Brazilian Journal of Probability and Statistics
2013, Vol. 27, No. 1, 31–53
DOI: 10.1214/11-BJPS149

The exponentiated Kumaraswamy distribution and its log-transform

Artur J. Lemonte[a], Wagner Barreto-Souza[a] and Gauss M. Cordeiro[b]

[a]*Universidade de São Paulo*
[b]*Universidade Federal de Pernambuco*

Abstract. The paper by Kumaraswamy (*Journal of Hydrology* **46** (1980) 79–88) introduced a probability distribution for double bounded random processes which has considerable attention in hydrology and related areas. Based on this distribution, we propose a generalization of the Kumaraswamy distribution refereed to as the exponentiated Kumaraswamy distribution. We derive the moments, moment generating function, mean deviations, Bonferroni and Lorentz curves, density of the order statistics and their moments. We also present a related distribution, so-called the log-exponentiated Kumaraswamy distribution, which extends the generalized exponential (*Aust. N. Z. J. Stat.* **41** (1999) 173–188) and double generalized exponential (*J. Stat. Comput. Simul.* **80** (2010) 159–172) distributions. We discuss maximum likelihood estimation of the model parameters. In applications to real data sets, we show that the log-exponentiated Kumaraswamy model can be used quite effectively in analyzing lifetime data.

1 Introduction

The beta distribution has been utilized extensively in statistical theory and practice for over one hundred years. The beta distribution is very flexible to model data restricted to any finite interval since it can take an amazingly great variety of forms depending on the values of the index parameters. Many of the finite range distributions encountered in practice can be easily transformed into the standard beta distribution. In econometrics, several types of data are modeled by finite range distributions. The application turns to be more interesting when the interval used is the standard unit interval $(0, 1)$, since the data can be interpreted as rates or proportions. The beta density is defined by

$$g_{\mathrm{B}}(x; a, b) = \frac{1}{B(a, b)} x^{a-1}(1-x)^{b-1}, \qquad x \in (0, 1),$$

where its two shape parameters a and b are positive and $B(\cdot, \cdot)$ is the beta function. Beta densities are unimodal, uniantimodal, increasing, decreasing or constant depending on the values of a and b.

Key words and phrases. Beta distribution, Kumaraswamy distribution, maximum likelihood estimation, mean deviation, order statistic.

Received October 2010; accepted March 2011.

Beta distributions are very versatile and a variety of uncertainties can be usefully modeled by them. This flexibility encourages its empirical use in a wide range of applications. On the other hand, Kumaraswamy (1980) argues that the beta distribution does not faithfully fit hydrological random variables such as daily rainfall, daily stream flow, etc. Also, according to Jones (2009, p. 70) "the beta distribution is fairly tractable, but in some ways not fabulously so. In particular, its distribution function is an incomplete beta function ratio and its quantile function the inverse thereof." The cumulative distribution function (c.d.f.) of the Kumaraswamy (1980) distribution, say $\mathrm{K}(\alpha, \beta)$, is given by

$$G_{\mathrm{K}}(x; \alpha, \beta) = 1 - (1 - x^{\alpha})^{\beta}, \qquad x \in (0, 1), \tag{1.1}$$

where $\alpha > 0$ and $\beta > 0$ are shape parameters. This compares extremely favorably in terms of simplicity with the beta c.d.f. which is given by the incomplete beta function ratio. The probability density function (p.d.f.) corresponding to (1.1) is

$$g_{\mathrm{K}}(x; \alpha, \beta) = \alpha\beta x^{\alpha-1}(1 - x^{\alpha})^{\beta-1}, \qquad x \in (0, 1). \tag{1.2}$$

The K density is unimodal, uniantimodal, increasing, decreasing or constant depending (in the same way as the beta distribution) on the values of its parameters. It can be shown that the K distribution has the same basic shape properties of the beta distribution (Kumaraswamy, 1980): $\alpha > 1$ and $\beta > 1$ (unimodal); $\alpha < 1$ and $\beta < 1$ (uniantimodal); $\alpha > 1$ and $\beta \leq 1$ (increasing); $\alpha \leq 1$ and $\beta > 1$ (decreasing); $\alpha = \beta = 1$ (constant). For a detailed survey of the K distribution, the reader is referred to Jones (2009).

The K distribution 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 has been widely appreciated. Jones (2009) explored the background and genesis of the K distribution and, more importantly, made clear some similarities and differences between the beta and K distributions. He highlighted several advantages of the K distribution over the beta distribution: the normalizing constant is very simple; simple explicit formulae for the distribution and quantile functions which do not involve any special functions; a simple formula for random variate generation; explicit formulae for moments of order statistics and L-moments. Further, according to Jones (2009), the beta distribution has the following advantages over the K distribution: simpler formulae for moments and moment generating function (m.g.f.); a one-parameter sub-family of symmetric distributions; simpler moment estimation and more ways of generating the distribution via physical processes.

In hydrology and related areas, for example, the K distribution has received considerable interest, see Sundar and Subbiah (1989), Fletcher and Ponnambalam (1996), Seifi et al. (2000), Ponnambalam et al. (2001) and Ganji et al. (2006). According to Nadarajah (2008), many papers in the hydrological literature have used this distribution because it is deemed as a "better alternative" to the beta distribution, see, for example, Koutsoyiannis and Xanthopoulos (1989).

It is well-known that both beta and K distributions are special cases of the three-parameter density function

$$g_{\mathrm{GB}}(x; p, q, \delta) = \frac{p}{B(q,\delta)} x^{qp-1}(1 - x^p)^{\delta-1}, \qquad x \in (0, 1), \tag{1.3}$$

where $p > 0$. This is the generalized beta distribution (McDonald, 1984). Beta and K distributions correspond to the special cases $(p, q, \delta) = (1, a, b)$ and $(p, q, \delta) = (\alpha, 1, \beta)$, respectively.

Generalized distributions have been widely studied in statistics and numerous authors have developed various types of generalizations. In this article, we propose two new distributions: the exponentiated Kumaraswamy (EK) distribution and its log-transform. Some mathematical properties of both distributions are derived and maximum likelihood estimation of their parameters is discussed.

The rest of the paper is organized as follows. In Section 2, we introduce the EK distribution. We give expansions for the moments and moment generating function in Sections 3 and 4, respectively. In Section 5, we obtain the mean deviations and Bonferroni and Lorentz curves. Section 6 deals with order statistics and their moments, L-moments and entropy. In Section 7, we discuss maximum likelihood estimation and inference. In Section 8, we introduce the log-exponentiated Kumaraswamy (log-EK) distribution, which extends the generalized exponential (Gupta and Kundu, 1999) and double generalized exponential (Barreto-Souza et al., 2010) distributions, and provide some mathematical properties. In Section 9, we apply the log-EK distribution to two real data sets to show that it can be used quite effectively in analyzing lifetime data. Finally, concluding remarks are addressed in Section 10.

2 Exponentiated Kumaraswamy distribution

The construction of the exponentiated distribution is rather simple and is based on the observation that by raising an arbitrary c.d.f. $G(x)$ to an arbitrary power $\gamma > 0$, a new c.d.f. $F(x) = G(x)^\gamma$ emerges with one additional parameter. The parameter γ characterizes the skewness, kurtosis and tails of the F distribution. In this construction, $G(x)$ is the baseline distribution and $F(x)$ may be refereed to as the exponentiated G distribution. The relation between the corresponding density functions is $f(x) = \gamma G(x)^{\gamma-1} g(x)$. We note that for $\gamma > 1$ and $\gamma < 1$ and for larger values of x, the multiplicative factor $\gamma G(x)^{\gamma-1}$ is greater and smaller than one, respectively. The reverse assertion is also true for smaller values of x. The latter immediately implies that the ordinary moments associated with the density $f(x)$ are strictly larger (smaller) than those associated with the density $g(x)$ when $\gamma > 1$ ($\gamma < 1$).

Since 1995, the exponentiated distributions have been widely studied in statistics and numerous authors have developed various classes of these distributions.

Mudholkar et al. (1995) proposed the exponentiated Weibull distribution. Its properties have been studied in more detail by Mudholkar and Hutson (1996) and Nassar and Eissa (2003). Gupta and Kundu (1999) introduced the exponentiated exponential distribution as a generalization of the standard exponential distribution. Nadarajah and Kotz (2006) proposed, based on the same idea, four more exponentiated type distributions to extend the standard gamma, standard Weibull, standard Gumbel and standard Fréchet distributions. Barreto-Souza and Cribari-Neto (2009) developed the exponentiated exponential-Poisson distribution, whereas Silva et al. (2010) proposed the exponentiated exponential-geometric distribution. More recently, Lemonte and Cordeiro (2011) introduced the exponentiated generalized inverse Gaussian distribution. Here, in the same way, we generalize the K distribution.

