Journal of Data Science 12(2014), 661-698

The Kummer Beta Generalized Gamma Distribution

Gauss M. Cordeiro[1], Rodrigo R. Pescim[2*], Clarice G.B. Demétrio[2] and Edwin M.M. Ortega[2]

[1] *Departamento de Estatística, UFPE*
[2] *Departamento de Ciências Exatas, ESALQ - USP*

Abstract: A new extension of the generalized gamma distribution with six-parameter called the Kummer beta generalized gamma distribution is introduced and studied. It contains at least 28 special models such as the beta generalized gamma, beta Weibull, beta exponential, generalized gamma, Weibull and gamma distributions and thus could be a better model for analyzing positive skewed data. The new density function can be expressed as a linear combination of generalized gamma densities. Various mathematical properties of the new distribution including explicit expressions for the ordinary and incomplete moments, generating function, mean deviations, entropy, density function of the order statistics and their moments are derived. The elements of the observed information matrix are provided. We discuss the method of maximum likelihood and a Bayesian approach to fit the model parameters. The superiority of the new model is illustrated by means of three real data sets.

Key words: Bayesian analysis, Generalized gamma distribution, Kummer beta generalized distribution, Lifetime data, Maximum likelihood estimation.

1. Introduction

The generalized gamma (GG) distribution (Stacy, 1962) is an important lifetime model since it includes as special models the exponential, Weibull, gamma and Rayleigh distributions, among others. It is suitable for modeling data with hazard rate function (hrf) of different forms (increasing, decreasing, bathtub and unimodal) and then it is useful for estimating individual hazard functions and both relative hazards and relative times (Cox 2008). The GG distribution has been used in several research areas such as engineering, environment, hydrology and survival analysis. For example, Ortega *et al*. (2003) discussed influence diagnostics in GG regression models, Nadarajah and Gupta (2007) applied this distribution to drought data, Cox *et al*. (2007) presented a parametric survival analysis based on GG hazard functions and Cox (2008) discussed and compared the F-generalized family with the GG model. More recently, Barkauskas *et al*. (2009) modeled the noise part of a spectrum as an autoregressive moving average (ARMA) model with the innovations following the GG distribution Malhotra *et al*.

* Corresponding author.

(2009) provided a unifed analysis for wireless system over generalized fading channels that is modeled by a two parameter GG model and Xie and Liu (2009) analyzed three-moment auto conversion parametrization based on this model. Further Ortega *et al.* (2009) proposed a modifed GG regression model to allow the possibility that long-term survivors may be presented in the data and Cordeiro *et al.* (2011b) studied the exponentiated generalized gamma (EGG) distribution.

Let γ_1 (k, x/α) be the cumulative distribution function (cdf) of the standard gamma distribution where $\gamma_1(\cdot,\cdot)$ is the incomplete gamma function ratio defned by $\gamma_1(k, x) = \gamma(k, x)/\Gamma(k)$, $\gamma(k, x) = \int_0^x w^{k-1}e^{-w}dw$ and $\Gamma(\cdot)$ are the incomplete and complete gamma functions. The probability density function (pdf) of the GG distribution, with three parameters $\alpha > 0$, $\beta > 0$ and $k > 0$, defned by Stacy (1962), has the form

$$g(x;\alpha,\beta,k) = \frac{\beta}{\alpha\Gamma(k)}\left(\frac{x}{\alpha}\right)^{\beta k-k} exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right], x > 0. \tag{1}$$

In the density function (1), $\alpha > 0$ is a scale parameter and $\beta > 0$ and $k > 0$ are shape parameters. The cdf corresponding to (1) is

$$G(x;\ \alpha,\beta,k) = \gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]. \tag{2}$$

For an arbitrary baseline distribution G(x; γ) with parameter vector γ and density function g(x; γ) Pescim *et al*. (2012) proposed the *Kummer beta generalized* (denoted by the prefix "KB-G" for short) cumulative function defned by

$$F_{K\beta g}(x) = K\int_0^{G(x;\gamma)} t^{a-1}(1-t)^{b-1}e^{-ct}dt, \tag{3}$$

where $a > 0$ and $b > 0$ are shape parameters which induce skewness, and thereby promote weight variation of the tails, whereas the parameter $-\infty < c < \infty$ “squeezes” the pdf to the left or right, i.e., it gives weights to the extremes of the density functions. Here,

$$K^{-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}{}_1F_1(a;a+b;-c)$$

and

$${}_1F_1(a;a+b;-c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 t^{a-1}(1-t)^{b-1}e^{-ct}dt = \sum_{k=0}^{\infty}\frac{(\mathrm{a})_\mathrm{k}(-\mathrm{c})^\mathrm{k}}{(\mathrm{a}+\mathrm{b})_\mathrm{k}\mathrm{k}!}$$

is the confluent hypergeometric function (Abramowitz and Stegun, 1968) and $(d)_k = d(d+1)\ldots(d+k-1)$ denotes the ascending factorial. The density function corresponding to (3) can be expressed as

$$f_{K\beta g}(x) = K_{\mathrm{g}}(x;\gamma)G(x;\gamma)^{a-1}[1-G(x;\gamma)]^{b-1}\exp[-cG(x;\gamma)]. \tag{4}$$

Clearly, the Kummer beta distribution (Ng and Kotz, 1995) is a basic exemplar of equation (4) for $G(x;\gamma) = x$, where $x \in (0,1)$. Equation (4) will be most tractable when both functions $G(x;\gamma)$ and $g(x;\gamma)$ have simple analytic expressions. Its major benefit is to offer more flexibility to extremes (right and/or left) of the density functions and therefore it becomes suitable for analyzing data with high degree of asymmetry.

The class of distributions (4) includes two important special cases: the beta-generalized (BG) and exponentiated generalized (EG) distributions when $c = 0$ and $c = 0$ and $b = 1$ respectively. We can note that the BG distributions can be limited in one aspect. They have only two additional shape parameters and so they can add only a limited structure to the generated distribution. For instance, a BG distribution may have problems to capture the behavior of random variables with symmetric but highly leptokurtic distributions. While the beta parameters offer explicit control over skewness when the parent is symmetric they have less control over higher moments such as kurtosis. Further the EG distribution still introduces only one extra shape parameter whereas three may be required to control both tail weights and the distribution of weight in the center. Hence the generated distribution (4) is a more flexible model since it has one more shape parameter than the classical beta generator.

In this paper we study a new six-parameter model called the *Kummer beta generalized gamma* (KBGG) distribution which contains at least 28 special models. The main motivation for this extension is that the new model is a highly flexible lifetime distribution which admits different degrees of kurtosis and asymmetry. The KBGG density function is defined from (4) by taking (1) and (2) as the baseline model. The six-parameter KBGG density function can be expressed as

$$\begin{aligned} f(x) &= K\,\frac{\beta}{\alpha\,\Gamma(k)}\left(\frac{x}{\alpha}\right)^{\beta k-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]^{a-1} \\ &\quad\times\left\{1-\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{b-1}\exp\left\{-c\,\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}. \end{aligned} \tag{5}$$

The corresponding hrf to (5) becomes

$$\tau(x) = \frac{K\,\beta\left(\frac{x}{\alpha}\right)^{\beta k-1}\left\{1-\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{b-1}}{\alpha\,\Gamma(k)\,[1-F(x)]\;\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]^{1-a}}\exp\left[-\left\{c\,\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]+\left(\frac{x}{\alpha}\right)^{\beta}\right\}\right]. \tag{6}$$

Hereafter, we denote by $\boldsymbol{X}$ a random variable following (5), say $\boldsymbol{X} \sim$ KBGG($a, b, c, \alpha, \beta, k$). This density has five shape parameters a, b, c, β and k which allow for a high degree of flexibility. The parameter c controls tail weights to the extremes of the distribution. The study of the new distribution is important since it extends some distributions previously considered in the literature. In fact, the generalized gamma (GG) model is clearly a basic exemplar for $a = b = 1$ and $c = 0$ with a continuous crossover towards models with different shapes (e.g. a specifized combination of skewness and kurtosis). The KBGG model contains as sub-models the beta generalized gamma (BGG) (Cordeiro *et al*., 2013a) and the exponentiated generalized gamma (EGG) (Cordeiro *et al*., 2011b) distributions when $c = 0$ and $b = 1$ in addition to $c = 0$, respectively. Plots of the new density function for selected parameter values are displayed in

Figure 1. It is evident that this density function is much more flexible than the GG distribution. The KBGG model is very flexible and hence can be used in many practical situations. In fact, it can be symmetric, asymmetric and also exhibit bimodality. We also provide a comprehensive description of some of its mathematical properties with the hope that it will attract wider applications in reliability, engineering, environment and in other areas of research.

The paper is outlined as follows. In Section 2, we derive more than 28 special distributions from the KBGG model. In Section 3, we demonstrate that the KBGG density function can be expressed as a linear combination of EGG density functions. This is an important result to provide some mathematical properties of the KBGG distribution. We obtain explicit expressions for the moments and generating function (Section 4), incomplete moments (Section 5), mean deviations and Rényi entropy (Section 6) and order statistics (Section 7). In Section 8, we discuss some statistical inference such as maximum likelihood method and Bayesian approach. Three applications given in Section 9 reveal the usefulness of the new distribution for analyzing real data. Concluding remarks are addressed in Section 10.

2. Special distributions

The following well-known distributions are special models of the KBGG distribution.

2.1 Kummer Beta Generator

- For $k=1$, the KBGG distribution reduces to the Kummer beta Weibull (KBW) distribution. If $k=1$ and $\beta=1$, it yields the Kummer beta exponential (KBE) distribution. If $\beta=2$ in addition to $k=1$, it gives the Kummer beta Rayleigh (KBR) distribution. For $\alpha=\sqrt{2}\sigma$, $\beta=1$ and $k=p/2$, the KBGG distribution reduces to the Kummer beta scaled chi-square (KBSChi) distribution. For $\alpha=\sqrt{\theta}$, $\beta=2$ and $k=3/2$, the KBGG distribution coincides with the Kummer beta Maxwell (KBMa) distribution.

- For $\beta=1$, the KBGG distribution coincides to the five parameter Kummer beta gamma (KBGa) distribution. Taking $\alpha=2$, $\beta=1$ and $k=p/2$, we obtain the Kummer beta chi-square (KBChi) distribution. Moreover if $\alpha=2^{\frac{1}{2\gamma}}\theta$, $\beta=2\gamma$ and $k=1/2$, the KBGG distribution becomes the Kummer beta generalized half-normal (KBGHN) distribution. If $\alpha=2^{\frac{1}{2}}\theta$, $\beta=2$ and $k=1/2$, the KBGG model gives the Kummer beta half-normal (KBHN) distribution. Finally, if $\alpha=\sqrt{\omega/\mu}$, $\beta=2$ and $k=\mu$, it yields the Kummer beta Nakagami (KBNa) distribution.

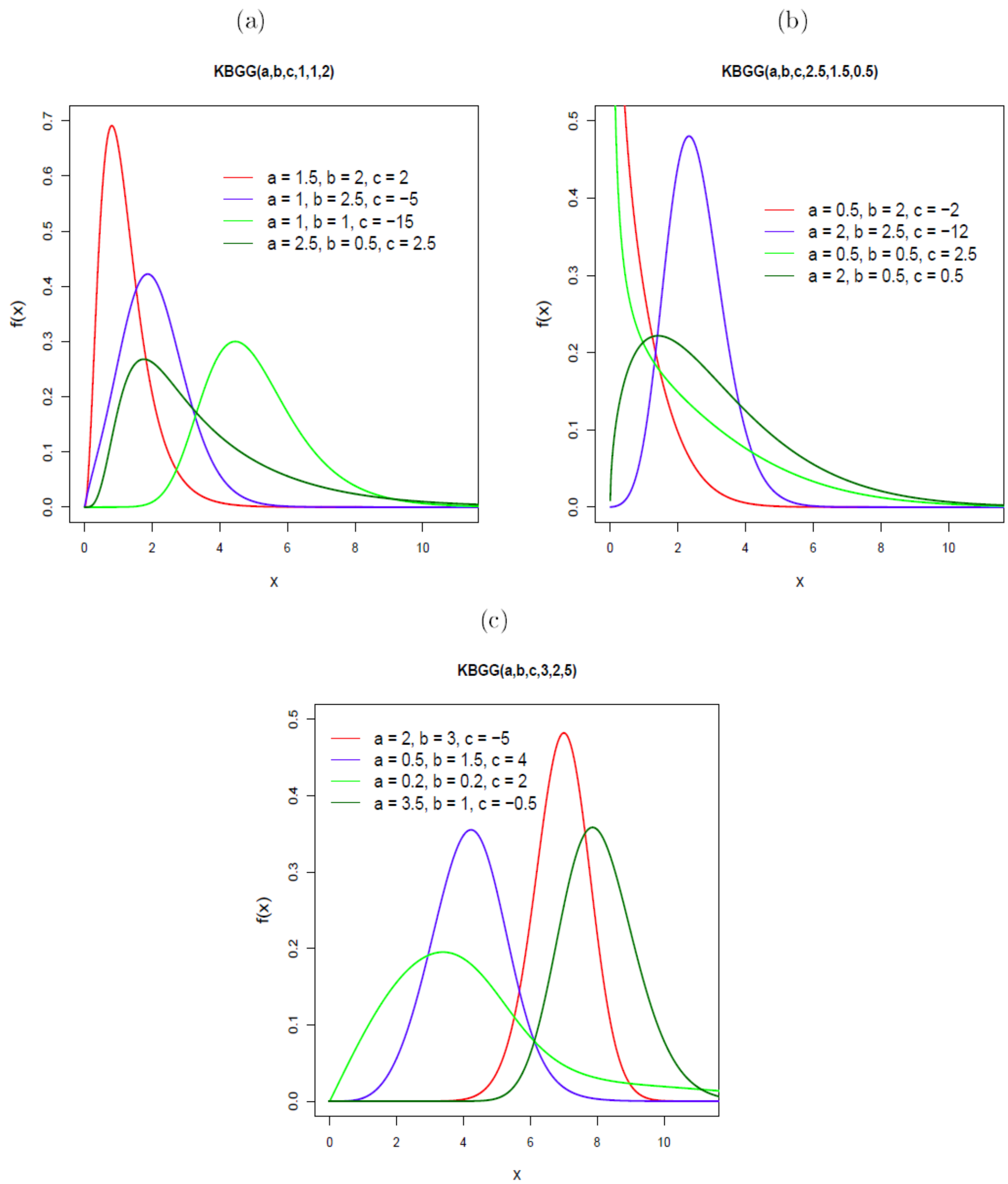

Figure 1: Plots of the density function (5) for some parameter values.

