

## The Kumaraswamy Pareto distribution

Marcelo Bourguignon\*, Rodrigo B. Silva, Luz M. Zea and Gauss M. Cordeiro

*Departamento de Estatística, Universidade Federal de Pernambuco.*

Received 29 February 2012

Accepted 2 October 2012

### Abstract

The modeling and analysis of lifetimes is an important aspect of statistical work in a wide variety of scientific and technological fields. For the first time, the called Kumaraswamy Pareto distribution, is introduced and studied. The new distribution can have a decreasing and upside-down bathtub failure rate function depending on the values of its parameters. It includes as special sub-models the Pareto and exponentiated Pareto (Gupta *et al.*, 1998) distributions. Some structural properties of the proposed distribution are studied including explicit expressions for the moments and generating function. We provide the density function of the order statistics and obtain their moments. The method of maximum likelihood is used for estimating the model parameters and the observed information matrix is derived. A real data set is used to compare the new model with widely known distributions.

*Keywords:* Hazard function; Kumaraswamy distribution; moments; maximum likelihood estimation; Pareto distribution.

### 1. Introduction

The Pareto distribution is a very popular model named after a professor of economics: Vilfredo Pareto. The various forms of the Pareto distribution are very versatile and a variety of uncertainties can be usefully modelled by them. For instance, they arise as tractable ‘lifetime’ models in actuarial sciences, economics, finance, life testing and climatology, where it usually describes the occurrence of extreme weather.

The random variable  $X$  has the Pareto distribution if its cumulative distribution function (cdf) for  $x \geq \beta$  is given by

$$G(x; \beta, k) = 1 - \left( \frac{\beta}{x} \right)^k, \quad (1)$$

where  $\beta > 0$  is a scale parameter and  $k > 0$  is a shape parameter. The probability density function (pdf) corresponding to (1) is

$$g(x; \beta, k) = \frac{k\beta^k}{x^{k+1}}.$$

Several generalized forms of the Pareto distribution can be found in the literature. The term “generalized Pareto” (GP) distribution was first used by Pickands (1975) when making statistical inferences about the upper tail of a distribution function. As expected, the Pareto distribution can be seen as a special case of the GP

---

\*Corresponding author. E-mail: m.p.bourguignon@gmail.com. Address for correspondence: Universidade Federal de Pernambuco, DE/CCEN, Cidade Universitária, s/n, 50740-540, Pernambuco-PE, Brasil.

distribution. It can also be obtained as a special case of another extended form generated by compounding a heavy-tailed skewed conditional gamma density function with parameters  $\alpha$  and  $\beta^{-1}$ , where the weighting function for  $\beta$  has a gamma distribution with parameters  $k$  and  $\theta$  (Hogg *et al.*, 2005).

Gupta *et al.* (1998) extended the Pareto distribution by raising (1) to a positive power. In this note, we refer to this extension as the exponentiated Pareto (EP) distribution. Recently, many authors have considered various exponentiated-type distributions based on some known distributions such as the exponential, Rayleigh, Weibull, gamma and Burr distributions; see, for example, Gupta and Kundu (2001), Kundu and Raqab (2005) and Silva *et al.* (2010). The methods of moments and maximum likelihood have been used to fit these models.

Further, Akinsete *et al.* (2008) and Mahmoudi (2011) extended the Pareto and GP distributions by defining the beta Pareto (BP) and beta generalized Pareto (BGP) distributions, respectively, based on the class of generalized (so-called “beta-G”) distributions introduced by Eugene *et al.* (2002). The generalized distributions are obtained by taking any parent G distribution in the cdf of a beta distribution with two additional shape parameters, whose role is to introduce skewness and to vary tail weight. Following the same idea, many beta-type distributions were introduced and studied, see, for example, Barreto-Souza *et al.* (2010) and Silva *et al.* (2010).

In this context, we propose an extension of the Pareto distribution based on the family of Kumaraswamy generalized (denoted with the prefix “Kw-G” for short) distributions introduced by Cordeiro and de Castro (2011). Nadarajah *et al.* (2011) studied some mathematical properties of this family. The Kumaraswamy (Kw) distribution is not very common among statisticians and has been little explored in the literature. Its cdf (for  $0 < x < 1$ ) is  $F(x) = 1 - (1 - x^a)^b$ , where  $a > 0$  and  $b > 0$  are shape parameters, and the density function has a simple form  $f(x) = abx^{a-1}(1 - x^a)^{b-1}$ , which can be unimodal, increasing, decreasing or constant, depending on the parameter values. It does not seem to be very familiar to statisticians and has not been investigated systematically in much detail before, nor has its relative interchangeability with the beta distribution been widely appreciated. However, in a very recent paper, Jones (2009) explored the background and genesis of this distribution and, more importantly, made clear some similarities and differences between the beta and Kw distributions.

In this note, we combine the works of Kumaraswamy (1980) and Cordeiro and de Castro (2011) to derive some mathematical properties of a new model, called the Kumaraswamy Pareto (Kw-P) distribution, which stems from the following general construction: if G denotes the baseline cumulative function of a random variable, then a generalized class of distributions can be defined by

$$F(x) = 1 - [1 - G(x)^a]^b, \quad (2)$$

where  $a > 0$  and  $b > 0$  are two additional shape parameters which govern skewness and tail weights. Because of its tractable distribution function (2), the Kw-G distribution can be used quite effectively even if the data are censored. Correspondingly, its density function is distributions has a very simple form

$$f(x) = abg(x)G(x)^{a-1}[1 - G(x)^a]^{b-1}. \quad (3)$$

The density family (3) has many of the same properties of the class of beta-G distributions (see Eugene *et al.*, 2002), but has some advantages in terms of tractability, since it does not involve any special function such as the beta function.

Equivalently, as occurs with the beta-G family of distributions, special Kw-G distributions can be generated as follows: the Kw-normal distribution is obtained by taking  $G(x)$  in (2) to be the normal cumulative function. Analogously, the Kw-Weibull (Cordeiro *et al.* (2010), Kw-generalized gamma (Pascoa *et al.*, 2011), Kw-Birnbaum-Saunders (Saulo *et al.*, 2012) and Kw-Gumbel (Cordeiro *et al.*, 2011) distributions are obtained by taking  $G(x)$  to be the cdf of the Weibull, generalized gamma, Birnbaum-Saunders and Gumbel distributions, respectively, among several others. Hence, each new Kw-G distribution can be generated from a specified G distribution.

