances for the other sub-models are less than 0.0001 expect of the sub-model TWLM, however, its p-value is not high (K-S= 0.146, p-value=0.091). We also show formally that the size-bias in this data is present by preforming the likelihood ratio test for testing the following hypothesis $\mathsf{H}_0 : \alpha = 0$ (there is no size-bias) versus $\mathsf{H}_a : \alpha \neq 0$ (there is size-bias). The likelihood ration test is $\Lambda(x) = -2\left(\mathcal{L}_f(\alpha, c, k) - \mathcal{L}_r(0, c, k)\right) = 200.6$ with p-value approximately equal to 0. The graphs of the estimated densities of WWBXII, and BXII given in Figure 5 which show clearly that WWBXII distribution fits much better than the BXII distributions.

Table 5: Goodness-of-fit tests for Guinea pigs data.

Model	K-S	p-VALUE	Statistics	
			W_n	A_n
WWBXII	0.101	0.4581	0.14	0.75
TWLM	0.146	0.091	0.20	1.11
TWLL	0.444	< 0.0001	0.14	20.82
BXII	0.480	< 0.0001	5.02	23.49
SBXII	0.410	< 0.0001	3.73	18.26
LL	0.721	< 0.0001	12.00	57.27
SLL	0.447	< 0.0001	4.41	21.00
LM	0.485	< 0.0001	5.08	23.74
SLM	0.452	< 0.0001	4.46	21.18

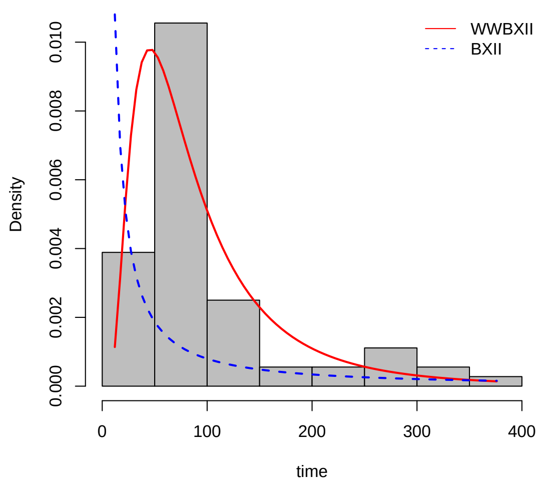

Figure 5: Fitted WWBXII and BXII densities and the histogram for Guinea pigs data

We further consider three well-known three-parameter distributions to fit this data.

1. The generalized gamma distribution, GG(α, c, k); $f(x; \alpha, c, k) = \frac{\alpha}{c\Gamma(k)}(\frac{x}{c})^{\alpha k-1}e^{-(\frac{x}{c})^{\alpha}}$,

where α, $c\ k > 0$

2. The size-biased generalized gamma distribution $\mathsf{GSGG}(\alpha, c, k)$; $f(x; \alpha, c, k) = \frac{\alpha x}{c^2\Gamma(k+\frac{1}{\alpha})}\left(\frac{x}{c}\right)^{\alpha k-1} e^{-\left(\frac{x}{c}\right)^{\alpha}}$, $\alpha, c, k > 0$.

3. The size-biased Weibull distribution $\mathsf{GSBW}(\alpha, c, k)$ $f(x; \alpha, c, k) = \alpha \frac{x^k\left(\frac{x}{c}\right)^{\alpha-1} e^{-\left(\frac{x}{c}\right)^{\alpha}}}{c^{k+1}\Gamma(\frac{k}{\alpha}+1)}$, where α, $c, k > 0$.

Table 6 lists the MLEs of the unknown parameters of the above distributions whereas Table 7 lists the corresponding values of the statistics K-S, W_n^*, and A_n^*. By examining the values of statistics and measures given in these table, we conclude that that the WWBXII distribution provides better fit than all the distributions considered in Table 6. The plots of fitted densities of these models and the histogram of the Gunia pigs data is given in Figure 6(left panel) and the plots of the fitted distributions of these models are given in Figure 6(right panel) show that the WWBXII distribution produces better fit than the other models. Additionally, the estimated hazard rate function is displayed in Figure 6(bottom panel) show that the estimated hazard rate function is unimodal which reflects the actual behavior of these data. Gupta and Kundu (2009) analyzed this data by using the weighted exponential distribution (WE) and they compared it with the Weibull (W), gamma (G) and generalized exponential (GE) distributions. They observed that the K-S distances of WE, W, G, and GE are 0.1173,0.149, 0.139 and 0.135 and the corresponding p-values are 0.275, 0.082, 0.112 and 0.135, respectively. Additionally, Kharazmi et al. (2015) fitted this data by using the generalized weighted exponential (GWE) distribution and they compared the GWE distribution with GE distribution and the two-parameter weighted exponential (TWE) distribution due to Shakhatreh (2012). They found that the K-S distances of GWE and TWE are 0.11 and 0.114, and the corresponding p-values are 0.35 and 0.31 respectively.

