Research Article

The odd log-logistic generalized gamma model: properties, applications, classical and bayesian approach

Abstract

We propose a new lifetime model called the odd log-logistic generalized gamma distri¬bution that can be easily interpreted. Some of its special models are discussed. We obtain general mathematical properties of this distribution including the ordinary moments, and quantile functions. We discuss parameter estimation by the maximum likelihood method and a Bayesian approach, where Gibbs algorithms along with metropolis steps are used to obtain the posterior summaries of interest for survival data with right censoring. Further, for different parameter settings, sample sizes and censoring percentages, we perform various simulations and evaluate the behavior of the estimators. The potentiality of the new distri¬bution is proved by means of two real data sets. In fact, the new distribution can produce better fits than some well-known distributions

Keywords: censored data, exponentiated distribution, generalized gamma distribution, moments, survival analysis

Fábio Prataviera,[1] Gauss M. Cordeiro,[2] Adriano K. Suzuki,[3] Edwin M. M. Ortega[4]

[1]Departamento de Ciências Exatas, Universidade de São Paulo, Brazil
[2]Departamento de Estat´ıstica, Universidade Federal de Pernambuco, Brazil
[3]Departamento de Matemática Aplicada e Estat´ıstica, Universidade de São Paulo, Brazil
[4]Departamento de Ciências Exatas, Universidade de São Paulo, Brazil

Correspondence: Edwin M. M. Ortega, Departamento de Ciências Exatas, Universidade de São Paulo, Piracicaba, SP, Brazil, Email edwin@usp.br

Introduction

The statistics literature is filled with hundreds of continuous univariate distributions. Recent developments focus on new techniques for building meaningful models. More recently, several methods of introducing one or more parameters to generate new distributions have been proposed. Among these methods, the compounding of some discrete and important lifetime distributions has been in the vanguard of lifetime modeling. So, several families of distributions were investigated by compounding some useful lifetime and truncated discrete distributions. The log-logistic (LL) distribution with a shape parameter $\lambda > 0$ is a useful model for survival analysis and it is an alternative to the log-normal distribution. Unlike the more commonly used Weibull distribution, the LL distribution has a non-monotonic hazard rate function (hrf), which makes it suitable for modeling cancer survival data. For $\lambda > 1$, the hrf is unimodal and when $\lambda = 1$, the hazard decreases monotonically. The fact that its cumulative distribution function (cdf) has a closed-form is particularly useful for analysis of survival data with censoring.

The odd log-logistic (OLL) family of distributions was pioneered by Gleaton and Lynch;[1] they called this family the generalized log-logistic (GLL) family. Recently, Braga et al.,[2] studied the odd log-logistic normal distribution, da Cruz et al.,[3] proposed the odd log-logistic Weibull distribution and Cordeiro et al.,[2] proposed the beta odd log-logistic generalized family. We develop a similar methodology to propose a new model based on the generalized gamma (GG) distribution. The GG distribution plays a very important role in statistical inferential problems. When modeling monotone hazard rates, the Weibull distribution may be an initial choice because of its negatively and positively skewed density shapes. However, the Weibull distribution does not provide a reasonable parametric fit for modeling phenomenon with bathtub shaped and unimodal failure rates, which are common in biological and reliability studies. Alternatively, other extensions of the GG distribution were developed for modeling

lifetime data. For example, Cordeiro et al.,[4] defined the exponentiated generalized gamma with applications, Pascoa et al.,[5] introduced the Kumaraswamy generalized gamma distribution, Ortega et al.,[6] proposed the generalized gamma geometric distribution, Cordeiro et al.,[7] studied the beta generalized gamma distribution and, more recently, Lucena et al.,[8] defines the transmuted generalized gamma distribution and Silva et al.,[9] proposed the generalized gamma power series class.

Given a continuous baseline cdf $G(t;\boldsymbol{\xi})$ with a parameter vector $\boldsymbol{\xi}$, the cdf of the *odd log-logistic*-G ("OLL-G" for short) distribution with an extra shape parameter $\lambda > 0$ is defined by

$$F(t) = \int_0^{\frac{G(t;\boldsymbol{\xi})}{\overline{G}(t;\boldsymbol{\xi})}} \frac{\lambda x^{\lambda-1}}{\left(1+x^\lambda\right)^2} dx = \frac{G(t;\boldsymbol{\xi})^\lambda}{G(t;\boldsymbol{\xi})^\lambda + \overline{G}(t;\boldsymbol{\xi})^\lambda}. \qquad (1)$$

We can write

$$\lambda = \frac{\log\left[\dfrac{F(t)}{\overline{F}(t)}\right]}{\log\left[\dfrac{G(t)}{\overline{G}(t)}\right]} \text{ and } \overline{G}(t;\xi) = 1 - G(t;\xi)$$

So, the parameter λ represents the quotient of the log odds ratio for the generated and baseline distributions. We note that there is no complicated function in equation (1) in contrast with the beta generalized family (Eugene et al.,[10]), which includes two extra parameters and also involves the beta incomplete function. The baseline cdf $G(t;\boldsymbol{\xi})$ is clearly a special case of (1) when $G(t;\boldsymbol{\xi})$. If $G(t;\boldsymbol{\xi}) = t/(1+t)$, it becomes the LL distribution. Several distributions can be generated from equation (1). For example, the odd log-logistic Fréchet and odd log-logistic gamma distributions are obtained by

taking $G(t;\boldsymbol{\xi})$ to be the Fréchet and gamma cumulative distributions, respectively. The probability density function (pdf) of the new family is given by

$$f(t) = \frac{\lambda g(t;\boldsymbol{\xi})\{G(t;\boldsymbol{\xi})[1-G(t;\boldsymbol{\xi})]\}^{\lambda-1}}{\{G(t;\boldsymbol{\xi})^{\lambda}+[1+G(t;\boldsymbol{\xi})^{\lambda}]\}^{2}} \qquad (2)$$

The OLL-G family of densities (2) allows for greater flexibility of its tails and can be widely applied in many areas of engineering and biology. We can study some of its mathematical properties because it extends several well-known distributions.

The inferential part of this model is carried out using the asymptotic distribution of the maximum likelihood estimators (MLEs), which in situations when the sample size is small or moderate, might lead to poor inference on the model parameters. Hence, in this paper, we also explore the Markov Chain Monte Carlo (MCMC) techniques to develop a Bayesian inference as an alternative analysis for the model. So, we discuss the inference aspects of the OLL-G model following both a classical and a Bayesian approach.

The rest of the paper is organized as follows. In Section 2, we define the odd log-logistic generalized gamma (OLLGG) distribution and present some special cases. Section 3 provides a useful linear representation for the OLLGG density function. We derive in Section 4 some structural properties of the new distribution. Considering censored data, we adopt a classic analysis for the parameters of the model in Section 5. In Section 6, the Bayesian approach is considered using MCMC with Metropolis-Hasting algorithms steps to obtain the posterior summaries of interest. In Section 7, we present results from various simulation studies displayed graphically and commented. Two applications to real data are performed in Section 8. Some concluding remarks are given in Section 9.

The OLLGG distribution

The gamma distribution is the most popular model for analyzing skewed data. The generalized gamma distribution (GG) was introduced by Stacy [11] and includes as special models: the exponential, Weibull, gamma and Rayleigh distributions, among others. It is suitable for modeling data with different forms of the hazard rate function (hrf): increasing, decreasing, bathtub and unimodal. This characteristic is useful for estimating individual hrfs and both relative hazards and relative times. The GG distribution has been used in several research areas such as engineering, hydrology and survival analysis.

The cdf and pdf of the $GG(\alpha,\tau,k)$ distribution (Stacy,[10]) are given by

$$G(t;\alpha,\tau,k) = \gamma_1\left(k,\left(\frac{t}{\alpha}\right)^{\tau}\right) = \frac{\gamma\left(k,(t/\alpha)^{\tau}\right)}{\Gamma(k)}, t>0 \qquad (3)$$

$$g(t;\alpha,\tau,k) = \frac{\tau}{\alpha\Gamma(k)}\left(\frac{t}{\alpha}\right)^{\tau k-1}\exp\left[-\left(\frac{t}{\alpha}\right)^{\tau}\right] \qquad (4)$$

where $\alpha>0, T>0, k>0, \gamma(k,x) = \int_0^x w^{k-1}e^{-w}dw$ is the incomplete gamma function and $\Gamma(.)$ is the gamma function. Basic properties of the GG distribution are given by Stacy and Mihram[12] and Lawless.[13] The OLLGG distribution (for $t>0$) is defined by substituting $G(t;\alpha,\tau,k)$ and $g(t;\alpha,\tau,k)$ in equations (1) and (2), respectively.

Hence, its density function with four positive parameters $\boldsymbol{\xi} = (\alpha,\tau,k)^T$ and $\lambda>0$ has the form

$$f(t) = \frac{\lambda\tau(t/\alpha)^{\tau\kappa-1}\exp\left[-(t/\alpha)^{\tau}\right]\left\{\gamma_1\left(k,(t/\alpha)^{\tau}\right)\left[1-\gamma_1\left(k,(t/\alpha)^{\tau}\right)\right]\right\}^{\lambda-1}}{\alpha\Gamma(\kappa)\left\{\gamma_1^{\lambda}\left(k,(t/\alpha)^{\tau}\right)+\left[1-\gamma_1\left(k,(t/\alpha)^{\tau}\right)\right]^{\lambda}\right\}^{2}}, t>0 \qquad (5)$$

where α is a scale parameter and the other positive parameters τ, k and $\ddot{e}$ are shape parameters. One major benefit of (5) is its ability of fitting skewed data that can not be properly fitted by existing distributions. The OLLGG density allows for greater flexibility of its tails and can be widely applied in many areas of engineering and biology.

The Weibull and GG distributions are the most important sub-models of (5) for $\lambda=k=1$ and $\lambda=1$, respectively. The OLLGG distribution approaches the log-normal (LN) distribution when $\lambda=1$ and $k\to\infty$. Other sub-models are listed in Table 2: OLL-Gamma, OLL-Chi-Square, OLL-Exponential, OLL-Weibull, OLL-Rayleigh, OLL-Maxwell, OLL-Folded normal, among others.