A random variable X has the EK distribution with parameters α, β and γ, say EK(α, β, γ), if its cumulative function is

$$F_{\text{EK}}(x; \alpha, \beta, \gamma) = [1 - (1 - x^{\alpha})^{\beta}]^{\gamma}, \qquad x \in (0, 1), \tag{2.1}$$

where α, β and γ are positive shape parameters. Clearly, if $\gamma = 1$, the EK distribution reduces to the K distribution. The EK density function can be written as

$$f_{\text{EK}}(x; \alpha, \beta, \gamma) = \alpha\beta\gamma x^{\alpha-1}(1 - x^{\alpha})^{\beta-1}[1 - (1 - x^{\alpha})^{\beta}]^{\gamma-1}, \quad x \in (0, 1). \tag{2.2}$$

The distribution (2.2) provides more options for analyzing data restricted to the interval $(0, 1)$. Plots of the EK density for selected choices of the parameters α, β, γ are given in Figure 1.

The inverse of the distribution function (2.1) yields a very simple quantile function

$$Q(y) = \{1 - (1 - y^{1/\gamma})^{1/\beta}\}^{1/\alpha}, \qquad y \in (0, 1),$$

which facilitates ready quantile-based statistical modeling (Gilchrist, 2001). In addition, $Q(y)$ gives a trivial random variable generation. If $U \sim \mathcal{U}(0, 1)$, then $X \sim$ EK(α, β, γ) is given by

$$X = \{1 - (1 - U^{1/\gamma})^{1/\beta}\}^{1/\alpha}.$$

This compares extremely favorably with the sophisticated algorithms preferred to generate random variates from the beta distribution (see, e.g., Jones, 2009).

The EK density (2.2) can be written as a linear combination of K densities. For $a > 0$ real noninteger, a series representation for $(1 - z)^{a-1}$ which holds for $|z| < 1$ is

$$(1 - z)^{a-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a)}{\Gamma(a - j) j!} z^j, \tag{2.3}$$

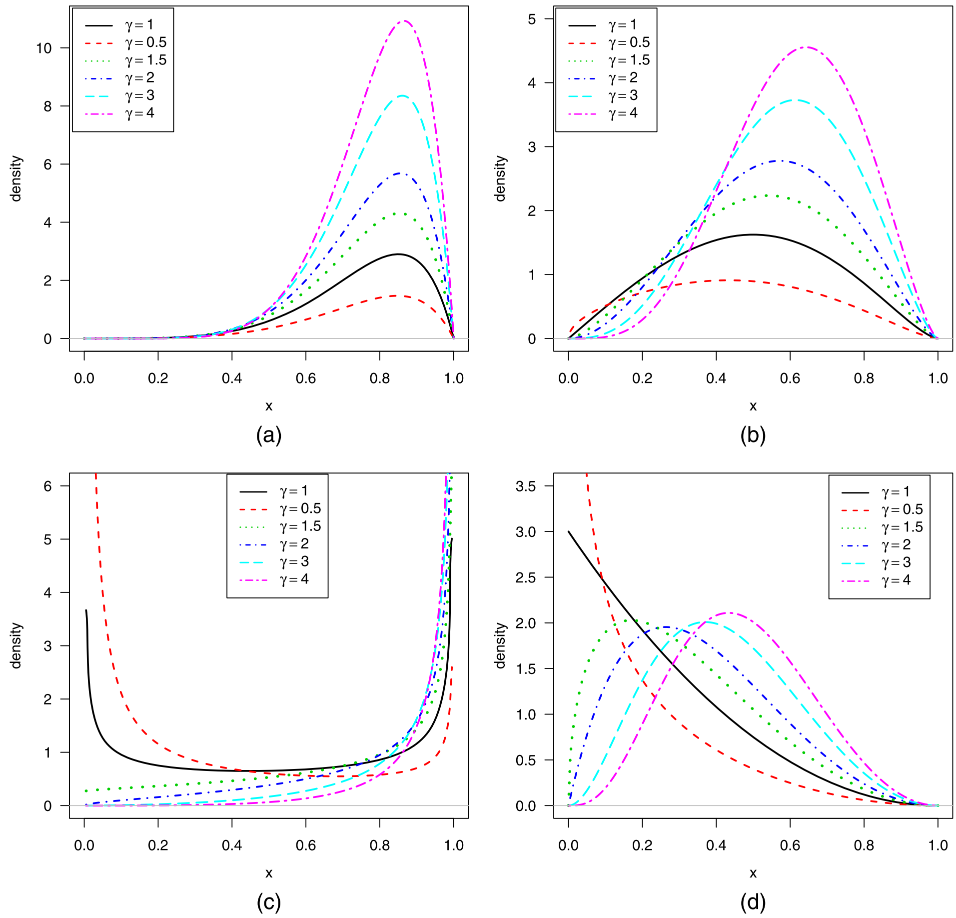

Figure 1 *EK densities for selected values of* γ: (a) $\alpha = 5$ *and* $\beta = 2$; (b) $\alpha = 2$ *and* $\beta = 2.5$; (c) $\alpha = 0.5$ *and* $\beta = 0.5$; (d) $\alpha = 1$ *and* $\beta = 3$.

where $\Gamma(\cdot)$ is the gamma function. For $\gamma > 0$ real noninteger, using the series expansion (2.3) in (2.2), yields

$$f_{\text{EK}}(x; \alpha, \beta, \gamma) = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\gamma + 1)}{\Gamma(\gamma - j)(j + 1)!} g_{\text{K}}\big(x; \alpha, \beta(j + 1)\big). \tag{2.4}$$

If $\gamma > 0$ is an integer, the index j in the above sum stops at $\gamma - 1$. Equation (2.4) means in turn that the EK density is an infinite (or finite) linear combination of K densities and, therefore, some of its mathematical properties can be obtained directly from those of the K distribution, as for example, ordinary, inverse and factorial moments, m.g.f., characteristic function, etc.

Equation (2.4) (and others expansions in this article) can be computed numerically in software such as MAPLE (Garvan, 2002), MATLAB (Sigmon and Davis,

2002) and MATHEMATICA (Wolfram, 2003). These symbolic software have currently the ability to deal with analytic expressions of formidable size and complexity. In numerical applications, a large natural number, N say, can be used in the sums instead of infinity.

3 Moments

We hardly need to emphasize the necessity and importance of the moments in any statistical analysis especially in applied work. Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness and kurtosis). Let X be a random variable with density (2.2). The rth moment of the EK distribution is $\mathrm{E}(X^r) = \alpha\beta\gamma \int_0^1 x^{r+\alpha-1}(1-x^\alpha)^{\beta-1}[1-(1-x^\alpha)^\beta]^{\gamma-1}\,\mathrm{d}x$. Setting $t = 1-(1-x^\alpha)^\beta$ in the integral, yields $\mathrm{E}(X^r) = \gamma\int_0^1 [1-(1-t)^{1/\beta}]^{r/\alpha} t^{\gamma-1}\,\mathrm{d}t$. For $r/\alpha > 0$ real noninteger, we obtain from expansion (2.3)

$$\mathrm{E}(X^r) = \gamma\Gamma\left(1+\frac{r}{\alpha}\right)\sum_{j=0}^{\infty}\frac{(-1)^j B(\gamma, 1+j/\beta)}{\Gamma(1+r/\alpha-j)\,j!}. \tag{3.1}$$

For $r/\alpha > 0$ integer, the index j in the above sum stops at r/α. If $\gamma = 1$, equation (3.1) reduces to $\mathrm{E}(X^r) = \beta B(1+r/\alpha, \beta)$, which agrees with the rth moment of the K distribution; see, for example, equation (3.5) in Jones (2009). An alternative expansion to (3.1) follows from the linear combination (2.4) (for γ real noninteger) given by

$$\mathrm{E}(X^r) = \beta\Gamma(\gamma+1)\sum_{j=0}^{\infty}\frac{(-1)^j B(1+r/\alpha, \beta(j+1))}{\Gamma(\gamma-j)\,j!}. \tag{3.2}$$

If γ is a positive integer, the upper limit in the above sum is $\gamma - 1$. This result shows a useful application of the infinite linear combination for the EK density. Plots of the skewness and kurtosis of the EK distribution as a function of γ for selected values of α and $\beta = 1.2$ are given in Figure 2.