This paper is outlined as follows. In section 2, we define the Kw-P distribution and provide expansions for its cumulative and density functions. In addition, we study the limit behavior of its pdf and hazard rate function. A range of mathematical properties of this distribution is considered in sections 3-7. These include quantile function, simulation, skewness and kurtosis, order statistics, generating and characteristic functions, incomplete moments, L-moments and mean deviations. The Rényi entropy is calculated in section 8. Maximum likelihood estimation is performed and the observed information matrix is determined in section 9. In section 10, we provide an application of the Kw-P distribution to a flood data set. Finally, some conclusions are addressed in section 11.

## 2. The Kw-P distribution

If  $G(x; \beta, k)$  is the Pareto cumulative distribution with parameters  $\beta$  and  $k$ , then equation (2) yields the Kw-P cumulative distribution (for  $x \geq \beta$ )

$$F(x; \beta, k, a, b) = 1 - \left\{ 1 - \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^a \right\}^b, \quad (4)$$

where  $\beta > 0$  is a scale parameter and the other positive parameters  $k, a$  and  $b$  are shape parameters. The corresponding pdf and hazard rate function are

$$f(x; \beta, k, a, b) = \frac{abk\beta^k}{x^{k+1}} \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}, \quad (5)$$

and

$$\tau(x; \beta, k, a, b) = \frac{abk\beta^k \left[ 1 - (\beta/x)^k \right]^{a-1}}{x^{k+1} \{ 1 - [1 - (\beta/x)^k]^a \}}, \quad (6)$$

respectively.

The Kw-P distribution is not in fact very tractable. However, its heavy tail can adjust skewed data that cannot be properly fitted by existing distributions. Furthermore, the cumulative and hazard rate functions are simple.

In Figures 1 and 2, we plot the density and failure rate functions of the Kw-P distribution for selected parameter values, respectively. We can verify that this distribution can have a decreasing and upside-down bathtub failure rate function depending on the values of its parameters.

### 2.1. Expansions for the cumulative and density functions

Here, we give simple expansions for the Kw-P cumulative distribution. By using the generalized binomial theorem (for  $0 < a < 1$ )

$$(1+a)^v = \sum_{i=0}^{\infty} \binom{v}{i} a^i, \quad (7)$$

where

$$\binom{v}{i} = \frac{n(n-1) \dots (v-i+1)}{i!},$$

in equation (4), we can write

$$F(x; \beta, k, a, b) = 1 - \sum_{i=0}^{\infty} (-1)^i \binom{b}{i} \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^{ai} = 1 - \sum_{i=0}^{\infty} \eta_i H(x; \beta, k, ia),$$

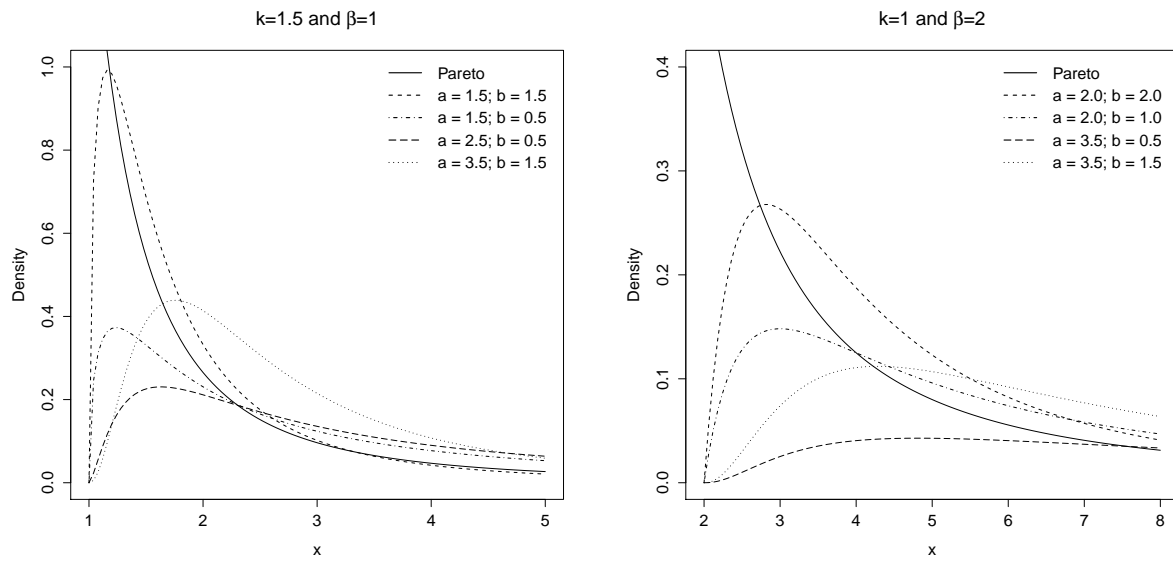


Figure 1: Plots of the Kw-P density function for some parameter values.