2.2 Beta Generator (for c = 0)

- For $c = 0$, the KBGG distribution reduces to the five parameter beta generalized gamma (BGG) distribution. If $k = 1$, the BGG distribution corresponds to the beta Weibull (BW) distribution pioneered by Famoye *et al.* (2005). If $\beta = 1$ and $k = 1$, it gives the beta exponential (BE):distribution (see, Nadarajah and Kotz, 2005). If $\beta = 2$ in addition to $k = 1$, it yields the beta Rayleigh (BR) distribution (Cordeiro *et al*., 2013b). For $\alpha = \sqrt{\theta}$, $\beta = 1$ and $k = 3/2$, the BGG distribution gives the beta Maxwell (BMa) distribution.

- For $\beta = 1$, the BGG distribution yields the four parameter beta gamma (BGa4) distribution. If $\alpha = 1$ in addition to $\beta = 1$, the special case corresponds to the beta gamma (BGa3) distribution. Further, if $\alpha = 2^{\frac{1}{2\gamma}}\theta$, $\beta = 2\gamma$ and $k = 1/2$, the BGG distribution becomes the beta generalized half-normal (BGHN) distribution defined by Pescim *et al.* (2010). If $\alpha = 2^{\frac{1}{2}}\theta$, $\beta = 2\gamma$ and $k = 1/2$, the BGG model reduces to the distribution called the beta half-normal (BHN) (see, for example, Pescim *et al.*, 2010).

- Finally, if $\alpha = \sqrt{\omega/\mu}$, $\beta = 2$ and $k = \mu$, the BGG distribution becomes the Beta Nakagami (BNa) distribution.

2.3 Exponentiated Generator (for b = 1 and c = 0)

- For $b = 1$ and $c = 0$ we obtain from (5) the EGG density function. If $k = 1$, the EGG distribution reduces to the exponentiated Weibull (EW) distribution introduced by Mudholkar *et al.* (1995). If $\beta = 1$ in addition to $k = 1$, the special case corresponds to the exponentiated exponential (EE) distribution (see, Gupta and Kundu, 2001). If $\beta = 2$ in addition to $k = 1$, the special case corresponds to the generalized Rayleigh (GR) distribution (Kundu and Raqab, 2005). For $\alpha = \sqrt{\theta}$, $\beta = 2$ and $k = 3/2$, the EGG distribution becomes the exponentiated Maxwell (EMa) distribution.

- For $\beta = 1$, the EGG distribution reduces to the three parameter exponentiated gamma (EGa3) distribution. If $\alpha = 1$ in addition to $\beta = 1$, the special case corresponds to the exponentiated gamma (EGa2) distribution. If $\alpha = 2^{\frac{1}{2\gamma}}\theta$, $\beta = 2\gamma$ and $k = 1/2$, the EGG distribution becomes the exponentiated generalized half-normal (EGHN) distribution.

- If $\alpha = 2^{\frac{1}{2}}\theta$, $\beta = 2$ and $k = 1/2$, the EGG model reduces to the exponentiated half-normal (EHN). Finally, if $\alpha = \sqrt{\omega/\mu}$, $\beta = 2$ and $k = \mu$, the EGG distribution becomes the exponentiated Nakagami (ENa) distribution.

2.4 Baseline distributions (for a = b = 1 and c = 0)

- For $a = b = 1$ and $c = 0$, the new model reduces to the three parameter generalized gamma (GG) distribution. The case $k = 1$ gives the classical two parameter Weibull (W) distribution. If $\beta = 1$and $\beta = 2$, in addition to $k = 1$, the special cases coincide with the exponential (E) and Rayleigh (R) distributions, respectively. For $\alpha = \sqrt{2}\sigma$, $\beta = 1$ and $k = p/2$, the special case gives the scaled chi-square (SChi) distribution. If $\alpha = \sqrt{\theta}$ in addition to $\beta = 2$ and $k = 3/2$, it reduces to the Maxwell (Ma) distribution (see, for example, Bekker and Roux, 2005).

- Setting $\beta = 1$, the special case gives the classical gamma (Ga) distribution. If $\alpha = 2$, in addition to $\beta = 1$ and $k = p/2$, we obtain the chi-square (Chi) distribution. If $\alpha = 2^{1/(2\gamma)}\theta$ in addition to $\beta = 2\gamma$, $k = 1/2$, it coincides with the generalized half-normal (GHN) distribution pioneered by Cooray and Ananda (2008). Taking $\alpha = 2^{\frac{1}{2}}\theta$ in addition to $\beta = 2$ and $k = 1/2$, it reduces to the well-known half-normal (HN) distribution. Further, if $\alpha = \sqrt{\omega/\mu}$ in addition to $\beta = 2$ and $k = \mu$, the special case corresponds to the Nakagami (Na) distribution.

Several special KBGG models are listed in Table 1.

Table 1: Some special cases of the KBGG distribution

Case	α	β	k	Distribution	References
			$a=b=1$ and $c=0$		
(1)	α	β	k	Generalized gamma	Stacy (1962)
(2)	α	β	1	Weibull	Fréchet (1927)
(3)	α	1	k	Gamma	Laplace (1836)
(4)	α	1	1	Exponential	Kondo (1930)
(5)	α	2	1	Rayleigh	Rayleigh (1880)
(6)	$\sqrt{\theta}$	2	3/2	Maxwell	Maxwell (1860)
(7)	$2^{\frac{1}{2\gamma}}\theta$	2γ	1/2	Generalized half-normal	Cooray and Ananda (2008)

$b=1$ and $c=0$

Case	α	β	k	a	Distribution	References
(8)	α	β	k	a	Exponentiated generalized gamma	Cordeiro *et al.* (2011)
(9)	α	β	1	a	Exponentiated Weibull	Mudholkar *et al.* (1995)
(10)	α	1	k	a	Exponentiated gamma	New
(11)	α	1	1	a	Exponentiated exponential	Gupta and Kundu (2001)
(12)	α	2	1	a	Exponentiated Rayleigh	Cordeiro *et al.* (2013b)
(13)	$\sqrt{\theta}$	2	3/2	a	Exponentiated Maxwell	New
(14)	$2^{\frac{1}{2\gamma}}\theta$	2γ	1/2	a	Exponentiated generalized half-normal	Pescim *et al.* (2010)

$c=0$

Case	α	β	k	a	b	Distribution	References
(15)	α	β	k	a	b	Beta generalized gamma	Cordeiro *et al.* (2013a)
(16)	α	β	1	a	b	Beta Weibull	Famoye *et al.* (2005)
(17)	α	1	k	a	b	Beta gamma	Kong *et al.* (2007)
(18)	α	1	1	a	b	Beta exponential	Nadarajah and Kotz (2006)
(19)	α	2	1	a	b	Beta Rayleigh	Cordeiro *et al.* (2013b)
(20)	$\sqrt{\theta}$	2	3/2	a	b	Beta Maxwell	Amusan (2010)
(21)	$2^{\frac{1}{2\gamma}}\theta$	2γ	1/2	a	b	Beta generalized half-normal	Pescim *et al.* (2010)

Case	α	β	k	a	b	c	Distribution	References
(22)	α	β	1	a	b	c	Kummer beta Weibull	Pescim *et al.* (2012)
(23)	α	1	k	a	b	c	Kummer beta gamma	Pescim *et al.* (2012)
(24)	α	1	1	a	b	c	Kummer beta exponential	New
(25)	α	2	1	a	b	c	Kummer beta Rayleigh	New
(26)	$\sqrt{\theta}$	2	3/2	a	b	c	Kummer beta Maxwell	New
(27)	$2^{\frac{1}{2\gamma}}\theta$	2γ	1/2	a	b	c	Kummer beta generalized half-normal	New
(28)	$2^{\frac{1}{2}}\theta$	2	1/2	a	b	c	Kummer beta half-normal	New

3. Expansion for the density function

A useful expansion for equation (5) can be derived using the concept of exponentiated-G ("EG" for short) distributions. First, we use an expansion for the density function (4) given by a linear combination of EG densities. The properties of some exponentiated distributions have been studied by several authors, see Mudholkar and Srivastava (1993) and Mudholkar *et al*. (1995) for exponentiated Weibull (EW), Gupta *et al*. (1998) for exponentiated Pareto (EPa), Gupta and Kundu (2001) for exponentiated exponential (EE) and, more recently, Cordeiro *et al*. (2011) for exponentiated generalized gamma (EGG) distribution.

Pescim *et al*. (2012) demonstrated that

$$f(x) = \sum_{r=0}^{\infty} c_r v_{r+1}(x), \tag{7}$$

where the coefficients (for $r = 0, 1 \ldots$) are $c_r = \sum_{i,j=0}^{\infty} \sum_{k=r+1}^{\infty} t_{i,j,k,r+1}$,

$$t_{i,j,k,r} = t_{i,j,k,r}(a,b,c) = \frac{K(-1)^{i+j+k+r} c^i}{i!\,(a+i+j)} \binom{a+i+j}{k} \binom{k}{r} \binom{b-1}{j}$$

and $v_{r+1}(x) = (r+1) g(x) G(x)^r$ denotes the EG density function with power parameter $r + 1$.

Equation (7) reveals that the KB-G density function is a linear combination of EG densities. This result is important to derive some mathematical properties of the KBGG distribution from those of the EGG distribution. This equation holds for any real non-integers a, b and c.

Replacing (1) and (2) in $v_{r+1}(x)$, we obtain the EGG(α, β, k, $r+1$) density function

$$v_{r+1}(x) = \frac{(r+1)\,\beta}{\alpha\,\Gamma(k)} \left(\frac{x}{\alpha}\right)^{\beta k - 1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^r. \tag{8}$$

We require a power series for the incomplete gamma function ratio in (8) given by

$$\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right] = \frac{\left(\frac{x}{\alpha}\right)^{\beta k}}{\Gamma(k)} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{x}{\alpha}\right)^{\beta m}}{(k+m)m!}. \tag{9}$$

By application of an align in Section 0.314 of Gradshteyn and Ryzhik (2007) for a power series raised to a positive power, we obtain for any r positive integer

$$\left(\sum_{m=0}^{\infty} a_m\, x^m\right)^r = \sum_{m=0}^{\infty} d_{r,m}\, x^m, \tag{10}$$

where the coefficients $d_{r,m}$ (for $m = 1, 2, \ldots$) satisfy the recurrence relation

$$d_{r,m} = (m\, a_0)^{-1} \sum_{p=1}^{m} (p(r+1) - m)\, a_p\, d_{r,m-p} \tag{11}$$

and $d_{r,0} = a_0^r$. The coefficient $d_{r,m}$ comes from $d_{r,0}, \ldots, d_{r,m-1}$ and hence from $a_0, \ldots, a_m$. The coefficients $d_{r,m}$ can also be written explicitly as functions of the quantities a_m.