4 Moment generating function

Let $X \sim \mathrm{EK}(\alpha, \beta, \gamma)$. The m.g.f. of X, $M(t)$ say, and characteristic function (c.h.f.), $\phi(t)$ say, are given by $M(t) = \sum_{r=0}^{\infty}\frac{t^r}{r!}\mathrm{E}(X^r)$ and $\phi(t) = \sum_{r=0}^{\infty}\frac{(\mathrm{i}t)^r}{r!}\mathrm{E}(X^r)$, respectively, where $\mathrm{i} = \sqrt{-1}$ and $\mathrm{E}(X^r)$ is obtained from (3.1) or (3.2). We now derive new representations for $M(t)$ and $\phi(t)$. The m.g.f. of X can be expressed from equation (2.4) as

$$M(t) = \sum_{j=0}^{\infty}\frac{(-1)^j\Gamma(\gamma+1)}{\Gamma(\gamma-j)(j+1)!}M_{\mathrm{K}}\big(t; \alpha, \beta(j+1)\big), \tag{4.1}$$

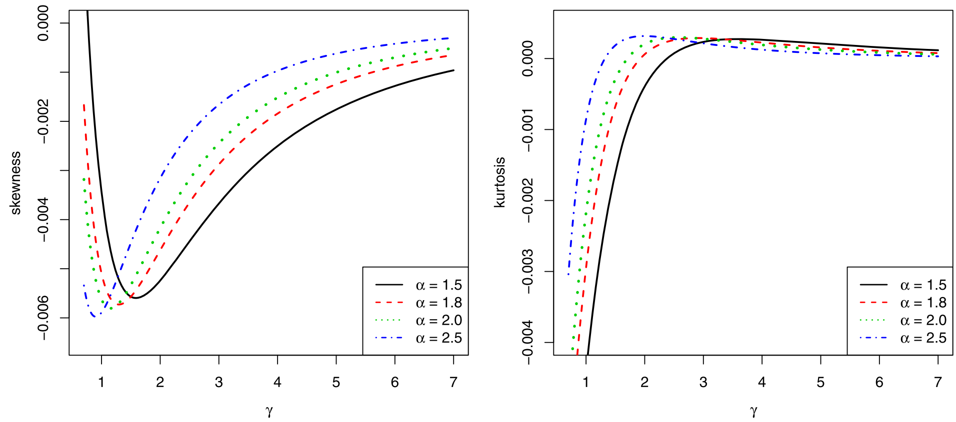

Figure 2 *Plots of the skewness and kurtosis of the EK distribution as a function of γ for some values of α and $\beta = 1.2$.*

where M_{K} is the m.g.f. of the K distribution. It comes from equation (1.2) as

$$M_{\mathrm{K}}(t) = M_{\mathrm{K}}(t;\alpha,\beta) = \alpha\beta \int_0^1 \exp(tx) x^{\alpha-1}(1-x^{\alpha})^{\beta-1}\,\mathrm{d}x.$$

Using the representation (2.3), the last equation can be further expanded as

$$M_{\mathrm{K}}(t) = \alpha\beta \sum_{s=0}^{\infty} \frac{(-1)^s \Gamma(\beta)}{\Gamma(\beta-s)s!} \int_0^1 \exp(tx) x^{\alpha(s+1)-1}\,\mathrm{d}x. \tag{4.2}$$

However, the m.g.f. of the beta distribution with parameters $(\alpha(s+1), 1)$ is just given by the above integral divided by the beta function $B(\alpha(s+1), 1)$. From this fact we conclude, after some algebra, that

$$M_{\mathrm{K}}(t) = \alpha\beta \sum_{s=0}^{\infty} \frac{(-1)^s \Gamma(\beta)}{\Gamma(\beta-s)\alpha(s+1)!} {}_1F_1\big(\alpha(s+1), \alpha(s+1)+1; t\big),$$

where ${}_1F_1$ is the confluent hypergeometric function defined by

$${}_1F_1(a,b;t) = \sum_{m=0}^{\infty} \frac{(a)_m}{(b)_m} \frac{t^m}{m!},$$

and $(a)_m$ is the Pochhammer symbol given by

$$(a)_m = a(a+1)\cdots(a+m-1) = \frac{\Gamma(a+m)}{\Gamma(a)} = \frac{(-1)^m \Gamma(1-a)}{\Gamma(1-a-m)}.$$

By combining equations (4.1) and (4.2), $M(t)$ can be expressed as

$$M(t) = \alpha\beta \sum_{j,s=0}^{\infty} \frac{(-1)^{j+s}\Gamma(\gamma+1)\Gamma(\beta(j+1))}{\Gamma(\gamma-j)\Gamma(\beta(j+1)-s)\alpha j!(s+1)!} \times {}_1F_1\big(\alpha(s+1), \alpha(s+1)+1; t\big). \tag{4.3}$$

The m.g.f. of the EK distribution is then a double infinite series of confluent hypergeometric functions. Equation (4.3) is the main result of this section. Setting $\gamma = 1$, we obtain the m.g.f. of the K distribution, which seems not to be known in the literature. Its c.h.f. of the EK distribution follows from (4.3) by substituting t by $\mathrm{i}t$.

5 Mean deviations and Bonferroni and Lorentz curves

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. If X has the EK distribution with c.d.f. $F(x)$, we can derive the mean deviations about the mean $\nu = \mathrm{E}(X)$ and about the median m from the relations $\delta_1 = \int_0^\infty |x-\nu| f(x)\,\mathrm{d}x$ and $\delta_2 = \int_0^\infty |x-m| f(x)\,\mathrm{d}x$, respectively. The median is $m = [1-(1-2^{-\gamma})^{1/\beta}]^{1/\alpha}$. Defining the integral $I(a) = \int_0^a x f(x)\,\mathrm{d}x$, these measures can be calculated from $\delta_1 = 2\nu F(\nu) - 2I(\nu)$ and $\delta_2 = \nu - 2I(m)$, where $F(\nu) = [1-(1-\nu^\alpha)^\beta]^\gamma$.

We now derive a formula to obtain the integral $I(a)$. From equation (2.4) and using (1.2) and (2.3), we obtain for $\gamma > 0$ real noninteger

$$I(a) = \alpha\beta \sum_{r,j=0}^{\infty} \frac{(-1)^{j+r}\Gamma(\gamma+1)\Gamma(\beta(j+1))a^{\alpha(r+1)+1}}{j!r!\Gamma(\gamma-j)\Gamma(\beta(j+1)-r)(\alpha(r+1)+1)}. \tag{5.1}$$

If γ is a positive integer, the index j in the above sum stops at $\gamma - 1$. The mean deviations for the K distribution follow from equation (5.1) with $\gamma = 1$.

Bonferroni and Lorenz curves have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medicine. They are defined by $B(p) = I(q)/(p\nu)$ and $L(p) = I(q)/\nu$, respectively, where $\nu = \mathrm{E}(X)$ and $q = F^{-1}(p) = Q(p) = \{1-(1-p^{1/\gamma})^{1/\beta}\}^{1/\alpha}$. These measures can be calculated immediately from equation (5.1) evaluated at $a = q$.

6 Order statistics, L-moments and entropy

We now give the density of the ith order statistic $X_{i:n}$, say $f_{i:n}(x)$, in a random sample of size n from the EK distribution. From equations (2.1) and (2.2), we can write $f_{i:n}(x)$ as

$$f_{i:n}(x) = \frac{\alpha\beta\gamma x^{\alpha-1}(1-x^\alpha)^{\beta-1}\{1-[1-(1-x^\alpha)^\beta]^\gamma\}^{n-i}}{B(i,n+1-i)\{1-(1-x^\alpha)^\beta\}^{1-\gamma i}}.$$

Using the binomial expansion, the density of the ith order statistic can be expressed as a finite linear combination of EK densities

$$f_{i:n}(x) = \frac{1}{B(i,n+1-i)} \sum_{k=0}^{n-i} \frac{(-1)^k\binom{n-i}{k}}{k+i} f_{\mathrm{EK}}\big(x;\alpha,\beta,\gamma(k+i)\big). \tag{6.1}$$

Additionally, the c.d.f. of the ith order statistic $X_{i:n}$, say $F_{i:n}(x)$, is a finite linear combination of EK cumulative functions

$$F_{i:n}(x) = \frac{1}{B(i,n+1-i)} \sum_{k=0}^{n-i} \frac{(-1)^k \binom{n-i}{k}}{k+i} F_{\mathrm{EK}}\big(x;\alpha,\beta,\gamma(k+i)\big).$$

The density of the ith order statistic $X_{i:n}$ can be represented as an infinite linear combination of K densities

$$f_{i:n}(x) = \sum_{j=0}^{\infty}\sum_{k=0}^{n-i} \frac{(-1)^{(k+j)}\binom{n-i}{k}\Gamma(\gamma(k+i)+1)}{(k+i)\Gamma(\gamma(k+i)-j)(j+1)!} \frac{g_{\mathrm{K}}(x;\alpha,\beta(j+1))}{B(i,n+1-i)}. \tag{6.2}$$

Using equation (6.2), we can provide some mathematical characteristics of the EK order statistics directly as linear functions of similar characteristics of K distributions.