Table 1 Some new OLL-G sub-models

Distribution	α	τ	k	λ
OLL-gamma	α	1	k	λ
OLL-weibull	α	τ	1	λ
OLL-exponential	α	1	1	λ
OLL-chi-square	2	1	$\frac{n}{2}$	λ

Table continued

OLL-chi	$\sqrt{2}$	2	$\dfrac{n}{2}$	λ
OLL-rayleigh	α	2	1	λ
OLL-maxwell	α	2	$\dfrac{3}{2}$	λ
OLL-folded normal	$\sqrt{2}$	2	$\dfrac{1}{2}$	λ
OLL-circular normal	$\sqrt{2}$	2	1	λ
OLL-spherical normal	$\sqrt{2}$	2	$\dfrac{3}{2}$	λ

If T is a random variable with density function (5), we write $T \sim \text{OLLGG}(\alpha, \tau, k, \lambda)$. The survival and hazard rate functions corresponding to (5) a

$$S(t) = 1 - F(t) = \frac{\left[1 - \gamma_1\left(k, \left(t/\alpha\right)^\tau\right)\right]^\lambda}{\gamma_1^{\breve{}}\left(k, \left(t/\alpha\right)^\tau\right) + \left[1 - \gamma_1\left(k, \left(t/\alpha\right)^\tau\right)\right]^\lambda} \qquad (6)$$

$$h(t) = \frac{\lambda\tau\left(t/\alpha\right)^{\tau k - 1}\exp\left[-\left(t/\alpha\right)^\tau\right]\gamma_1^{\lambda-1}\left(k, \left(t/\alpha\right)^\tau\right)\left\{\gamma_1^\lambda\left(k, \left(t/\alpha\right)^\tau\right) + \left[1 - \gamma_1\left(k, \left(t/\alpha\right)^\tau\right)\right]^\lambda\right\}}{\alpha\Gamma(k)\left\{\gamma_1^\lambda\left(k, \left(t/\alpha\right)^\tau\right) + \left[1 - \gamma_1\left(k, \left(t/\alpha\right)^\tau\right)\right]^\lambda\right\}^2\left[1 - \gamma_1\left(k, \left(t/\alpha\right)^\tau\right)\right]} \qquad (7)$$

respectively. Plots of the OLLGG density function for selected parameter values are given in Figure 1. We note that the OLLGG density function can be symmetrical, left-skewed, right-skewed, unimodal and bimodal shaped.

The hrf (7) is quite flexible for modeling survival data as indicated by the plots for selected parameter values in Figure 2. The hrf can be increasing, decreasing, unimodal, bathtub and have other forms.

The OLLGG model is easily simulated by inverting (1) as follows:

$$t = Q_{GG}\left(\frac{\upsilon^{1/\lambda}}{\left(1-\upsilon\right)^{1/\lambda} + \upsilon^{1/\lambda}}, \alpha, \tau, k\right), \qquad (8)$$

where u has a uniform $U(0,1)$ distribution and $Q_{GG}(.) = G^{-1}(.)$ is the baseline quantile function

(qf).

Some properties of the OLLGG distribution are:

If $T \sim OLLGG(\alpha, \tau, k, \lambda) \Rightarrow bTOLLGG(b\alpha, \tau, k, \lambda), \forall b > 0.$

If

$T \sim OLLGG(\alpha, \tau, k, \lambda) \Rightarrow T^m OLLGG(\alpha^m, \tau/m, k, \lambda), \forall m \neq 0.$

So, the new distribution is closed under power transformation.

Linear representation for the OLLGG distribution

First, we define the exponentiated-generalized gamma ("Exp-GG") distribution, say $W \sim Exp^c(GG)$ with power parameter $c > 0$, if W has cdf and pdf given by

$$H_c(t) = G(t; \alpha, \tau, k)^c \text{ and}$$

$$h_c(t) = \frac{c\tau}{\alpha\Gamma(k)}\left(\frac{t}{\alpha}\right)^{\tau k - 1}G(x; \alpha, \tau, k)^{c-1},$$

respectively. In a general context, the properties of the exponentiated-G (Exp-G) distributions have been studied by several authors for some baseline G models, see Mudholkar and Srivastava[14] and Mudholkar et al.,[15] for exponentiated Weibull, Nadarajah[16] for exponentiated Gumbel, Shirke and Kakade[17] for exponentiated log-normal and Nadarajah and Gupta[18] for exponentiated gamma distributions. See, also, Nadarajah and Kotz,[19] among others.

First, we obtain an expansion for $F(t; \alpha, \tau, k, \lambda)$ using a power series for $G(t; \alpha, \tau, k)^\lambda$ ($\lambda > 0$ real)

$$G\left(t;\alpha,\tau,\mathrm{k}\right)^{\lambda} = \sum_{j=0}^{\infty} a_j G\left(t;\alpha,\tau,\mathrm{k}\right)^{j}, \qquad (9)$$

where

$$a_j = a_j\left(\lambda\right) \sum_{\upsilon=j}^{\infty} \left(-1\right)^{j+\upsilon} \binom{\lambda}{\upsilon}\binom{\upsilon}{j}.$$

For any real $\lambda > 0$, we consider the generalized binomial expansion

$$\left[1 - G\left(t;\alpha,\tau,k\right)\right]^{\lambda} = \sum_{j=0}^{\infty} \left(-1\right)^{j} \binom{\lambda}{j} G\left(t;\alpha,\tau,k\right)^{j}. \qquad (10)$$

Inserting (9) and (10) in equation (1), we obtain

$$F\left(t;\alpha,\tau,k,\lambda\right) = \frac{\sum_{j=0}^{\infty} a_j G\left(t;\alpha,\tau,k\right)^{j}}{\sum_{j=0}^{\infty} b_j G\left(t;\alpha,\tau,k\right)^{j}},$$

where $b_j = a_j + \left(-1\right)^{j} \binom{\lambda}{j}$ for $j \geq 0$.

The ratio of the two power series can be expressed as

$$F\left(t;\alpha,\tau,k,\lambda\right) = \sum_{j=0}^{\infty} c_j G\left(t;\alpha,\tau,k\right)^{j}, \qquad (11)$$

where $c_0 = a_0/b_0$ and the coefficients c_j's (for $j \geq 1$) are determined from the recurrence equation

$$c_j = b_0^{-1} \left(a_j - \sum_{r=1}^{j} b_r c_{j-r}\right).$$

The pdf of T is obtaining by differentiating (11) as

$$f\left(t;\alpha,\tau,k,\lambda\right) = \sum_{j=0}^{\infty} c_{j+1} h_{j+1}\left(t\right), \qquad (12)$$

where

$$h_{j+1}\left(t\right) = \frac{\left(j+1\right)\tau}{\alpha\Gamma\left(k\right)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[\left(\frac{t}{\alpha}\right)^{\tau}\right] G\left(t;\alpha,\tau,k\right)^{j}$$

is the Exp-GG density function with power parameter $j+1$.

For $j \geq 0$, we can write

$$h_{j+1}\left(t\right) = \frac{\left(j+1\right)\tau}{\alpha\Gamma\left(k\right)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\tau}\right] \gamma_1\left(k,(t/\alpha)^{\tau}\right)^{j},$$

$$\qquad (13)$$

where $\gamma_1\left(k,(t/\alpha)^{\tau}\right) = \gamma\left(k,\left(t/\alpha\right)^{\tau}\right)/\Gamma\left(k\right)$.

By application of an equation in Section 0.314 of Gradshteyn and Ryzhik[20] for a power series raised to a power, we obtain for any j positive integer

$$\left(\sum_{i=0}^{\infty} a_i x^{i}\right)^{j} = \sum_{i=0}^{\infty} d_{j,i} x^{i}, \qquad (14)$$

where the coefficients $d_{j,i}$ (for $i=1,2,.....$) satisfy the recurrence relation

$$d_{j,i} = \left(ia_0\right)^{-1} \sum_{p=1}^{i} \left[j\left(p+1\right)-i\right] a_p d_{j,i-p} \qquad (15)$$

and $d_{j,0} = a_0^{j}$. The coefficient $d_{j,i}$ can be expressed explicitly from $d_{j,0,...,}d_{j,i-1}$ and then from $a_{0,...,}a_{i,}$, although it is not necessary for programming numerically our expansions using any software with numerical facilities.

Further, using equation (14), we can write (for $j \geq 1$)

$$\gamma_1\left(k,(t/\alpha)^{\tau}\right)^{j} = \frac{(t/\alpha)^{jk\tau}}{\Gamma\left(k\right)^{j}} \sum_{i=0}^{\infty} d_{j,i}(t/\alpha)^{i\tau}, \qquad (16)$$

where the coefficients $d_{j,i}$ (for $i\geq1$) are determined from (15) with $a_p = \left(-1\right)^{p}/\left[\left(k+p\right)p!\right]$. Based upon equation (16), we can write the Exp-GG density (for $j \geq 1$) from (13) as

$$h_{j+1}\left(t\right) = \frac{\left(j+1\right)\tau}{\alpha\Gamma\left(k\right)^{j+1}} \exp\left[-\left(\frac{t}{\alpha}\right)^{\tau}\right] \sum_{i=0}^{\infty} \left(\frac{t}{\alpha}\right)^{\left[i+(j+1)k\right]\tau-1}.$$

The last density can be expressed in terms of the GG density functions. By noting the form of (4), we can write (for $j \geq 1$)

$$h_{j+1}\left(t\right) = \sum_{i=0}^{\infty} e_{j,i} g\left(t;\alpha,\tau,\left[i+\left(j+1\right)k\right]\right), \qquad (17)$$

where $g\left(t;\alpha,\tau,\left[i+\left(j+1\right)k\right]\right)$ is the GG density function with parameters α,τ and $i+\left(j+1\right)k$ and

$$e_{j,i} = \frac{\left(j+1\right)\Gamma\left(\left[i+\left(j+1\right)k\right]\right)}{\Gamma\left(k\right)^{j+1}} d_{j,i}. \qquad (18)$$

For $j = 0$, we have from (13)

$$h_1\left(t\right) = \frac{\tau}{\alpha\Gamma\left(k\right)} \left(t/\alpha\right)^{\tau k-1} \exp\left[\left(\frac{t}{\alpha}\right)^{\tau}\right] = g\left(t;\alpha,\tau,k\right).$$

