

METHODOLOGY

Open Access



The odd log-logistic logarithmic generated family of distributions with applications in different areas

Morad Alizadeh¹, S. M. T. K. MirMostafae², Edwin M. M. Ortega^{3*}, Thiago G. Ramires³ and Gauss M. Cordeiro⁴

*Correspondence: edwin@usp.br
³Departamento de Ciências Exatas,
Universidade de São Paulo,
Piracicaba, Brazil
Full list of author information is
available at the end of the article

Abstract

We introduce and study general mathematical properties of a new generator of continuous distributions with three extra parameters called the odd log-logistic logarithmic generated family of distributions. We present some special models and investigate the asymptotes and shapes. The new density function can be expressed as a linear combination of exponentiated densities based on the same baseline distribution. Explicit expressions for the ordinary and incomplete moments, quantile and generating functions, Shannon and Rényi entropies and order statistics, which hold for any baseline model, are determined. We discuss the estimation of the model parameters by maximum likelihood. Further, we introduce the new family in long-term survival models. We illustrate the potentiality of the proposed models by means of four applications to real data.

Keywords: Generated family, Maximum likelihood, Moment, Order statistic, Quantile function, Survival analysis

AMS Subject Classification: 97K50; 62N01; 62N02

1 Introduction

Statistical distributions are very useful in describing and predicting real world phenomena. Numerous extended distributions have been extensively used over the last decades for modeling data in several areas. Recent developments focus on defining new families that extend well-known distributions and at the same time provide greater flexibility in modeling data in practice. Hence, several classes to generate new distributions by adding one or more parameters have been proposed in the statistical literature. Some well-known generators are the Marshall-Olkin generated (MO-G) by Marshall and Olkin (1997), beta-G by Eugene et al. (2002), Kumaraswamy-G (Kw-G) by Cordeiro and de Castro (2011), Weibull-G by Bourguignon et al. (2014), exponentiated half-logistic-G by Cordeiro et al. (2014a), Lomax-G by Cordeiro et al. (2014b), among others.

Let $G(x; \xi)$ be a baseline cumulative distribution function (cdf) and ξ be the vector of associated parameters. Recently, Gleaton and Lynch (2004, 2006, 2010) defined the cdf of the *odd log-logistic* family with one extra shape parameter $\alpha > 0$ by

$$H(x) = \frac{G(x; \xi)^\alpha}{G(x; \xi)^\alpha + \bar{G}(x; \xi)^\alpha}, \quad (1)$$

where $\bar{G}(x; \xi) = 1 - G(x; \xi)$. More precisely, they showed the following facts:

- The set of generalized log-logistic (GLL) transformations form an Abelian group with the binary operation of composition;
- The transformation group partitions the set of all lifetime distributions into equivalence classes, so that any two distributions in an equivalence class are related through a GLL transformation;
- Either every distribution in an equivalence class has a moment generating function (mgf), or none does;
- Every distribution in an equivalence class has the same number of moments;
- Each equivalence class is linearly ordered according to the transformation parameter, with larger values of this parameter corresponding to smaller dispersion of the distribution about the common median class; and
- Within an equivalence class, the Kullback-Leibler information is an increasing function of the ratio of the transformation parameters.

In addition, Gleaton and Rahman (2010, 2014) obtained asymptotic results for the maximum likelihood estimates (MLEs) of the parameters of these two distributions. They proved that for distributions generated from either a two-parameter Weibull or a two-parameter inverse Gaussian distributions by a GLL transformation, the joint MLEs of the parameters are asymptotically normal and efficient, provided the GLL transformation parameter exceeds three.

We define the cdf of the *odd log-logistic logarithmic-G* (OLLL-G) family by

$$F(x) = \frac{\log \left[1 - \frac{\beta G(x; \xi)^\alpha}{G(x; \xi)^\alpha + \bar{G}(x; \xi)^\alpha} \right]}{\log(1 - \beta)}, \tag{2}$$

where $G(x; \xi)$ is the baseline cdf depending on a parameter vector ξ and $\alpha > 0$ and $0 < \beta < 1$ are two additional shape parameters. It includes the odd log-logistic-G (OLL-G) family (Gleaton and Lynch 2004, 2006) and the logarithmic-G family. Some special models are given in Table 1.

This paper is organized as follows. In Section 2, we provide a physical interpretation of the OLLL-G family and define two special cases. In Section 3, two useful linear representations are derived. In Section 4, we obtain explicit expressions for the moments and generating function. In Section 5, general expressions for the Rényi and Shannon entropies and order statistics are presented. Estimation of the model parameters by maximum likelihood is investigated in Section 6. We also present the performance of the MLEs through a simulation study. In Section 7, the OLLL-G model is modified for possible presence of long-term survivors in the data. Four applications to real data illustrate the performance of the proposed models in Section 8. The paper is concluded in Section 9.

Table 1 Some special models

α	β	$G(x)$	Reduced distribution
-	$\uparrow 1$	$G(x)$	OLL-G family [Gleaton and Lynch (2004, 2006)]
1	-	$G(x)$	Logarithmic-G family baseline distribution
1	$\uparrow 1$	$G(x)$	$G(x)$

2 Motivation and special cases

The density function corresponding to (2) is given by

$$f(x; \alpha, \beta, \xi) = \frac{\alpha\beta g(x, \xi)G(x, \xi)^{\alpha-1}\bar{G}(x, \xi)^{\alpha-1}}{-[G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha][(1-\beta)G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha]\log(1-\beta)}, \quad (3)$$

where $g(x; \xi)$ is the baseline pdf. Hereafter, a random variable X with density function (3) is denoted by $X \sim \text{OLLL-G}(\alpha, \beta, \xi)$. Further, we can omit sometimes the dependence on the vector ξ of the parameters and write simply $G(x) = G(x; \xi)$.

A motivation of this family can be explained as follows. Suppose that a parallel system is made up of N components and the lifetimes of the components are independent and identically distributed (iid) random variables, denoted by Z_1, \dots, Z_N , with common cdf (1). Then, the system fails as soon as the last component fails, namely the lifetime of the whole system is represented by $X = \max\{Z_1, \dots, Z_N\}$. In many survival parallel systems, it is almost impossible to have a fixed number of components because some of them get lost or censored for various reasons. Therefore we may assume that N is a discrete random variable. Suppose that N has the logarithmic distribution with probability mass function given by

$$P(N = n) = \frac{-1}{\log(1-\beta)} \frac{\beta^n}{n}, \quad n = 1, 2, \dots, \quad 0 < \beta < 1.$$

Then, the cdf of the life length of the whole system, X , is obtained as

$$\begin{aligned} F(x) &= \sum_{n=1}^{\infty} P(X \leq x | N = n)P(N = n) \\ &= \sum_{n=1}^{\infty} \left[\frac{G(x, \xi)^\alpha}{G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha} \right]^n \frac{-1}{\log(1-\beta)} \frac{\beta^n}{n} = \frac{-\log \left[1 - \frac{\beta G(x, \xi)^\alpha}{G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha} \right]}{-\log(1-\beta)} \end{aligned}$$

which is identical to (2).