From equations (3.2) and (6.1), the rth moment of the ith order statistic $X_{i:n}$ is given by

$$\mathrm{E}(X_{i:n}^r) = \beta \sum_{j=0}^{\infty}\sum_{k=0}^{n-i} \frac{(-1)^{j+k}\binom{n-i}{k}\Gamma(\gamma(k+i)+1)}{\Gamma(\gamma(k+i)-j)(k+i)j!} \frac{B(1+r/\alpha,\beta(j+1))}{B(i,n+1-i)}. \tag{6.3}$$

The L-moments (Hosking, 1990) 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

$$\tau_{r+1} = (r+1)^{-1}\sum_{k=0}^{r}(-1)^k\binom{r}{k}\mathrm{E}(X_{r+1-k:r+1}), \qquad r=0,1,2,\ldots.$$

The first four L-moments are: $\tau_1 = \mathrm{E}(X_{1:1})$, $\tau_2 = \frac{1}{2}\mathrm{E}(X_{2:2} - X_{1:2})$, $\tau_3 = \frac{1}{3}\mathrm{E}(X_{3:3} - 2X_{2:3} + X_{1:3})$ and $\tau_4 = \frac{1}{4}\mathrm{E}(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$. 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 (6.3), we can obtain expansions for the L-moments of the EK distribution.

The entropy of a random variable is a measure of uncertainty variation. The Rényi entropy is defined as $I_R(\delta) = (1-\delta)^{-1}\log\{\int_{\mathbb{R}} f^{\delta}(x)\,\mathrm{d}x\}$, where $\delta > 0$ and $\delta \neq 1$. For $\delta(\gamma-1)+1 > 0$ real noninteger (if integer, we consider the binomial expansion), using the series representation (2.3), we can write

$$f_{\mathrm{EK}}^{\delta}(x) = (\alpha\beta\gamma)^{\delta} x^{\delta(\alpha-1)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\delta(\gamma-1)+1)}{\Gamma(\delta(\gamma-1)+1-j)j!}(1-x^{\alpha})^{\delta(\beta-1)+\beta j}.$$

The density $f_{\mathrm{EK}}^{\delta}(x)$ can be expressed as an infinite (a finite) mixture of generalized beta densities as defined in equation (1.3)

$$f_{\mathrm{EK}}^{\delta}(x) = (\beta\gamma)^{\delta}\alpha^{\delta-1}\sum_{j=0}^{\infty}\frac{(-1)^{j}\Gamma(\delta(\gamma-1)+1)}{\Gamma(\delta(\gamma-1)+1-j)j!} \times \frac{g_{\mathrm{GB}}(x;\alpha,\delta(1-1/\alpha)+1/\alpha,\delta(\beta-1)+\beta j+1)}{[B(\delta(1-1/\alpha)+1/\alpha,\delta(\beta-1)+\beta j+1)]^{-1}}.$$

Under these conditions, the Rényi entropy of the EK distribution follows as

$$I_R(\delta) = \frac{1}{1-\delta}\log\Bigg\{(\beta\gamma)^{\delta}\alpha^{\delta-1}\sum_{j=0}^{\infty}\frac{(-1)^{j}\Gamma(\delta(\gamma-1)+1)}{\Gamma(\delta(\gamma-1)+1-j)j!} \times B\big(\delta(1-1/\alpha)+1/\alpha,\delta(\beta-1)+\beta j+1\big)\Bigg\}.$$

For furthers details, see Song (2001).

7 Estimation and inference

The estimation of the model parameters is investigated by the method of maximum likelihood. Let $\mathbf{x} = (x_1,\ldots,x_n)^{\top}$ be a random sample of the EK distribution with unknown parameter vector $\boldsymbol{\theta} = (\alpha,\beta,\gamma)^{\top}$. The log-likelihood function $\ell(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$ is

$$\ell(\boldsymbol{\theta}) = n\log(\alpha\beta\gamma) + (\alpha-1)\sum_{i=1}^{n}\log(x_i) + (\beta-1)\sum_{i=1}^{n}\log(1-x_i^{\alpha}) + (\gamma-1)\sum_{i=1}^{n}\log\{1-(1-x_i^{\alpha})^{\beta}\}.$$

The components of the score vector $\mathbf{U}(\boldsymbol{\theta}) = (U_\alpha, U_\beta, U_\gamma)^{\top}$ are

$$U_\alpha = \frac{n}{\alpha} + \sum_{i=1}^{n}\log(x_i) + (1-\beta)\sum_{i=1}^{n}\frac{x_i^{\alpha}\log(x_i)}{1-x_i^{\alpha}} + \beta(\gamma-1)\sum_{i=1}^{n}\frac{x_i^{\alpha}(1-x_i^{\alpha})^{\beta-1}\log(x_i)}{1-(1-x_i^{\alpha})^{\beta}},$$

$$U_\beta = \frac{n}{\beta} + \sum_{i=1}^{n}\log(1-x_i^{\alpha}) + (1-\gamma)\sum_{i=1}^{n}\frac{(1-x_i^{\alpha})^{\beta}\log(1-x_i^{\alpha})}{1-(1-x_i^{\alpha})^{\beta}}$$

and

$$U_\gamma = \frac{n}{\gamma} + \sum_{i=1}^{n}\log\{1-(1-x_i^{\alpha})^{\beta}\}.$$

Setting these expressions to zero, $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$, and solving them simultaneously yields the maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma})^\top$ of the three parameters. These equations cannot be solved analytically and statistical software can be used to solve them numerically via iterative techniques such as the Newton–Raphson algorithm.

For interval estimation of α, β and γ, and tests of hypotheses on these parameters, we obtain the observed information matrix, since the expected information matrix is very complicated and requires numerical integration. The 3×3 observed information matrix $\mathbf{J}(\boldsymbol{\theta})$ is obtained in the form $\mathbf{J}(\boldsymbol{\theta}) = -\partial^2 \ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^\top$, whose elements are given in Appendix A. The multivariate normal $\mathcal{N}_3(\mathbf{0}, \mathbf{J}(\widehat{\boldsymbol{\theta}})^{-1})$ distribution can be used to construct approximate confidence intervals and confidence regions for the parameters. In fact, asymptotic $100(1-\eta)\%$ confidence intervals for α, β and γ are given, respectively, by $\widehat{\alpha} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{\alpha})]^{1/2}$, $\widehat{\beta} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{\beta})]^{1/2}$ and $\widehat{\gamma} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{\gamma})]^{1/2}$, where $\mathrm{var}(\cdot)$ is the diagonal element of $\mathbf{J}(\widehat{\boldsymbol{\theta}})^{-1}$ corresponding to each parameter, and $z_{\eta/2}$ is the quantile $(1-\eta/2)$ of the standard normal distribution.

The likelihood ratio (LR) statistic is useful for testing the goodness-of-fit of the EK model and for comparing it with the K model. We can easily check if the fit using the EK model is statistically "superior" to a fit using the K model for a given data set by computing $w = 2\{\ell(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}) - \ell(\widetilde{\alpha}, \widetilde{\beta}, 1)\}$, where $\widehat{\alpha}, \widehat{\beta}$ and $\widehat{\gamma}$ are the unrestricted MLEs and $\widetilde{\alpha}$ and $\widetilde{\beta}$ are the restricted estimates. Also, w is asymptotically distributed under the null model as χ_1^2. The LR test rejects the null hypothesis if $w > \xi_\eta$, where ξ_η denotes the upper $100\eta\%$ point of the χ_1^2 distribution.

8 The log-EK distribution

For the first time, we introduce the log-exponentiated Kumaraswamy (log-EK) distribution which can be useful to model lifetime data. If $X \sim \mathrm{EK}(\alpha, \beta, \gamma)$, we define the standard log-EK distribution by $Y = -\log(1-X)$. Its density function is

$$g_{\text{log-EK}}(y;\alpha,\beta,\gamma) = \alpha\beta\gamma \frac{\mathrm{e}^{-y}(1-\mathrm{e}^{-y})^{\alpha-1}[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta-1}}{\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{1-\gamma}}, \quad y > 0. \tag{8.1}$$

The corresponding c.d.f. and hazard rate function are, respectively,

$$G_{\text{log-EK}}(y;\alpha,\beta,\gamma) = \{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma}$$

and

$$h(y) = \alpha\beta\gamma \frac{\mathrm{e}^{-y}(1-\mathrm{e}^{-y})^{\alpha-1}[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta-1}\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma-1}}{1-\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma}}.$$