Further, combining equations (9) and (10), we obtain

$$\left\{\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{r} = \frac{x^{\beta k r}}{\alpha^{\beta k r}\,\Gamma(k)^{r}}\left[\sum_{m=0}^{\infty}\frac{(-1)^{m}\left(\frac{x}{\alpha}\right)^{\beta m}}{(k+m)m!}\right]^{r} = \frac{x^{\beta k r}}{\alpha^{\beta k r}\,\Gamma(k)^{r}}\sum_{m=0}^{\infty} d_{r,m}\left(\frac{x}{\alpha}\right)^{\beta m}, \tag{12}$$

where the coefficients $d_{r,m}$ are obtained from equation (11) with $a_p = (-1)^p/(k+p)p!$. Combining (8) and (12) and after some algebra manipulations, we can rewrite the EGG density function as

$$v_{r+1}(x) = \sum_{m=0}^{\infty} e_{r,m}\; g_{\alpha,\beta,k^{\star}}(x), \tag{13}$$

where

$$e_{r,m} = \frac{d_{r,m}\,\Gamma\left(k^{\star}\right)}{\Gamma(k)^{r+1}\,(r+1)^{-1}},$$

$k^{\star} = k(r + 1) + m$ and $g_{\alpha,\beta,k^{\star}}(x)$ is the density function of the GG(α, β, $k^{\star}$) distribution.

Combining (7) and (13), we obtain

$$f(x) = \sum_{r,m=0}^{\infty} \eta_{r,m}\; g_{\alpha,\beta,k^{\star}}(x), \tag{14}$$

where $\eta_{r,m} = e_{r,m}\, c_r$.

Equation (14) reveals that the KBGG density function can be expressed as a linear combination of GG densities. This equation is the main result of this section. It plays an important role in this paper. In the next sections, based on this equation, we obtain some KBGG structural properties including explicit expressions for the ordinary and incomplete moments, generating function, mean deviations and order statistics.

4. Moments and generating function

The *s*th moment of ***X*** can be expressed from (14) as

$$\mu'_s = \mathrm{E}(X^s) = \sum_{r,m=0}^{\infty} \eta_{r,m}\int_0^{\infty} x^s\, g_{\alpha,\beta,k^{\star}}(x)dx$$

and then

$$\mu'_s = \sum_{r,m=0}^{\infty} \eta_{r,m}\; \mathrm{E}(X_{k^{\star}}^{s}), \tag{15}$$

where $X_{k^{\star}} \sim \mathrm{GG}(\alpha,\beta,k^{\star})$.

Equation (15) is an important result since it provides the moments of the KBGG distribution as a linear combination of GG moments. So, we have

$$\mathrm{E}(X_{k^\star}^s) \;=\; \frac{\beta}{\alpha\,\Gamma(k^\star)}\int_0^\infty x^s\,\left(\frac{x}{\alpha}\right)^{\beta k^\star-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]dx.$$

Next, on setting $u=\left(\frac{x}{\alpha}\right)^{\beta}$ in last integral, $\mathrm{E}(X_{k^\star}^s)$ reduces to

$$\mathrm{E}(X_{k^\star}^s)=\alpha^s\,\frac{\Gamma\left[k(r+1)+m+s/\beta\right]}{\Gamma(k(r+1)+m)}.$$

Replacing the last result in (15), we obtain the *s*th moment of ***X*** as

$$\mu_s'=\alpha^s\sum_{r,m=0}^{\infty}\eta_{r,m}\,\frac{\Gamma\left[k(r+1)+m+s/\beta\right]}{\Gamma(k(r+1)+m)}, \qquad (16)$$

where $\eta_{r,m}$ is defined by (14).

Equation (16) is readily computed numerically using standard statistical software. It (and other expansions in this paper) can also be evaluated in symbolic computation software such as Mathematica and Maple. In numerical applications, a large natural number *N* can be used in the sums instead of infinity. Several mathematical quantities of *X* (central, incomplete and factorial moments, variance, skewness and kurtosis) can be derived from this result.

The skewness and kurtosis measures can be determined from the ordinary moments using well-known relationships. Plots of the skewness and kurtosis of the KBGG distribution as functions of *c* for selected values of *a* and *b* for α = 0.5, β = 1.0 and *k* = 2.0 are displayed in Figures 2 and 3, respectively. Figures 2a and 2b indicate that the additional parameter *c* promotes high levels of asymmetry.

(a) (b)

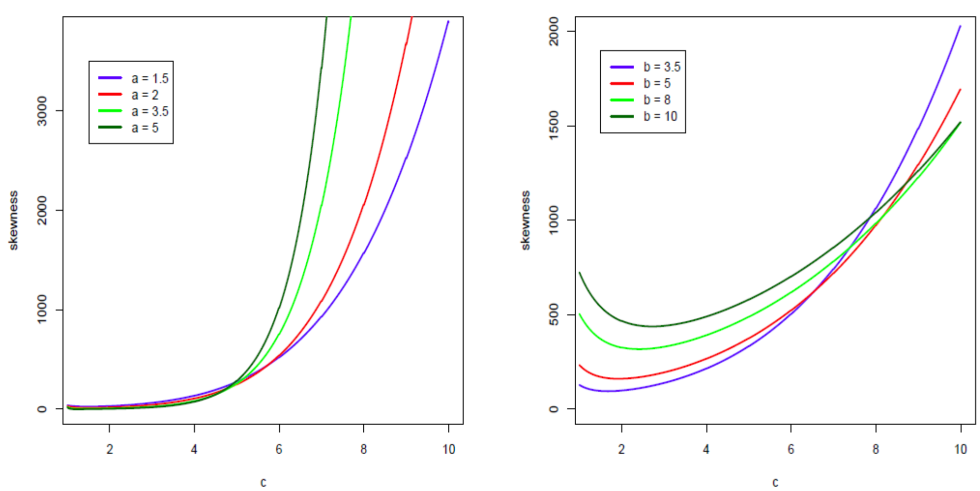

Figure 2: Skewness of the KBGG distribution as a function of c for some values of *a* and *b* for α = 0.5, β = 1.0 and *k* = 2.0. (a) *b* = 2.0 and (b) *a* = 1.2

Further, we provide a representation for the moment generating function (mgf) of X, say $M(t)$ = E[exp(tX)], from the linear combination of GG generating functions. From equation (14), we have

$$M(t) = \sum_{r,m=0}^{\infty} \eta_{r,m}\ M_{\alpha,\beta,k^\star}(t), \tag{17}$$

Where $M_{\alpha,\beta,k^\star}(t)$ denotes the mgf of the GG(α, β, $k^\star$) distribution.

We can derive $M_{\alpha,\beta,k^\star}(t)$ as

$$M_{\alpha,\beta,k^\star}(t) = \frac{\beta}{\alpha\Gamma(k^\star)} \int_{0}^{\infty} \exp(tx) \left(\frac{x}{\alpha}\right)^{\beta k^\star - 1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] dx.$$

Using the power series for the exponential function and replacing $u = (x/\alpha)^{\beta}$ in this integral, $M_{\alpha,\beta,k^\star}(t)$ reduces to

$$M_{\alpha,\beta,k^\star}(t) = \frac{1}{\Gamma(k^\star)} \sum_{\nu=0}^{\infty} \frac{(\alpha t)^{\nu}}{\nu!} \int_{0}^{\infty} u^{\frac{\nu}{\beta}+k^\star-1}\, \mathrm{e}^{-\mathrm{u}}\, \mathrm{du}. \tag{18}$$

Computing the integral in (18), we obtain

$$M_{\alpha,\beta,k^\star}(t) = \frac{1}{\Gamma(k^\star)} \sum_{\nu=0}^{\infty} \Gamma\left(\frac{\nu}{\beta} + k^\star\right) \frac{(\alpha t)^{\nu}}{\nu!}.$$

Consider the Wright generalized hypergeometric function defined by

$${}_p\Psi_q \left[\begin{array}{c} (\alpha_1, A_1), \ldots, (\alpha_p, A_p) \\ (\beta_1, B_1), \ldots, (\beta_q, B_q) \end{array} ; x \right] = \sum_{n=0}^{\infty} \frac{\prod_{j=1}^{p} \Gamma(\alpha_j + A_j n)}{\prod_{j=1}^{q} \Gamma(\beta_j + B_j n)} \frac{x^n}{n!}.$$

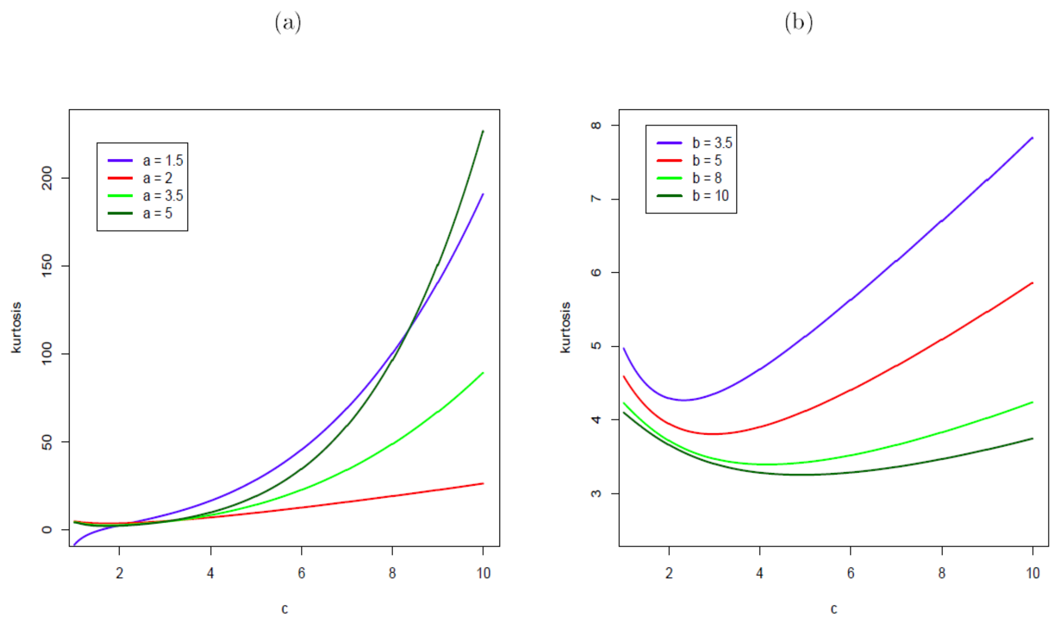

Figure 3: Kurtosis of the KBGG distribution as a function of c for some values of a and b for α = 0.5, β= 1.0 and k = 2.0. (a) b = 2.0 and (b) a = 1.2

Combining the last two results, we can rewrite the mgf of the GG distribution as

$$M_{\alpha,\beta,k^\star}(t) = \frac{1}{\Gamma(k^\star)}\, {}_1\Psi_0\left[\begin{array}{c}(k^\star,\beta^{-1})\\ -\end{array}\;;\alpha t\right], \tag{19}$$

provided that $\beta > 1$.

The KBGG generating function follows by inserting (19) in equation (17). For $\beta > 1$, we have

$$M(t) = \sum_{r,m=0}^{\infty} \eta_{r,m}\, {}_1\Psi_0\left[\begin{array}{c}(k^\star,\beta^{-1})\\ -\end{array}\;;\alpha t\right]. \tag{20}$$

Equations (16) and (20) are the main results of this section. The mgf of any KBGG sub-model, as those discussed in Section 2, can be determined from (20) by substitution of known parameters.

5. Incomplete moments

The answers to many important questions in economics require more than just knowing the mean of the distribution, but its shape as well. This is obvious not only in the study of econometrics but in other areas as well. Incomplete moments of the income distribution form natural building blocks for measuring inequality: for example, the Lorenz and Bonferroni curves and Pietra and Gini measures of inequality depend upon the incomplete moments of the income distribution. The sth incomplete moment of $\boldsymbol{X}$ is defined by $m_s(y) = \int_0^y x^s f(x)dx$. From the linear combination (14), we have

$$m_s(y) = \sum_{r,m=0}^{\infty} \eta_{r,m}\ t_s^{\star}(y), \tag{21}$$

where $t_s^{\star}(y) = \int_0^y x^s g_{\alpha,\beta,k^{\star}}(x)dx$ denotes the sth incomplete moment of the GG distribution with parameters α, β and $k^{\star}$ given by

$$t_s^{\star}(y) = \frac{\beta}{\alpha\,\Gamma(k^{\star})} \int_0^y x^s \left(\frac{x}{\alpha}\right)^{\beta k^{\star}-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] dx.$$

Calculating the integral above, $t_s^{\star}(y)$ reduces to

$$t_s^{\star}(y) = \alpha^s\, \frac{\gamma\left(k(r+1)+m+s/\beta, (y/\alpha)^{\beta}\right)}{\Gamma(k(r+1)+m)}.$$

Substituting the last equation in (21), we obtain

$$m_s(y) = \alpha^s \sum_{r,m=0}^{\infty} \eta_{r,m}\, \frac{\gamma\left(k(r+1)+m+s/\beta, (y/\alpha)^{\beta}\right)}{\Gamma(k(r+1)+m)}. \tag{22}$$

6. Other Measures

Here, we derive the means deviations, Lorenz and Bonferroni curves and the Rényi entropy of the KBGG distribution.