Combining the result (17) (for $j \geq 1$) and that one for $j = 0$, we can write $f\left(t\right) = f\left(t;\alpha,\tau,k,\right)$ in (12) as

$$f\left(t\right) = c_1 g\left(t;\alpha,\tau,k\right) + \sum_{j=1}^{\infty} \sum_{i=0}^{\infty} e_{j,i} c_{j+1} g\left(t;\alpha,\tau,\left[i+(j+1)k\right]\right).$$

$$\qquad (19)$$

Equation (19) reveals that the OLLGG density function is a linear combination of Exp-GG densities. Hence, some mathematical properties of the OLLGG distribution can follow directly from those properties of the GG distribution. For example, the ordinary, central, fac¬torial moments and the moment generating function (mgf) of the proposed distribution can be obtained from the same weighted infinite linear combination of the corresponding quantities for the GG distribution. This equation is the main result of this section.
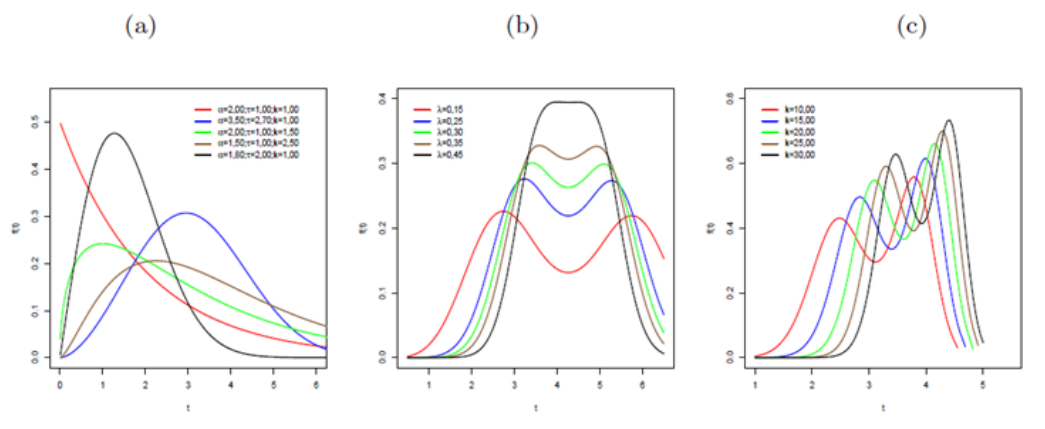
Figure 1 Plots of the OLLGG density function for some parameter values. (a) Fixed $ë = 1$. (b) Fixed $\alpha = 2$, $\tau = 3$ and $k = 10$. (c) Fixed $\alpha = 2$, $ë = 0,15$ and $ë = 0,15$.
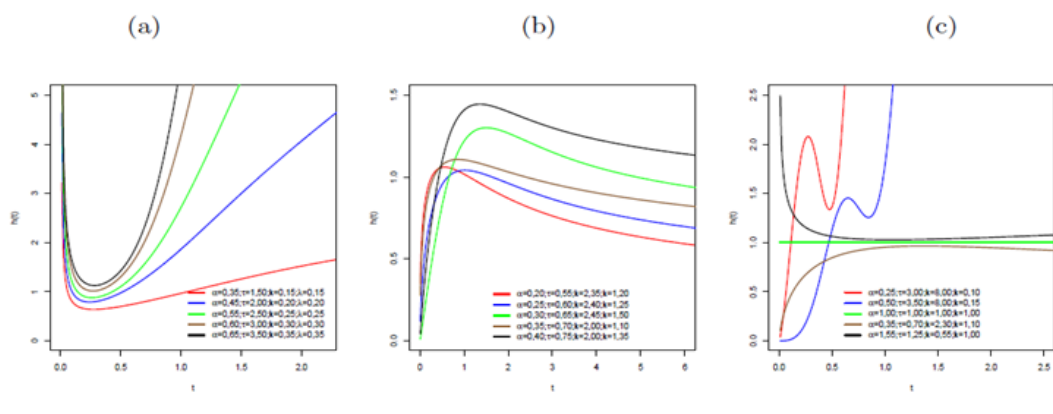
Figure 2 The OLLGG hrf. (a) Bathtub. (b) Unimodal. (c) Increasing, decreasing and other forms.

Mathematical properties

Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness and kurtosis). In this section, we give two different expansions for calculating the moments of the EGG distribution.

First, we obtain an infinite sum representation for the r th ordinary moment μ_r' of the EGG distribution based on the equation (19). The r th moment of the $GG(\alpha, \tau, k)$ distribution is well known to be

$$\mu_{r,GG}' = \frac{\alpha^r \Gamma(k+r/\tau)}{\Gamma(k)}$$

Equation (19) then immediately gives

$$\mu_r' = \frac{c_1 \alpha^r \Gamma(k+r/\tau)}{\Gamma(k)} + \frac{\alpha^r}{\Gamma(k)} \sum_{j=1}^{\infty} \sum_{i=0}^{\infty} e_{j,i} c_{j+1} \Gamma\left(\left[i+(j+1)k\right]+r/\tau\right).$$

$$(20)$$

Equation (20) reveals that the moment μ_r' does have the inconvenient of depending on the quantities $e_{j,i}$ given by (18).

We now derive another infinite sum representation for μ_r' by computing the r th moment directly without requiring the quantities $e_{j,i}$. We readily obtain

$$\mu_r' = \frac{\beta \alpha^{r-1}}{\Gamma(k)} \int_0^{\infty} \left(\frac{t}{\alpha}\right)^{\beta k + r - 1} \exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\} \left\{\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^{\beta}\right]\right\}^{\lambda-1} dt$$

and then $x = (t/\alpha)^{\beta}$ gives

$$\mu_r' = \frac{ë\alpha^r}{\Gamma(k)^{\lambda}} \int_0^{\infty} x^{k+r/\beta-1} e^{-x} \gamma(k,x)^{\lambda-1} dx.$$

Using expansion (16) for $\gamma(k,x)$ leads to

$$\gamma(k,x)^{\lambda-1} = \sum_{j=0}^{\infty} \sum_{m=0}^{j} (-1)^{j+m} \binom{\lambda-1}{j} \binom{j}{m} \gamma(k,x)^m.$$

Inserting the last equation in the expression for μ_r' and interchanging terms, we obtain

$$\mu_r' = \frac{\lambda \alpha^r}{\Gamma(k)^r} \sum_{j=0}^{\infty} \sum_{m=0}^{j} (-1)^{j+m} \binom{\lambda-1}{j} \binom{j}{m} I(k, r/\beta, m), \qquad (21)$$

where

$$I\left(k, \text{r}/\beta, m\right) = \int_0^\infty x^{k+r/\beta-1} e^{-x} \gamma\left(k, x\right)^m dx .$$

For calculating the last integral, the series expansion (16) for the incomplete gamma function gives

$$I\left(k, \text{r}/\beta, m\right) = \int_0^\infty x^{k+r/\beta-1} e^{-x} \left\{ x^k \sum_{p=0}^\infty \frac{(-x)^p}{(k+p)p!} \right\}^m dx.$$

Now this integral can be obtained from equations (24) and (25) of Nadarajah[21] in terms of the Lauricella function of type A (Exton;[22] Aarts,[23]) defined by

$$F_A^{(n)}\left(a; b_1, \ldots, b_n; c_1, \ldots, c_n; x_1, \ldots, x_n\right) =$$

$$\sum_{m_1=0}^\infty \cdots \sum_{m_n=0}^\infty \frac{(a)_{m_1+\ldots+m_n} (b_1)_{m_1} \ldots (b_n)_{m_n}}{(c_1)_{m_1} \ldots (c_n)_{m_n}} \frac{x_1^{m_1} \ldots x_n^{m_n}}{m_1! \ldots m_n!},$$

where $(a)_i$ is the ascending factorial defined by (with the convention that $(a)_0 = 1$)
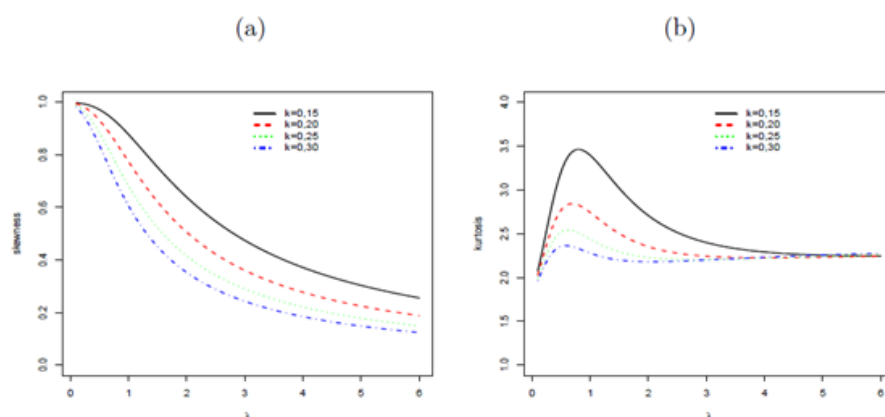
$$(a)_i = a\left(a+1\right) \ldots \left(a+i-1\right).$$

Numerical routines for the direct computation of the Lauricella function of type A are available, see Exton[22] and Mathematica (Trott,[24]). We obtain

$$I\left(k, r/\beta, m\right) = k^{-m} \Gamma\left(r/\beta+k\left(m+1\right)\right) F_A^{(m)}\left(r/\beta+k\left(m+1\right); k, \ldots k; k+1, \ldots, k+1; -1, \ldots -1\right).$$

(22) Hence, as an alternative way to equation (20), the rth moment of the EGG distribution follows from both formulae (21) and (22) as an infinite weighted sum of the Lauricella functions of type A. In Figures 3 and 3, we display plots of the skewness and kurtosis the OLGG distribution for some parameter values.

Maximum likelihood estimation

Let Ti be a random variable following (5) with the vector of parameters $\boldsymbol{\theta} = \left(T, k, \lambda\right)^T$. The data encountered in survival analysis and reliability studies are often censored. A very simple random censoring mechanism that is often realistic is one in which each individual i is assumed to have a lifetime T_i and a censoring time C_i, where T_i and C_i are independent random variables. Suppose that the data consist of n independent observations $t_i = \min\left(T_i, C_i\right)$ for $i = 1, \ldots, n$.