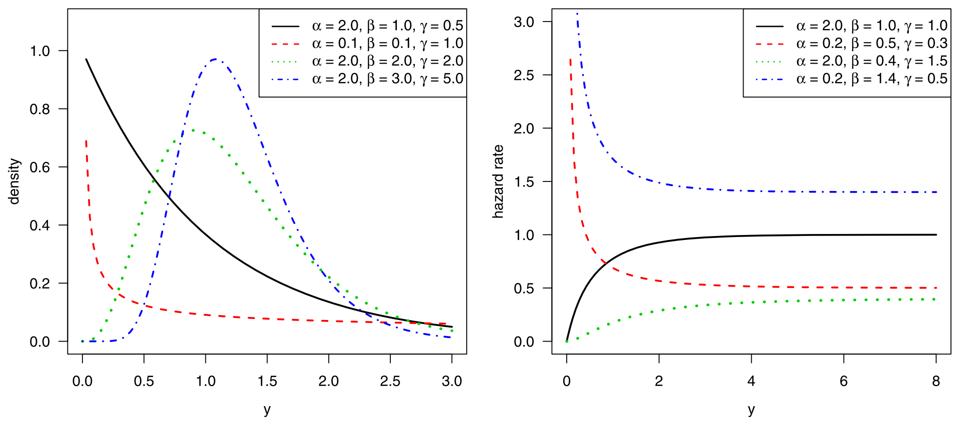

Figure 3 *Plots of the density and hazard rate functions of the standard log-EK distribution.*

We introduce a scale parameter $\sigma > 0$ in the density (8.1) using the scale transformation $Z = \sigma Y$. We say that Z has the log-EK distribution and denote by $Z \sim \text{log-EK}(\sigma, \alpha, \beta, \gamma)$. The log-EK distribution is very flexible and it contains the following special sub-models:

- Double generalized exponential (DGE) distribution (Barreto-Souza et al., 2010) for $\gamma = 1$;
- Generalized exponential (GE) distributions (Gupta and Kundu, 1999), namely $\text{GE}(\sigma^{-1}, \alpha\beta)$, $\text{GE}(\sigma^{-1}\beta, \gamma)$ and $\text{GE}(\sigma^{-1}, \alpha)$ for the choices $\gamma = 1$, $\alpha = 1$ and $\beta = \gamma = 1$, respectively.

Plots of the density and hazard rate functions of the standard log-EK distribution for selected parameter values are given in Figure 3.

8.1 Order statistics

Here, we demonstrate that the log-EK density can be expressed as an infinite (or finite) linear combination of beta generalized exponential (BGE) densities. The BGE distribution was first introduced and studied by Barreto-Souza et al. (2010). It is denoted by $\text{BGE}(a, b, \lambda, \alpha)$ (for $a, b, \lambda, \alpha > 0$) for which the density function is

$$f_{\text{BGE}}(w; a, b, \lambda, \alpha) \\ = \frac{\alpha\lambda}{B(a,b)} \mathrm{e}^{-\lambda z}(1 - \mathrm{e}^{-\lambda w})^{\alpha a - 1}\{1 - (1 - \mathrm{e}^{-\lambda w})^{\alpha}\}^{b-1}, \qquad w > 0.$$

For $\gamma > 0$ real noninteger, we use equation (2.3) to obtain

$$\{1 - [1 - (1 - \mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma - 1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\gamma)}{\Gamma(\gamma - j) j!} [1 - (1 - \mathrm{e}^{-y})^{\alpha}]^{\beta j}.$$

Hence, the density (8.1) can be expressed as an infinite (or finite) linear combination of BGE densities

$$g_{\text{log-EK}}(y;\alpha,\beta,\gamma)=\sum_{j=0}^{\infty}\frac{(-1)^j\Gamma(\gamma+1)}{\Gamma(\gamma-j)(j+1)!}f_{\text{BGE}}\big(y;1,\beta(j+1),1,\alpha\big),\qquad(8.2)$$

with $y>0$. If $\gamma>0$ is integer, the upper limit in the above sum is $\gamma-1$. So, a treatment of some mathematical properties of the log-EK distribution can be easily derived from similar properties of the BGE distribution.

Let $Y_1,\ldots,Y_n$ be a random sample following the log-EK distribution with parameters $\alpha,\beta,\gamma>0$. The density of the ith order statistic is given by

$$\begin{aligned}g_{i:n}(y)=\alpha\beta\gamma\mathrm{e}^{-y}(1-\mathrm{e}^{-y})^{\alpha-1}[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta-1}\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma i-1}\\ /\big(B(i,n+i-1)\big\{1-\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma}\big\}^{i-n}\big),\end{aligned}$$

with $y>0$, for $i=1,\ldots,n$. From the binomial expansion, we obtain

$$\big\{1-\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma}\big\}^{n-i}=\sum_{k=0}^{n-i}(-1)^k\binom{n-i}{k}\{1-[1-(1-\mathrm{e}^{-y})^{\alpha}]^{\beta}\}^{\gamma k}.$$

Thus, the density of the ith order statistic can be rewritten as a finite linear combination of log-EK densities

$$g_{i:n}(y)=\frac{1}{B(i,n+1-i)}\sum_{k=0}^{n-i}\frac{(-1)^k}{k+i}\binom{n-i}{k}g_{\text{log-EK}}\big(y;\alpha,\beta,\gamma(k+i)\big),\qquad(8.3)$$

where $y>0$.

Alternatively, combining equations (8.2) and (8.3), the density of the log-EK order statistics can be rewritten as an infinite linear combination of BGE densities

$$\begin{aligned}g_{i:n}(y)=\sum_{j=0}^{\infty}\sum_{k=0}^{n-i}\frac{(-1)^{k+j}\Gamma(\gamma(k+i)+1)\binom{n-i}{k}}{(k+i)\Gamma(\gamma(k+i)-j)(j+1)!}\\ \times\frac{f_{\text{BGE}}(y;1,\beta(j+1),1,\alpha)}{B(i,n+1-i)}.\end{aligned}\qquad(8.4)$$

Some of the results for the BGE distribution (Barreto-Souza et al., 2010) are readily applicable to the distributions of the log-EK order statistics using equation (8.4).

8.2 Moments

Consider the random variable $Y\sim\text{log-EK}(\alpha,\beta,\gamma)$, with γ and β real non-integers. We now provide expansions for the moments and m.g.f. of a random variable following the standard log-EK distribution. Using (8.2) and the results by Barreto-Souza et al. (2010), the m.g.f. of Y can be expressed as

$$\mathrm{E}(e^{tY}) = \alpha\beta\Gamma(\gamma+1)\sum_{k,j=0}^{\infty}\frac{(-1)^{j+k}\Gamma(\beta(j+1))B(1-t,k+1)}{\Gamma(\gamma-j)\Gamma(\beta(j+1)-k)j!k!} \tag{8.5}$$

for $t<1$. Hence, the rth moment of the standard log-EK distribution can be obtained from $\mathrm{E}(Y^r) = \mathrm{d}^r\mathrm{E}(\mathrm{e}^{tY})/\mathrm{d}t^r|_{t=0}$. Thus, it follows for $\gamma>0$ and $\beta>0$ real nonintegers that

$$\mathrm{E}(Y^r) = \alpha\beta\Gamma(\gamma+1) \times \sum_{k,j=0}^{\infty}\frac{(-1)^{j+k+r}\Gamma(\beta(j+1))}{\Gamma(\gamma-j)\Gamma(\beta(j+1)-k)j!k!}\frac{\mathrm{d}^r B(p,\alpha(k+1))}{\mathrm{d}p^r}\bigg|_{p=1}. \tag{8.6}$$

If $\gamma>0$ is integer, the index j in the sums (8.5) and (8.6) stops at $\gamma-1$. If $\beta>0$ is integer, for each j, the index k in the sums (8.5) and (8.6) k stops at $\beta(j+1)-1$. Plots of the skewness and kurtosis of the standard log-EK distribution are given in Figure 4 for selected parameter values.

Expressions for the m.g.f. and moments of the ith order statistic from a random sample of the log-EK distribution can be easily obtained from equation (8.3) and some results of this section. For example, the rth moment of the ith order statistic is given by

$$\mathrm{E}(Y_{i:n}^r) = \frac{\alpha\beta}{B(i,n+1-i)}\sum_{j,k=0}^{\infty}\sum_{k=0}^{n-i}\frac{(-1)^{l+j+k+r}\binom{n-i}{k}\Gamma(\gamma(l+i)+1)}{\Gamma(\gamma(l+i)-j)\Gamma(\beta(j+1)-k)} \times \frac{\Gamma(\beta(j+1))}{(k+i)j!k!}\frac{\mathrm{d}^r B(p,\alpha(k+1))}{\mathrm{d}p^r}\bigg|_{p=1}.$$