6.1 Mean deviations

We can derive the mean deviations about the mean $\mu_1'(\delta_1)$ and about the median $M = (\delta_2)$ in terms of the first incomplete moment. The median is obtained by inverting $F(M) = K\int_0^{\gamma_1[k,\left(\frac{M}{\alpha}\right)^{\beta}]} t^{a-1}(1-t)^{b-1}e^{-ct}dt = 1/2$ numerically. They can be expressed as

$$\delta_1 = 2[\mu_1' F(\mu_1') - m_1(\mu_1')] \text{ and } \delta_2 = \mu_1' - 2m_1(M),$$

where $m_1(\cdot)$ is the first incomplete moment of $\boldsymbol{X}$ given by (22) with $s = 1$. We have

$$m_1(\omega) = \alpha \sum_{r,m=0}^{\infty} \eta_{r,m}\, \frac{\gamma\left(k(r+1)+m+1/\beta, (\omega/\alpha)^{\beta}\right)}{\Gamma(k(r+1)+m)}. \tag{23}$$

The measures δ_1 and δ_2 are calculated from (23) by setting $\omega = (\mu_1')$ and $\omega = M$, respectively.

Bonferroni and Lorenz curves are useful in fields such as reliability, economics, demography, insurance and medicine. For the KBGG distribution, these curves can be obtained (for given $0 < \pi < 1$) from $B(\pi) = (\pi\mu_1')^{-1}m_1(q)$ and $L(\pi) = (\mu_1')^{-1}m_1(q)$, respectively, where $\mu_1' = E(X)$, $q = F^{-1}(\pi)$ can be computed for a given probability π by inverting (3) numerically when $G(x;\pi)$ is the GG cdf. These curves determined from equation (23) and have applications in several fields.

6.2 Rényi Entropy

Entropy has been used in various situations in science and engineering and numerous measures of entropy have been studied and compared in the literature. The Rényi entropy is defined by

$$\mathcal{J}_R(\xi) = \frac{1}{1-\xi}\log\left[\int f^{\xi}(x)dx\right],\ \xi > 0 \ \text{and}\ \xi \neq 1.$$

Note that the integral above is obtained from (5) as

$$I(\xi) = \int_0^\infty f^{\xi}(x)\,dx = \left(\frac{K\,\beta}{\alpha\,\Gamma(k)}\right)^{\xi}\int_0^\infty \left(\frac{x}{\alpha}\right)^{\xi(\beta k-1)}\exp\left[-\xi\left(\frac{x}{\alpha}\right)^{\beta}\right]\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)^{\xi(a-1)} \times\left[1-\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{\xi(b-1)}\exp\left[-c\,\xi\,\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]dx. \quad (24)$$

Using the exponential and binomial expansions in (24), we obtain

$$I(\xi) = \left[\frac{K\,\beta}{\alpha\,\Gamma(k)}\right]^{\xi}\sum_{i,j=0}^{\infty}\frac{(-1)^{i+j}}{i!\,(c\,\xi)^{-i}}\binom{\xi(b-1)}{j} \times\int_0^\infty\left(\frac{x}{\alpha}\right)^{\xi(\beta k-1)}\exp\left[-\xi\left(\frac{x}{\alpha}\right)^{\beta}\right]\left[\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{\xi(a-1)+i+j}dx. \quad (25)$$

Noting that $\xi > 0$ and $a > 0$ are real non-integers, we can expand $\left[\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{\xi(a-1)+i+j}$ as

$$\left[\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{\xi(a-1)+i+j} = \left\{1-\left[1-\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]\right\}^{\xi(a-1)+i+j} = \sum_{p=0}^{\infty}(-1)^p\binom{\xi(a-1)+i+j}{p}\left\{1-\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right\}^{p}$$

and then

$$\left[\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{\xi(a-1)+i+j}=\sum_{p=0}^{\infty}\sum_{r=0}^{p}(-1)^{p+r}\binom{\xi(a-1)+i+j}{p}\binom{p}{r}\left[\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{r}.$$

Replacing $\sum_{p=0}^{\infty}\sum_{r=0}^{\infty}$ by $\sum_{r=0}^{\infty}\sum_{p=r}^{\infty}$, quantity, $I(\xi)$ can be expressed in the form

$$I(\xi) = \left(\frac{K\,\beta}{\alpha\,\Gamma(k)}\right)^{\xi}\sum_{r=0}^{\infty}\rho_r\int_0^{\infty}\left(\frac{x}{\alpha}\right)^{\xi(\beta k-1)}\exp\left[-\xi\left(\frac{x}{\alpha}\right)^{\beta}\right]\left[\gamma_1\left(k,\left(\frac{x}{\alpha}\right)^{\beta}\right)\right]^{r}dx, \quad (26)$$

where

$$\rho_r=\sum_{i,j=0}^{\infty}\sum_{p=r}^{\infty}\frac{(-1)^{i+j+p+r}}{i!\,(c\,\xi)^{-i}}\binom{\xi(b-1)}{j}\binom{\xi(a-1)+i+j}{p}\binom{p}{r}.$$

Using expansion (12) in (26), we obtain

$$I(\xi) = \left[\frac{K\,\beta}{\alpha\,\Gamma(k)}\right]^{\xi}\sum_{r,m=0}^{\infty}\frac{d_{r,m}}{\Gamma(k)^{r}}\,\rho_r\int_0^{\infty}\left(\frac{x}{\alpha}\right)^{\beta[k(r+\xi)+m]-\xi}\exp\left[-\xi\left(\frac{x}{\alpha}\right)^{\beta}\right]dx. \quad (27)$$

Calculating the integral in (27), we have

$$I(\xi) = \frac{K^{\xi}\,\beta^{\xi-1}}{\Gamma(k)^{\xi}\,\alpha^{\xi-1}}\sum_{r,m=0}^{\infty}\rho^{\star}_{r,m}\,\Gamma\left(k(r+\xi)+m+\frac{(1-\xi)}{\beta}\right),$$

where

$$\rho^{\star}_{r,m}=\frac{d_{r,m}\;\rho_r}{\Gamma(k)^{r}\,\xi^{k(r+\xi)+m+(1-\xi)/\beta}}.$$

Finally, the Rényi entropy reduces to

$$\begin{aligned}\mathcal{J}_R(\xi) &= (1-\xi)^{-1}\{\xi\,[\log(K)-\log\Gamma(k)]+(\xi-1)\,[\log(\beta)-\log(\alpha)]\\ &\quad+\log\left[\sum_{r,m=0}^{\infty}\rho^{\star}_{r,m}\,\Gamma\left(k(r+\xi)+m+\frac{(1-\xi)}{\beta}\right)\right]\right\}.\end{aligned}$$

7. Order statistics

Here, we derive an explicit expression for the density function of the ith order statistic $X_{i:n}$, say $f_{i:n}(x)$, in a random sample of size n from $\boldsymbol{X}\sim$ KBGG(a, b, c, α, β, k). It is well-known that

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

and using the binomial expansion, we obtain

$$f_{i:n}(x) = \frac{n!\, f(x)}{(i-1)!\,(n-i)!} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \qquad (28)$$

We demonstrate that $f_{i:n}(x)$ can be expressed as a linear combination of GG densities. First, we provide an expansion for the KBGG cdf. Pescim *et al.* (2012) demonstrated that

$$F(x) = \sum_{r=0}^{\infty} b_r\, G(x;\gamma)^r, \qquad (29)$$

where the coefficient $b_r = \sum_{i,j=0}^{\infty} \sum_{k=r}^{\infty} t_{i,j,k,r}$ denotes a sum of constants and $t_{i,j,k,r}$ is defined in (7).

Equation (29) gives the KBGG cdf as an infinite weighted power series of the baseline cdf. Inserting (2) in (29), we have

$$F(x) = \sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r. \qquad (30)$$

Combining (7) and (30), the pdf of the ith order statistic reduces to

$$f_{i:n}(x) \;=\; \sum_{j=0}^{n-i} \frac{n!\,(-1)^j}{(i-1)!(n-i)!} \binom{n-i}{j} \left[\sum_{r=0}^{\infty} c_r\, v_{r+1}(x) \right] \left[\sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r \right]^{i+j-1} \qquad (31)$$

Applying the identity (10) in (31), we have

$$\left[\sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r \right]^{i+j-1} = \sum_{r=0}^{\infty} d^{\star}_{i+j-1,r} \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r, \qquad (32)$$

where $d^{\star}_{i+j-1,r}$ can be obtained from (11) as $d^{\star}_{i+j-1,r} = (rb_0)^{-1} \sum_{p=1}^{m} [p(i+j) - r] b_p\, d^{\star}_{i+j-1,r-p}$ for $r \geq 1$ and $d^{\star}_{i+j-1,0} = b_0^{i+j-1}$ Inserting (12) in equation (32), we obtain

$$\left[\sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r \right]^{i+j-1} = \sum_{r,m=0}^{\infty} \frac{d_{r,m}\, d^{\star}_{i+j-1,r}}{\Gamma(k)^r} \left(\frac{x}{\alpha}\right)^{\beta(kr+m)}. \qquad (33)$$

Substituting (13) and (33) in equation (31), we can write

$$f_{i:n}(x) = \sum_{r,m=0}^{\infty}\sum_{j=0}^{n-i}\frac{(-1)^j\binom{n-i}{j}\,n!\,c_r\,e_{r,m}\,d_{r,m}\,d^{\star}_{i+j-1,r}\,\Gamma(k^{\star\star})}{(i-1)!\,(n-i)!\,\Gamma[k(r+1)+m]\,\Gamma(k)^r}\,g_{\alpha,\beta,k^{\star\star}}(x), \tag{34}$$

where $k^{\star\star} = k(2r+1)+2m$ and

$$g_{\alpha,\beta,k^{\star\star}}(x) = \frac{\beta}{\alpha\,\Gamma(k^{\star\star})}\left(\frac{x}{\alpha}\right)^{\beta k^{\star\star}-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]$$

denotes the GG(α, β, $k^{\star\star}$) density function.

Equation (34) reveals that the density function of the KBGG order statistics is an infinite linear combination of GG densities. Hence, ordinary moments of order statistics can be determined directly from those quantities of the GG distribution.

The sth moment of $X_{i:n}$ comes from (34) as

$$\mathrm{E}(X_{i:n}^s) = \sum_{r,m=0}^{\infty}\sum_{j=0}^{n-i}\frac{(-1)^j\binom{n-i}{j}\,n!\,c_r\,e_{r,m}\,d_{r,m}\,d^{\star}_{i+j-1,r}\,\Gamma(k^{\star\star})}{(i-1)!\,(n-i)!\,\Gamma[k(r+1)+m]\,\Gamma(k)^r}\,\mathrm{E}(X_{r,m}^s), \tag{35}$$

where $X_{r,m} \sim$ GG(α, β, $k^{\star\star}$). Equation (35) is the main result of this section.

Based upon these moments, we can derive expansions for the L-moments as infinite weighted linear combinations of suitable KBGG means. 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 Hosking (1990) and are relatively robust to the effects of outliers.