The distribution of C_i does not depend on any of the unknown parameters of T_i. Parametric inference for such data are usually based on likelihood methods and their asymptotic theory. The censored log-likelihood $l\left(\boldsymbol{\theta}\right)$ for the model parameters is given by

$$l\left(\boldsymbol{\theta}\right) = r \log\left[\frac{\lambda_T}{\alpha \Gamma\left(k\right)}\right] - \sum_{i \in F} \left(\frac{t_i}{\alpha}\right)^t + \left(Tk-1\right) \sum_{i \in F} \log\left(\frac{t_i}{\alpha}\right) + \left(\lambda-1\right) \sum_{i \in F} \log\left(u_i\right) +$$

$$\left(\lambda-1\right) \sum_{i \in F} \log\left(1-u_i\right) - 1 \sum_{i \in F} \log\left[u_i^\lambda + \left(1-u_i\right)^\lambda\right] + \lambda \sum_{i \in c} \log\left(1-u_i\right) - \sum_{i \in c} \log\left[u_i^\lambda + \left(1-u_i\right)^\lambda\right], \quad (23)$$

Figure 3 Skewness and kurtosis of the OLLGG distribution as a function of λ for some values of k with $\alpha = 2$ and $\tau = 1$.
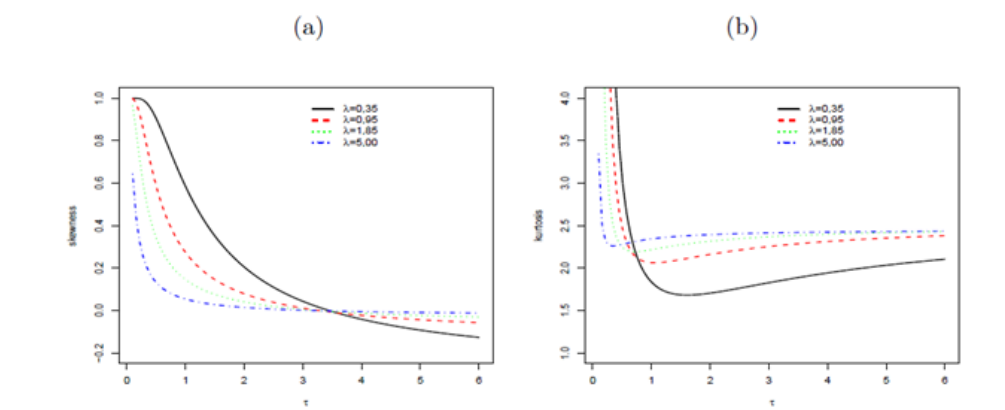
Figure 4 Skewness and kurtosis of the OLLGG distribution as a function of τ for some values of λ with $\alpha = 2$ and $k = 1$.

Where $\mu_i = \gamma_1\left(k, \left(\frac{t_i}{\alpha}\right)^t\right)$, r is the number of failures and F and C denote the uncensored and censored sets of observations, respectively.

The score components corresponding to the parameters in θ are:

$$U_\alpha(\boldsymbol{\theta}) = -\frac{r\tau k}{\alpha} + \tau\alpha^{\tau-1}\sum_{i\in F} t_i^\tau + (\lambda-1)\sum_{i\in F}\frac{[\dot{u}_i]_\alpha}{\mu_i} - (\lambda-1)\sum_{i\in F}\frac{[\dot{u}_i]_\alpha}{1-\mu_i}$$

$$-2\lambda\sum_{i\in F}\frac{[\dot{u}_i]_\alpha\left[u_i^{\lambda-1}\right]-(1-u_i)^{\lambda-1}}{\left[u_i^\lambda+(1-u_i)^\lambda\right]} - \lambda\sum_{i\in C}\frac{[\dot{u}_i]_\alpha}{(1-u_i)} - \lambda\sum_{i\in C}\frac{[\dot{u}_i]_\alpha\left[u_i^{\lambda-1}\right]-(1-u_i)^{\lambda-1}}{\left[u_i^\lambda+(1-u_i)^\lambda\right]},$$

$$U_\tau(\boldsymbol{\theta}) = -\frac{r}{\tau} - \sum_{i\in F}\left(\frac{t_i}{\alpha}\right)^\tau\log\left(\frac{t_i}{\alpha}\right) + k\sum_{i\in F}\left(\frac{t_i}{\alpha}\right)\log + (\lambda-1)\sum_{i\in F}\frac{[\dot{u}_i]\tau}{\mu_i} - (\lambda-1)\sum_{i\in F}\frac{[\dot{u}_i]\tau}{1-\mu_i}$$

$$-2\sum_{i\in F}\frac{[\dot{u}_i]\tau\left[u_i^{\lambda-1}-(1-u_i)^{\lambda-1}\right]}{\left[u_i^\lambda+(1-u_i)^\lambda\right]} - \lambda\sum_{i\in C}\frac{[\dot{u}_i]\tau}{(1-\mu_i)} - \lambda\sum_{i\in C}\frac{[\dot{u}_i]\tau\left[u_i^{\lambda-1}-(1-u_i)^{\lambda-1}\right]}{\left[u_i^\lambda+(1-u_i)^\lambda\right]},$$

$$U_k(\boldsymbol{\theta}) = -r\psi(k) + \tau\sum_{i\in F}\log\left(\frac{t_i}{\alpha}\right) + (\lambda-1)\sum_{i\in F}\frac{[\dot{u}_i]k}{u_i} - (\lambda-1)\sum_{i\in F}\frac{[\dot{u}_i]k}{1-u_i}$$

$$-2\lambda\sum_{i\in F}\frac{[\dot{u}_i]k\left[u_i^{\lambda-1}-1(1-u_i)^{\lambda-1}\right]}{\left[u_i^\lambda+(1-u_i)^\lambda\right]} - \lambda\sum_{i\in C}\frac{[\dot{u}_i]k}{(1-u_i)} - \lambda\sum_{i\in C}\frac{[\dot{u}_i]k\left[u_i^{\lambda-1}-(1-u_i)^{\lambda-1}\right]}{\left[u_i^\lambda+(1-u_i)^\lambda\right]},$$

$$U_\lambda(\boldsymbol{\theta}) = \frac{r}{\lambda}\sum_{i\in F}\log(u_i) + \sum_{i\in F}\log(1-u_i) - 2\sum_{i\in F}\frac{u_i^\lambda\log(u_i)+(1-u_i)^\lambda\log(1-u_i)}{\left[u_i^\lambda+(1-u_i)^\lambda\right]}$$

$$+ \sum_{i\in F}\log(1-u_i) - \sum_{i\in F}\frac{\mu_i^\lambda\log(u_i)+(1-u_i)^\lambda\log(1-u_i)}{\left[\mu_i^\lambda+(1-u_i)^\lambda\right]},$$

Where

$$[\dot{u}_i]\alpha + \frac{\partial_{\gamma 1}\left(k,\left(\frac{t_i}{\alpha}\right)^\tau\right)}{\partial\alpha} [\dot{u}_i]\tau\frac{\partial_{\gamma 1}\left(k,\left(\frac{t_i}{\alpha}\right)^\tau\right)}{\partial\alpha} [\dot{u}_i]k\frac{\partial_{\gamma 1}\left(k,\left(\frac{t_i}{\alpha}\right)^\tau\right)}{\partial k}$$

$\psi(.)$ is the digamma function and $i = 1,...,n$.

The *MLE* $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ can be obtained numerically from the nonlinear equations $U_\tau(\boldsymbol{\theta}) = U_k(\boldsymbol{\theta}) = U_\lambda(\boldsymbol{\theta}) = 0$. For interval estimation and hypothesis tests on the model parameters, we require the 4×4 unit observed information matrix $J = J(\boldsymbol{\theta})$, whose elements are evaluated numerically. Under general regularity conditions, the asymptotic distribution of $(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is $N_4\left(0,(\boldsymbol{\theta})^{-1}\right)$, where $I(\boldsymbol{\theta})$ is the expected information matrix. This matrix can be replaced by $J(\hat{\boldsymbol{\theta}})$, i.e., the observed information matrix evaluated at $\hat{\boldsymbol{\theta}}$. The multivariate normal $N_4\left(0,(\hat{\boldsymbol{\theta}})^{-1}\right)$ distribution can be used to construct approximate confidence intervals for the individual parameters. Further, the likelihood ratio (LR) statistic can be adopted for comparing this distribution with some of its special models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to construct LR statistics for testing some sub-models of the OLLGG distribution. For example, the test of $H_0 : \lambda = 1$ versus $H : H_0$ is *not true* is equivalent to compare the OLLGG and GG distributions and the LR statistic reduces to

$$w = 2\left\{l\left(\hat{\alpha},\hat{\tau},\hat{k},\hat{\lambda}\right)-l\left(\hat{\alpha},\hat{\tau},\hat{k},1\right)\right\},$$

where $\hat{\alpha},\hat{\tau},\hat{k}$, and $\hat{\lambda}$ are the MLEs under H and $\hat{\alpha},\hat{\tau}$ and $\hat{k}$ are the estimates under H_0.

Bayesian inference

In this section we briefly discuss the inference from a Bayesian viewpoint. We making a change in the parameters to $\boldsymbol{\xi} = \left(\log(\lambda),\ \log(k),\ \log(\tau),\log(\alpha)\right)$, so that the parameter space

is transformed into R^4 (necessary for the work with the proposed Gaussian densities). We assume that λ, k, τ and α are prior independent, that is,

$$\log(\lambda) \sim N(0, \sigma^2),\ \log(k) \sim N(0, \sigma^2),\ \log(\tau) \sim N(0, \sigma^2)\ \text{and}\ \log(\alpha) \sim N(0, \sigma^2)\ \text{and}\ N(\mu, \sigma^2)$$ denotes the normal distribution with

mean μ and variance σ^2. All the hyper-parameters σ^2 have been specified to express non-informative priors.

Regarding the Jacobian of this transformation, our joint posterior density (or target density) reduces to

$$\pi(\xi \mid D) \propto L(\xi; D) \exp\left\{-\frac{1}{2}\left[\frac{\log(\lambda)}{\sigma^2} + \frac{\log(k)}{\sigma^2} + \frac{\log(\tau)}{\sigma^2} + \frac{\log(\alpha)}{\sigma^2}\right]\right\} \tag{25}$$

where $L(\hat{\imath}; D)$ is the likelihood function.