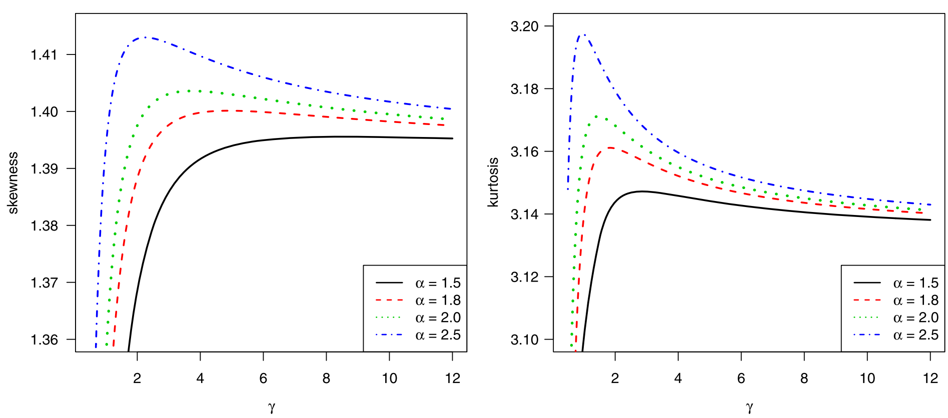

Figure 4 *Plots of the skewness and kurtosis of the standard log-EK distribution as a function of γ for selected values of α and $\beta = 1.2$.*

8.3 Rényi entropy

The log-EK density, if $\delta(\gamma - 1) + 1 > 0$ is real noninteger (if integer, we use the binomial expansion), has the series representation

$$g_{\text{log-EK}}^{\delta}(y) = (\alpha\beta\gamma)^{\delta} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\delta(\gamma - 1) + 1)}{\Gamma(\delta(\gamma - 1) + 1 - j) j!} \times \mathrm{e}^{-\delta y} (1 - \mathrm{e}^{-y})^{\delta(\alpha - 1)} [1 - (1 - \mathrm{e}^{-y})^{\alpha}]^{(\beta - 1)\delta + \beta j}.$$

Setting $y = -\log(1 - u)$, the Rényi entropy of the standard log-EK distribution reduces to

$$I_R(\delta) = (1 - \delta)^{-1} \log\left\{ (\alpha\beta\gamma)^{\delta} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\delta(\gamma - 1) + 1)}{\Gamma(\delta(\gamma - 1) + 1 - j) j!} c_j(\alpha, \beta, \delta) \right\}.$$

Here, $c_j(\alpha, \beta, \delta) = \int_0^1 (1 - u^{\alpha})^{\delta(\beta - 1) + \beta j} u^{\delta(\alpha - 1)} (1 - u)^{\delta - 1} \, \mathrm{d}u$. We can easily calculate $c_j(\alpha, \beta, \delta)$ from the integral of the beta distribution. We have

$$c_j(\alpha, \beta, \delta) = \Gamma\big(\delta(\beta - 1) + \beta j + 1\big) \sum_{r=0}^{\infty} \frac{(-1)^r B(\delta(\alpha - 1) + r\alpha + 1, \delta)}{\Gamma(\delta(\beta - 1) + \beta j + 1 - r) r!},$$

and then we obtain the Rényi entropy as desired. If $\delta(\beta - 1) + \beta j + 1$ is an integer, the index r in the above sum stops at $\delta(\beta - 1) + \beta j + 1$.

8.4 Maximum likelihood estimation

Let $\mathbf{z} = (z_1, \ldots, z_n)^{\top}$ be a random sample of the log-EK distribution with unknown parameter vector $\boldsymbol{\theta} = (\sigma, \alpha, \beta, \gamma)^{\top}$. The log-likelihood function for $\boldsymbol{\theta}$ is

$$\ell(\boldsymbol{\theta}) = n \log\left(\frac{\alpha\beta\gamma}{\sigma}\right) + (\alpha - 1) \sum_{i=1}^{n} \log(1 - \mathrm{e}^{-z_i/\sigma}) - \frac{1}{\sigma} \sum_{i=1}^{n} z_i + (\beta - 1) \sum_{i=1}^{n} \log\big(1 - (1 - \mathrm{e}^{-z_i/\sigma})^{\alpha}\big) + (\gamma - 1) \sum_{i=1}^{n} \log\big\{1 - \big(1 - (1 - \mathrm{e}^{-z_i/\sigma})^{\alpha}\big)^{\beta}\big\}.$$

The components of the score vector $\mathbf{U}(\boldsymbol{\theta}) = (U_{\sigma}, U_{\alpha}, U_{\beta}, U_{\gamma})^{\top}$ are

$$U_{\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^2} \sum_{i=1}^{n} z_i - \frac{(\alpha - 1)}{\sigma^2} \sum_{i=1}^{n} \frac{z_i \mathrm{e}^{-z_i/\sigma}}{1 - \mathrm{e}^{-z_i/\sigma}} + \frac{\alpha(\beta - 1)}{\sigma^2} \sum_{i=1}^{n} \frac{z_i \mathrm{e}^{-z_i/\sigma} (1 - \mathrm{e}^{-z_i/\sigma})^{\alpha - 1}}{1 - (1 - \mathrm{e}^{-z_i/\sigma})^{\alpha}}$$

$$- \frac{\alpha\beta(\gamma-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i e^{-z_i/\sigma}(1-e^{-z_i/\sigma})^{\alpha-1}[1-(1-e^{-z_i/\sigma})^{\alpha}]^{\beta-1}}{1-[1-(1-e^{-z_i/\sigma})^{\alpha}]^{\beta}},$$

$$U_\alpha = \frac{n}{\alpha} + \sum_{i=1}^{n}\log(1-e^{-z_i/\sigma}) + (1-\beta)\sum_{i=1}^{n}\frac{(1-e^{-z_i/\sigma})^{\alpha}\log(1-e^{-z_i/\sigma})}{1-(1-e^{-z_i/\sigma})^{\alpha}}$$
$$+ \beta(\gamma-1)\sum_{i=1}^{n}\frac{(1-e^{-z_i/\sigma})^{\alpha}(1-(1-e^{-z_i/\sigma})^{\alpha})^{\beta-1}\log(1-e^{-z_i/\sigma})}{1-(1-(1-e^{-z_i/\sigma})^{\alpha})^{\beta}},$$

$$U_\beta = \frac{n}{\beta} + \sum_{i=1}^{n}\log(1-(1-e^{-z_i/\sigma})^{\alpha})$$
$$+ (1-\gamma)\sum_{i=1}^{n}\frac{(1-(1-e^{-z_i/\sigma})^{\alpha})^{\beta}\log(1-(1-e^{-z_i/\sigma})^{\alpha})}{1-(1-(1-e^{-z_i/\sigma})^{\alpha})^{\beta}}$$

and

$$U_\gamma = \frac{n}{\gamma} + \sum_{i=1}^{n}\log\{1-(1-(1-e^{-z_i/\sigma})^{\alpha})^{\beta}\}.$$

The MLE $\widehat{\boldsymbol\theta} = (\widehat\sigma, \widehat\alpha, \widehat\beta, \widehat\gamma)^\top$ of $\boldsymbol\theta = (\sigma, \alpha, \beta, \gamma)^\top$ is obtained by solving the nonlinear equations $\mathbf{U}(\boldsymbol\theta) = \mathbf{0}$. These equations cannot be solved analytically but statistical software can be used to solve them numerically.

The approximate multivariate normal $\mathcal{N}_4(\mathbf{0}, \mathbf{J}(\widehat{\boldsymbol\theta})^{-1})$ distribution can be used to construct confidence intervals for the parameters (see discussion of Section 7). The elements of the 4×4 observed information matrix $\mathbf{J}(\boldsymbol\theta) = -\partial^2\ell(\boldsymbol\theta)/\partial\boldsymbol\theta\,\partial\boldsymbol\theta^\top$ is given in Appendix B. In addition, we can compare the fits of the log-EK distribution with its sub-models for a given data set via the LR test. For example, to test $\sigma = 1$, the LR statistic becomes $w = 2\{\ell(\widehat\sigma, \widehat\alpha, \widehat\beta, \widehat\gamma) - \ell(1, \widetilde\alpha, \widetilde\beta, \widetilde\gamma)\}$, where $\widehat\sigma$, $\widehat\alpha$, $\widehat\beta$ and $\widehat\gamma$ are the unrestricted estimates and $\widetilde\alpha$, $\widetilde\beta$ and $\widetilde\gamma$ are the restricted estimates. The LR test rejects the null hypothesis if $w > \xi_\eta$, where ξ_η denotes the upper $100\eta\%$ point of the χ^2_1 distribution.

9 Applications

In this section, we compare the results of fitting the log-EK$(\sigma, \alpha, \beta, \gamma)$, DGE$(\sigma, \alpha, \beta)$ and GE(σ, α) distributions to two real data sets. All the computations were done using the R programming language (R Development Core Team, 2009).