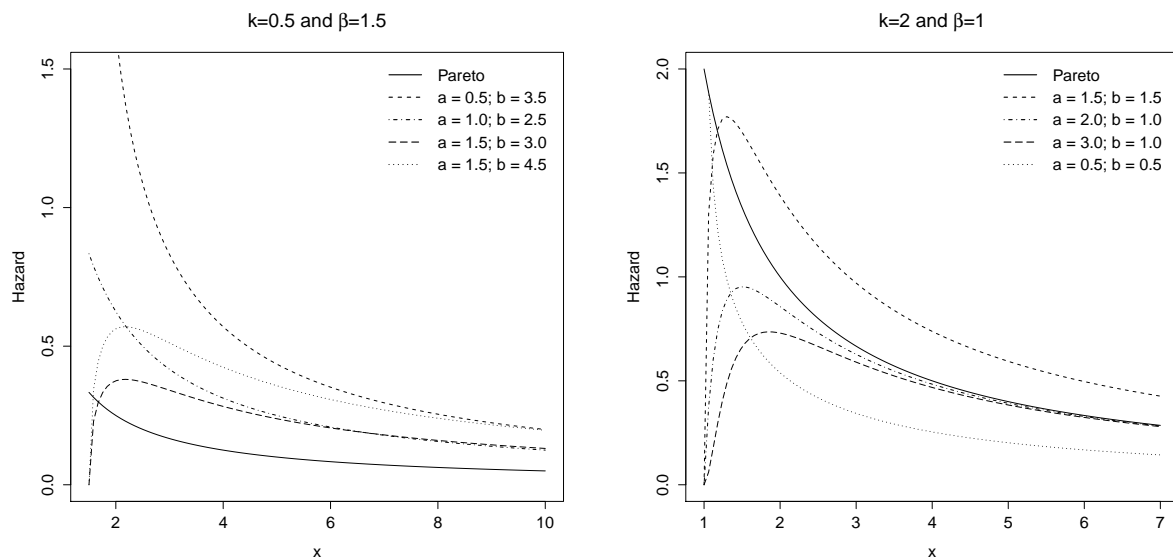


Figure 2: Plots of the Kw-P hazard function for some parameter values.

where  $\eta_i = (-1)^i \binom{b}{i}$  and  $H(x; \beta, k, ia)$  denotes the EP cumulative distribution (with parameters  $\beta$ ,  $k$  and  $ia$ ) given by

$$H(x; \beta, k, \alpha) = \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^\alpha.$$

Now, using the power series (7) in the last term of (5), we obtain

$$f(x; \beta, k, a, b) = \frac{abk\beta^k}{x^{k+1}} \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^{a(i+1)-1} = \sum_{j=0}^{\infty} w_j g(x; \beta, k(j+1)),$$

where

$$w_j = \frac{ab}{(j+1)} \sum_{i=0}^{\infty} (-1)^{i+j} \binom{b-1}{i} \binom{a(i+1)-1}{j}$$

and  $g(x; \beta, k(j+1))$  denotes the Pareto density function with parameters  $\beta$  and  $k(j+1)$  and cumulative distribution as in (1). Thus, the Kw-P density function can be expressed as an infinite linear combination of Pareto densities. Thus, some of its mathematical properties can be obtained directly from those properties of the Pareto distribution. For example, the ordinary, inverse and factorial moments, moment generating function (mgf) and characteristic function of the Kw-P distribution follow immediately from those quantities of the Pareto distribution.

## 2.2. Limiting behaviour of Kw-P density and hazard functions

**Lemma 1.** The limit of the Kw-P density function as  $x \rightarrow \infty$  is 0 and the limit as  $x \rightarrow \beta$  are

$$\lim_{x \rightarrow \beta} f(x; \beta, k, a, b) = \begin{cases} \infty, & \text{for } 0 < a < 1, \\ \frac{bk}{\beta}, & \text{for } a = 1, \\ 0, & \text{for } a > 1. \end{cases}$$

*Proof.* It is easy to demonstrate the result from the density function (5).

**Lemma 2.** The limit of the Kw-P hazard function as  $x \rightarrow \infty$  is 0 and the limit as  $x \rightarrow \beta$  are

$$\lim_{x \rightarrow \beta} \tau(x; \beta, k, a, b) = \begin{cases} \infty, & \text{for } 0 < a < 1, \\ \frac{bk}{\beta}, & \text{for } a = 1, \\ 0, & \text{for } a > 1. \end{cases}$$

*Proof.* It is straightforward to prove this result from (6).

## 3. Moments and generating function

Here and henceforth, let  $X$  be a Kw-P random variable following (5).

### 3.1. Moments

The  $r$ th moment of  $X$  can be obtained from (8) as

$$\begin{aligned} E(X^r) &= \sum_{j=0}^{\infty} w_j \int_{\beta}^{\infty} x^r g(x; \beta, k(j+1)) dx \\ &= k\beta^r \sum_{j=0}^{\infty} \frac{(j+1)w_j}{[k(j+1)-r]}, \end{aligned} \quad (8)$$

for  $r < bk$ . In particular, setting  $r = 1$  in (8), the mean of  $X$  reduces to

$$\mu = E(X) = k\beta \sum_{j=0}^{\infty} \frac{(j+1)w_j}{[k(j+1)-1]}, \quad \text{for } bk > 1. \quad (9)$$

Setting  $a = b = 1$ , we have

$$w_j = \begin{cases} 1, & \text{for } j = 0, \\ 0, & \text{for } j \geq 1. \end{cases}$$

Then, equation (9) reduces to (for  $k > 1$ )

$$E(X) = \frac{k\beta}{k-1},$$

which is precisely the mean of the Pareto distribution.

### 3.2. Incomplete moments

If  $Y$  is a random variable with a Pareto distribution with parameters  $\beta$  and  $k$ , the  $r$ th incomplete moment of  $Y$ , for  $r < k$ , is given by

$$M_r(z) = \int_{\beta}^z y^r g(y; \beta, k) dy = \frac{k\beta^r}{(k-r)} \left[ 1 - \left( \frac{\beta}{z} \right)^{k-r} \right].$$

From this equation, we note that  $M_r(z) \rightarrow E(Y^r)$  when  $z \rightarrow \infty$ , whenever  $k > r$ . Let  $X \sim \text{Kw-P}(\beta, k, a, b)$ . The  $r$ th incomplete moment of  $X$  is then equal to

$$M_r(z) = \int_{\beta}^z x^r f(x; \beta, k, a, b) dx = k\beta^r \sum_{j=0}^{\infty} \frac{(j+1)w_j}{[k(j+1)-r]} \left[ 1 - \left( \frac{\beta}{z} \right)^{k(j+1)-r} \right], \quad (10)$$

which provided that  $r < bk$ .