This joint posterior density is analytically intractable. Therefore, we based our inference on the MCMC simulation methods. No closed-form is available for any of the full conditional distributions necessary for the implementation of the Gibbs sampler. Then, we have resorted to the Metropolis–Hastings algorithm. To implement this algorithm, we proceed as follows:

(1) Start with any point $\xi_{(10)}$ and stage indicator $j = 0$;

(2) Generate a point ξ'' according to the transitional kernel $Q(\xi', \xi_j) = N_4(\xi_j, \tilde{\Sigma})$, where $\tilde{\Sigma}$ is the covariance matrix of $\hat{\imath}$, which is the same in any stage;

(3) Update $\xi_{(j)}$ to $\xi_{(j+1)} = \xi''$ with probability $p_j = \min\left\{1, \pi(\xi' \mid D)/\pi(\xi_{(j)} \mid D)\right\}$, or keep $\theta_{(j)}$ with probability $1 - p_{(j)}$;

(4) Repeat steps (2) and (3) by increasing the stage indicator until the process has reached a stationary distribution.

In this scheme, we consider 30,000 sample burn-in, and we use every tenth sample from the 200,000 MCMC posterior samples to reduce the autocorrelations and yield better convergence results, thus obtaining an effective sample of size 20,000 from which the posterior is based on. We monitor the convergence of the Metropolis-Hasting algorithm using the method proposed by Geweke (1992), as well as trace plots. All computations are performed in the *R* software (*R* Development Core Team, 2011).

Bayesian model comparison

In the literature, a variety of Bayesian methodologies can be applied for comparing of several competing models for a given data set and selection of the best one to fit the data. In this paper, we use the deviance information criterion (DIC) proposed by Spiegelhalter et al.,[25] the expected Akaike information criterion (EAIC)given by Brooks,[26] and the expected Bayesian (or Schwarz) information criterion (EBIC) discussed by Carlin and Louis.[27]

$$\pi(\theta) = \pi(\lambda)\pi(k)\pi(\tau)\pi(\alpha),$$

where

They are based on the posterior mean of the deviance, which can be approximated by

$$\bar{d} = \sum_{q=1}^{Q} d(\theta_q)/Q, where\ d(\theta) = -2\sum_{i=1}^{n}\log\left[f(t_i \mid \theta)\right].$$ The

DIC criterion can be estimated using the MCMC output by $\widehat{DIC} = \bar{d} + \widehat{\rho d} = 2\bar{d} - \hat{d}$, where ρD is the effective number of parameters given by $E\{d(\theta)\} - d\{E(\theta)\}, and\ d\{E(\theta)\}$ is the deviance evaluated at the posterior mean. Similarly, the EAIC and EBIC criteria can be estimated by means of $\widehat{EAIC} = \bar{d} + 2\,\#(\theta)$ and $\widehat{EAIC} = \bar{d} + \#(\theta)\log(n), where\ \#(\theta)$ is the number of the model parameters.

Simulation study

We evaluate some properties of the MLEs using the classical and Bayesian analysis by means of a simulation study. We simulate the OLLGG distribution considering modality form from equation (8) by using a random variable U having a uniform distribution in (0, 1).

We take n=50, 150 and 350 and, for each replication, we calculate the MLEs $\hat{\alpha}, \hat{\tau}, \hat{k}\ and\ \hat{\lambda}$. We repeat this process 1, 000 times and determine the average estimates (AEs), biases and means squared errors (MSEs). In this study, we consider two scenarios. In the first scenario, we take $\alpha = 2, \tau = 5, k = 10, \lambda = 0.5$. In the second scenario, we use the values fitted in the adjustment to the temperature data set in Section 8 $(\alpha = 21.2911, \tau = 13.0661, k = 2.8755, \lambda = 0.2882)$. The estimates of $\alpha, \tau, k, and\ \lambda$ are determined by solving the nonlinear equations $U_\alpha(\theta) = 0, U_T(\theta) = 0, U_k(\theta) = 0, U_\lambda(\theta) = 0$. The results of the Monte Carlo study under maximum likelihood and Bayesian estimation are given in Tables 2 and 3, respectively. They indicate that the MSEs of the MLEs of $\alpha, \tau, k, and\ \lambda$ decay toward zero as the sample size increases, as expected under first-order asymptotic theory. The same results are obtained using the Bayesian approach. In Figures 5 and 6, we present the estimated densities based on 1,000 samples of the AEs of the parameters $\alpha, \tau, k, and\ \lambda$, respectively and n = 50, 150 and 350 for both scenarios. These plots are in agreement with the first-order asymptotic theory for the MLEs and reveal a fast convergence even for small sample sizes.

Simulation study of random censored values

Similarly, we also consider a simulation study in the presence of censored data. The censoring times C_i are sampled from the uniform distribution in the interval $(0, v), where\ v$ denotes the proportion of censored observations. In this study, the proportions of censored observations are approximately equal to 10% and 30%. In this scenario, we take the values of the parameters as $\alpha = 2, \tau = 5, k = 10, \lambda = 0.15$. Table 4 lists the averages of the MLEs (Mean) and the MSEs.

The figures in this table indicate that the MSEs increase when the censoring percentage increases. Further, the MSEs of the MLEs of $\alpha, \tau, k,$ and λ decay toward zero as the sample size increases, as expected under first-order asymptotic theory.

Table 5 lists the posterior means (Mean) and the MSEs. We can note that increasing the sample size and decreasing the percentage of censure, the estimates are closer to the true values with lower MSEs.

Table 2 AEs, biases and MSEs for the estimates of the OLLGG parameters

Scenario 1				
n	Parameters	AEs	Biases	MSEs
50	α	2.0404	-0.0404	0.1984
	τ	5.3257	-0.3257	1.8523
	k	10.7653	-0.7653	2.9000
	λ	0.1708	-0.0208	0.0115
150	α	2.0393	-0.0393	0.0242
	τ	5.1585	-0.1585	0.2070
	k	9.8491	0.1509	1.9955
	λ	0.1528	-0.0028	0.0011
350	α	2.0065	-0.0065	0.0024
	τ	5.0417	-0.0417	0.0276
	k	10.012	-0.0012	0.2220
	λ	0.1511	-0.0011	0.0001
Scenario 2				
n	Parameters	AEs	Biases	MSEs
50	α	21.1422	0.1489	7.7557
	τ	15.5491	-2.483	64.7128
	k	4.5288	-1.6533	22.2571
	λ	0.3400	-0.0518	0.0685
150	α	21.3407	-0.0496	2.1903
	τ	13.8973	-0.8312	9.9415
	k	3.2666	-0.3911	3.3779
	λ	0.3060	-0.0178	0.0167
350	α	21.2908	0.0003	0.8393
	τ	13.3138	-0.2477	3.0814
	k	3.0593	-0.1838	1.2018
	λ	0.2956	-0.0074	0.0058

In Figures 7 and 8, we present the estimated densities based on 1,000 samples of the AEs of the parameters $\alpha, \tau, k,$ and λ respectively, and n = 50, 150 and 350 for both scenarios with 10% and 30% of censored. These plots are in agreement with the first-order asymptotic theory for the MLEs and indicate a fast convergence even for small sample sizes and considering censored data.

Applications

In this section, we provide two applications to real data to prove empirically the flexibility of the OLLG model. The computations are performed using the R software and NLMixed procedure in SAS. In the first application, we give an application for bimodal data comparing the OLLGG, GG and Weibull models. In the second application, we prove the usefulness of the new distribution for censored data.

Temperature data

The first data set refers to daily temperatures $\left(^{0}C\right)$ $(n=365)$ in the period from January 1 to December 31, 2011 in the city of Piracicaba obtained from the Department of Biosystems Engi-neering of the Luiz de Queiroz Superior School of Agriculture (ESALQ), part of the University of São Paulo (USP).

We show the superiority of the OLLGG distribution as compared to some of its sub-mo¬dels and also to the following non-nested models: the exponentiated generalized gamma (EGG) proposed by Cordeiro et al.,[28] and beta Weibull (BW) distributions. The BW cdf (Famoye et al.,[29]) is given by

Table 3 Posterior means, biases and MSEs for the estimates of the OLLGG parameters

Scenario 1				
n	**Parameters**	**Means**	**Biases**	**MSEs**
50	α	1.8130	0.1870	0.0703
	τ	4.1719	0.8281	1.1152
	k	9.9011	0.0989	0.0601
	λ	0.2795	-0.1295	0.0319
150	α	1.8891	0.1109	0.0240
	τ	4.4648	0.5352	0.4132
	k	9.9893	0.0107	0.0824
	λ	0.2005	-0.0505	0.0031
350	α	1.9283	0.0717	0.0128
	τ	4.6425	0.3575	0.2232
	k	9.9929	0.0071	0.0913
	λ	0.1812	-0.0312	0.0014
Scenario 2				
n	**Parameters**	**Means**	**Biases**	**MSEs**
50	α	19.4002	1.8909	6.0127
	τ	10.6098	2.4563	15.3457
	k	5.2667	-2.3912	6.8778
	λ	0.4200	-0.1318	0.0536
150	α	20.4151	0.8760	1.5490
	τ	11.5327	1.5334	5.6849
	k	4.1478	-1.2723	2.3679
	λ	0.3344	-0.0462	0.0070
350	α	21.3516	-0.0605	0.1011
	τ	13.2395	-0.1734	0.2929
	k	3.0900	-0.2145	0.2465
	λ	0.3040	-0.0158	0.0020

$$F(t) = \frac{1}{B(a,b)} \int_0^{\left\{1-\exp\left[-\left(t/\alpha\right)^\gamma\right]\right\}} w^{a-1} (1-w)^{b-1} \, dw.$$

The Kumaraswamy generalized gamma (KumGG) distribution (for t > 0) is defined by Pascoa et al.,[5] Its density function with five positive parameters $\alpha, \tau, k, \lambda \text{ and } \varphi$ is given by

$$f(t) = \frac{\lambda \varphi \tau}{\alpha \Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k - 1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left\{\gamma_1\left[k\left(\frac{t}{\alpha}\right)^\tau\right]\right\}^{\lambda-1} \left(1-\left\{\gamma_1\left[k,\left(\frac{t}{\alpha}\right)^\tau\right]\right\}^\lambda\right)^{\varphi-1}$$
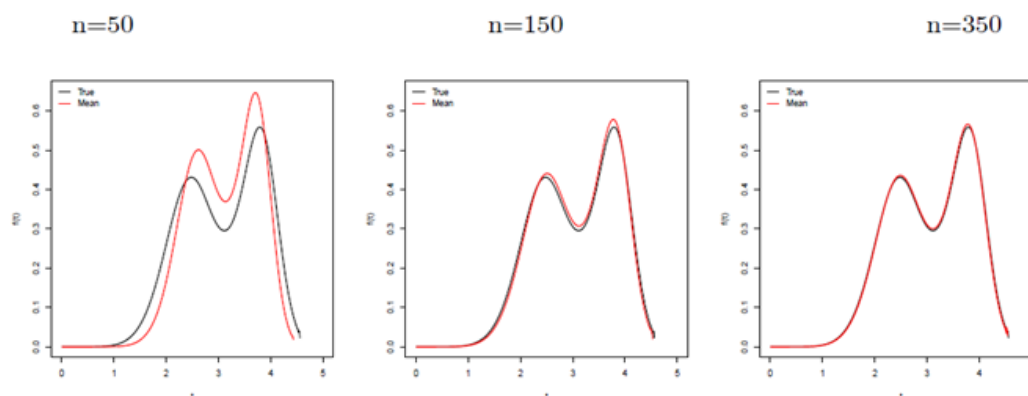
(26)