First, we consider the data set consisting of the length of intervals between the times at which vehicles pass a point on a road. The data are given in Table 1, and their source is Jørgensen (1982). The MLEs of the model parameters (standard errors in parentheses) and the values of the AIC (Akaike Information Criterion), CAIC (Consistent Akaike Information Criterion) and BIC (Bayesian Information

Table 1 *Traffic data*

2.50	2.60	2.60	2.70	2.80	2.80	2.90	3.00	3.00	3.10	3.20	3.40
3.70	3.90	3.90	3.90	4.60	4.70	5.00	5.60	5.70	6.00	6.00	6.10
6.60	6.90	6.90	7.30	7.60	7.90	8.00	8.30	8.80	8.80	9.30	9.40
9.50	10.1	11.0	11.3	11.9	11.9	12.3	12.9	12.9	13.0	13.8	14.5
14.9	15.3	15.4	15.9	16.2	17.6	20.1	20.3	20.6	21.4	22.8	23.7
23.7	24.7	29.7	30.6	31.0	34.1	34.7	36.8	40.1	40.2	41.3	42.0
44.8	49.8	51.7	55.7	56.5	58.1	70.5	72.6	87.1	88.6	91.7	119.8

Criterion) are given in Table 2. From the values of these statistics, we note that the log-EK model is better than the DGE and GE models fitted to these data. The LR statistic to test the hypotheses H_0: DGE against H_1: log-EK and H_0: GE against H_1: log-EK are 12.606 (p-value: < 0.001) and 16.5324 (p-value: < 0.001), respectively. Thus, we reject the null hypotheses in favor of the log-EK distribution using any usual significance level. Therefore, the log-EK distribution is significantly better than the DGE and GE distributions based on the LR statistic.

As a second application, we analyze a real data set on the active repair times (hours) for an airborne communication transceiver. The data are given in Table 3, and their source is Jørgensen (1982). Table 4 gives the MLEs of the model parameters (standard errors in parentheses) and the values of AIC, CAIC and BIC. Again, from the values of these statistics, we conclude that the log-EK model provides a better fit to these data than the DGE and GE models. The LR statistic to test the hypotheses H_0: DGE against H_1: log-EK and H_0: GE against H_1: log-EK are 6.428 (p-value: 0.011) and 8.3712 (p-value: 0.015), respectively. Thus, the log-EK distribution is significantly better than the DGE and GE distributions.

Plots of the estimated densities of the log-EK, DGE and GE models fitted for the data sets corresponding to Table 1 and 3, respectively, are given in Figure 5. These plots give evidence that the log-EK distribution is superior to the DGE and GE distributions in terms of model fit.

Table 2 *MLEs of the model parameters*; *traffic data*

	Estimates				Statistic		
Distribution	σ	α	β	γ	AIC	CAIC	BIC
log-EK	3.2817 (0.1176)	7.7912 (3.2225)	0.1101 (0.0256)	0.4225 (0.0889)	675.13	675.64	672.68
DGE	3.4139 (0.1071)	2.0049 (0.5296)	0.1759 (0.0209)		685.74	686.04	683.04
GE	20.0627 (2.8509)	1.1198 (0.1672)			687.67	687.81	685.05

Table 3 *Active repair times (hours)*

0.50	0.60	0.60	0.70	0.70	0.70	0.80	0.80
1.00	1.00	1.00	1.00	1.10	1.30	1.50	1.50
1.50	1.50	2.00	2.00	2.20	2.50	2.70	3.00
3.00	3.30	4.00	4.00	4.50	4.70	5.00	5.40
5.40	7.00	7.50	8.80	9.00	10.20	22.00	24.50

10 Concluding remarks

We have introduced a three parameter distribution, so-called the exponentiated Kumaraswamy distribution, as a simple extension of the Kumaraswamy distribution (Kumaraswamy, 1980). We also proposed a related distribution, refereed to as the log-exponentiated Kumaraswamy (log-EK) distribution, which extends the generalized exponential (Gupta and Kundu, 1999) and double generalized exponential (Barreto-Souza et al., 2010) distributions. We provide a mathematical treatment of both distributions including moments, moment generating function, densities of the order statistics and their moments. We discuss maximum likelihood estimation of the parameters and obtain the observed information matrix for both models. Applications of the log-EK distribution to a real data sets show that this distribution can yield a better fit than some known models. We hope that this generalization may attract wider applications in reliability and lifetime analysis.

Appendix A

The elements of the observed information matrix

$$\mathbf{J}(\boldsymbol{\theta}) = -\begin{pmatrix} U_{\alpha\alpha} & U_{\alpha\beta} & U_{\alpha\gamma} \\ \cdot & U_{\beta\beta} & U_{\beta\gamma} \\ \cdot & \cdot & U_{\gamma\gamma} \end{pmatrix},$$

Table 4 *MLEs of the model parameters; active repair times (hours)*

	Estimates				Statistic		
Distribution	σ	α	β	γ	AIC	CAIC	BIC
log-EK	0.6667 (0.1653)	5.1479 (2.5172)	0.1366 (0.0397)	0.5072 (0.0209)	190.55	191.69	188.10
DGE	0.7219 (0.0601)	1.5906 (0.6136)	0.1995 (0.0361)		194.97	195.64	192.27
GE	3.7349 (0.7820)	1.1137 (0.2446)			194.92	195.24	192.30

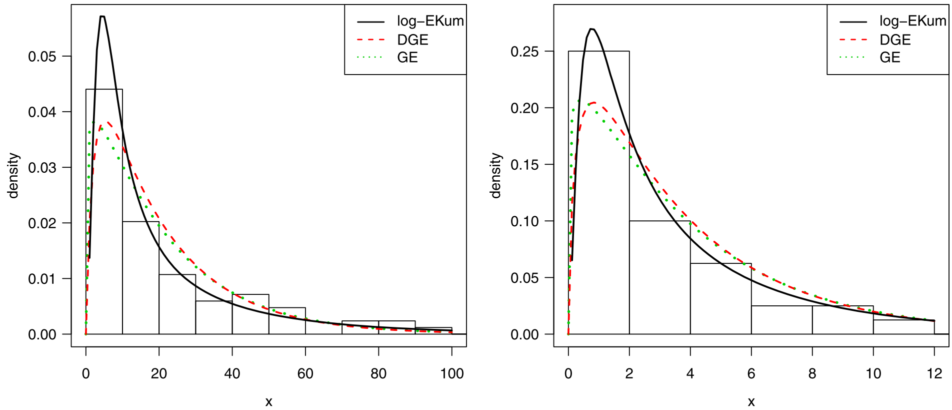

Figure 5 *Fitted densities of the log-EK, DGE and GE distributions for the data sets corresponding to Tables 1 and 3, respectively.*

are given by

$$U_{\alpha\alpha} = -\frac{n}{\alpha^2} - (\beta - 1)\sum_{i=1}^{n} \frac{x_i^{\alpha}[\log(x_i)]^2}{(1 - x_i^{\alpha})^2} + \beta(\gamma - 1)\sum_{i=1}^{n} \frac{x_i^{\alpha}(1 - x_i^{\alpha})^{\beta-2}[\log(x_i)]^2}{1 - (1 - x_i^{\alpha})^{\beta}} - \beta^2(\gamma - 1)\sum_{i=1}^{n} \frac{x_i^{2\alpha}(1 - x_i^{\alpha})^{\beta-2}[\log(x_i)]^2}{[1 - (1 - x_i^{\alpha})^{\beta}]^2},$$

$$U_{\alpha\beta} = -\sum_{i=1}^{n} \frac{x_i^{\alpha}\log(x_i)}{1 - x_i^{\alpha}} + (\gamma - 1)\sum_{i=1}^{n} \frac{x_i^{\alpha}(1 - x_i^{\alpha})^{\beta-1}\log(x_i)}{1 - (1 - x_i^{\alpha})^{\beta}} + \beta(\gamma - 1)\sum_{i=1}^{n} \frac{x_i^{\alpha}(1 - x_i^{\alpha})^{\beta-1}\log(x_i)\log(1 - x_i^{\alpha})}{[1 - (1 - x_i^{\alpha})^{\beta}]^2},$$

$$U_{\alpha\gamma} = \beta\sum_{i=1}^{n} \frac{x_i^{\alpha}(1 - x_i^{\alpha})^{\beta-1}\log(x_i)}{1 - (1 - x_i^{\alpha})^{\beta}},$$

$$U_{\beta\beta} = -\frac{n}{\beta^2} + (1 - \gamma)\sum_{i=1}^{n} \frac{(1 - x_i^{\alpha})^{\beta}(\log(1 - x_i^{\alpha}))^2}{[1 - (1 - x_i^{\alpha})^{\beta}]^2},$$

$$U_{\beta\gamma} = -\sum_{i=1}^{n} \frac{(1 - x_i^{\alpha})^{\beta}\log(1 - x_i^{\alpha})}{1 - (1 - x_i^{\alpha})^{\beta}} \quad \text{and} \quad U_{\gamma\gamma} = -\frac{n}{\gamma^2}.$$