### 3.3. Generating function

First, the mgf  $M_Y(t)$  corresponding to a random variable  $Y$  with Pareto distribution with parameters  $\beta$  and  $k$  is only defined for non-positive values of  $t$ . It is given by

$$M_Y(t) = k(-\beta t)^k \Gamma(-k, -\beta t), \quad \text{if } t < 0,$$

where  $\Gamma(\cdot, \cdot)$  denotes the incomplete gamma function

$$\Gamma(s, x) = \int_x^{\infty} t^{s-1} e^{-t} dt.$$

Thus, using  $M_Y(t)$  and (8), we can write for  $t < 0$

$$\begin{aligned} M_X(t) &= \sum_{j=0}^{\infty} w_j \int_{\beta}^{\infty} e^{tx} g(x; \beta, k(j+1)) dx \\ &= k \sum_{j=0}^{\infty} (j+1) (-\beta t)^{k(j+1)} w_j \Gamma(-k(j+1), -\beta t). \end{aligned} \quad (11)$$

#### 4. Quantile function and simulation

We present a method for simulating from the Kw-P distribution (5). The quantile function corresponding to (4) is

$$Q(u) = F^{-1}(u) = \frac{\beta}{\{1 - [1 - (1-u)^{1/b}]^{1/a}\}^{1/k}}. \quad (12)$$

Simulating the Kw-P random variable is straightforward. Let  $U$  be a uniform variate on the unit interval  $(0, 1)$ . Thus, by means of the inverse transformation method, we consider the random variable  $X$  given by

$$X = \frac{\beta}{\{1 - [1 - (1-U)^{1/b}]^{1/a}\}^{1/k}},$$

which follows (5), i.e.,  $X \sim \text{Kw-P}(\beta, k, a, b)$ .

The plots comparing the exact Kw-P densities and histograms from two simulated data sets for some parameter values are given in Figure 3. These plots indicate that the simulated values are consistent with the Kw-P theoretical density function.

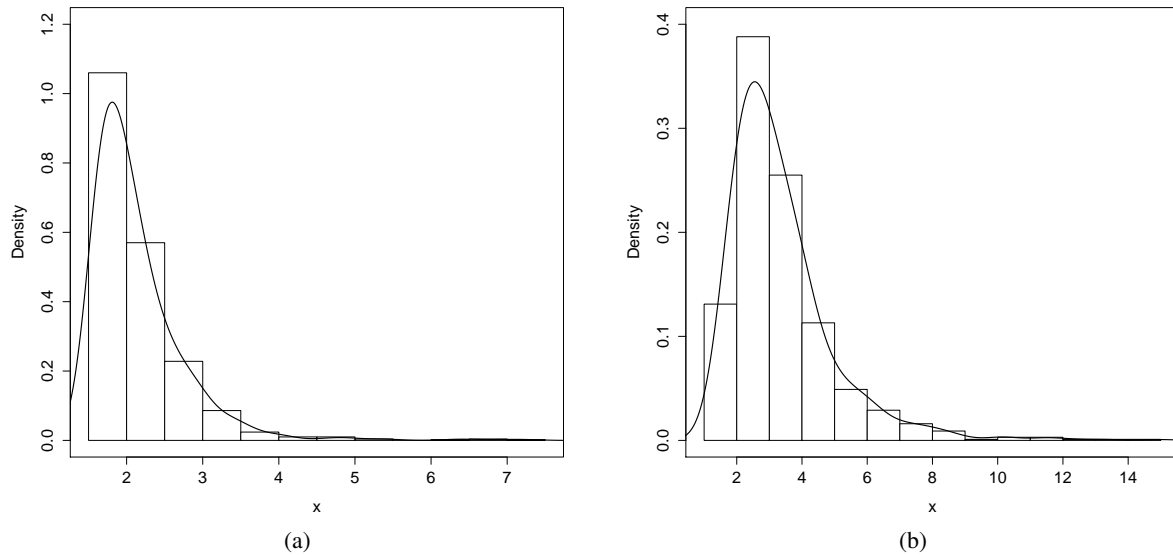


Figure 3: Plots of the Kw-P densities for simulated data sets: (a)  $a = 1.5, b = 3.5, \beta = 1.5, k = 1.5$  and (b)  $a = 5.0, b = 3.0, \beta = 1.5, k = 1.5$ .

## 5. Skewness and Kurtosis

The shortcomings of the classical kurtosis measure are well-known. There are many heavy-tailed distributions for which this measure is infinite. So, it becomes uninformative precisely when it needs to be. Indeed, our motivation to use quantile-based measures stemmed from the non-existence of classical kurtosis for many of the Kw distributions.

The Bowley's skewness (see Kenney and Keeping, 1962) is based on quartiles:

$$B = \frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)}$$

and the Moors' kurtosis (see Moors, 1998) is based on octiles:

$$M = \frac{Q(7/8) - Q(5/8) - Q(3/8) + Q(1/8)}{Q(6/8) - Q(2/8)},$$

where  $Q(\cdot)$  represents the quantile function.

Plots of the skewness and kurtosis for some choices of the parameter  $b$  as function of  $a$ , and for some choices of the parameter  $a$  as function of  $b$ , for  $\beta = 1.0$  and  $k = 1.5$ , are shown in Figure 4. These plots show that the skewness and kurtosis decrease when  $b$  increases for fixed  $a$  and when  $a$  increases for fixed  $b$ .

## 6. Order statistics

Moments of order statistics play an important role in quality control testing and reliability, where a practitioner needs to predict the failure of future items based on the times of a few early failures. These predictors are often based on moments of order statistics. We now derive an explicit expression for the density function of the  $i$ th order statistic  $X_{i:n}$ , say  $f_{i:n}(x)$ , in a random sample of size  $n$  from the Kw-P distribution. We can write

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} f(x) F^{i-1}(x) [1 - F(x)]^{n-i},$$

where  $f(\cdot)$  and  $F(\cdot)$  are the pdf and cdf of the Kw-P distribution, respectively. From the above equation and using the series representation (7) repeatedly, we obtain a useful expression for  $f_{i:n}(x)$  given by

$$f_{i:n}(x) = \sum_{r=0}^{\infty} c_{i:n}^{(r)} g(x; k(r+1), \beta), \quad (13)$$

where

$$c_{i:n}^{(r)} = \frac{n! a b}{(i-1)!(n-i)!} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{l+m+r}}{r+1} \binom{i-1}{l} \binom{b(n+l+1-i)-1}{m} \binom{a(m+1)-1}{r}$$

and  $g(x; k(r+1), \beta)$  denotes the Pareto density function with parameters  $k(r+1)$  and  $\beta$ . So, the density function of the order statistics is simply an infinite linear combination of Pareto densities. The pdf of the  $i$ th order statistic from a random sample of the Pareto distribution comes by setting  $a = b = 1$  in (13). Evidently, equation (13) plays an important role in the derivation of the main properties of the Kw-P order statistics.