8. Inference and estimation

8.1 The Classical Inference

Here, the estimation of the model parameters of the KBGG distribution is investigated by the maximum likelihood method. Let $X = (X_1, \ldots, X_n)$ be a random sample of the new distribution with unknown parameter vector $\theta = (a, b, c, \alpha, \beta, k)^{\mathrm{T}}$. The total log-likelihood function for θ is

$$\begin{aligned}\ell(\theta) = {} & n\log\left[\frac{K\,\beta}{\alpha\,\Gamma(k)}\right] + (\beta k - 1)\sum_{i=1}^{n}\log\left(\frac{x_i}{\alpha}\right) - \sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta} + (a-1)\sum_{i=1}^{n}\log\left\{\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}\\ & + (b-1)\sum_{i=1}^{n}\log\left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\} - c\sum_{i=1}^{n}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right].\end{aligned} \tag{36}$$

The elements of score vector are given by

$$U_a(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial a} + \sum_{i=1}^{n}\log[\gamma_1(k,u_i)]\,, \qquad U_b(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial b} + \sum_{i=1}^{n}\log[1-\gamma_1(k,u_i)],$$

$$U_c(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial c} - \sum_{i=1}^{n}\gamma_1(k,u_i),$$

$$\begin{aligned} U_\alpha(\boldsymbol{\theta}) &= -\frac{n}{\alpha} - \frac{n(\beta k-1)}{\alpha} + \frac{\beta}{\alpha}\sum_{i=1}^{n}u_i - \frac{\beta(a-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)|} \\ &+ \frac{\beta(b-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)} + \frac{\beta\,c}{\alpha\Gamma(k)}\sum_{i=1}^{n}v_i, \end{aligned}$$

$$\begin{aligned} U_\beta(\boldsymbol{\theta}) &= \frac{n}{\beta} + k\sum_{i=1}^{n}s_i - \sum_{i=1}^{n}u_i\,s_i + (a-1)\sum_{i=1}^{n}\frac{v_i\,s_i}{\gamma(k,u_i)} \\ &+ (1-b)\sum_{i=1}^{n}\frac{v_i\,s_i}{\Gamma(k)-\gamma(k,u_i)} - \frac{c}{\Gamma(k)}\sum_{i=1}^{n}v_i\,s_i, \end{aligned}$$

$$\begin{aligned} U_k(\boldsymbol{\theta}) &= -n\,\psi(k) + \beta\sum_{i=1}^{n}s_i - n\,(a-1)\,\psi(k) + (a-1)\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\gamma(k,u_i)} \\ &+ (1-b)\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} + \psi(k)\,(b-1)\sum_{i=1}^{n}\frac{\gamma(k,u_i)}{\Gamma(k)-\gamma(k,u_i)} \\ &- \frac{c}{\Gamma(k)}\sum_{i=1}^{n}\gamma'(k,u_i)|_k + \frac{c\,\psi(k)}{\Gamma(k)}\sum_{i=1}^{n}\gamma(k,u_i), \end{aligned}$$

where $u_i = \left(\frac{x_i}{\alpha}\right)^\beta$, $v_i = \left(\frac{x_i}{\alpha}\right)^\beta exp\left[-\left(\frac{x_i}{\alpha}\right)^\beta\right]$, $s_i = log\left(\frac{x_i}{\alpha}\right)$, $\gamma'(k,u_i)|_k = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}J(u_i, k + n - 1, 1)$, $\psi(\cdot)$ is the digamma function and $J(u_i, k + n - 1, 1)$ is defined in Appendix A. The partial derivatives of K in relation to a, b and c are given by

$$\frac{\partial K}{\partial a} = -\frac{\left\{[\psi(a) - \psi(a+b)]\ {}_1F_1(a, a+b, -c) + \frac{\partial\, {}_1F_1(a,a+b,-c)}{\partial a}\right\}}{B(a,b)\left[{}_1F_1(a, a+b, -c)\right]^2},$$

$$\frac{\partial K}{\partial b} = -\frac{\left\{[\psi(b) - \psi(a+b)]\ {}_1F_1(a, a+b, -c) + \frac{\partial\, {}_1F_1(a,a+b,-c)}{\partial b}\right\}}{B(a,b)\left[{}_1F_1(a, a+b, -c)\right]^2},$$

$$\frac{\partial K}{\partial c} = \frac{a\, {}_1F_1(a+1, a+b+1, -c)}{(a+b)B(a,b){}_1F_1(a, a+b, -c)}, \text{ where}$$

$$\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} = -\left[\psi(a) - \psi(a+b)\right]\ {}_1F_1(a, a+b, -c) - \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{k!(a+b)_k}\left[\psi(a+b+k) - \psi(a+k)\right]$$

and

$$\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} = \psi(a+b)\, {}_1F_1(a, a+b, -c) + \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{k!(a+b)_k}\psi(a+b+k).$$

Maximization of (36) can be performed using well established routines such as the nlm routine or optimize in the R statistical package. Setting these equations to zero, $U_a(\boldsymbol{\theta}) = U_b(\boldsymbol{\theta}) = U_c(\boldsymbol{\theta}) = U_\alpha(\boldsymbol{\theta}) = U_\beta(\boldsymbol{\theta}) = U_k(\boldsymbol{\theta}) = 0$, and solving them simultaneously yields the maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$. These equations cannot be solved analytically and statistical software can be used to solve them numerically by means of iterative techniques such as the Newton-Raphson algorithm.

For interval estimation and hypothesis tests on the parameters in $\boldsymbol{\theta}$, we require the 6 × 6 total observed information matrix $\mathbf{J}(\boldsymbol{\theta}) = -\{U_{rs}\}$, where the elements U_{rs} for $r, s = \alpha, \beta, k, a, b, c$ are given in Appendix A. The estimated asymptotic multivariate normal $N_6(\boldsymbol{\theta}, \mathbf{J}(\widehat{\boldsymbol{\theta}})^{-1})$ distribution of $\widehat{\boldsymbol{\theta}}$ can be used to construct approximate condence regions for the parameters. An asymptotic condence interval (ACI) with signicance level γ for each parameter θ_r is given by

$$\mathrm{ACI}(\theta_r, 100(1-\gamma)\%) = (\hat{\theta}_r - z_{\gamma/2}\sqrt{\hat{\kappa}^{\theta_r,\theta_r}}, \hat{\theta}_r + z_{\gamma/2}\sqrt{\hat{\kappa}^{\theta_r,\theta_r}}),$$

where $\hat{\kappa}^{\theta_r,\theta_r}$ is the rth diagonal element of $\mathbf{J}(\boldsymbol{\theta})^{-1}$ estimated at $\widehat{\boldsymbol{\theta}}$, for $r = 1, \ldots, 4$, and $z_{\gamma/2}$ is the quantile $1 - \gamma/2$ of the standard normal distribution.

We can compute the maximum values of the unrestricted and restricted log-likelihoods to construct likelihood ratio (LR) statistics for testing some sub-models of the KBGG distribution. For example, we may use LR statistics to check if the fit using the KBGG distribution is statistically "superior" to the fits using the KBW, BGHN, EW and GG distributions for a given data set. In any case, considering the partition $\boldsymbol{\theta} = \left(\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T\right)^T$, tests of hypotheses of the type H_0 : $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ versus H_A : $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$ can be performed using the LR statistic $\omega = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\}$, where $\widehat{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}$ are the estimates of $\boldsymbol{\theta}$ under H_A and H_0, respectively. Under the null hypothesis H_0, $\omega \overset{d}{\rightarrow} \mathcal{X}_q^2$, where q is the dimension of the vector $\boldsymbol{\theta_1}$ of interest. The LR test rejects H_0 if $\omega > \xi_r$, where ξ_r denotes the upper 100γ% point of the χ_q^2 distribution.

8.2 The Bayesian Inference

As is well-known, the Bayesian approach allows the incorporation of previous knowledge of the parameters through informative prior density functions. When this information is not available, we can consider a non-informative prior. In the Bayesian context, the information referring to the model parameters is obtained through a posterior marginal distribution. Thus, two difficulties usually arise. The first refers to attaining marginal posterior distribution, and the second to the calculation of the moments of interest. Both cases require numerical integration that, many times, do not present an analytical solution. To overcome these problems, we use the simulation method based on the Markov Chain Monte Carlo (MCMC), such as the Gibbs sampler and Metropolis-Hastings algorithms.

Since we have no prior information from historical data or from previous experiment, we assign conjugate but weakly informative prior distributions to the parameters. Since we assume informative (but weakly) prior distribution, the posterior distribution is a well-defined proper distribution. We suppose that the elements of the parameter vector $\boldsymbol{\theta} = (a, b, c, \alpha, \beta, k)^T$ are independent and consider that the joint prior distribution of all unknown parameters has a density function given by

$$\pi(a,b,c,\alpha,\beta,k) \propto \pi(a) \times \pi(b) \times \pi(c) \times \pi(\alpha) \times \pi(\beta) \times \pi(k), \quad (37)$$

where, $a \sim \Gamma(a_1, b_1)$, a_1 and b_1 known; $b \sim \Gamma(a_2, b_2)$, a_2 and b_2 known; $c \sim \mathrm{N}(\mu_0, \sigma_0^2)$, μ_0 and σ_0^2 known; $\alpha \sim \Gamma(a_3, b_3)$, a_3 and b_3 known; $\beta \sim \Gamma(a_4, b_4)$, a_4 and b_4 known; $k \sim \Gamma(a_5, b_5)$, a_5 and a_5 known; where $\Gamma(a_i, b_i)$ denotes the gamma distribution with mean a_i/b_i, variance a_i/b_i^2 for $a_i > 0$ and $b_i > 0$, and $N(\mu_0, \sigma_0^2)$ denotes the normal distribution with mean μ_0, variance σ_0^2 for $\mu_0 \in \mathbb{R}$ and $\sigma_0^2 > 0$. We note that gamma and normal priors are most commonly used priors for positive and real-values parameters.

Combining the likelihood function (36) and the prior distribution (37), the joint posterior distribution for a, b, c, α, β and k reduces to

$$\begin{aligned}\pi(a,b,c,\alpha,\beta,k|x) \quad \propto \quad & \left[\frac{K\beta}{\alpha\Gamma(k)}\right]^n \exp\left\{-c\sum_{i=1}^{n}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^\beta\right] - \sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^\beta\right\} \\ & \times \prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^\beta\right]^{a-1}\left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^\beta\right]\right\}^{b-1} \quad (38) \\ & \times \pi(a,b,c,\alpha,\beta,k).\end{aligned}$$

The joint posterior density (38) is analytically intractable because the integration of the joint posterior density is not easy to perform. So, the inference can be based on MCMC simulation methods such as the Gibbs sampler and Metropolis-Hastings algorithm, which can be used to draw samples, from which features of the marginal distributions of interest can be inferred. In this direction, we first obtain the full conditional distributions of the unknown quantities given by

$$\pi(a|x,b,c,\alpha,\beta,k) \propto K^n \prod_{i=1}^{n} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1} \times \pi(a),$$

$$\pi(b|x,a,c,\alpha,\beta,k) \propto K^n \prod_{i=1}^{n} \left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1} \times \pi(b),$$

$$\pi(c|x,a,b,c,\alpha,\beta,k) \propto K^n \exp\left\{-c\sum_{i=1}^{n} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\} \times \pi(c),$$

$$\begin{aligned}\pi(\alpha|x,a,b,c,\beta,k) \propto\ & \frac{1}{\alpha^n} \exp\left\{-c\sum_{i=1}^{n} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right] - \sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right\} \\ & \times \prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1} \left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1} \times \pi(\alpha),\end{aligned}$$

$$\begin{aligned}\pi(\beta|x,a,b,c,\alpha,k) \propto\ & \beta^n \exp\left\{-c\sum_{i=1}^{n} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right] - \sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right\} \\ & \times \prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1} \left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1} \times \pi(\beta)\end{aligned}$$

and

$$\begin{aligned}\pi(k|x,a,b,c,\alpha,\beta) \propto\ & \frac{1}{\Gamma(k)^n} \exp\left\{-c\sum_{i=1}^{n} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\} \\ & \times \prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1} \gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1} \left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1} \times \pi(k).\end{aligned}$$

Since the full conditional distributions do not have explicit expressions, we require the use of the Metropolis-Hastings algorithm to generate the variables a, b, c, α, β and k for the KBGG distribution.

9. Applications

In this section, we use three real data sets which come from diverse fields such as actuarial sciences (D1), environment (D2) and engineering (D3) to compare the fits of the KBGG distribution with those of three sub-models (i.e. BGG, EGG and GG distributions) and also to the following non-nested model: the Kumaraswamy generalized gamma (KwGG) distribution (Pascoa *et al*., 2011). The primary reason for choosing these data is that they allow us to show

how in different fields it is necessary to have positively skewed distributions with non-negative support. Moreover, these data sets present different degrees of variability, skewness and kurtosis.

9.1 Applications

Description of the data sets

D1 **Actuarial sciences:** It is important for the Mexican Institute of Social Security (IMSS) to study the distributional behaviour of the mortality of retired people on disability because it enables the calculation of long and short term financial estimation, such as the assessment of the reserve required to pay the minimum pensions. The data set corresponding to 280 lifetimes (in years) of retired women with temporary disabilities, which are incorporated in the Mexican insurance public system and who died during 2004 were reported and analyzed by Balakrishnan *et al*. (2009).

D2 **Environmental sciences:** These data were analyzed by Leiva *et al.* (2009) and correspond to daily ozone level measurements in New York in May-September, 1973, from the New York State Department of Conservation.

D3 **Engineering:** Failures can occur in microcircuits because of the movement of atoms in the conductors in the circuit, which is referred to the electromigration. The data set refers to an accelerated life test of 59 conductors reported by Lawless (1982).

Table 2 gives a descriptive summary for these data and suggest positively skewed distributions with different degrees of variability, skewness and kurtosis.

9.2 Maximum likelihood estimation

First, in order to estimate the model parameters, we consider the maximum likelihood estimation method discussed in Section 8.1. We take the estimates of α, β and k from the fitted GG distribution as starting values for the numerical iterative procedure. All the computations were performed using the R statistical software. Table 3 lists the MLEs of the parameters and the values of the following statistics for some models: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC). The results indicate that the KBGG model has the smallest values of the statistics (AIC and CAIC) among all fitted models. So, it could be chosen as the more suitable model.