The hazard rate function (hrf) of X becomes

$$h(x; \alpha, \beta, \xi) = \frac{-\alpha\beta g(x, \xi)G(x, \xi)^{\alpha-1}\bar{G}(x, \xi)^{\alpha-1} [G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha]^{-1}}{[(1-\beta)G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha] \left\{ \log \left[\frac{1-\beta}{1 - \frac{\beta G(x, \xi)^\alpha}{G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha}} \right] \right\}}. \quad (4)$$

The OLLL-G family of distributions is easily simulated by inverting (2) as follows: if U has a uniform $U(0, 1)$ distribution, then

$$Q(U) = Q_G \left(\frac{[1 - (1-\beta)U]^{\frac{1}{\alpha}}}{[1 - (1-\beta)U]^{\frac{1}{\alpha}} + [(1-\beta)U - (1-\beta)]^{\frac{1}{\alpha}}} \right), \quad (5)$$

has the density function (3), where $Q_G(u) = G^{-1}(u)$ is the quantile function (qf) of the baseline G .

Remark 1 Although, we have stated that $\beta \in (0, 1)$, Eq. (2) is still a cdf if $\beta < 0$. Hence, we can consider the OLLL-G family defined for any $\beta \in (-\infty, 0) \cup (0, 1)$.

In Appendix 1, we present the asymptotes and shapes of the OLLL-G model.

2.1 Special OLLL-G distributions

The OLLL-G density function (3) allows for greater flexibility of its tails and can be widely applied in many areas of engineering and biology. It will be most tractable when $G(x; \xi)$ and $g(x; \xi)$ have simple analytic expressions. We now present and discuss some special cases of this family because it extends several widely-known distributions in the literature.

2.1.1 Odd log-logistic logarithmic Weibull (OLLL-W) model

Let $G(x; \xi) = 1 - e^{-(bx)^a}$ the Weibull cdf, with scale parameter $b > 0$ and shape parameter $a > 0$, where $\xi = (a, b)$. The OLLL-W density function (for $x > 0$) is given by

$$f(x; a, b, \alpha, \beta) = \frac{a \alpha \beta b^a e^{-\alpha(bx)^a} [1 - e^{-(bx)^a}]^{\alpha-1}}{-\{e^{-\alpha(bx)^a} + [1 - e^{-(bx)^a}]^\alpha\} \{e^{-\alpha(bx)^a} + (1 - \beta) [1 - e^{-(bx)^a}]^\alpha\} \log(1 - \beta)}$$

Figures 1 and 2 display some shapes of the OLLL-W density and hazard functions, respectively. The plots in Fig. 1 indicate that this density function can be decreasing, unimodal and bimodal. Moreover, Fig. 2 reveals that the hrf of the OLLL-W model can be decreasing, increasing, increasing-decreasing-increasing, bathtub shaped and upside bathtub shaped.

2.1.2 Odd log-logistic logarithmic normal (OLLLN) model

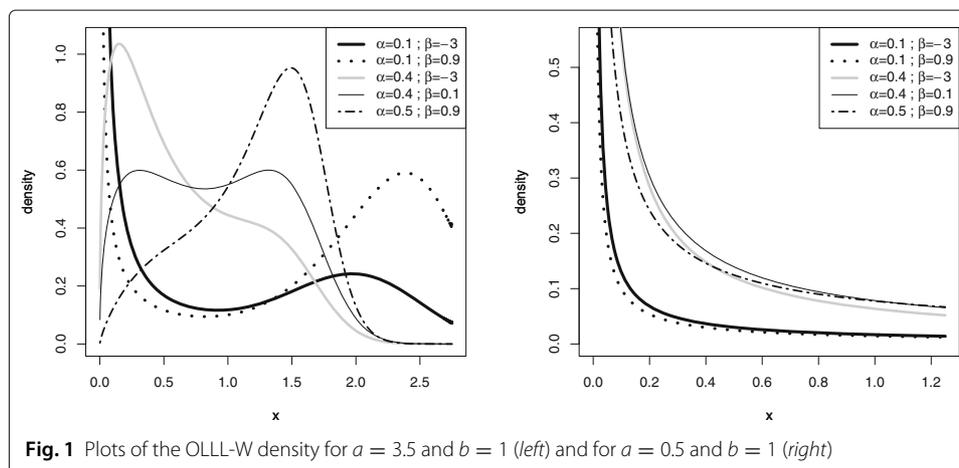
Consider the normal model with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$, whose pdf and cdf (for $x \in \mathbb{R}$) are given by

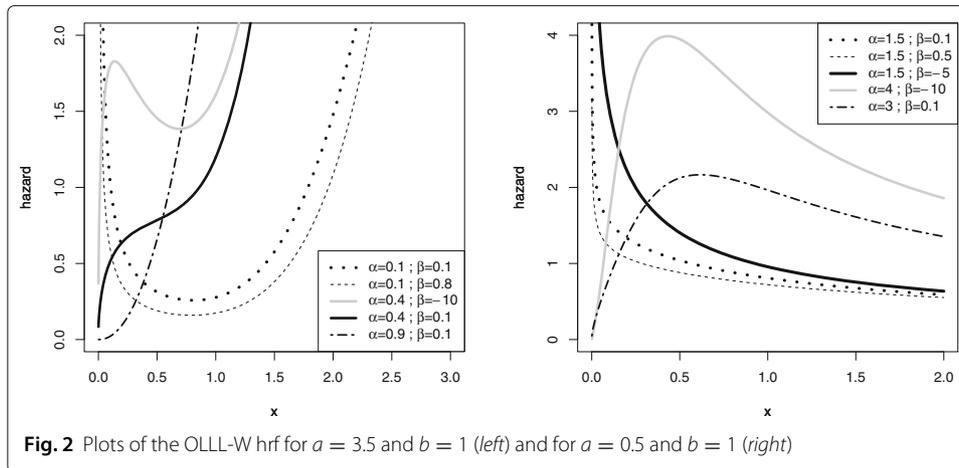
$$g(x; \xi) = \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) \quad \text{and} \quad G(x; \xi) = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

respectively, where $\xi = (\mu, \sigma)^T$. Inserting these expressions in (3), the OLLLN pdf is given by

$$f(x; \mu, \sigma, a, b) = \frac{\alpha \beta \phi\left(\frac{x - \mu}{\sigma}\right) \Phi\left(\frac{x - \mu}{\sigma}\right)^{\alpha-1} [1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]^{\alpha-1} [\log(1 - \beta)]^{-1}}{-\sigma \left\{ \Phi\left(\frac{x - \mu}{\sigma}\right)^\alpha + [1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]^\alpha \right\} \left\{ (1 - \beta) \Phi\left(\frac{x - \mu}{\sigma}\right)^\alpha + [1 - \Phi\left(\frac{x - \mu}{\sigma}\right)]^\alpha \right\}}$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, α and β are the shape and scale parameters, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal





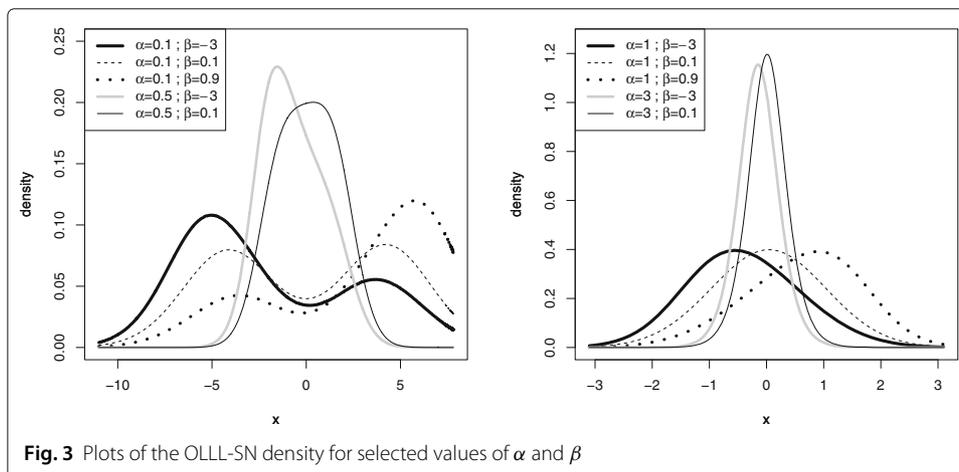
distribution, respectively. For $\mu = 0$ and $\sigma = 1$, we obtain the OLLL-standard normal (OLLL-SN) distribution. Plots of the OLLL-SN density function for selected parameter values are displayed in Fig. 3. We note that this model is suitable for unimodal and bimodal data sets.

3 Linear representations

Let $A(u) = \frac{u^\alpha}{u^\alpha + (1-u)^\alpha}$ be the cdf of the odd log-logistic uniform (Gleaton and Lynch 2004, 2006, 2010) distribution. For $\beta \in (0, 1)$, we have $0 < \beta A(u) < 1$. Then, we can apply the power series reported in Appendix 2 by taking $u = G(x)^\alpha$, since they are always convergent in the interval $(0, 1)$, i.e., the power series are valid in the support of X . Henceforth, we consider that $0 < \beta < 1$, which is not a restrictive assumption since it is in agreement with the logarithm distribution defined for compounding the OLLL-G family.

Based on the power series $-\log(1 - u) = \sum_{i=1}^{\infty} \frac{u^i}{i}$ (which converges for $|u| < 1$), the OLLL-G family cdf can be expanded as

$$F(x) = \frac{-1}{\log(1 - \beta)} \sum_{i=1}^{\infty} \frac{\beta^i}{i} \left[\frac{G(x)^\alpha}{G(x)^\alpha + [1 - G(x)]^\alpha} \right]^i.$$



By using Eq. (24) given in Appendix 2, we obtain

$$F(x) = \frac{-1}{\log(1 - \beta)} \sum_{i=1}^{\infty} \frac{\beta^i}{i} \sum_{k=0}^{\infty} h_k^*(\alpha, i) G(x)^k,$$

where the coefficients $h_k^*(\alpha, i)$ can be determined from the recursive formula given after (24).

Further, we define the exponentiated-G (“Exp-G”) distribution for an arbitrary parent distribution G , say $W \sim \text{Exp}^c G$, if W has cdf and pdf given by $H_c(x) = G(x)^c$ and $h_c(x) = c g(x) G(x)^{c-1}$, respectively. This transformed model is also called the Lehmann type I distribution, say $\text{Exp}^c(G)$.

Then, we can rewrite $F(x)$ as

$$F(x) = \sum_{k=0}^{\infty} d_k H_k(x), \tag{6}$$

where $d_k = \sum_{i=1}^{\infty} \frac{-\beta^i h_k^*(\alpha, i)}{i \log(1-\beta)}$ and $H_k(x)$ is the Exp-G cdf with power parameter k .

By differentiating (6), the pdf of X follows as

$$f(x) = \sum_{k=0}^{\infty} d_{k+1} h_{k+1}(x), \tag{7}$$

where $h_{k+1}(x) = (k + 1) G(x)^k g(x)$ is the Exp-G density function with power parameter $(k + 1)$.

Equation (7) reveals that the OLLL-G density function is a linear combination of Exp-G densities. Some structural properties of the new family such as the ordinary and incomplete moments and generating function can be determined from well-established properties of the Exp-G distribution. The properties of Exp-G distributions have been studied by many authors in recent years, see Mudholkar and Srivastava (1993) and Mudholkar et al. (1995) for exponentiated Weibull, Gupta and Kundu (1999) for exponentiated exponential and Nadarajah (2006) for exponentiated Gumbel, among others. The linear representations (6) and (7) are the main results of this section.

4 Moments and generating function

Let Y_k be a exp-G random variable with power parameter $k + 1$, i.e., having density $h_{k+1}(x)$. The n th ordinary moment of $X \sim \text{OLLL-G}$ follows from (7) as

$$E(X^n) = \sum_{k=0}^{\infty} d_{k+1} E(Y_k^n) = \sum_{k=0}^{\infty} (k + 1) d_{k+1} \tau(n, k), \tag{8}$$

where $\tau(n, k) = \int_{-\infty}^{\infty} x^n G(x)^k g(x) dx = \int_0^1 Q_G(u)^n u^k du$. In fact, it is possible to exchange the infinite sum and the integral using the dominated convergence theorem for series.

Expressions for moments of several exp-G distributions are given by Nadarajah and Kotz (2006), which can be used to obtain $E(X^n)$. Cordeiro and Nadarajah (2011) determined $\tau(n, k)$ for some well-known distributions such as the normal, beta, gamma and Weibull distributions.