For example, the  $s$ th raw moment of  $X_{i:n}$  can be expressed as

$$E(X_{i:n}^s) = k \beta^s \sum_{r=0}^{\infty} \frac{(r+1) c_{i:n}^{(r)}}{k(r+1) - s}. \quad (14)$$



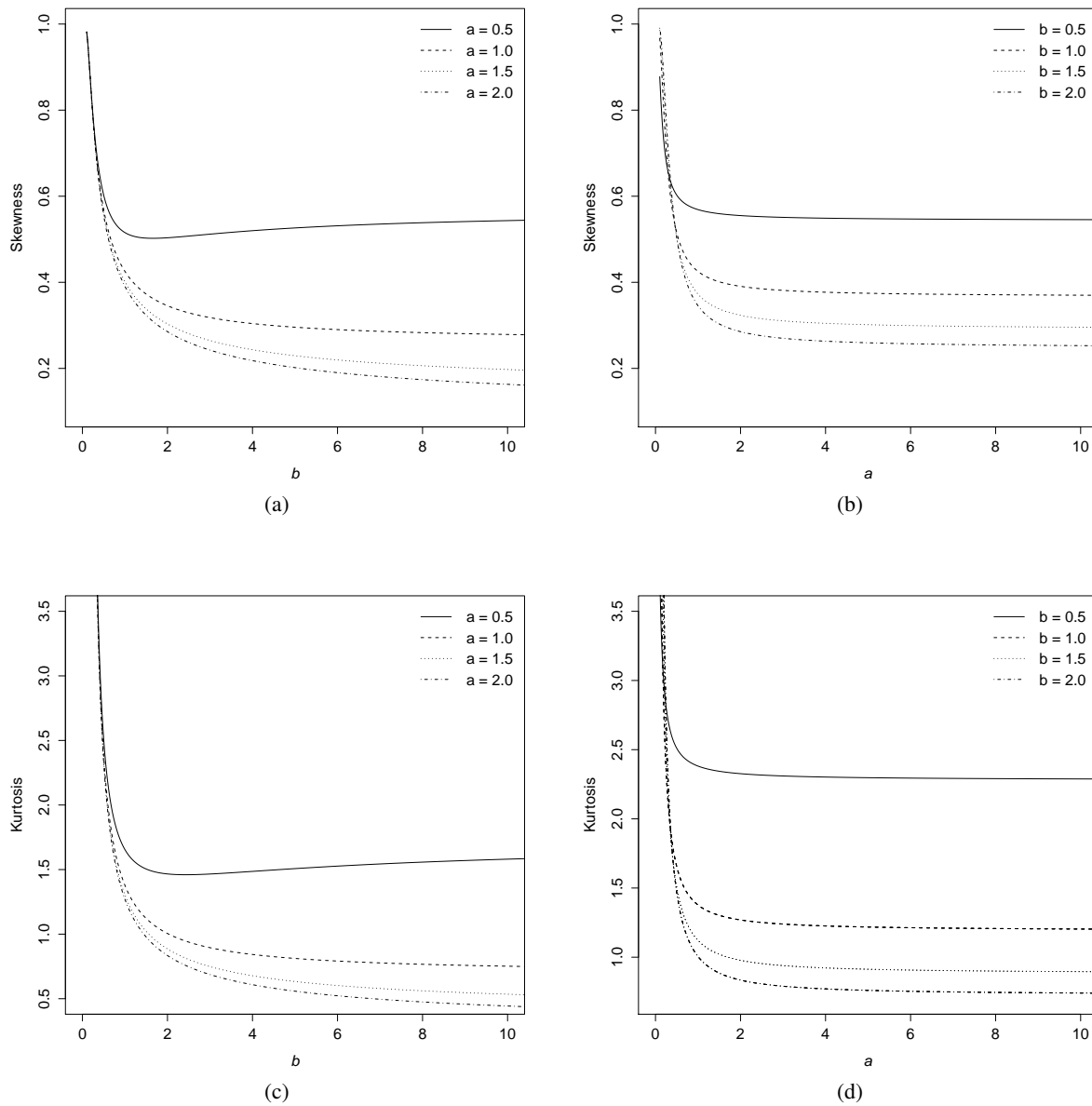


Figure 4: Plots of the Kw-P skewness and kurtosis as a function of  $a$  for selected values of  $b$  and as a function of  $b$  for selected values of  $a$ .

The L-moments are analogous to the ordinary moments but can be estimated by linear combinations of order statistics. They are linear functions of expected order statistics defined by

$$\lambda_{m+1} = \frac{1}{m+1} \sum_{k=0}^m (-1)^k \binom{m}{k} E(X_{m+1-k:m+1}), \quad m = 0, 1, \dots$$

The first four L-moments are:  $\lambda_1 = E(X_{1:1})$ ,  $\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$ ,  $\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$  and  $\lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$ . The L-moments have the advantage that they exist whenever the mean of the distribution exists, even though some higher moments may not exist, and are relatively robust to the effects of outliers. From equation (14) with  $s = 1$ , we can easily obtain explicit expressions for the L-moments of  $X$ .