Table 6: MLEs of the model parameters (standard deviations), and the measures $\mathcal{L}$, AIC and BIC for Guinea pigs data

	MLEs of the parameters			Measures		
Model	$\widehat{\alpha}$	$\widehat{c}$	$\widehat{k}$	$-\mathcal{L}(\widehat{\boldsymbol{\theta}})$	AIC$(\widehat{\boldsymbol{\theta}})$	BIC$(\widehat{\boldsymbol{\theta}})$
WWBXII(α, c, k)	31.1736	0.3532	107.8556	390.25	786.5	793.33
	(6.4808)	(0.0846)	(26.7517)			
GSBW(c, k)	0.5214	1.9823	3.2494	391.74	791.2	798.03
	(0.2026)	(6.0124)	(1.6061)			
GG(c, k)	0.5883	4.7225	5.650	392.03	790.06	796.89
	(0.2571)	(13.2308)	(4.8493)			
GSBGG(c)	0.5928	3.6796	4.3364	392.06	790.12	796.95
	(0.5887)	(31.4817)	(9.2221)			

Table 7: Goodness-of-fit tests for Guinea pigs data.

Model	K-S	p-VALUE	Statistics W_n	A_n
WWBXII	0.101	0.46	0.14	0.75
GSBW	0.115	0.29	0.24	1.27
GG	0.117	0.28	0.25	1.34
GSBGG	0.118	0.27	0.25	1.32

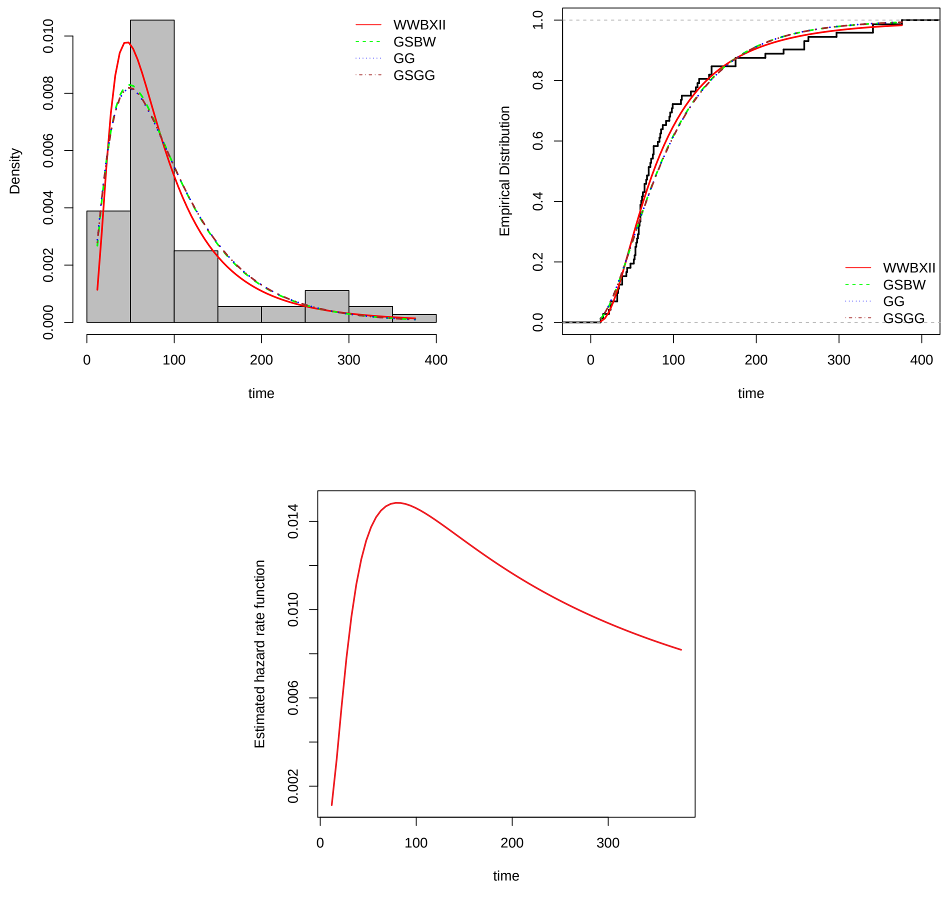

Figure 6: plots of estimated probability densities with the histogram of the data (left panel), estimated distribution (right panel), and estimated hazard rate (bottom panel) functions for Guinea pigs data