The variance, skewness and kurtosis measures are given by

$$\begin{aligned} \text{Var}(X) &= E(X^2) - [E(X)]^2, \\ \text{Skewness}(X) &= \frac{E(X^3) - 3E(X)E(X^2) + 2[E(X)]^3}{[\text{Var}(X)]^{\frac{3}{2}}}, \\ \text{Kurtosis}(X) &= \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4}{[\text{Var}(X)]^2}. \end{aligned}$$

Plots of the skewness and kurtosis of the OLLL-W distribution as functions of α and β for $a = 2$ and $b = 1$ are displayed in Fig. 4.

For empirical purposes, the shape of many distributions can be usefully described by what we call the incomplete moments. These types of moments play an important role for measuring inequality, for example, income quantiles and Lorenz and Bonferroni curves, which depend upon the incomplete moments of a distribution. The n th incomplete moment of X is calculated as

$$m_n(y) = \int_{-\infty}^y x^n f(x) dx = \sum_{k=0}^{\infty} (k+1) d_{k+1} \int_0^{G(y)} Q_G(u)^n u^k du. \tag{9}$$

The last integral can be determined analytically or numerically for most baseline distributions. Equation (9) can be used to determine conditional moments, mean deviations and Bonferroni and Lorenz curves of X .

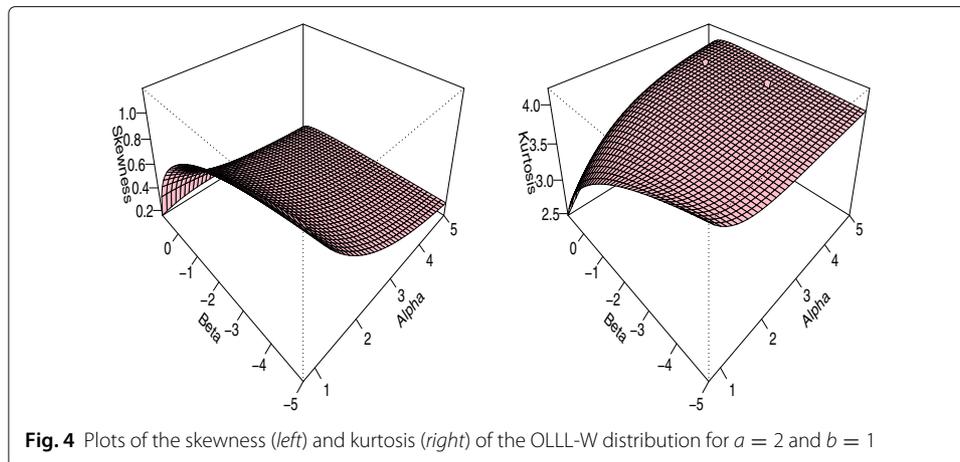
Let $M(t) = E(e^{tX})$ be the mgf of X . We can obtain $M(t)$ from (7) as

$$M(t) = \sum_{k=0}^{\infty} d_{k+1} M_k(t) = \sum_{i=0}^{\infty} (k+1) d_{k+1} \rho(t, k), \tag{10}$$

where $M_k(t)$ is the mgf of Y_k and $\rho(t, k) = \int_{-\infty}^{\infty} e^{tx} G(x)^k g(x) dx = \int_0^1 \exp[t Q_G(u)] u^k du$.

We can determine the mgfs for several OLLL-G distributions directly from Eq. (10).

We present some mathematical properties of the odd log-logistic logarithmic exponential (OLLL-E) distribution in Appendix 3 to illustrate the applicability of the previous results.



5 Other properties

We hardly need to emphasize the necessity and importance of entropies and order statistics in any statistical analysis especially in applied work.

5.1 Entropies

An entropy is a measure of variation or uncertainty of a random variable X . Two popular entropy measures are the Rényi and Shannon entropies (Shannon 1951; Rényi 1961). The Rényi entropy of a random variable with pdf $f(x)$ is defined by

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \left(\int_0^\infty f^\gamma(x) dx \right),$$

for $\gamma > 0$ and $\gamma \neq 1$. The Shannon entropy of a random variable X is defined by $E \{-\log [f(X)]\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$. Direct calculation gives

$$\begin{aligned} E \{-\log [f(X)]\} &= -\log \left[\frac{\alpha\beta}{-\log(1-\beta)} \right] - E \{\log [g(X; \xi)]\} + (1-\alpha)E \{\log [G(X; \xi)]\} \\ &\quad + (1-\alpha)E \{\log [\bar{G}(X; \xi)]\} + 2E \{\log [G(X; \xi)^\alpha + \bar{G}(X; \xi)^\alpha]\} \\ &\quad + E \left\{ \log \left[1 - \frac{\beta G(X, \xi)^\alpha}{G(X, \xi)^\alpha + \bar{G}(X, \xi)^\alpha} \right] \right\}. \end{aligned}$$

First, we define

$$A(a_1, a_2, a_3, a_4; \alpha, \beta) = \int_0^1 \frac{u^{a_1}(1-u)^{a_2}}{[u^\alpha + (1-u)^\alpha]^{a_3} \left[1 - \frac{\beta u^\alpha}{u^\alpha + (1-u)^\alpha} \right]^{a_4}} du.$$

By using the binomial expansion, we have

$$A(a_1, a_2, a_3, a_4; \alpha, \beta) = \sum_{i,j=0}^\infty (-1)^{i+j} \binom{-a_4}{i} \binom{a_2}{j} \int_0^1 \frac{\beta^i u^{a_1+\alpha i+j}}{[u^\alpha + (1-u)^\alpha]^{a_3+i}} du.$$

Second, we have the power series from Eq. (28) given in Appendix 2

$$\left[\frac{1}{u^\alpha + (1-u)^\alpha} \right]^{a_3+i} = \sum_{k=0}^\infty s_k^*(\alpha, a_3+i, 0) u^k,$$

where $s_k^*(\alpha, a_3+i, 0)$ is defined there. Then,

$$A(a_1, a_2, a_3, a_4; \alpha, \beta) = \sum_{i,j,k=0}^\infty \frac{(-1)^{i+j} \beta^i \binom{-a_4}{i} \binom{a_2}{j} s_k^*(\alpha, a_3+i, 0)}{k+1+a_1+\alpha i+j}.$$