## 7. Mean deviations

The mean deviations about the mean and the median can be used as measures of spread in a population. Let  $\mu = E(X)$  and  $m$  be the mean and the median of the Kw-P distribution, respectively. The mean deviations about the mean and about the median can be calculated as

$$D(\mu) = E(|X - \mu|) = \int_{\beta}^{\infty} |x - \mu| f(x) dx$$

and

$$D(m) = E(|X - m|) = \int_{\beta}^{\infty} |x - m| f(x) dx,$$

respectively. We obtain

$$D(\mu) = \int_{\beta}^{\infty} |x - \mu| f(x) dx = 2\mu F(\mu) - 2M_1(\mu),$$

where  $M_1(\mu)$  denotes the first incomplete moment calculated from (10) for  $r = 1$ . Similarly, the mean deviation about the median follows as

$$D(m) = \int_{\beta}^{\infty} |x - m| f(x) dx = \mu - 2M_1(m).$$

## 8. Rényi entropy

The entropy of a random variable  $X$  is a measure of uncertainty variation. The Rényi entropy is defined as

$$I_R(\delta) = \frac{1}{1-\delta} \log [I(\delta)],$$

where  $I(\delta) = \int_{\mathbb{R}} f^{\delta}(x) dx$ ,  $\delta > 0$  and  $\delta \neq 1$ . We have

$$I(\delta) = a^{\delta} b^{\delta} k^{\delta} \beta^{k\delta} \int_{\beta}^{\infty} \frac{1}{x^{\delta(k+1)}} \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^{\delta(a-1)} \left\{ 1 - \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^a \right\}^{\delta(b-1)} dx.$$

By expanding the last term of the above integrand as in equation (7), we obtain

$$I(\delta) = a^{\delta} b^{\delta} k^{\delta} \beta^{k\delta} \sum_{j=0}^{\infty} (-1)^j \binom{\delta(b-1)}{j} \int_{\beta}^{\infty} \frac{1}{x^{\delta(k+1)}} \left[ 1 - \left( \frac{\beta}{x} \right)^k \right]^{a(\delta+j)-\delta} dx.$$

Transforming variables, this equation becomes

$$I(\delta) = a^\delta b^\delta k^{\delta-1} \beta^{\delta+1} \sum_{j=0}^{\infty} (-1)^j \binom{\delta(b-1)}{j} B\left(a(\delta+j) - \delta + 1, \frac{\delta(k+1) - 1}{k}\right),$$

where

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$$

denotes the beta function. Hence, the Rényi entropy reduces to

$$I_R(\delta) = \frac{\delta \log(ab)}{1-\delta} - \log k + \log \beta + \frac{1}{1-\delta} \log \sum_{j=0}^{\infty} (-1)^j \binom{\delta(b-1)}{j} B\left(a(\delta+j) - \delta + 1, \frac{\delta(k+1) - 1}{k}\right).$$

## 9. Estimation and information matrix

In this section, we discuss maximum likelihood estimation and inference for the Kw-P distribution. Let  $x_1, \dots, x_n$  be a random sample from  $X \sim \text{Kw-P}(\beta, k, a, b)$  and let  $\theta = (\beta, k, a, b)^\top$  be the vector of the model parameters. The log-likelihood function for  $\theta$  reduces to

$$\begin{aligned} \ell(\theta) &= n \log a + n \log b + n \log k + nk \log \beta - (k+1) \sum_{i=1}^n \log(x_i) \\ &+ (a-1) \sum_{i=1}^n \log \left[ 1 - \left( \frac{\beta}{x_i} \right)^k \right] + (b-1) \sum_{i=1}^n \log \left\{ 1 - \left[ 1 - \left( \frac{\beta}{x_i} \right)^k \right]^a \right\}. \end{aligned} \quad (15)$$

The score vector is  $U(\theta) = (\partial \ell / \partial k, \partial \ell / \partial a, \partial \ell / \partial b)^\top$ , where the components corresponding to the model parameters are calculated by differentiating (15). By setting  $z_i = 1 - (\beta/x_i)^k$ , we obtain

$$\begin{aligned} \frac{\partial \ell}{\partial k} &= \frac{n}{k} + \frac{1}{k} \sum_{i=1}^n \log(1 - z_i) - \frac{(a-1)}{k} \sum_{i=1}^n \frac{(1 - z_i) \log(1 - z_i)}{z_i} \\ &+ \frac{a(b-1)}{k} \sum_{i=1}^n \frac{z_i^{a-1} (1 - z_i) \log(1 - z_i)}{(1 - z_i^a)}, \quad \frac{\partial \ell}{\partial a} = \frac{n}{a} + \sum_{i=1}^n \log z_i - (b-1) \sum_{i=1}^n \frac{z_i^a \log z_i}{1 - z_i^a} \end{aligned}$$

and

$$\frac{\partial \ell}{\partial b} = \frac{n}{b} + \sum_{i=1}^n \log(1 - z_i^a).$$

The maximum likelihood estimates (MLEs) of the parameters are the solutions of the nonlinear equations  $\nabla \ell = 0$ , which are solved iteratively. The observed information matrix given by

$$\mathbf{J}_n(\theta) = n \begin{bmatrix} J_{kk} & J_{ka} & J_{kb} \\ J_{ak} & J_{aa} & J_{ab} \\ J_{bk} & J_{ba} & J_{bb} \end{bmatrix},$$

whose elements are

$$\begin{aligned}
 J_{kk} &= -\frac{n}{k^2} - \frac{2(a-1)}{k} \sum_{i=1}^n \frac{(1-z_i) \log(1-z_i)}{z_i^2} \\
 &\quad + \frac{2a(b-1)}{k} \sum_{i=1}^n \frac{z_i^{a-1} (1-z_i) \log(1-z_i) [a - (a-1)(z_i^a + z_i^{-1}) + (a-2)z_i^{a-1}]}{(1-z_i^a)^2}, \\
 J_{ka} &= \frac{(b-1)}{k} \sum_{i=1}^n \frac{z_i^{a-1} (1-z_i) \log(1-z_i) [1 - z_i^a - a \log z_i]}{(1-z_i^a)^2} - \frac{1}{k} \sum_{i=1}^n \frac{(1-z_i) \log(1-z_i)}{z_i}, \\
 J_{kb} &= \frac{a}{k} \sum_{i=1}^n \frac{z_i^{a-1} (1-z_i) \log(1-z_i)}{1-z_i^a}, J_{aa} = -\frac{n}{a^2} - 2(b-1) \sum_{i=1}^n \frac{z_i^a \log z_i}{(1-z_i^a)^2}, J_{ab} = -\sum_{i=1}^n \frac{z_i^a \log z_i}{1-z_i^a} \text{ and } J_{bb} = -\frac{n}{b^2}.
 \end{aligned}$$

## 10. Simulation study and application

In this section, we illustrate the usefulness of the Kw-P distribution.