7 Conclusions

In this paper, we consider the weighted Burr-XII distribution and investigate many of its mathematical and statistical properties. The WWBXII distribution generalizes many interesting distributions including the two-parameter weighted lomax, two-parameter weighted log-logistic, and BurrXII distributions. We provide rigorous analysis to the shape of probability density and hazard rate functions of the WWBXII. We derive several structural properties such as moment generating function, entropies, mean residual life, extreme values and order statistics, and stochastic ordering. We investigate the estimation of the model parameters by maximum likelihood thoroughly and we provide sufficient conditions about the existence of these estimates. Also, we conduct some simulations to assess the finite sample behavior of the MLEs. Additionally, we provide a test concerning the existence of size-bias in the sample. Finally, the usefulness of the WWBXII distribution and practical relevance are demonstrated using the well-known data given byBjerkedal (1960).

Acknowledgement

We are very grateful to the Editor and the two anonymous reviewer for useful comments and suggestions on an early version of this paper.

References

Aarset, M. (1987). How to identify bathtub hazard rate. *IEEE Transactions on Reliability*, 36:106–108.

Afify, A., G.M. Cordeiro, E. O., Yousof, H., and Butt, N. (2018). The four-parameter Burr XII distribution: properties, regression model and applications. *Communications in Statistics - Theory and Methods*, 47:2605–2624.

AL-Hussaini, E. (1991). A characterization of the Burr type XII distribution. *Applied Mathematics Letters*, 4:59–61.

Arnold, B., Balakrishnan, N., and Nagaraja, H. (1992). *First Course in Order Statistics.* John Wiley, New york.

Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *American Journal of Hygiene*, 72:130—-148.

Burr, I. (1942). Cumulative frequency functions. *The Annals of mathematical statistics*, 13:215–232.

Burr, I. and Cislak, P. (1968). On a general system of distributions I. Its curve-shape characteristics II. The sample median. *Journal of the American Statistical Association*, 63:627–635.

Chen, C. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal Quality Technology*, 27:154—-161.

Cordeiro, G. M., Mead, M. E., Afify, A. Z., Suzuki, A. K., and El-Gaied, A. A. K. A. (2017). An extended Burr XII distribution:properties, inference and applications. *Pakistan Journal of Statistics and Operation Research*, 13:809–228.

Ducey, M. and Gove, J. (2015). Size-biased distributions in the generalized beta distribution family, with applications to forestry. *Forestry*, 88:143–151.

Fisher, R. (1934). The effect of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6:13–25.

Glaser, R. (1980). Bathtub and related failure rate characterization. *Journal of American Statistical Association*, 75:667–672.

Gove, J. (2003). Moment and maximum likelihood estimators for Weibull distributions under length- and area-biased sampling. *Environmental and Ecological Statistics*, 10:455—-467.

Gove, J. and Patil, G. (1998). Modeling the basal area-size distribution of forest stands: a compatible approach. *Forest Science*, 10:285—-297.

Guess, F., Nam, K. H., and Park, D. H. (1998). Failure rate and mean residual life with trend changes. *Asia-Pacific Journal of Operational Research*, 15:239—-244.

Gupta, D. and Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, 43(6):621–634.

Gupta, R. and Kirmani, S. (1990). The role of weighted distributions in stochastic modeling. *Communications in Statistics-Theory and methods*, 19:3147–3162.

Jain, K., Singh, H., and Bagai, I. (2007). Relations for reliability measures of weighted distributions. *Communications in Statistics - Theory and Methods*, 18:4393–4412.

Kharazmi, O., Mahdavi, A., and Fathizadeh, M. (2015). Generalized weighted exponential distribution. *Communications in Statistics - Simulation and Computation*, 44:1557–1569.

Leiva, V., Sanhueza, A., and Angulo, J. (2009). A length-biased version of the Birnbaum-Saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment*, 23:299–307.

Paranaíba, P., Ortega, E., Cordeiro, G., and de Pascoa, M. (2013). The Kumaraswamy Burr XII distribution: theory and practice. *Journal of Statistical Computation and Simulation*, 83:2117–2143.

Paranaíba, P., Ortega, E., Cordeiro, G., and Pescim, R. (2011). The beta Burr XII distribution with application to lifetime data. *Computational Statistics and Data Analysis*, 55:1118–1136.

Patil, G. (2002). Weighted distributions. *In: El-Shaarawi AH, Piegorsch WW (eds) Encyclopedia of environmetrics*, 4:2369–2377.