After some algebraic manipulations, we can write

$$E \{\log [G(X)]\} = \frac{\alpha\beta}{-\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha+t-1, \alpha-1, 2, 1; \alpha, \beta) \Big|_{t=0},$$

$$E \{\log [\bar{G}(X)]\} = \frac{\alpha\beta}{-\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha-1, \alpha+t-1, 2, 1; \alpha, \beta) \Big|_{t=0},$$

$$E \{\log \{G(X; \xi)^\alpha + \bar{G}(X; \xi)^\alpha\}\} = \frac{\alpha\beta}{-\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha-1, \alpha-1, 2-t, 1; \alpha, \beta) \Big|_{t=0},$$

$$E \left\{ \log \left[1 - \frac{\beta G(X, \xi)^\alpha}{G(X, \xi)^\alpha + \bar{G}(X, \xi)^\alpha} \right] \right\} = \frac{\alpha\beta}{-\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha-1, \alpha-1, 2, 1-t; \alpha, \beta) \Big|_{t=0}.$$

Then, the Shannon entropy of X reduces to

$$\begin{aligned}
 E \{ -\log[f(X)] \} &= -\log \left[\frac{\alpha\beta}{-\log(1-\beta)} \right] - E \{ \log [g(X; \xi)] \} \\
 &\quad - \frac{\alpha(1-\alpha)\beta}{\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha+t-1, \alpha-1, 2, 1; \alpha, \beta) \Big|_{t=0} \\
 &\quad - \frac{\alpha(1-\alpha)\beta}{\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha-1, \alpha+t-1, 2, 1; \alpha, \beta) \Big|_{t=0} \\
 &\quad - \frac{2\alpha\beta}{\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha-1, \alpha-1, 2-t, 1; \alpha, \beta) \Big|_{t=0} \\
 &\quad - \frac{\alpha\beta}{\log(1-\beta)} \frac{\partial}{\partial t} A(\alpha-1, \alpha-1, 2, 1-t; \alpha, \beta) \Big|_{t=0}.
 \end{aligned}$$

For the Rényi entropy, after some algebraic developments, we obtain

$$I_R(\gamma) = \frac{\gamma}{1-\gamma} \log \left[\frac{\alpha\beta}{-\log(1-\beta)} \right] + \frac{1}{1-\gamma} \log [B(\alpha, \beta, \gamma)],$$

where

$$B(\alpha, \beta, \gamma) = \int_0^1 \frac{g^{\gamma-1}(G^{-1}(u)) u^{\gamma(\alpha-1)} (1-u)^{\gamma(\alpha-1)}}{[u^\alpha + (1-u)^\alpha]^{2\gamma} \left[1 - \frac{\beta u^\alpha}{u^\alpha + (1-u)^\alpha} \right]^\gamma} du.$$

By using the generalized binomial expansion, we have

$$B(\alpha, \beta, \gamma) = \sum_{i,j=0}^{\infty} (-1)^{i+j} \beta^i \binom{-\gamma}{i} \binom{\gamma(\alpha-1)}{j} \int_0^1 \frac{g^{\gamma-1}(G^{-1}(u)) u^{\gamma(\alpha-1)+\alpha i+j}}{[u^\alpha + (1-u)^\alpha]^{2\gamma+i}} du.$$

Further, we can write from Eq. (28)

$$\frac{u^{-\gamma}}{[u^\alpha + (1-u)^\alpha]^{2\gamma+i}} = \sum_{k=0}^{\infty} s_k^*(\alpha, 2\gamma+i, \gamma) u^k,$$

where $s_k^*(\alpha, 2\gamma+i, \gamma)$ is defined there. Finally,

$$B(\alpha, \beta, \gamma) = \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j} \beta^i}{k+1+\alpha(\gamma+i)+j} \binom{-\gamma}{i} \binom{\gamma(\alpha-1)}{j} s_k^*(\alpha, 2\gamma+i, \gamma) E_{Y_k} \{ g^{\gamma-1}(G^{-1}(Y)) \},$$

where $Y_k \sim \text{Beta}(k+1+\alpha(\gamma+i)+j, 1)$. Figure 5 displays plots of the Rényi entropy versus γ for $a = 3.5, b = 1$ and selected values of α and β .

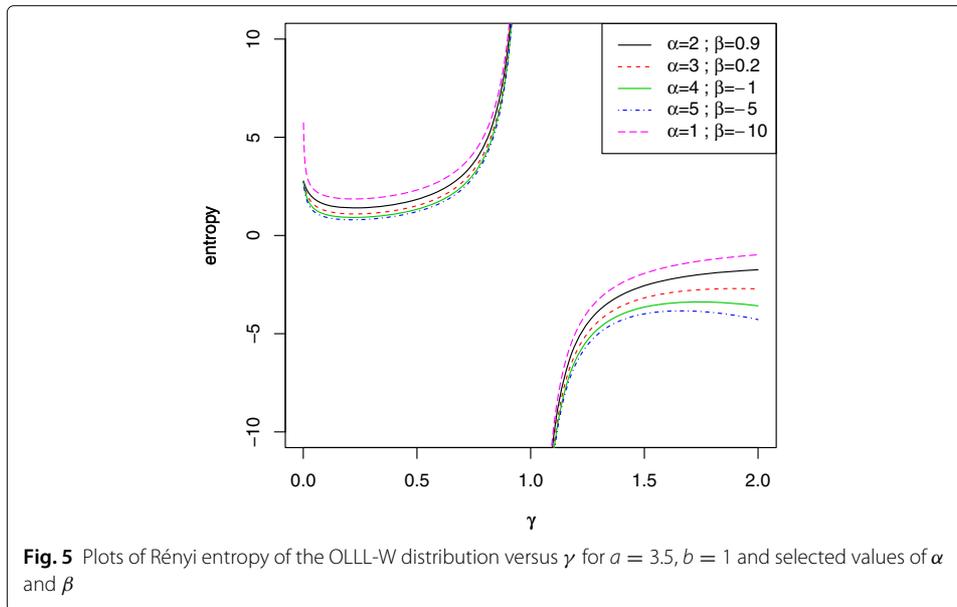
5.2 Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose that X_1, \dots, X_n is a random sample from the OLLL-G distribution. Let $X_{i:n}$ denote the i th order statistic. From Eqs. (6) and (7), the pdf of $X_{i:n}$ can be written as

$$f_{i:n}(x) = K \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \left[\sum_{r=0}^{\infty} d_{r+1} (r+1) G(x)^r g(x) \right] \left[\sum_{k=0}^{\infty} d_k G(x)^k \right]^{j+i-1},$$

where $K = n! / [(i-1)! (n-i)!]$. By using a result of Gradshteyn and Ryzhik (2000, Section 0.314) for a power series raised to a positive integer number, we obtain

$$\left[\sum_{k=0}^{\infty} d_k G(x)^k \right]^{j+i-1} = \sum_{k=0}^{\infty} e_{j+i-1,k} G(x)^k,$$