### 10.1. Simulation study

We conduct Monte Carlo simulation studies to assess on the finite sample behavior of the MLEs of  $\beta, k, a$  and  $b$ . All results were obtained from 1000 Monte Carlo replications and the simulations were carried out using the statistical software package R. In each replication, a random sample of size  $n$  is drawn from the Kw-P( $\beta, k, a, b$ ) distribution and the BFGS method has been used by the authors for maximizing the total log-likelihood function  $\ell(\theta)$ . The Kw-P random number generation was performed using the inversion method. The true parameter values used in the data generating processes are  $\beta = 1.5, k = 1.0, a = 0.5$  and  $b = 2.5$ . Table 1 lists the means of the MLEs of the four parameters that index the Kw-P distribution along with the respective biases for sample sizes  $n = 30, n = 50$  and  $n = 100$ . The figures in Table 1 indicate that the biases of the MLEs of  $\beta, k, a$ , and  $b$  decay toward zero as the sample size increases, as expected.

Table 1: Mean estimates and biases of the MLEs of  $\beta, k, a$  and  $b$ .

$n$	Parameter	Mean	bias
30	$\beta$	0.0002	1.4998
	$k$	1.6956	-0.6956
	$a$	4.9281	-4.4281
	$b$	0.8979	1.6021
50	$\beta$	1.5002	-0.0002
	$k$	0.7398	0.2602
	$a$	0.6686	-0.1686
	$b$	2.3205	0.1795
100	$\beta$	1.5001	-0.0001
	$k$	0.9063	0.0937
	$a$	0.6923	-0.1923
	$b$	2.4386	0.0614

## 10.2. The Wheaton River data

The data correspond to the exceedances of flood peaks (in  $\text{m}^3/\text{s}$ ) of the Wheaton River near Carcross in Yukon Territory, Canada. The data consist of 72 exceedances for the years 1958–1984, rounded to one decimal place. They were analysed by Choulakian and Stephens (2001) and are listed in Table 2. The distribution is highly skewed to the left. Recently, Akinsete *et al.* (2008) and Mahmoudi (2011) analysed these data using the BP and BGP distributions, respectively. We fit the Kw-P distribution to these data and compare the results with those by fitting some of its sub-models such as the EP and Pareto distributions, as well as the non-nested BP distribution. The required numerical evaluations are implemented using the SAS (PROCNLMIXED) and R softwares.

Table 2: Exceedances of Wheaton River flood data

1.7	2.2	14.4	1.1	0.4	20.6	5.3	0.7	1.9	13.0	12.0	9.3
1.4	18.7	8.5	25.5	11.6	14.1	22.1	1.1	2.5	14.4	1.7	37.6
0.6	2.2	39.0	0.3	15.0	11.0	7.3	22.9	1.7	0.1	1.1	0.6
9.0	1.7	7.0	20.1	0.4	2.8	14.1	9.9	10.4	10.7	30.0	3.6
5.6	30.8	13.3	4.2	25.5	3.4	11.9	21.5	27.6	36.4	2.7	64.0
1.5	2.5	27.4	1.0	27.1	20.2	16.8	5.3	9.7	27.5	2.5	27.0

Tables 3 and 4 provide some descriptive statistics and the MLEs (with corresponding standard errors in parentheses) of the model parameters. Since  $x \geq \beta$ , the MLE of  $\beta$  is the first-order statistic  $x_{(1)}$ , accordingly to Akinsete *et al.* (2008). Since the values of the Akaike information criterion (AIC), Bayesian information criterion (BIC) and consistent Akaike information criterion (CAIC) are smaller for the Kw-P distribution compared with those values of the other models, the new distribution seems to be a very competitive model to these data.

Table 3: Descriptive statistics

Min.	$Q_1$	$Q_2$	Mean	$Q_3$	Max.	Var.
0.100	2.125	9.500	12.200	20.120	64.000	151.221

Table 4: MLEs of the model parameters, the corresponding SEs (given in parentheses) and the statistics AIC, BIC and CAIC

Model	Estimates				Statistic		
	$a$	$b$	$k$	$\beta$	AIC	BIC	CAIC
Kw-P	2.8553 (0.3371)	85.8468 (60.4213)	0.0528 (0.0185)	0.1 —	548.4	555.3	548.8
BP	3.1473 (0.4993)	85.7508 (0.0001)	0.0088 (0.0015)	0.1 —	573.4	580.3	573.8
EP	2.8797 (0.4911)	1 —	0.4241 (0.0463)	0.1 —	578.6	583.2	578.8
Pareto	1 —	1 —	0.2438 (0.0287)	0.1 —	608.2	610.4	608.2

Plots of the estimated pdf and cdf of the Kw-P, BP, EP and Pareto models fitted to these data are given in

Figure 5. They indicate that the Kw-P distribution is superior to the other distributions in terms of model fitting.

Table 5 gives the values of the Kolmogorov-Smirnov (K-S) statistic and of  $-2\ell(\hat{\theta})$ . From these figures, we conclude that the Kw-P distribution provides a better fit to these data than the BP, EP and Pareto models.

Table 5: K-S and  $-2\ell(\hat{\theta})$  statistics for the exceedances of flood peaks data

Model	Kw-P	BP	EP	Pareto
K-S	0.1700	0.1747	0.1987	0.3324
$-2\ell(\hat{\theta})$	542.4	567.4	574.6	606.2