Table 2: Descriptive statistics for the three data sets.

Data	Mean	Median	SD	Variance	Skewness	Kurtosis	Min.	Max.
D1	47.78	49	57.0	108.62	0.06	3.05	22.0	86.0
D2	42.12	31.50	32.98	1088.20	1.22	4.18	1.0	168.0
D3	6.98	6.92	1.61	2.60	0.19	3.08	2.997	11.038

Table 3: MLEs of the model parameters for the three data sets and the corresponding AIC, CAIC and BIC statistics.

Data	Model	α	β	k	a	b	c	AIC	CAIC	BIC
D1	KBGG	7.4489	1.8564	35.6798	0.1621	0.4310	-0.9485	2104.0	2104.5	2125.8
	BGG	32.7248	1.9705	3.9984	0.9872	3.8361	0	2115.1	2115.5	2133.2
	EGG	33.8389	2.8185	3.1002	0.9056	1	0	2113.3	2113.3	2127.9
	GG	34.1730	2.8048	2.8693	1	1	0	2111.3	2111.6	2122.2
	Model	α	τ	k	λ	ϕ	-	AIC	CAIC	BIC
	KwGG	33.5340	1.4296	2.6525	2.2376	9.2101	-	2114.8	2115.2	2133.0
D2	KBGG	3.0409	1.0760	20.2422	0.0807	0.1598	-0.2154	1067.1	1068.5	1083.6
	BGG	4.0775	1.1376	17.5934	0.0923	0.1749	0	1087.7	1088.7	1101.5
	EGG	3.7038	0.6370	4.9592	0.7285	1	0	1090.2	1090.9	1101.2
	GG	3.1291	0.5924	4.3440	1	1	0	1088.3	1088.8	1096.6
	Model	α	τ	k	λ	ϕ	-	AIC	CAIC	BIC
	KwGG	0.6009	0.5508	11.2001	0.4059	0.7496	-	1091.9	1093.0	1105.7
D3	KBGG	7.0954	8.1282	2.0878	0.3840	0.1030	2.7935	221.7	224.5	234.1
	BGG	4.720	2.0391	3.0389	1.3445	2.1157	0	232.6	234.8	243.0
	EGG	0.0200	0.5933	28.3765	2.3890	1	0	234.0	235.6	242.3
	GG	4.1439	2.3300	3.6446	1	1	0	228.6	229.8	234.9
	Model	α	τ	k	λ	ϕ	-	AIC	CAIC	BIC
	KwGG	4.1410	1.8808	3.4611	1.3199	2.1071	-	232.6	234.8	243.0

A comparison of the proposed distribution with some of its sub-models using LR statistics is given in Table 4. The *p*-values indicate that the proposed model yields the best fit to the three data sets. This gives a clear evidence of the potential of the three parameters when modeling real data.

In order to assess if the model is appropriate, Figure 4 displays histograms and the estimated KBGG density functions for these data sets, respectively. We can conclude that the new distribution is a very suitable model to fit the three data sets.

9.3 Bayesian analysis

For the three data sets, the following independent priors were considered to perform the Metropolis-Hastings algorithm: $\alpha \sim \Gamma(0.01, 0.01)$, $\beta \sim \Gamma(0.01, 0.01)$, $k \sim \Gamma(0.01, 0.01)$, $a \sim \Gamma(0.01, 0.01)$, $b \sim \Gamma(0.01, 0.01)$ and $c \sim N(0,100)$, so that we have vague prior distributions. Considering these prior density functions, we generate two parallel independent runs of the Metropolis-Hastings with size 300, 000 for each parameter, disregarding the first 30, 000 iterations to eliminate the effect of the initial values and, to avoid correlation problems, we consider a spacing of size 10, obtaining a sample of size 27,000 from each chain. To monitor the convergence of the Metropolis-Hastings algorithm, we perform the methods suggested by Cowles and Carlin (1996) using the between and within sequence information, following the approach developed in Gelman and Rubin (1992) to obtain the potential scale reduction, $\hat{R}$. In all cases, these values were close to one, indicating the convergence of the chain.

Table 4: LR statistics for the three data sets.

Data	Model	Hypotheses	Statistic w	p-value
D1	KBGG vs BGG	$H_0 : c = 0$ vs $H_1 : H_0$ is false	13.10	0.00029
	KBGG vs EGG	$H_0 : c = 0$ and $b = 1$ vs $H_1 : H_0$ is false	13.36	0.00124
	KBGG vs GG	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	13.38	0.00387
D2	KBGG vs BGG	$H_0 : c = 0$ vs $H_1 : H_0$ is false	22.57	< 0.0001
	KBGG vs EGG	$H_0 : c = 0$ and $b = 1$ vs $H_1 : H_0$ is false	27.03	< 0.0001
	KBGG vs GG	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	27.17	< 0.0001
D3	KBGG vs BGG	$H_0 : c = 0$ vs $H_1 : H_0$ is false	12.92	0.00032
	KBGG vs EGG	$H_0 : c = 0$ and $b = 1$ vs $H_1 : H_0$ is false	16.32	0.00028
	KBGG vs GG	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	12.96	0.00471

The approximate posterior marginal density functions for the parameters are displayed in Figures 5, 6 and 7 for the first, second and third data sets, respectively. In Table 5, we report posterior summaries for the parameters of the KBGG model for the three data sets. We note that the values for the means a posteriori (Table 5) are quite close (as expected) to the MLEs obtained for the KBGG model given in Table 3. “SD” denotes the standard deviation from the posterior distributions of the parameters and “HPD” denotes the 95% highest posterior density intervals.

Table 5: Posterior summaries for the parameters from the KBGG model for the three data sets.

D1				
Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
α	7.4399	0.0099	(7.4201; 7.4590)	1.0005
β	1.8499	0.0099	(1.8301; 1.8689)	1.0014
k	35.6701	0.01003	(35.6500; 35.6892)	1.0004
a	0.1594	0.0098	(0.1407; 0.1792)	1.0003
b	0.4301	0.01008	(0.4103; 0.4499)	0.9997
c	-0.9401	0.0099	(-0.9594; -0.9204)	1.0002
D2				
Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
α	3.0399	0.0099	(3.0201; 3.0590)	1.0002
β	1.0599	0.0009	(1.0580; 1.0618)	1.0009
k	20.2401	0.01002	(20.2200; 20.2592)	1.0002
a	0.0798	0.0098	(0.0611; 0.0997)	1.0011
b	0.1502	0.01009	(0.1304; 0.1700)	0.9996
c	-0.2100	0.0009	(-0.2119; -0.2080)	1.0005
D3				
Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
α	7.0599	0.0099	(7.0401; 7.0790)	1.0001
β	8.1499	0.0099	(8.1301; 8.1689)	1.0011
k	2.0694	0.0100	(2.0491; 2.0883)	1.0002
a	0.3697	0.0098	(0.3510; 0.3896)	0.9997
b	0.0998	0.0100	(0.0798; 0.1192)	0.9998
c	2.7998	0.0009	(2.7980; 2.8019)	1.0006

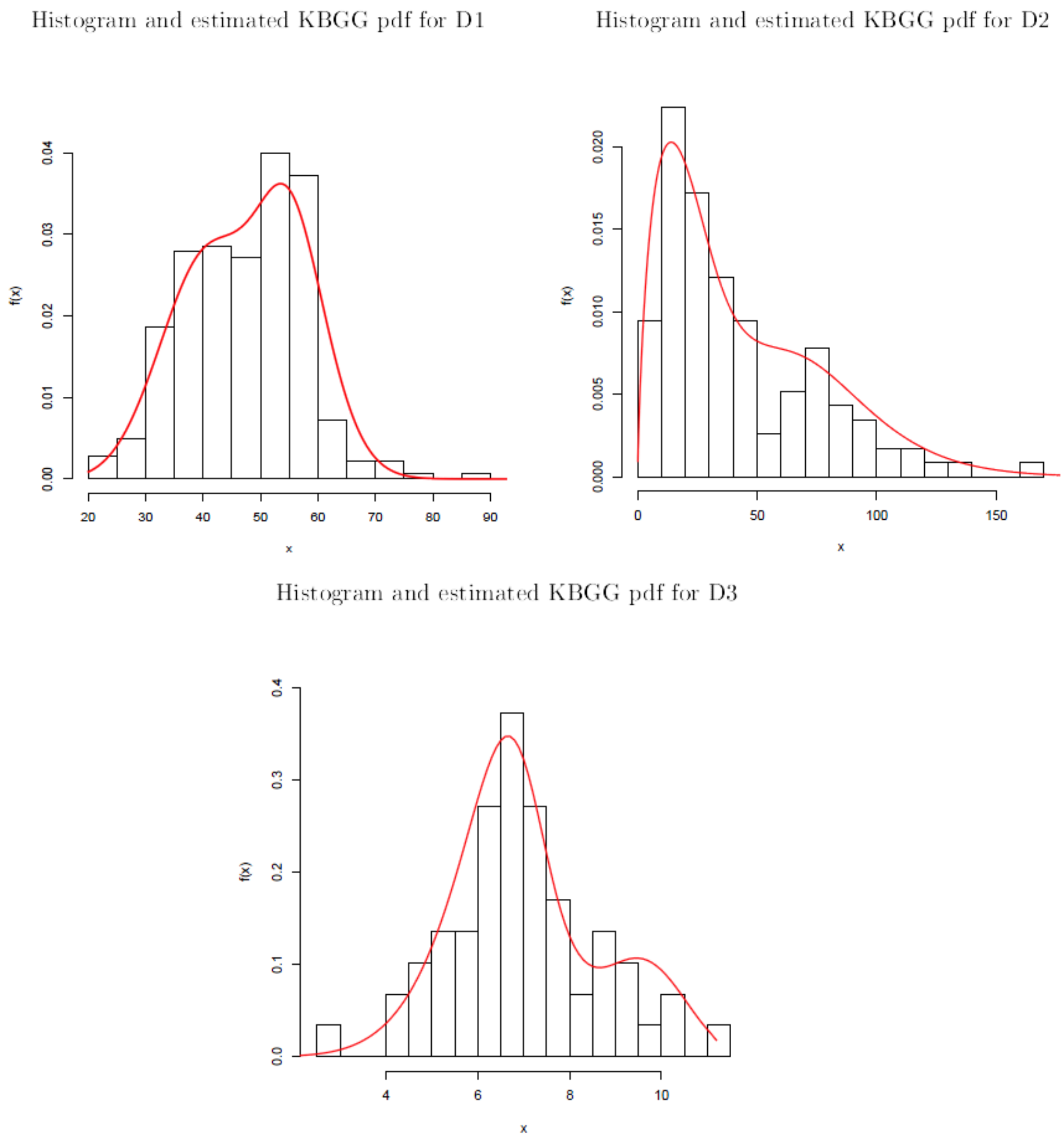

Figure 4: Histograms and the estimated KBGG density functions for the current data sets.

10. Concluding remarks

We introduce the Kummer beta generalized gamma (KBGG) distribution with three additional shape parameters because of the wide usage of the GG distribution and the fact that the current generalization provides extensions to its continuous extension to still more complex situations. The new distribution unifies more than 28 distributions and yields a general overview of these distributions for theoretical studies. In fact, the KBGG distribution (5) generalizes the Weibull, gamma, exponentiated Weibull, exponentiated gamma, beta Weibull, beta gamma, Kummer beta Weibull and Kummer beta gamma distributions and other important lifetime models. The KBGG density function can be expressed as a linear combination of GG density functions which allow us to derive some of its mathematical properties. The estimation of the model parameters is approached by the method of maximum likelihood and the Bayesian analysis. We consider the likelihood ratio (LR) statistic and other criteria to compare the KBGG model with its sub-models and other non-nested model. The potentiality of the KBGG distribution is illustrated in three applications to real data sets. The new model provides a rather flexible mechanism for fitting a wide spectrum of real world lifetime data in reliability, biology and other areas.