where $e_{j+i-1,0} = d_0^{j+i-1}$ and, for $k \geq 1$,

$$e_{j+i-1,k} = (k d_0)^{-1} \sum_{q=1}^k [q(j+i) - k] d_q e_{j+i-1,k-q}.$$

Setting $d_r^* = (r + 1)d_{r+1}$ and using a result of Gradshteyn and Ryzhik (2000, Section 0.316) for multiplying two power series, we have

$$\begin{aligned} f_{i:n}(x) &= K \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} g(x) \left[\sum_{r=0}^{\infty} d_r^* G(x)^r \right] \left[\sum_{k=0}^{\infty} e_{j+i-1,k} G(x)^k \right] \\ &= K \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} g(x) \sum_{k=0}^{\infty} e_k^* G(x)^k, \end{aligned}$$

where $e_k^* = \sum_{q=0}^k e_{j+i-1,q} d_{k-q}^*$. Hence, we can write

$$f_{i:n}(x) = \sum_{k=0}^{\infty} s_k h_{k+1}(x), \tag{11}$$

where (for $k \geq 0$)

$$s_k = \frac{k}{k+1} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} e_k^*.$$

Equation (11) is the main result of this section. It reveals that the pdf of the OLLL-G order statistics is a linear combination of Exp-G densities. So, several mathematical quantities of the OLLL-G order statistics such as ordinary, incomplete and factorial moments, mgf, mean deviations, among others, can be obtained from those quantities of the Exp-G distribution.

6 Estimation

In this section, we determine the MLEs of the model parameters of the new family from complete samples. Let x_1, \dots, x_n be the observed values from the OLLL-G distribution

with parameters α, β and ξ . Let $\theta = (\alpha, \beta, \xi)^\top$ be the parameter vector. The total log-likelihood function for θ is given by

$$\begin{aligned} \ell_n = \ell_n(\theta) = n \log \left[\frac{\alpha\beta}{-\log(1-\beta)} \right] + \sum_{i=1}^n \log[g(x_i; \xi)] + (\alpha - 1) \sum_{i=1}^n \log[G(x_i; \xi)\bar{G}(x_i; \xi)] \\ - \sum_{i=1}^n \log [G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha] - \sum_{i=1}^n \log [(1 - \beta)G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha]. \end{aligned} \quad (12)$$

The log-likelihood function can be maximized either directly or by solving the nonlinear likelihood equations obtained by differentiating (12). We use the goodness.fit function in R (R Development Core Team 2013) and NLMixed procedure in SAS to obtain the MLEs. The components of the score function $U_n(\theta) = (\partial \ell_n / \partial \alpha, \partial \ell_n / \partial \beta, \partial \ell_n / \partial \xi)^\top$ are

$$\begin{aligned} \frac{\partial \ell_n}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \log [G(x_i; \xi)\bar{G}(x_i; \xi)] \\ - \sum_{i=1}^n \frac{G(x_i; \xi)^\alpha \log [G(x_i; \xi)] + \bar{G}(x_i; \xi)^\alpha \log [\bar{G}(x_i; \xi)]}{G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha} \\ - \sum_{i=1}^n \frac{(1 - \beta)G(x_i; \xi)^\alpha \log [G(x_i; \xi)] + \bar{G}(x_i; \xi)^\alpha \log [\bar{G}(x_i; \xi)]}{(1 - \beta)G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha}, \end{aligned}$$

$$\frac{\partial \ell_n}{\partial \beta} = \frac{n}{\beta} + \frac{n}{(1 - \beta) \log(1 - \beta)} + \sum_{i=1}^n \frac{G(x_i; \xi)^\alpha}{(1 - \beta)G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha}$$

and

$$\begin{aligned} \frac{\partial \ell_n}{\partial \xi} = \sum_{i=1}^n \frac{g^{(\xi)}(x_i, \xi)}{g(x_i, \xi)} + (\alpha - 1) \sum_{i=1}^n \frac{G^{(\xi)}(x_i, \xi)}{G(x_i, \xi)} + (1 - \alpha) \sum_{i=1}^n \frac{G^{(\xi)}(x_i, \xi)}{\bar{G}(x_i, \xi)} \\ - \alpha \sum_{i=1}^n \frac{G^{(\xi)}(x_i, \xi) [G(x_i; \xi)^{\alpha-1} - \bar{G}(x_i; \xi)^{\alpha-1}]}{G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha} \\ - \alpha \sum_{i=1}^n \frac{G^{(\xi)}(x_i, \xi) [(1 - \beta)G(x_i; \xi)^{\alpha-1} - \bar{G}(x_i; \xi)^{\alpha-1}]}{(1 - \beta)G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha}, \end{aligned}$$

where $h^{(\xi)}(\cdot)$ means the derivative of the function h with respect to ξ .

6.1 Simulation study

We simulate the OLLL-SN distribution (with $\mu = 0, \sigma = 1, \alpha = 0.2, 0.5, \beta = -0.5, 0.7$) from Eq. (5) by using a random variable U having a uniform distribution in $(0, 1)$. We simulate $n = 50, 150$ and 300 variates and, for each replication, we evaluate the MLEs $\hat{\mu}, \hat{\sigma}, \hat{\alpha}$ and $\hat{\beta}$. We repeat this process 2000 times and determine the average estimates (AEs), biases and means squared errors (MSEs). The results are reported in Table 2.

The figures in Table 2 indicate that the MSEs and the AEs of the estimates of μ, σ, α and β decay toward zero when the sample size increases, as expected under first-order asymptotic theory. As n increases, the AEs of the parameters tend to be closer to the true parameter values. This fact supports that the asymptotic normal distribution provides an adequate approximation to the finite sample distribution of the MLEs.