Figure 5 Some OLLGG density functions at the true parameter values and at the AEs for scenario 1.
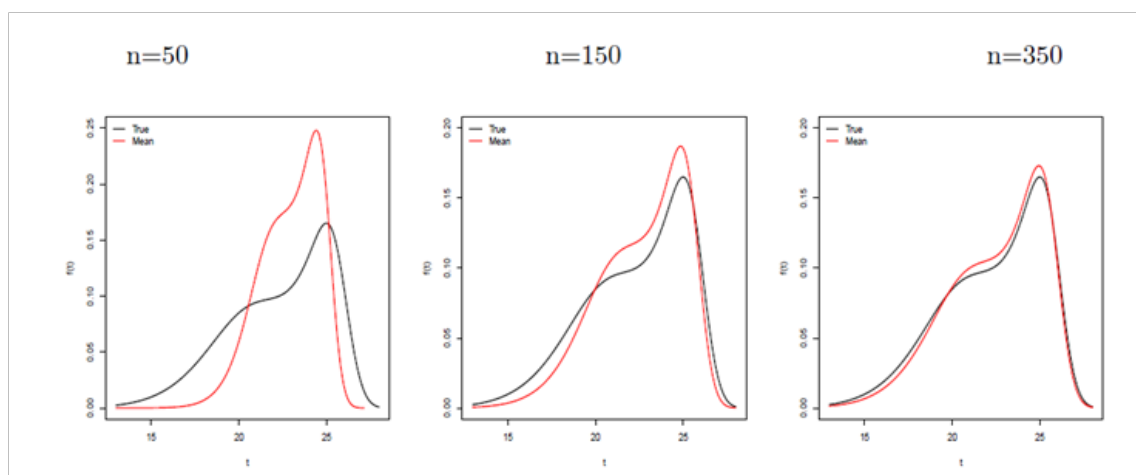

Figure 6 Some OLLGG density functions at the true parameter values and at the AEs for scenario 2.

Table 4 MLEs and (MSEs) for the estimates of the OLLGG parameters

n	Parameters	Actual values	0%	10%	30%
50	α	2.00	2.0404 (0.1984)	2.0366(0.2257)	2.0441 (0.2836)
	τ	5.00	5.3257 (1.8523)	5.395 (3.3121)	5.5626 (4.3955)
	k	10.00	10.7653 (2.9900)	10.9566 (3.20461)	11.2739 (3.63055) (3.63055)
	λ	0.15	0.1708 (0.0115)	0.1708 (0.0149)	0.1736 (0.0201)
150	α	2.00	2.0393 (0.0242)	2.0382 (0.03220)	2.0427 (0.0621)
	τ	5.00	5.1585 (0.2070)	5.1763 (0.2784)	5.2257 (0.5882)
	k	10.00	9.8491 (1.9955)	9.9663 (3.2201)	10.0686 (7.2771)
	λ	0.15	0.1528 (0.0011)	0.1521 (0.0015)	0.1539 (0.0022)

Table continued

n	Parameters	Actual values	0%	10%	30%
	α	2.00	2.0065 (0.0024)	2.0089 (0.0033)	2.0181 (0.0115)
350	τ	5.00	5.0417 (0.0276)	5.0483 (0.0315)	5.0823 (0.0969)
	k	10.00	10.0120 (0.2220)	9.9941 (0.3281)	9.9645 (1.3263)
	λ	0.15	0.1511 (0.0001)	0.1506 (0.0002)	0.1510 (0.0005)

Table 5 Posterior means and (MSEs) for the estimates of the OLLGG parameters

n	Parameteres	Actual values	0%	10%	30%
	α	2.00	1.8130 (0.0703)	1.7642 (0.1691)	1.6585 (0.3824)
	τ	5.00	4.1719 (1.1152)	3.9498 (1.8535)	3.6105 (2.7121)
50	k	10.00	9.9011 (0.0601)	9.6298 (3.3379)	10.2626 (6.3004)
	λ	0.15	0.2795 (0.0319)	0.3293 (0.0623)	0.4377 (0.3072)
	α	2.00	1.8891 (0.0240)	1.9183 (0.0474)	1.9070 (0.0548)
	τ	5.00	4.4648 (0.4132)	4.4970 (0.5506)	4.4131 (0.6805)
150	k	10.00	9.9893 (0.0824)	9.6082 (3.9058)	9.5601 (4.2357)
	λ	0.15	0.2005 (0.0031)	0.2169 (0.0061)	0.2397 (0.0119)
	α	2.00	1.9283 (0.0128)	1.9348 (0.0169)	1.9336 (0.0211)
	τ	5.00	4.6425 (0.2232)	4.6729 (0.2098)	4.6333 (0.2833)
350	k	10.00	9.9929 (0.0913)	10.0471 (1.1531)	9.9108 (1.1627)
	λ	0.15	0.1812 (0.0014)	0.1795 (0.0012)	0.1876 (0.0020)

where $\gamma_1(k,x) = \gamma(k,x) \big/ \Gamma(k)$ is the incomplete gamma function ratio, α is a scale parameter and the other positive parameters τ, k, φ and λ are shape parameters.

Next, we report the MLEs and their corresponding standard errors (SEs) in parentheses of the parameters and the values of the Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC). The lower the values of these criteria, the better the fit. In each case, the parameters are estimated by maximum likelihood using the NLMixed procedure in SAS.

We compute the MLEs of the model parameters and the AIC, CAIC and BIC statistics for each fitted model to these data. The OLLGG model was fitted and compared with the fits from two sub-models

cited before. The results are reported in Table 6. The three information criteria agree on the model's ranking. The lowest values of these criteria correspond to the OLLGG distribution, which could be preferred in this case.

We perform the LR tests to verify if the extra shape parameter λ is really necessary. We provide the histogram of the data and the fitted density functions. Formal tests for the skewness parameter in the generated distribution can be based on LR statistics. The LR statistics for comparing the fitted models are listed in Table 7. We reject the null hypotheses in the two tests in favor of the wider distribution. The rejection is extremely highly significant and it gives clear evidence of the potential need for the shape parameter λ when modeling real data. More information is provided by a visual comparison of the histogram of the data and the fitted density functions. The plots of the fitted OLLGG, GG and Weibull densities are displayed in Figure 9a. The estimated OLLGG density provides the closest fit to the histogram of the data.

Table 6 MLEs of the model parameters for the temperature data and information criteria

Model	α	τ	k	λ	φ	AIC	CAIC	BIC
OLLGG	21.2911	13.0661	2.8755	0.2882		1752.1	1752.2	1767.7
	(0.0012)	(0.0234)	(0.1095)	(0.0127)				
KumGG	25.3965	25.2759	12.8897	0.0243	2.3730	1780.6	1780.7	1800.1
	(1.6147)	(3.0850)	(0.6885)	(0.0079)	(2.4887)			
EGG	23.8850	22.9475	12.8766	0.0215	1	1777.6	1777.7	1793.2
	(2.8175)	(7.7331)	(9.1805)	(0.0019)				
GG	26.1868	33.1789	0.1888	1		1777.6	1777.7	1788.3
	(0.1877)	(7.5737)	(0.0514)	(-)				
Weibull	23.5808	9.4296	1	1		1796.4	1796.5	1804.2
	(0.1376)	(0.4038)	(-)	(-)				
	α	γ	a	b				
BW	25.0516	25.7636	0.2460	0.6159		1778.1	1778.3	1793.7
	(1.3335)	(8.2195)	(0.0858)	(0.4512)				

Table 7 LR tests

Models	Hypotheses	Statistic w	p-value
OLLGG vs GG		26.5	<0.0001
OLLGG vs Weibull	$H_0 : \lambda = k = 1\ vs\ H_1 : H_0\ is\ false$	48.3	<0.0001

In order to assess if the model is appropriate, plots of the fitted OLLGG, GG and Weibull cumulative distributions and the empirical cdf are displayed in Figure 9b. They indicate that the OLLGG distribution gives a good fit to these data.

Under a Bayesian approach, we also fit the OLLGG model and some models described above. For each fitted model to these data, the Bayesian estimates of the model parameters and the DIC, EAIC and EBIC statistics are shown in the Tables 8 and 9, respectively. According to the three Bayesian information criteria, the OLLGG model stands out as the best one.