Patil, G. and Ord, J. (1976). On size-biased sampling and related form-invariant weighted distributions. *Sankhyā: The Indian Journal of Statistics*, 38:48–61.

Rodriguez, R. (1977). A guide to the Burr type XII distributions. *Biometrika*, 64:129–134.

Sansgiry, P. and Akman, O. (2006). Reliability estimation via length-biased transformation. *Communications in Statistics - Theory and Methods*, 10:2473–2479.

Self, S. G. and Liang, K. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82:605–610.

Shakeda, M. and Shanthikumar, J. (2007). *Stochastic Orders*. Springer Science+Business Media, LLC, New york.

Shakhatreh, M. (2012). A two-parameter of weighted exponential distributions. *Statistics and Probability Letters*, 82(2):252–261.

Shao, Q. (2004). Notes on maximum likelihood estimation for the three-parameter BurrXII distribution. *Computational Statistics and Data Analysis*, 45:675–687.

Shao, Q., Wong, H., Xia, J., and Ip, W.-C. (2004). Models for extremes using the extended three parameter Burr XII system with application to flood frequency analysis. *Hydrological Sciences Journal*, 49:685–702.

Shao, Q. and Zhou, X. (2004). A new parametric model for survival data with long-term survivors. *Statistics in Medicine*, 23:3525—-3543.

Silvaa, G., E.M.M.Ortegaa, Canchob, V., and Barretoc, M. (2008). Log–Burr XII regression models with censored data. *Computational Statistics and Data Analysis*, 52:3820–3842.

Soliman, A. (2005). Estimation of parameters of life from progressively censored data using Burr XII model. *IEEE Transactions on Reliability*, 54:34–42.

Tadikamalla, P. (1980). A look at the Burr and related distributions. *International Statistical Review/Revue Internationale de Statistique*, 48:337–344.

Team, R. D. C. (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Usta, I. (2013). Different estimation methods for the parameters of the extended Burr XII distribution. *Journal of Applied Statistics*, 40:397–414.

Wang, F., Keats, J., and Zimmer, W. (1996). Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data. *Microelectronics Reliability*, 36:359–362.

Wingo, D. (1993). Maximum likelihood estimation of Burr XII distribution parameters under type II censoring. *Microelectronics Reliability*, 33:1251–1257.

Wu, J., Chen, Y., and Chang, C. (2007). Statistical inference based on progressively censored samples with random removals from the Burr type XII distribution. *Journal of Statistical Computation and Simulation*, 77:19–27.

Zimmer, W., Keats, J., and Wang, F. (1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, 30:386–394.

Appendix

The elements of the observed Fisher information matrix are given below.

$$\frac{\partial^2 \mathcal{L}}{\partial k^2} = -\frac{n}{k^2} + n\acute{\psi}(k) - n\acute{\psi}(k - \frac{\alpha}{c})$$

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha^2} = \frac{n}{\alpha^2} - \frac{n}{c^2}\,\acute{\psi}(k - \frac{\alpha}{c}) - \frac{n}{c^2}\acute{\psi}(\frac{\alpha}{c})$$

$$\frac{\partial^2 \mathcal{L}}{\partial c \partial k} = -\frac{n\alpha}{c^2}\acute{\psi}(k - \frac{\alpha}{c}) - \sum_{i=1}^{n} x_i^c\,(1 + x_i^c)^{-1}\ln^2(x_i)$$

$$\frac{\partial^2 \mathcal{L}}{\partial k \partial \alpha} = \frac{n}{c}\acute{\psi}(k - \frac{\alpha}{c})$$

$$\frac{\partial^2 \mathcal{L}}{\partial c \partial \alpha} = \frac{\alpha n\acute{\psi}(k - \frac{\alpha}{c}) - cn\psi(k - \frac{\alpha}{c})}{c^3} + \frac{n}{c^3}[c\psi(\frac{\alpha}{c}) + \alpha\acute{\psi}(\frac{\alpha}{c})]$$

$$\frac{\partial^2 \mathcal{L}}{\partial c^2} = -\frac{2n}{c^2} - \frac{\alpha n}{c^4}\,[\alpha\acute{\psi}(k - \frac{\alpha}{c}) - 2c\psi(k - \frac{\alpha}{c})] - \frac{\alpha n}{c^4}\,[2c\psi(\frac{\alpha}{c}) + \alpha\acute{\psi}(\frac{\alpha}{c})] - (k+1)\sum_{i=1}^{n} x_i^c(1 + x_i^c)^{-2}\ln^2(x_i)(x_i)$$