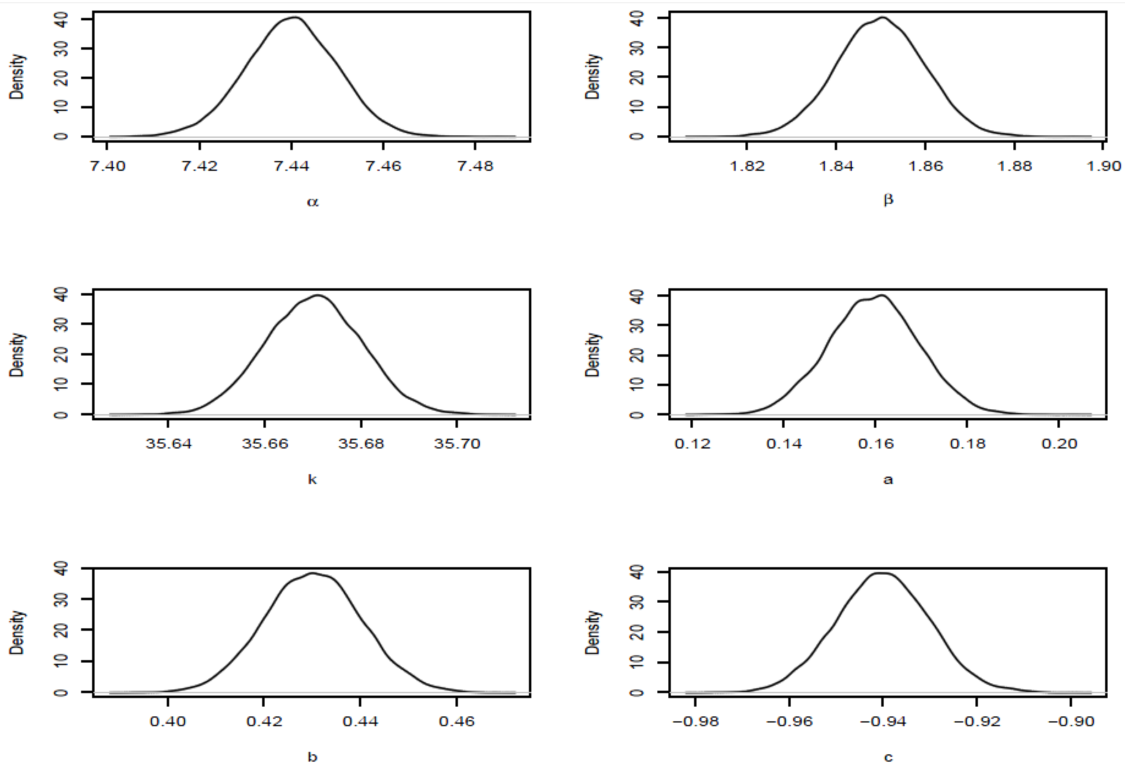

Figure 5: Approximate posterior marginal densities for the parameters of the KBGG model for the first data set.

Acknowledgments

The authors are grateful to the editor and the referees for helpful comments and suggestions. We gratefully acknowledge grants from CNPq, Brazil.

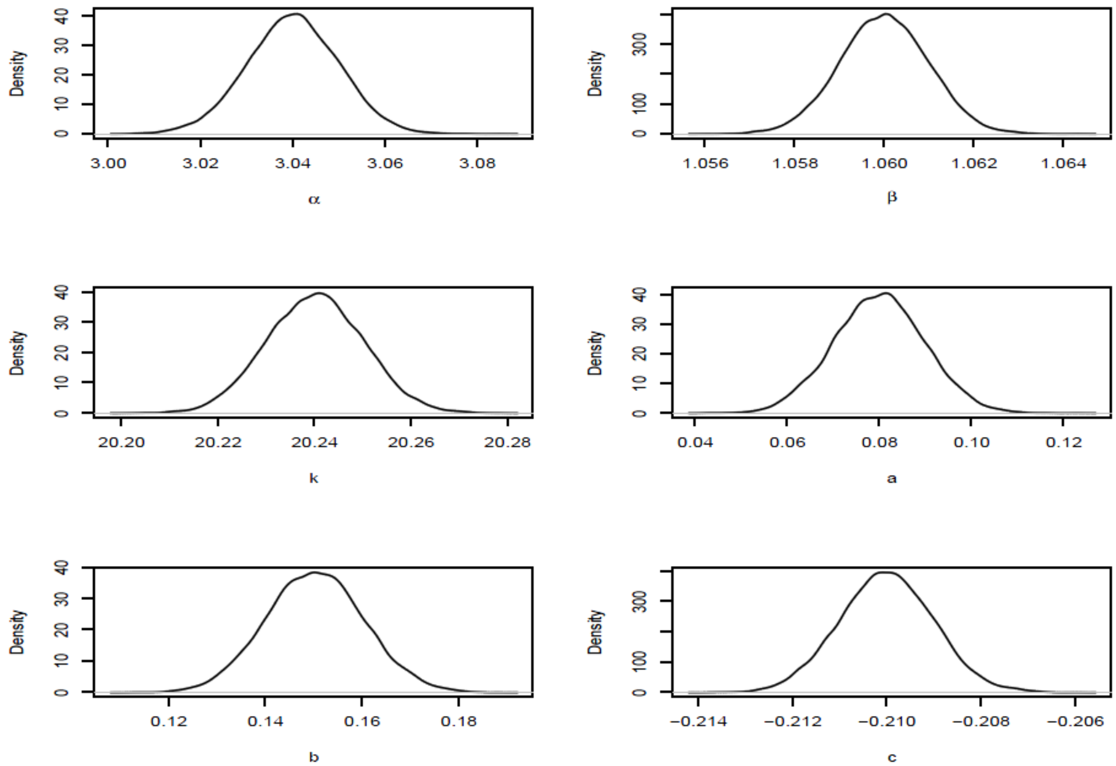

Figure 6: Approximate posterior marginal densities for the parameters of the KBGG model for the second data set.

Appendix A: Elements of the observed information matrix

The elements of the observed information matrix, $\mathbf{J}(\theta)$, for the parameters α, β, k, a, b and c are:

$$
\begin{aligned}
U_{\alpha\alpha} \;=\; & \frac{n}{\alpha^2} + \frac{n(\beta\,k-1)}{\alpha^2} - \frac{\beta(\beta+1)}{\alpha^2}\sum_{i=1}^{n} u_i + \frac{\beta(a-1)}{\alpha^2}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} \\
& - \frac{\beta(a-1)}{\alpha}\left\{ -\frac{\beta\,k}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} + \frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{u_i\,v_i}{\gamma(k,u_i)} + \frac{\beta}{\alpha}\sum_{i=1}^{n}\left[\frac{v_i}{\gamma(k,u_i)}\right]^2 \right\} \\
& - \frac{\beta(b-1)}{\alpha^2}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)} + \frac{\beta(b-1)}{\alpha}\left\{ -\frac{\beta\,k}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)} \right. \\
& \left. + \frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{u_i\,v_i}{\Gamma(k)-\gamma(k,u_i)} - \frac{\beta}{\alpha}\sum_{i=1}^{n}\left[\frac{v_i}{\Gamma(k)-\gamma(k,u_i)}\right]^2 \right\} \\
& - \frac{\beta\,c}{\alpha^2\,\Gamma(k)}\sum_{i=1}^{n} v_i + \frac{\beta\,c}{\alpha\Gamma(k)}\left[-\frac{\beta\,k}{\alpha}\sum_{i=1}^{n} v_i + \frac{\beta}{\alpha}\sum_{i=1}^{n} u_i\,v_i \right],
\end{aligned}
$$

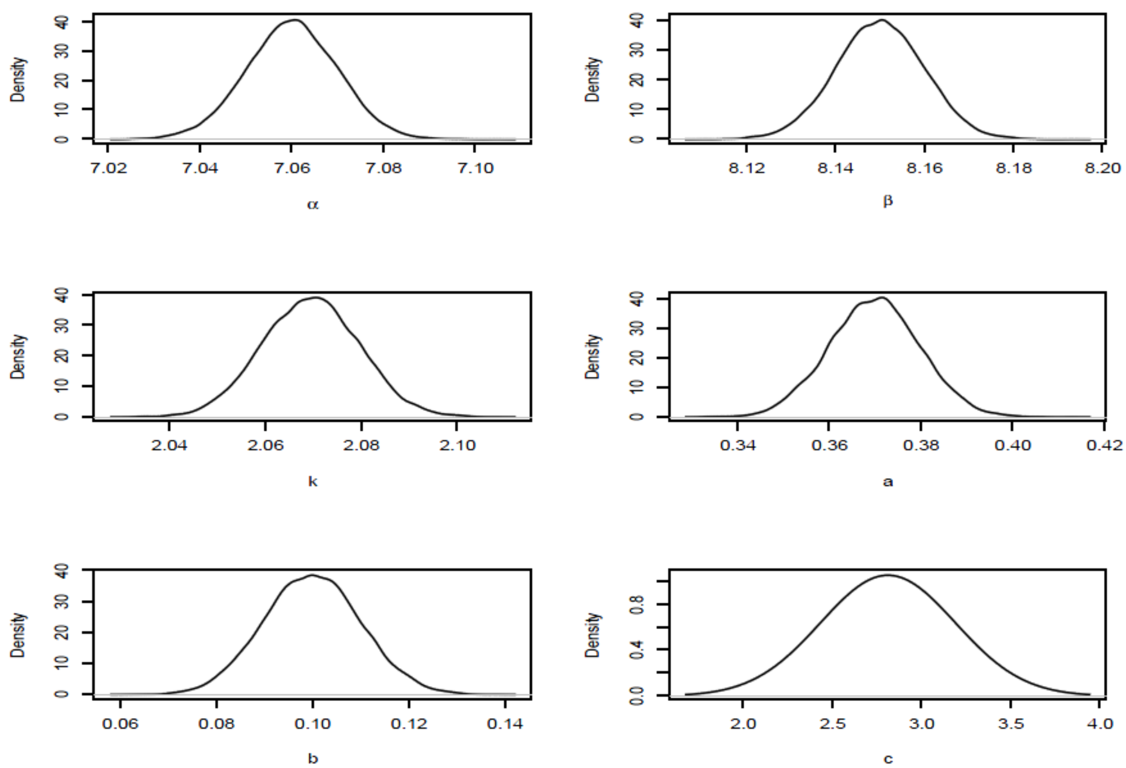

Figure 7: Approximate posterior marginal densities for the parameters of the KBGG model for the third data set.

$$
\begin{aligned}
U_{\alpha\beta} \;=\;& -\frac{n\,k}{\alpha} + \frac{1}{\alpha}\sum_{i=1}^{n} u_i + \frac{\beta}{\alpha}\sum_{i=1}^{n} u_i\, s_i - \frac{(a-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} - \frac{\beta(a-1)}{\alpha}\left\{k\sum_{i=1}^{n}\frac{v_i\, s_i}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{u_i\, v_i\, s_i}{\gamma(k,u_i)}\right.\\
&\left. - \sum_{i=1}^{n}\frac{v_i^2\, s_i}{[\gamma(k,u_i)]^2}\right\} + \frac{(b-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)} + \frac{\beta(b-1)}{\alpha}\left\{k\sum_{i=1}^{n}\frac{v_i\, s_i}{\Gamma(k)-\gamma(k,u_i)}\right.\\
&\left. - \sum_{i=1}^{n}\frac{u_i\, v_i\, s_i}{\Gamma(k)-\gamma(k,u_i)} + \sum_{i=1}^{n}\frac{v_i^2\, s_i}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\} + \frac{c}{\alpha\,\Gamma(k)}\sum_{i=1}^{n} v_i\\
&+ \frac{\beta\, c}{\alpha\,\Gamma(k)}\left\{k\sum_{i=1}^{n} v_i\, s_i - \sum_{i=1}^{n} u_i\, v_i\, s_i\right\},
\end{aligned}
$$

$$
\begin{aligned}
U_{\alpha k} \;=\;& -\frac{n\beta}{\alpha} - \frac{\beta(a-1)}{\alpha}\left\{\beta\sum_{i=1}^{n}\frac{v_i\, s_i}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{v_i\gamma'(k,u_i)|_k}{[\gamma(k,u_i)]^2}\right\} + \frac{\beta(b-1)}{\alpha}\left\{\beta\sum_{i=1}^{n}\frac{v_i\, s_i}{\Gamma(k)-\gamma(k,u_i)}\right.\\
&\left. - \sum_{i=1}^{n}\frac{\Gamma(k)\,\psi(k)\, v_i}{[\Gamma(k)-\gamma(k,u_i)]^2} + \sum_{i=1}^{n}\frac{v_i\,\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\} - \frac{\beta\, c\,\psi(k)}{\alpha\Gamma(k)}\sum_{i=1}^{n} v_i + \frac{\beta^2\, c}{\alpha\Gamma(k)}\sum_{i=1}^{n} v_i\, s_i,
\end{aligned}
$$

$$
\begin{aligned}
U_{\beta\beta} \;=\;& -\frac{n}{\beta^2} - \sum_{i=1}^{n} u_i\, s_i^2 + (a-1)\left\{k\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{u_i\, v_i\, s_i^2}{\gamma(k,u_i)} - \sum_{i=1}^{n}\left[\frac{u_i^{1/2}\, v_i\;\, s_i}{\gamma(k,u_i)}\right]^2\right\}\\
&-(b-1)\left\{k\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{u_i\, v_i\, s_i^2}{\Gamma(k)-\gamma(k,u_i)} + \sum_{i=1}^{n}\left[\frac{u_i^{1/2}\, v_i\;\, s_i}{\Gamma(k)-\gamma(k,u_i)}\right]^2\right\}\\
&- \frac{c}{\Gamma(k)}\left\{k\sum_{i=1}^{n} v_i\, s_i^2 - \sum_{i=1}^{n} u_i\, v_i\, s_i^2\right\},
\end{aligned}
$$