Table 2 The AEs, biases and MSEs of the OLLL distribution for $\mu=0, \sigma=1, \alpha = 0.2, 0.5$ and $\beta = -0.5, 0.7$ varying n

n	Parameter	$\alpha = 0.2 \beta = -0.5$			Parameter	$\alpha = 0.5 \beta = -0.5$		
		AE	Bias	MSE		AE	Bias	MSE
50	μ	-0.0499	-0.0499	0.3033	μ	-0.0602	-0.0602	0.2529
	σ	0.9025	-0.0975	0.2581	σ	1.1766	0.1766	42.5479
	α	0.1926	-0.0074	0.0358	α	0.6543	0.1543	20.1315
	β	-0.7939	-0.2939	1.6520	β	-1.0228	-0.5228	3.4332
150	μ	-0.0518	-0.0518	0.2303	μ	-0.0003	-0.0003	0.1117
	σ	0.8911	-0.1089	0.1107	σ	0.9765	-0.0235	0.1005
	α	0.1771	-0.0229	0.0137	α	0.4934	-0.0066	0.0487
	β	-0.5812	-0.0812	0.4428	β	-0.9119	-0.4119	2.1433
300	μ	-0.0395	-0.0395	0.1105	μ	0.0037	0.0037	0.0551
	σ	0.8787	-0.1213	0.0772	σ	0.9877	-0.0123	0.0452
	α	0.1737	-0.0263	0.0099	α	0.4973	-0.0027	0.0211
	β	-0.5532	-0.0532	0.1873	β	-0.7216	-0.2216	0.9713
n	Parameter	$\alpha = 0.2 \beta = 0.7$			Parameter	$\alpha = 0.5 \beta = 0.7$		
		AE	Bias	MSE		AE	Bias	MSE
50	μ	0.0436	0.0436	0.3052	μ	0.0792	0.0792	0.2650
	σ	0.9008	-0.0992	0.1778	σ	1.1175	0.1175	83.1835
	α	0.1933	-0.0067	0.0251	α	0.6144	0.1144	42.5964
	β	0.5943	-0.1057	0.1579	β	0.4014	-0.2986	0.8603
150	μ	-0.0064	-0.0064	0.2759	μ	0.0152	0.0152	0.1170
	σ	0.8930	-0.1070	0.1148	σ	0.9899	-0.0101	0.0938
	α	0.1864	-0.0136	0.0151	α	0.5029	0.0029	0.0474
	β	0.6668	-0.0332	0.0500	β	0.6043	-0.0957	0.1170
300	μ	0.0624	0.0624	0.1436	μ	-0.0038	-0.0038	0.0616
	σ	0.8393	-0.1607	0.0897	σ	0.9916	-0.0084	0.0460
	α	0.1713	-0.0287	0.0117	α	0.5012	0.0012	0.0227
	β	0.6797	-0.0203	0.0310	β	0.6603	-0.0397	0.0356

7 The OLLL-G family with long-term survival

Models for survival analysis typically consider that every subject in the population under study is susceptible to the event of interest and will eventually experience such event if follow-up is sufficiently long. However, there are situations when a fraction of individuals are not expected to experience the event of interest, that is, those individuals are cured or not susceptible. Cure rate models for survival data have been used to model time-to-event data for various types of cancers, including breast cancer, non-Hodgkins lymphoma, leukemia, prostate cancer and melanoma. These models have become very popular due to significant progress in treatment therapies leading to enhanced cure rates.

Models to accommodate a cured fraction have been widely developed. Perhaps the most popular type of cure rate models are the mixture models (MMs) developed by Boag (1949), Berkson and Gage (1952) and Farewell (1982). Some proposals have been made recently in the literature by more long term survival to model lifetimes with covariates. For example, Ortega et al. (2012) considered the problem of assessing local influence in the negative binomial beta Weibull regression model to predict the cure of prostate

cancer, Hashimoto et al. (2013) derived curvature quantities under various perturbation schemes in Neyman type A beta-Weibull model for long-term survivors, Fachini et al. (2014) adapted local influence methods to a bivariate regression model with cure fraction and, recently, Ortega et al. (2015) used local influence methods to the power series beta-Weibull regression model for predicting breast carcinoma. The MMs allow simultaneously estimating whether the event of interest will occur, which is called incidence, and when it will occur, given that it can occur, which is called latency. Let N_i (for $i = 1, \dots, n$) be the indicator denoting that the i th individual is susceptible ($N_i = 1$) or non-susceptible ($N_i = 0$), i.e., the population is divided into two sub-populations so that an individual either is cured with probability $0 < p < 1$, or has a proper survival function $S(x)$ with probability $(1 - p)$. In this work, we do not consider the regression structure, although future research using covariates on the probability p may be investigated. The MM can be represented by

$$S_{pop}(x_i) = p + (1 - p)S(x_i|N_i = 1), \tag{13}$$

where $S_{pop}(x_i)$ is the unconditional survival function of x_i for the entire population, $S(x_i|N_i = 1) = 1 - F(x_i|N_i = 1)$ is the survival function for the susceptible individuals and $p = P(N_i = 0)$ is the probability of cure of an individual. The pdf corresponding to (13) is given by

$$f_{pop}(x_i) = -\frac{dS_{pop}(x_i)}{dt} = (1 - p)f(x_i|N_i = 1), \tag{14}$$

where $f(x_i|N_i = 1)$ is the baseline pdf (see Section 2.1) for the susceptible individuals. Equations (13) and (14) are improper functions, since $S_{pop}(x)$ is not a proper survival function. We can omit sometimes the dependence on the indicator N_i and write simply $S(x_i|N_i = 1) = S(x), f(x_i|N_i = 1) = f(x)$, etc.

Inserting (3) in (14) and (2) in (13), the pdf and survival function of the OLLG cure rate family are given, respectively, by

$$f_{pop}(x; \alpha, \beta, \xi, p) = \frac{(1 - p)\alpha\beta g(x, \xi)G(x, \xi)^{\alpha-1}\bar{G}(x, \xi)^{\alpha-1}}{-[G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha] [(1 - \beta)G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha] \log(1 - \beta)} \tag{15}$$

and

$$S_{pop}(x_i) = p + \frac{(1 - p)}{\log(1 - \beta)} \left[\log(1 - \beta) - \log \left(1 - \frac{\beta G(x, \xi)^\alpha}{G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha} \right) \right]. \tag{16}$$

A random variable having density (15) is denoted by $X \sim \text{OLLGcr}(\alpha, \beta, \xi, p)$. The hrf of the OLLGcr model is given by $h_{pop}(t) = f_{pop}(t)/S_{pop}(t)$.