Survival data

Aids is a pathology that mobilizes its sufferers because of the implications for their interpersonal relationships and reproduction. Therapeutic advances have enabled seropositive women to bear children safely. In this respect, the pediatric immunology outpatient service and social service of Hospital das Cl´ınicas have a special program for care of newborns of seropositive mothers to provide orientation and support for antiretroviral therapy to allow these women and their babies to live as normally as possible. Here, we analyze a data set on the time to serum reversal of 148 children exposed to HIV by vertical transmission, born at Hospital das Cl´ınicas (associated with the Ribeirão Preto School of Medicine) from 1995 to 2001, where the mothers were not treated (Silva, [30]; Perdoná, [31]). Vertical HIV transmission can occur during gestation in around 35% of cases, during labor and birth itself in some 65% of cases, or during breast feeding, varying from 7% to 22% of cases. Serum reversal or serological reversal can occur in children of HIV-contaminated mothers. It is the process by which HIV antibodies disappear from the blood in an individual who tested positive for HIV infection. As the months pass, the maternal antibodies are eliminated and the child ceases to be HIV positive. The exposed newborns were monitored until definition of their serological condition, after administration of Zidovudin (AZT) in the first 24 hours and for the following 6 weeks. We assume that the lifetimes are independently distributed, and also independent from the censoring mechanism.
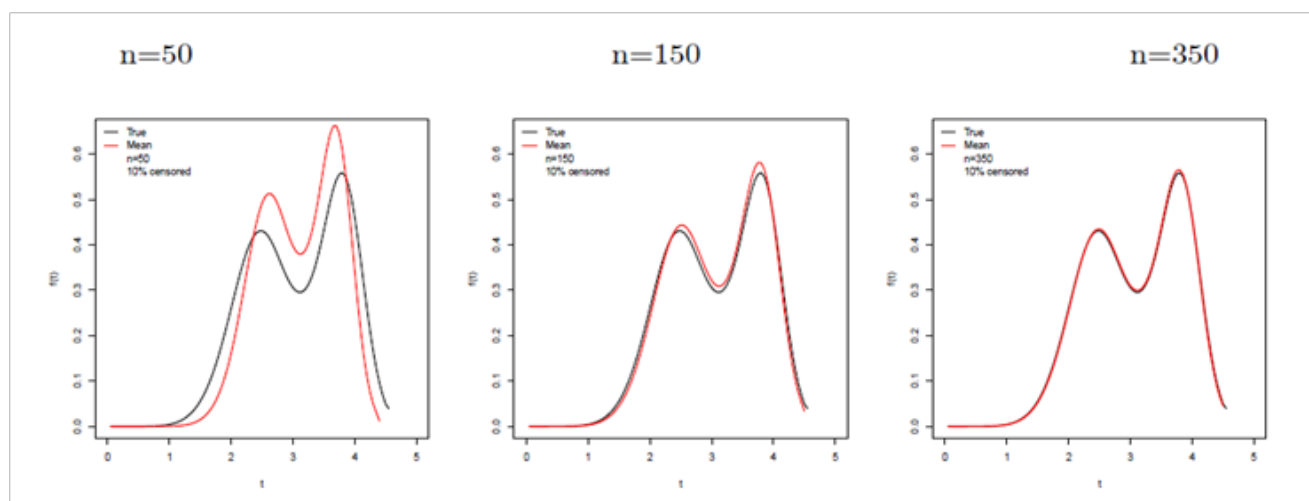
Figure 7 Some OLLGG density functions at the true parameter values and at the AEs for scenario 1 and censored data.
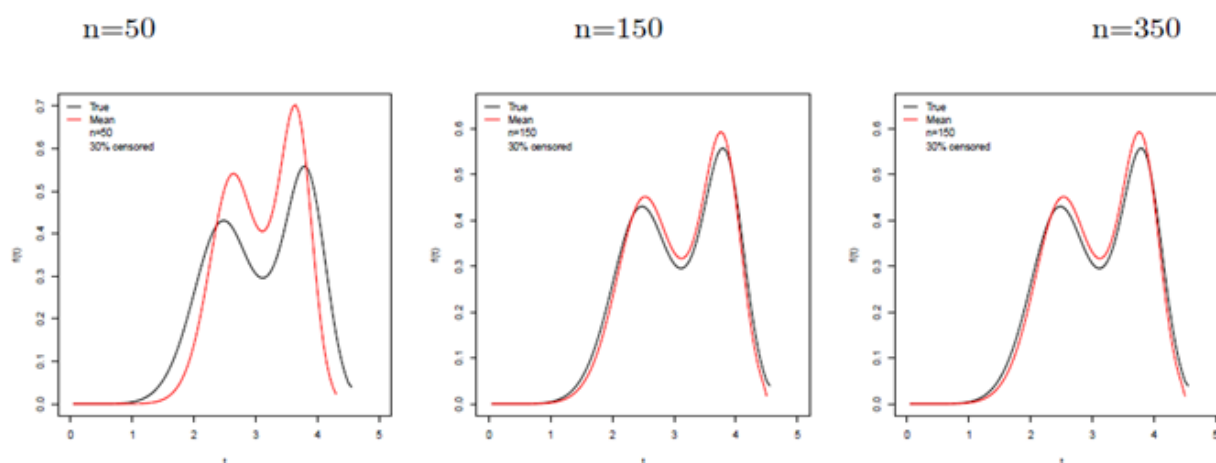
Figure 8 Some OLLGG density functions at the true parameter values and at the AEs for scenario 2 and censoringed data.

Table 8 Posterior mean (standard deviation) and 95% highest posterior density (HPD) interval of the model parameters

Model	α	τ	k	λ	φ
OLLGG	20.5189 (0.7529)	12.6685 (1.2244)	4.3121 (1.2345)	0.2245 (0.0456)	
	(18.8963, 21.8643)	(10.2667, 14.8841)	(2.0842, 6.8480)	(0.1573, 0.3168)	
KumGG	25.4831 (0.1970)	25.7606 (0.2195)	13.3406 (0.1878)	0.0226 (0.00109)	2.3779 (0.1234)
	(25.1727, 25.8274)	(25.3986, 26.1771)	(13.0556, 13.7286)	(0.0205, 0.0247)	(2.1607, 2.6765)
EGG	24.2333 (0.1250)	23.9107 (0.3303)	9.5727 (0.6326)	0.0278 (0.0022)	
	(23.9937, 24.4557)	(23.3771, 24.4652)	(8.6935, 10.9631)	(0.0237, 0.0322)	
GG	26.1305 (0.2334)	32.4133 (7.9028)	0.2104 (0.0669)		
	(25.6783, 26.5446)	(18.2675, 48.2415)	(0.0981, 0.3381)		
Weibull	23.5782 (0.1381)	9.3741 (0.4078)			
	(23.3033, 23.8465)	(8.5262, 10.1351)			

Citation: Prataviera, F, Cordeiro GM, Suzuki AK, et al. The odd log-logistic generalized gamma model: properties, applications, classical and bayesian approach. *Biom Biostat Int J.* 2017;6(4):388–405. DOI: 10.15406/bbij.2017.06.00174

Table 9 Bayesian information criteria

Model	DIC	EAIC	EBIC
OLLGG	1746.344	1752.546	1768.146
KumGG	1775.319	1783.009	1802.508
EGG	1773.724	1779.722	1795.322
GG	1774.657	1779.718	1791.418
Weibull	1796.501	1798.483	1806.283

Tables 10, 11 and 12 list, respectively, the MLEs and their corresponding SEs in paren¬theses and posterior mean (standard deviation) and 95% highest posterior density (HPD) interval for the parameters and the values of the model selection statistics. These results indicate that the OLLGG model has the lowest AIC, BIC, CAIC, DIC, EAIC e EBIC values among those of all fitted models, and hence it could be chosen as the best model.

Note that the KumGG model is competitive with the model OLLGG. However, the model KumGG has two disadvantages:

It does not model bimodal data.

It has five parameters, i.e. is less parsimonious.

Table 10 MLEs of the model parameters for the serum reversal data, the corresponding SEs (given in parentheses) and the AIC, BIC and CAIC statistics

| Model	α	τ	k	λ	φ	AIC	BIC	CAIC
OLLGG	352.0	46.9706	0.1043	0.4468		771.1	783.6	771.9
	(1.0590)	(1.4847)	(0.0324)	(0.0881)				
KumGG	350.05	49.8303	0.2176	0.1282	0.3424	770.7	785.7	771.1
	(1.5707)	(5.8895)	(0.0073)	(0.0236)	(0.0522)			
EGG	350.45	22.2991	1.0741	0.1072	1	798.1	810.1	798.3
	(2.4187)	(0.0375)	(0.0004)	(0.0113)				
GG	379.40	24.5312	0.0974	1	1	783.7	792.7	783.9
	(8.8211)	(10.3258)	(0.0402)					
Weibull	307.62	3.1132	1	1	1	808.0	814.0	808.1
	(12.3523)	(0.3250)						
	α	γ	a	b				
BW	349.99	6.3895	0.3944	0.9273		797.9	809.9	798.2
	(23.0923)	(0.7657)	(0.0468)	(0.3361)				

A comparison of the proposed distribution with some of its sub-models using LR statis¬tics is performed in Table 13. The figures in this table, specially the p-values, suggest that the OLLGG model yields a better fit to these data than the other three distributions. In order to assess if the model is appropriate, plots of the estimated survival functions of the KumGG, EGG, GG, Weibull and BW distributions and the empirical survival function are given in Figure 10. We conclude that the OLLGG distribution provides a good fit for these data.

Table 11 Posterior means (stantard deviations) and 95% HPD intervals for the model parameters in the serum reversal data