$$
\begin{aligned}
U_{\beta k} \;=\;& \sum_{i=1}^{n} s_i + (a-1)\left\{\beta\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{v_i\, s_i\,\gamma'(k,u_i)|_k}{[\gamma(k,u_i)]^2}\right\}\\
&-(b-1)\left\{\beta\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{v_i\, s_i\,\Gamma(k)\,\psi(k)}{[\Gamma(k)-\gamma(k,u_i)]^2} + \sum_{i=1}^{n}\frac{v_i\, s_i\,\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\}\\
&+ \frac{c}{\Gamma(k)}\left\{\psi(k)\sum_{i=1}^{n} v_i\, s_i - \beta\sum_{i=1}^{n} v_i\, s_i^2\right\},
\end{aligned}
$$

$$\begin{aligned} U_{kk} &= -n\,a\,\psi'(k) + (a-1)\left\{\sum_{i=1}^{n}\frac{\gamma''(k,u_i)|_k}{\gamma(k,u_i)} - \sum_{i=1}^{n}\left[\frac{\gamma'(k,u_i)|_k}{\gamma(k,u_i)}\right]^2\right\} \\ &\quad -(b-1)\left\{\sum_{i=1}^{n}\frac{\gamma''(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{\Gamma(k)\,\psi(k)\,\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2} + \sum_{i=1}^{n}\left[\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)}\right]^2\right\} \\ &\quad +(b-1)\sum_{i=1}^{n}\frac{\psi'(k)\,\gamma(k,u_i)}{\Gamma(k)-\gamma(k,u_i)} + \psi(k)\,(b-1)\left\{\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{\Gamma(k)\,\psi(k)\,\gamma(k,u_i)}{[\Gamma(k)-\gamma(k,u_i)]^2}\right. \\ &\quad \left. + \sum_{i=1}^{n}\frac{\gamma(k,u_i)\,\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\} + \frac{c}{\Gamma(k)}\left\{\psi(k)\sum_{i=1}^{n}\gamma'(k,u_i)|_k - \sum_{i=1}^{n}\gamma''(k,u_i)|_k\right\} \\ &\quad + \left\{\frac{c}{\Gamma(k)}\left[\psi'(k)-\psi^2(k)\right]\right\}\sum_{i=1}^{n}\gamma(k,u_i) + \frac{c\,\psi(k)}{\Gamma(k)}\sum_{i=1}^{n}\gamma'(k,u_i)|_k, \end{aligned}$$

$$U_{\alpha a} = -\frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)},\ U_{\alpha b} = \frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)},\ U_{\alpha c} = \frac{\beta}{\alpha\Gamma(k)}\sum_{i=1}^{n}v_i,\ U_{\beta a} = \sum_{i=1}^{n}\frac{v_i\,s_i}{\gamma(k,u_i)},$$

$$U_{\beta b} = -\sum_{i=1}^{n}\frac{v_i\,s_i}{\Gamma(k)-\gamma(k,u_i)},\ U_{\beta c} = -\frac{1}{\Gamma(k)}\sum_{i=1}^{n}v_i\,s_i,\ U_{ka} = -n\psi(k) + \sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\gamma(k,u_i)},$$

$$U_{kb} = -\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} + \psi(k)\sum_{i=1}^{n}\frac{\gamma(k,u_i)}{\Gamma(k)-\gamma(k,u_i)},\ U_{kc} = -\frac{1}{\Gamma(k)}\sum_{i=1}^{n}\gamma'(k,u_i)|_k + \frac{\psi(k)}{\Gamma(k)}\sum_{i=1}^{n}\gamma(k,u_i),$$

$$U_{aa} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a^2} - \frac{1}{K}\left[\frac{\partial K}{\partial a}\right]^2\right\},\ U_{bb} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial b^2} - \frac{1}{K}\left[\frac{\partial K}{\partial b}\right]^2\right\},\ U_{cc} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial c^2} - \frac{1}{K}\left[\frac{\partial K}{\partial c}\right]^2\right\},$$

$$U_{ab} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a\partial b} - \frac{1}{K}\frac{\partial K}{\partial a}\frac{\partial K}{\partial b}\right\},\ U_{ac} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a\partial c} - \frac{1}{K}\frac{\partial K}{\partial a}\frac{\partial K}{\partial c}\right\},\ U_{bc} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial b\partial c} - \frac{1}{K}\frac{\partial K}{\partial b}\frac{\partial K}{\partial c}\right\},$$

where

$$\gamma'(k,u_i)|_k = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\,J(u_i,k+n-1,1),$$

$$\gamma''(k,u_i)|_k = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\,J(u_i,k+n-1,2),$$

$\frac{\partial^2 K}{\partial a^2}, \frac{\partial^2 K}{\partial b^2}, \frac{\partial^2 K}{\partial c^2}, \frac{\partial^2 K}{\partial_a\partial_b}, \frac{\partial^2 K}{\partial_a\partial_c}$ and $\frac{\partial^2 K}{\partial_b\partial_c}$ are defined in Pescim *et al*. (2012). The $J(.,\ .,\ .)$ function can be determined from the integral given by Prudnikov *et al*. (1986, vol. 1, Section 2.6.3, integral 1)

$$J(a,p,1) = \int_0^a x^p \, \log(x) \, dx = \frac{a^{p+1}}{(p+1)} \left[(p+1) \, \log(a) - 1\right]$$

and

$$J(a,p,2) = \int_0^a x^p \, \log^2(x) \, dx = \frac{a^{p+1}}{(p+3)} \left\{2 - (p+1) \log(a) \left[2 - \log(a) \, (p+1)\right]\right\}.$$

References

[1] Abramowitz, M. and Stegun, I. A. (1968). *Handbook of Mathematical Functions*. Dover Publications, New York.

[2] Amusan, G. E. (2010). The Beta Maxwell Distribution. Thesis submitted to the Graduate College of Marshall University, Huntington, West Virginia.

[3] Balakrishnan, N., Leiva, V., Sanhueza, A. and Cabrera, E. (2009). Mixture inverse Gaussian distributions and its transformations, moments and applications. *Statistics*, **43**, 91-104.

[4] Barkauskas, D. A., Kronewitter, S. R., Lebrilla, C. B. and Rocke, D. M. (2009). Analysis of MALDI FT-ICR mass spectrometry data: a time series approach. *Analytica Chimica Acta*, **648**, 207-214.

[5] Bekker, A. and Roux, J. J. (2005). Reliability characteristics of the Maxwell distribution: a Bayes estimation study. *Communications in Statistics, Theory and Methods*, **34**, 2169-2178.

[6] Cooray, K. and Ananda, M. M. A. (2008). A generalization of the half-normal distribution with applications to lifetime data. *Communications in Statistics, Theory and Methods*, **37**, 1323-1337.

[7] Cordeiro, G. M., Castellares, F., Montenegro, L. C. and de Castro, M. (2013a). The beta generalized gamma distribution. *Statistics*, **47**, 888-900.

[8] Cordeiro, G. M., Cristino, C. T., Hashimoto, E. M. and Ortega, E. M. M. (2013b). The beta generalized Rayleigh distribution with applications to lifetime data. *Statistical Papers*, **54**, 133-161.

[9] Cordeiro, G. M., Ortega, E. M. M. and Silva, G. O. (2011). The exponentiated generalized gamma distribution with application to lifetime data. *Journal of Statistical Computation and Simulation*, **81**, 827-842.

[10] Cowles, M. K. and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, **91**, 133-169.

[11] Cox, C. (2008). The generalized F distribution: an umbrella for parametric survival analysis. *Statistics in Medicine*, **27**, 4301-4312.

[12] Cox, C., Chu, H., Schneider, M. F. and Muñoz, A. (2007). Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine*, **26**, 4352-4374.

[13] Famoye, F., Lee, C. and Olumolade, O. (2005). The beta-Weibull distribution. *Journal of Statistical Theory and Practice*, **4**, 121-136.

[14] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathematique*, **6**, 93-116.

[15] Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, **7**, 457-472.

[16] Gradshteyn, I. S. and Ryzhik, I. M. (2007). Table of integrals, series, and products. Academic Press, New York.

[17] Gupta, R. D. and Kundu, D. (2001). Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions. *Biometrical Journal*, **43**, 117-130.

[18] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communications Statistics, Theory and Methods*, **27**, 887-904.

[19] Hosking, J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B*, **52**, 105-124.

[20] Kondo, T. (1930). A theory of the sampling distribution of standard deviations. *Biometrika*, **22**, 36-64.

[21] Kong, L., Lee, C. and Sepanski, J. H. (2007). On the properties of beta-gamma distribution. *Journal of Modern Applied Statistical Methods*, **6**, 187-211.

[22] Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: dierent methods of estimation. *Computational Statistics and Data Analysis*, **49**, 187-200.

[23] Laplace, P. S. (1836). Theorie Analytique des Probabilities, (Supplement to 3rd ed.).

[24] Lawless, J. F. (1982). Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York.

[25] Leiva, V., Barros, M. and Paula, G.A. (2009). Generalized Birnbaum-Saunders Models using R. XI Escola de Modelos de Regressão, Recife, Brazil.

[26] Malhotra, J., Sharma, A. K. and Kaler, R. S. (2009). On the performance analysis of wireless receiver using generalized-gamma fading model. *Annals of Telecommunications*, **64**, 147-153.

[27] Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. *Philosophical Magazine*, **19**, 19-32.

[28] Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-real data. *IEEE Transaction on Reliability*, **42**, 299-302.

[29] Mudholkar, G .S, Srivastava, D. K. and Friemer, M. (1995). The exponential Weibull family: A reanalysis of the bus-motor failure data. *Technometrics*, **37**, 436-445.

[30] Nadarajah, S. and Gupta, A. K. (2007). A generalized gamma distribution with application to drought data. *Mathematics and Computer in Simulation*, **74**, 1-7.

[31] Nadarajah, S. and Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering and System Safety*, **91**, 689-697.

[32] Ng, K. W. and Kotz, S. (1995). Kummer-gamma and Kummer-beta univariate and multivariate distributions. *Research Report*, **84**, Department of Statistics, The University of Hong Kong, Hong Kong.

[33] Ortega, E. M. M., Bolfarine, H. and Paula, G. A. (2003). Inuence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, **42**, 165-186.

[34] Ortega, E. M. M., Cancho, V. G. and Paula, G. A. (2009). Generalized log-gamma regression models with cure fraction. *Lifetime Data Analysis*, **15**, 79-106.

[35] Pascoa, M. A. R., Ortega, E. M. M. and Cordeiro, G. M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, **8**, 411-433.

[36] Pescim, R. R., Cordeiro, G .M., Demétrio, C. G. B., Ortega, E. M. M. and Nadarajah, S. (2012). The new class of Kummer beta generalized distributions. *SORT*, **36**, 153-180.

[37] Pescim, R.R., Demétrio, C.G.B., Cordeiro, G.M., Ortega, E.M.M. and Urbano, M.R. (2010). The beta generalized half-normal distribution. *Computation Statistics and Data Analysis*, **54**, 945-957.

[38] Prudnikov, A.P., Brychkov, Y.A. and Marichev, O.I. (1986). Integrals and Series, vol. 1, Gordon and Breach Science Publishers, Amsterdam.

[39] Rayleigh, L. (1880). On the resultant of a large number vibrations of the same pitch and of arbitrary phase, *Phil*, **10**. 73-80.

[40] Stacy, E.W. (1962). A generalization of the gamma distribution, *Annals of Mathematical Statistics*, **33**, 1187-1192.

[41] Xie, X. and Liu, X. (2009). Analytical three-moment autoconversion parameterization based on generalized gamma distribution. *Journal of Geophysical Research*, **114**, D17201. doi:10.1029/2008JD011633.

Received October 6, 2013; accepted April 7, 2014.

Gauss M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco-UFPE
Cidade Universitária, 50740-540—Recife, PE, Brazil
gausscordeiro@uol.com.br

Rodrigo R. Pescim
Departamento de Ciências Exatas
Universidade de São Paulo-ESALQ - USP
Av. Pádua Dias 11 - Caixa Postal 9, 13418-900 Piracicaba - São Paulo - Brazil.
rrpescim@gmail.com

Clarice G.B. Demétrio
Departamento de Ciências Exatas
Universidade de São Paulo-ESALQ – USP
Av. Pádua Dias 11 - Caixa Postal 9, 13418-900 Piracicaba - São Paulo - Brazil.
clarice@esalq.usp.br

Edwin M.M. Ortega
Departamento de Ciências Exatas
Universidade de São Paulo-ESALQ – USP
Av. Pádua Dias 11 - Caixa Postal 9, 13418-900 Piracicaba - São Paulo - Brazil.
edwin@esalq.usp.br