Appendix B

The elements of the observed information matrix

$$\mathbf{J}(\boldsymbol{\theta}) = -\begin{pmatrix} U_{\sigma\sigma} & U_{\sigma\alpha} & U_{\sigma\beta} & U_{\sigma\gamma} \\ \cdot & U_{\alpha\alpha} & U_{\alpha\beta} & U_{\alpha\gamma} \\ \cdot & \cdot & U_{\beta\beta} & U_{\beta\gamma} \\ \cdot & \cdot & \cdot & U_{\gamma\gamma} \end{pmatrix},$$

are given by

$$\begin{aligned}
U_{\sigma\sigma} = {} & \frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{n} z_i + \frac{2(\alpha-1)}{\sigma^3}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}}{1-\mathrm{e}^{-z_i/\sigma}} - \frac{(\alpha-1)}{\sigma^4}\sum_{i=1}^{n}\frac{z_i^2\mathrm{e}^{-z_i/\sigma}}{(1-\mathrm{e}^{-z_i/\sigma})^2} \\
& - \frac{2\alpha(\beta-1)}{\sigma^3}\sum_{i=1}^{n}\frac{z_i\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}}{1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}} \\
& - \frac{\alpha^2(\beta-1)}{\sigma^4}\sum_{i=1}^{n}\frac{z_i^2\mathrm{e}^{-2z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-2}}{[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^2} \\
& + \frac{\alpha(\beta-1)}{\sigma^4}\sum_{i=1}^{n}\frac{z_i^2\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-2}}{1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}} \\
& + \frac{2\alpha\beta(\gamma-1)}{\sigma^3}\sum_{i=1}^{n}\frac{z_i\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}} \\
& - \frac{\alpha\beta(\gamma-1)}{\sigma^4}\sum_{i=1}^{n}\frac{z_i^2\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-2}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}} \\
& + \frac{\alpha^2\beta(\gamma-1)}{\sigma^4}\sum_{i=1}^{n}\frac{z_i^2\mathrm{e}^{-2z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-2}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-2}}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}} \\
& - \frac{\alpha^2\beta^2(\gamma-1)}{\sigma^4} \\
& \times \sum_{i=1}^{n}\frac{z_i^2\mathrm{e}^{-2z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{2(\alpha-1)}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-2}}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}^2},
\end{aligned}$$

$$\begin{aligned}
U_{\sigma\alpha} = {} & -\frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{z_i\mathrm{e}^{-z_i/\sigma}}{1-\mathrm{e}^{-z_i/\sigma}} + \frac{(\beta-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}}{1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}} \\
& - \frac{\beta(\gamma-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}} \\
& + \frac{\alpha(\beta-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i\mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}\log(1-\mathrm{e}^{-z_i/\sigma})}{[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^2}
\end{aligned}$$

$$- \frac{\alpha\beta(\gamma-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}\log(1-\mathrm{e}^{-z_i/\sigma})}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{2-\beta}}$$

$$+ \frac{\alpha\beta^2(\gamma-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{2\alpha-1}\log(1-\mathrm{e}^{-z_i/\sigma})}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}^2[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{2-\beta}},$$

$$U_{\sigma\beta} = \frac{\alpha}{\sigma^2}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}}{1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}}$$

$$- \frac{\alpha(\gamma-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}}$$

$$- \frac{\alpha\beta(\gamma-1)}{\sigma^2}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}\log[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}^2[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{1-\beta}},$$

$$U_{\sigma\gamma} = -\frac{\alpha\beta}{\sigma^2}\sum_{i=1}^{n}\frac{z_i \mathrm{e}^{-z_i/\sigma}(1-\mathrm{e}^{-z_i/\sigma})^{\alpha-1}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}},$$

$$U_{\alpha\alpha} = -\frac{n}{\alpha^2} - (\beta-1)\sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}[\log(1-\mathrm{e}^{-z_i/\sigma})]^2}{[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^2}$$

$$+ \beta(\gamma-1)\sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-2}[\log(1-\mathrm{e}^{-z_i/\sigma})]^2}{1-(1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha})^{\beta}}$$

$$- \beta^2(\gamma-1)$$

$$\times \sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{2\alpha}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-2}[\log(1-\mathrm{e}^{-z_i/\sigma})]^2}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}^2},$$

$$U_{\alpha\beta} = -\sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}\log(1-\mathrm{e}^{-z_i/\sigma})}{1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}}$$

$$+ (\gamma-1)\sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}\log(1-\mathrm{e}^{-z_i/\sigma})}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}}$$

$$+ \beta(\gamma-1)\sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}\log(1-\mathrm{e}^{-z_i/\sigma})\log(1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha})}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}^2[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{1-\beta}},$$

$$U_{\alpha\gamma} = \beta\sum_{i=1}^{n}\frac{(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta-1}\log(1-\mathrm{e}^{-z_i/\sigma})}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}},$$

$$U_{\beta\beta} = -\frac{n}{\beta^2} + (1-\gamma)\sum_{i=1}^{n}\frac{[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}[\log(1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha})]^2}{\{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}\}^2},$$

$$U_{\beta\gamma} = -\sum_{i=1}^{n} \frac{[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta} \log(1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha})}{1-[1-(1-\mathrm{e}^{-z_i/\sigma})^{\alpha}]^{\beta}} \quad \text{and}$$

$$U_{\gamma\gamma} = -\frac{n}{\gamma^2}.$$

Acknowledgments

We gratefully acknowledge grants from FAPESP, CAPES and CNPq (Brazil). The authors thank an anonymous referee for helpful comments.

References

Barreto-Souza, W. and Cribari-Neto, F. (2009). A generalization of the exponential-Poisson distribution. *Statistics and Probability Letters* **79**, 2493–2500. MR2556316

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. MR2603623

Fletcher, S. C. and Ponnambalam, K. (1996). Estimation of reservoir yield and storage distribution using moments analysis. *Journal of Hydrology* **182**, 259–275.

Ganji, A., Ponnambalam, K. and Khalili, D. (2006). Grain yield reliability analysis with crop water demand uncertainty. *Stochastic Environmental Research and Risk Assessment* **20**, 259–277. MR2297440

Garvan, F. (2002). *The Maple Book*. London: Chapman & Hall/CRC.

Gilchrist, W. G. (2001). *Statistical Modelling with Quantile Functions*. Boca Raton, FL: Chapman & Hall/CRC.

Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistic* **41**, 173–188. MR1705342

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Ser. B* **52**, 105–124. MR1049304

Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology* **6**, 70–81. MR2655540

Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution.* New York: Springer-Verlag. MR0648107

Koutsoyiannis, D. and Xanthopoulos, T. (1989). On the parametric approach to unit hydrograph identification. *Water Resources Management* **3**, 107–128.

Kumaraswamy, P. (1980). Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology* **46**, 79–88.

Lemonte, A. J. and Cordeiro, G. M. (2011). The exponentiated generalized inverse Gaussian distribution. *Statistics and Probability Letters* **81**, 506–517. MR2765171

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica* **52**, 647–664.

Mudholkar, G. S. and Hutson, A. D. (1996). The exponentiated Weibull family: Some properties and a flood data application. *Communication in Statistics—Theory and Methods* **25**, 3059–3083. MR1422323

Mudholkar, G. S., Srivastava, D. K. and Freimer, M. (1995). The exponentiated Weibull family. *Technometrics* **37**, 436–445.

Nadarajah, S. (2008). On the distribution of Kumaraswamy. *Journal of Hydrology* **348**, 568–569.

Nadarajah, S. and Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae* **92**, 97–111. MR2265333

Nassar, M. M. and Eissa, F. H. (2003). On the exponentiated Weibull distribution. *Communication in Statistics—Theory and Methods* **32**, 1317–1336. MR1985853

Ponnambalam, K., Seifi, A. and Vlach, J. (2001). Probabilistic design of systems with general distributions of parameters. *International Journal of Circuit Theory and Applications* **29**, 527–536.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Seifi, A., Ponnambalam, K. and Vlach, J. (2000). Maximization of manufacturing yield of systems with arbitrary distributions of component values. *Annals of Operations Research* **99**, 373–383. MR1837747

Sigmon, K. and Davis, T. A. (2002). *MATLAB Primer*, 6th ed. New York: Chapman & Hall/CRC.

Silva, R. B., Barreto-Souza, W. and Cordeiro, G. M. (2010). A new distribution with decreasing, increasing and upside-down bathtub failure rate. *Computational Statistics and Data Analysis* **54**, 935–944. MR2580928

Song, K. S. (2001). Rényi information, loglikelihood and an intrinsic distribution measure. *Journal of Statistical Planning and Inference* **93**, 51–69. MR1822388

Sundar, V. and Subbiah, K. (1989). Application of double bounded probability density-function for analysis of ocean waves. *Ocean Engineering* **16**, 193–200.

Wolfram, S. (2003). *The Mathematica Book*, 5th ed. New York: Cambridge Univ. Press.

A. J. Lemonte
W. Barreto-Souza
Departamento de Estatística
Universidade de São Paulo
São Paulo
SP
Brazil
E-mail: arturlemonte@gmail.com

G. M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco
Recife
PE
Brazil