Model	α	τ	k	λ	φ
OLLGG	348.9 (11.5813)	47.7542 (22.7428)	0.1741 (0.1443)	0.4342 (0.1619)	
	(324.1, 366.5)	(15.3289, 98.0374)	(0.0230, 0.4910)	(0.1331, 0.7222)	
KumGG	351 (1.0623)	42.8395 (1.4827)	0.0114 (0.00383)	3.0697 (0.5911)	0.3601 (0.0550)
	(349.0, 353.1)	(39.7113, 45.1984)	(0.0058, 0.0191)	(1.7862, 4.0205)	(0.2678, 0.4790)
EGG	348.6 (0.8519)	19.7657 (1.1290)	4.3776 (1.0638)	0.0309 (0.0097)	
	(347.3, 350.4)	(18.3590, 22.5768)	(2.5525, 6.0764)	(0.0177, 0.0505)	
GG	376.3 (6.7347)	44.2185 (16.2531)	0.0652 (0.0341)		
	(364.5, 389.9)	(15.9226, 71.1764)	(0.0279, 0.1302)		
Weibull	307.5 (12.6278)	3.0864 (0.3237)			
	(283.7, 333.4)	(2.4619, 3.7203)			
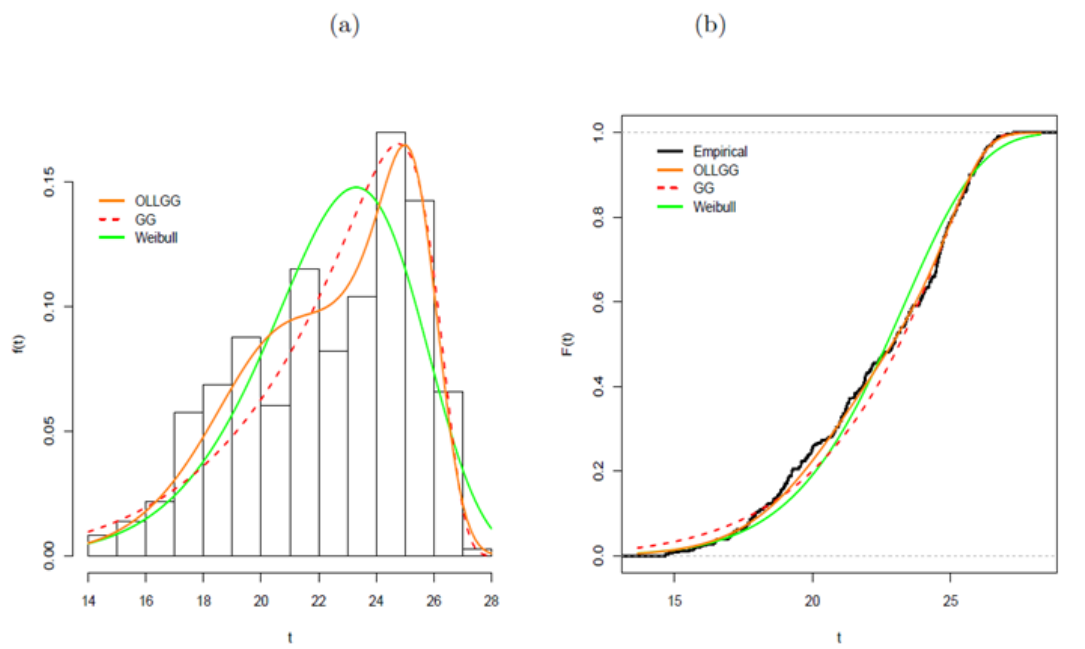
Figure 9 (a) Estimated densities of the OLLGG, GG and Weibull models for fibre data. (b) Estimated cumulative functions of the OLLGG, GG and Weibull models and the empirical cdf for temperature data.
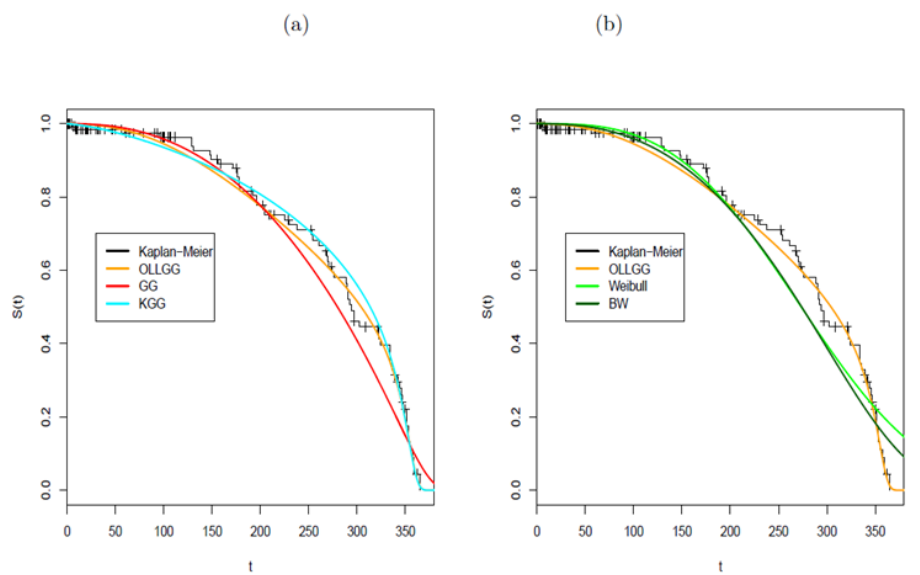
Figure 10 Estimated survival function by fitting the OLLGG distribution and some other models and the empirical survival for the serum reversal data. (a) OLLGG vs KGG and GG. (b) OLLGG vs BW and Weibull.

Table 12 Bayesian information criterion

Model	DIC	EAIC	EBIC
OLLGG	752.017	775.385	787.3738
KumGG	764.746	772.79	787.776
EGG	781.475	788.53	800.519
GG	776.599	783.425	792.417
Weibull	807.984	809.989	815.983

Table 13 LR statistics for the serum reversal data

Model	Hypotheses	Statistic w	*p*-value
OLLGG vs GG	$H_0 : \lambda = 1 \, vs \, H_1 : H_0 \, is \, false$	13.0	0.00031
OLLGG vs Weibull	$H_0 : \varphi = \lambda = k = 1 \, vs \, H_1 : H_0 \, is \, false$	40.3	<0.0001

Concluding remarks

The odd log-logistic generalized gamma (OLLGG) distribution provides a rather general and flexible framework for statistical analysis of positive data. It unifies some previously known distributions and yields a general overview of these distributions for theoretical studies. It also represents a rather flexible mechanism for fitting a wide spectrum of real world data sets. The OLLGG distribution is motivated by the wide use of the generalized gamma (GG) distribution in practice, and also for the fact that the generalization provides more flexibility to analyze skewed data. This extension provides a continuous cross over to other cases with different shapes (e.g. a particular combination of skewness and kurtosis). We derive an expansion for the density function as a linear combination of GG density functions. We obtain explicit expressions for the moments and moment generating function. The estimation of parameters is approached by the maximum likelihood method and a Bayesian approach, where the Gibbs algorithms along with metropolis steps are used to obtain the posterior summaries of interest for survival data with right censoring. Two applications of the OLLGG distribution to real data show that it could provide a better fit than other statistical models frequently used in lifetime data analysis.

Acknowledgements

None.

Conflicts of interest

None.

References

1. Gleaton JU, Lynch JD. On the distribution of the breaking strain of a bundle of brittle elastic fibers. *Advances in Applied Probability*. 2006;36(1):98–115.

2. Braga AS, Cordeiro GM, Ortega EMM, et al. The odd log-logistic normal distribution: Theory and applications in analysis of experiments. *Journal of Statistical Theory and Practice*. 2016;10(2):311–335.

3. da Cruz JN, Ortega EMM, Cordeiro GM. The log-odd log-logistic Weibull re-gression model: Modelling, estimation, influence diagnostics and residual analysis. *Journal of Statistical Computation and Simulatio*. 2016;86(8):1516–1538.

4. Cordeiro GM, Ortega EMM, Silva GO. The exponentiated generalized gamma distribution with application to lifetime data. *Journal of Statistical Computation and Simulation*. 2011;81(7):827–842.

5. Pascoa MAR, Ortega EMM, Cordeiro GM. The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*. 2011;8(5):411–433.

6. Ortega EMM, Cordeiro GM, Pascoa MAR. The generalized gamma geometric distribution. *Journal of Statistical Theory and Applications*. 2011;10(3):433–454.

7. Cordeiro GM, Castellares F, Montenegro LC, de Castro M. The beta generalized gamma distribution. *Statistics*. 2013;47(4):888–900.

8. Lucena SEF, Silva AH, Cordeiro GM. The transmuted generalized gamma distribution: Properties and application. *Journal of Data Science*. 2015;13(1):409–420.

9. Silva RB, Bourguignon M, Cordeiro GM. A new compounding family of dis- tributions: The generalized gamma power series distributions. Journal of Computational and Applied Mathematics. 2016;303(C):119–139.

10. http://www.tandfonline.com/doi/abs/10.1081/STA-120003130

11. Stacy EW. A generalization of the gamma distribution. Annals of Mathematical Statistics. 1962;33(3):1187–1192.

12. Stacy EW, Mihram GA. Parameter estimation for a generalized gamma distri- bution. *Technometrics*. 1965;7(3):349–358.

13. Lawless JF. Statistical Models and Methods for Lifetime Data. 2nd ed. New York, USA;2003.

14. Mudholkar GS, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*. 1993;42(2):299–302.

15. Mudholkar GS, Srivastava DK, Freimer M. The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*. 1995;37(4):436–445.

16. Nadarajah S. The exponentiated Gumbel distribution with climate application. *Environmetrics*. 2005;17(1):13–23.

17. Shirke DT, Kakade CS. On exponentiated lognormal distribution. International Journal of Agricultural and Statistical Sciences. 2006;2:319–326.

18. Nadarajah S, Gupta AK. The exponentiated gamma distribution with application to drought data. Calcutta Statistical *Association Bulletin*. 2007;59:29–54.

19. Nadarajah S, Kotz S. The exponetiated type distributions. *Acta Applicandae Mathematicae*. 2006;92(2):97–111.

20. Gradshteyn and Ryzhik's Table of Integrals, Series, and Products. 6th edn. Alan Jeffrey and Daniel Zwillinger; 2000.

21. Nadarajah S. Generalized gamma variables with drought application. *Journal of the Korean Statistical Society*. 2008;37(1):37–45.

22. Murphy JA. Handbook of hypergeometric integrals-theory, application, tables, computer programs, by Harold Exton, Ellis Horwood Limited, Chichester, 1978. No. of pages: 316, price £15. *International Journal for Numerical Methods in Engineering*. 1979;14(1):1–155.

23. Aarts RM. Lauricella functions. From MathWorld, A Wolfram Web Resource, created by Eric W Weisstein. 2000.

24. http://www.emeraldinsight.com/doi/pdfplus/10.1108/10610420610679593

25. Spiegelhalter DJ, Best NG, Carlin BP, et al. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series*. 2002;B64(4):583–639.

26. Brooks SP, Jim Smith, Aki Vehtari, et al. Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde. *Journal of the Royal Statistical Society*. Series B: Statistical Methodology. 2002;64(4):616–639.

27. Carlim BP, Louis TA. Bayes and empirical bayes methods for data analysis. 2nd edn. In: Chapman and Hall. Boca Raton, Florida, USA. 2001;85(503): 381-383.

28. Cordeiro GM, Alizadeh M, Tahir MH, et al. The beta odd log-logistic generalized family of distributions. Hacettepe Journal of Mathematics and Statistics. 2016;45: 1–28.

29. Famoye F, Lee C, Olumolade O. The beta-Weibull distribution. *Journal of Statistical Theory and Applications*. 2005;4:121–136.

30. Silva ANF. Estudo evolutivo das criancas expostas ao HIV enotificadas pelo nucleo de vigilancia epidemiologica do HCFMRP-USP. M.Sc. Thesis. University of Sao Paulo, Brazil. 2004.

31. Perdona GSC. Modelos de riscos aplicadosa analise de sobrevivencia. Doctoral Thesis, Institute of Computer Science and Mathematics, University of Sao Paulo, Brasil. 2006. 20,31,27