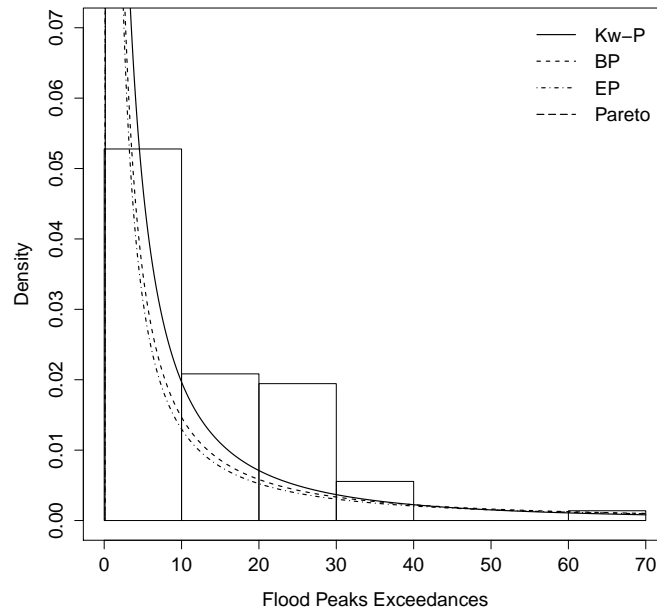
## 11. Concluding remarks

The well-known two-parameter Pareto distribution is extended by introducing two extra shape parameters, thus defining the Kumaraswamy Pareto (Kw-P) distribution having a broader class of hazard rate functions. This is achieved by taking (1) as the baseline cumulative distribution of the generalized class of Kumaraswamy distributions defined by Cordeiro and de Castro (2011). A detailed study on the mathematical properties of the new distribution is presented. The new model includes as special sub-models the Pareto and exponentiated Pareto (EP) distributions (Gupta *et al.*, 1998). We obtain the moment generating function, ordinary moments, order statistics and their moments and Rényi entropy. The estimation of the model parameters is approached by maximum likelihood and the observed information matrix is derived. An application to a real data set shows that the fit of the new model is superior to the fits of its main sub-models. We hope that the proposed model may attract wider applications in statistics.

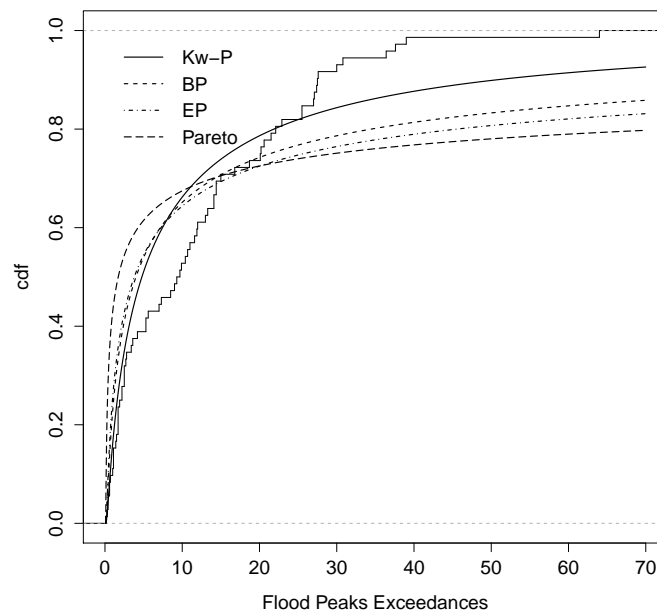
## Acknowledgments

We gratefully acknowledge the grants from CAPES and CNPq (Brazil).

1. Akinsete, A., Famoye, F. and Lee, C. (2008). The beta-Pareto distribution. *Statistics*, **42**, 6, 547–563.
2. Barreto-Souza, W., Santos, A.H.S. and Cordeiro, G.M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, **80**, 159–172.
3. Choulakian, V. and Stephens, M.A. (2001). Goodness-of-fit for the generalized Pareto distribution. *Technometrics*, **43**, 478–484.
4. Cordeiro G. M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, **81**, 883–898.
5. Cordeiro, G.M., Ortega, E.M.M. and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, **347**, 1399–1429.
6. Cordeiro, G.M., Nadarajah, S. and Ortega, E.M.M. (2012). The Kumaraswamy Gumbel distribution. *Statistical Methods and Applications*, **21**, 139–168.
7. Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics: Theory and Methods*, **31**, 497–512.
8. Gupta, R.D. and Kundu, D. (2001). Exponentiated exponential distribution: An alternative to gamma and Weibull distributions. *Biomet. J.*, **43**, 117–130.
9. Gupta, R.C., Gupta, R.D. and Gupta, P.L. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics: Theory and Methods*, **27**, 887–904.
10. Hogg, R.V., McKean, J.W. and Craig, A.T. (2005). Introduction to Mathematical Statistics, 6th ed. Pearson Prentice-Hall, New Jersey, 2005.
11. Jones, M. C. (2009). A beta-type distribution with some tractability advantages. *Statistical Methodology*, **6**, 70–81.
12. Keeping E.S. Kenney, J.F. (1962). *Mathematics of Statistics*. Part1.



(a)



(b)

Figure 5: Estimated pdf and cdf from the fitted Kw-P, BP, EP and Pareto models for the exceedances of flood peaks data.

13. Kumaraswamy, P. (1980). Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology*, **462**, 79–88.
14. Kundu, D. and Raqab, M.Z. (2005). Generalized Rayleigh distribution: Different methods of estimation. *Comput. Statist. Data Anal.*, **49**, 187–200.
15. Mahmoudi, E. (2011). The beta generalized Pareto distribution with application to lifetime data. *Mathematics and Computers in Simulation*, **81**, 11, 2414–2430.
16. Moors, J.J. (1998). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society D*, **37**, 25–32.
17. Nadarajah, S., Cordeiro, G.M. and Ortega, E.M.M. (2011). General results for the Kumaraswamy-G distribution. *Journal of Statistical Computation and Simulation*. DOI: 10.1080/00949655.2011.562504.
18. Pascoa, A.R.M., Ortega, E.M.M. and Cordeiro, G.M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, **8**, 411–433.
19. Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, **3**, 119–131.
20. Saulo, H., Leão, J. and Bourguignon, M. (2012). The Kumaraswamy Birnbaum-Saunders Distribution. *Journal of Statistical Theory and Practice*, **6**, 745–759.
21. Silva, G.O., Ortega, E.M.M. and Cordeiro, G.M. (2010). The beta modified Weibull distribution. *Lifetime Data Anal.*, **16**, 409–430.
22. Silva, R.B., Barreto-Souza, W. and Cordeiro, G.M. (2010). A new distribution with decreasing, increasing and upside-down bathtub failure rate. *Comput. Statist. Data Anal.*, **54**, 935–934.