7.1 Estimation

We consider the situation when the time-to-event is not completely observed and is subject to right censoring. Let c_i denote the censoring time. We observe $x_i = \min\{x_i, c_i\}$ and $\delta_i = I(x_i \leq c_i)$, where $\delta_i = 1$ if x_i is a time-to-event and $\delta_i = 0$ if x_i is right censored (for $i = 1, \dots, n$). From n pairs of times and censoring indicators $(x_1, \delta_1), \dots, (x_n, \delta_n)$, the log-likelihood function under non-informative censoring is given by $\ell_n(\theta) = \sum_{i \in F} \log f_{pop}(x_i; \theta) + \sum_{i \in C} \log S_{pop}(x_i; \theta)$, where $\theta = (\alpha, \beta, \xi, p)^T$ denotes the parameter vector and F and C denote the uncensored and censored sets of observations,

respectively. Replacing $f_{pop}(x_i; \theta)$ and $S_{pop}(x; \theta)$ by (15) and (16), respectively, the log-likelihood reduces to

$$\begin{aligned} \ell_n(\theta) = & r \log \left[\frac{(1-p)\alpha\beta}{-\log(1-\beta)} \right] + \sum_{i \in F} \log [g(x_i; \xi)] + (\alpha - 1) \sum_{i \in F} \log [G(x_i; \xi)\bar{G}(x_i; \xi)] \\ & - \sum_{i \in F} \log [G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha] - \sum_{i \in F} \log [(1-\beta)G(x_i; \xi)^\alpha + \bar{G}(x_i; \xi)^\alpha] \\ & + \sum_{i \in C} \log \left\{ p + \frac{(1-p)}{\log(1-\beta)} \left[\log(1-\beta) - \log \left(1 - \frac{\beta G(x, \xi)^\alpha}{G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha} \right) \right] \right\}, \end{aligned} \quad (17)$$

where r is the number of failures (uncensored observations). We can obtain the MLE $\hat{\theta}$ of θ by maximizing the log-likelihood (17) either directly in R using the `optim` function, in SAS using the `NLMixed` procedure and in other statistical software or by solving the nonlinear likelihood equations obtained by differentiating (17).

8 Applications

In this section, we provide four applications to real data. In the first three applications, we present some results by fitting special models defined in Section 2.1. In the fourth application, we present an application using the long-term survival model defined in Section 7.

For the first three applications, the goodness-of-fit statistics including the Cramér-von Mises (W^*) and Anderson-Darling (A^*) test statistics are used to compare the fitted models; see Chen and Balakrishnan (1995) for more details. The smaller the values of A^* and W^* , the better the fit to the data. We also consider the Kolmogorov-Smirnov (K-S) statistic (and its corresponding p -value) and minus the maximized log-likelihood ($\hat{\ell}_n$) for the sake of comparison. For the fourth application (censored data), we adopt the AIC and BIC statistics to compare the fitted models since the A^* and W^* statistics are not suitable for censored data.

For the next three applications, we consider the OLLLN distribution and, for the purpose of comparison, we fit the following models to the data sets described below:

- The normal distribution.
- The exponentiated normal (EN) distribution.
- The logarithmic normal (LN) distribution, the special case of the OLLLN distribution when $\alpha = 1$.
- The beta normal (BN) distribution (Eugene et al. 2002) with density

$$f_{BN}(x) = \frac{1}{\sigma B(\alpha, \beta)} \left[\Phi \left(\frac{x - \mu}{\sigma} \right) \right]^{\alpha-1} \left[1 - \Phi \left(\frac{x - \mu}{\sigma} \right) \right]^{\beta-1} \phi \left(\frac{x - \mu}{\sigma} \right).$$

- The gamma normal (GN) distribution (Alzaatreh et al. 2014) with density

$$f_{GN}(x) = \frac{\beta^\alpha}{\sigma \Gamma(\alpha)} \left[-\log \left\{ 1 - \Phi \left(\frac{x - \mu}{\sigma} \right) \right\} \right]^{\alpha-1} \left[1 - \Phi \left(\frac{x - \mu}{\sigma} \right) \right]^{\beta-1} \phi \left(\frac{x - \mu}{\sigma} \right).$$

- The Kumaraswamy normal (KN) distribution (Cordeiro and de Castro 2011) with density

$$f_{KN}(x) = \frac{\alpha\beta}{\sigma} \left\{ \Phi \left[\left(\frac{x - \mu}{\sigma} \right) \right] \right\}^{\alpha-1} \left\{ 1 - \left[\Phi \left(\frac{x - \mu}{\sigma} \right) \right]^\alpha \right\}^{\beta-1} \phi \left(\frac{x - \mu}{\sigma} \right).$$

- The odd log-logistic normal (OLL-N) distribution (special case of OLLLN distribution when $\beta \rightarrow 1$) with density (Braga et al. 2016)

$$f_{OLL-N}(x) = \frac{\alpha \phi\left(\frac{x-\mu}{\sigma}\right) [\Phi\left(\frac{x-\mu}{\sigma}\right)]^{\alpha-1} [1 - \Phi\left(\frac{x-\mu}{\sigma}\right)]^{\alpha-1}}{\sigma \{ [1 - \Phi\left(\frac{x-\mu}{\sigma}\right)]^{\alpha} + [\Phi\left(\frac{x-\mu}{\sigma}\right)]^{\alpha} \}^2},$$

where $x \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0, \beta > 0$ and $\sigma > 0$.

8.1 Application 1

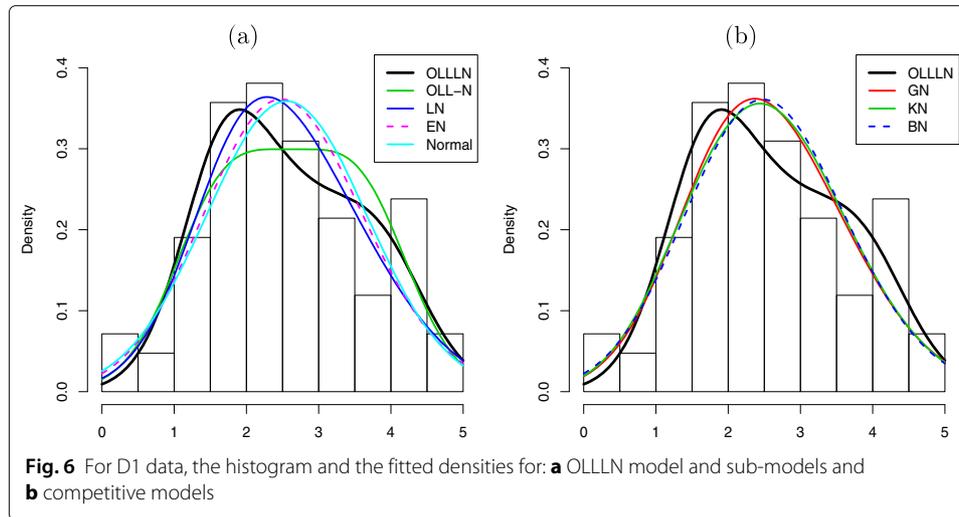
First, we consider the data set representing the failure times of a particular windshield device. These data were also studied by Blischke and Murthy (2000) and Murthy et al. (2004). The data, referred as D1, are: 0.040, 1.866, 2.385, 3.443, 0.301, 1.876, 2.481, 3.467, 0.309, 1.899, 2.610, 3.478, 0.557, 1.911, 2.625, 3.578, 0.943, 1.912, 2.632, 3.595, 1.070, 1.914, 2.646, 3.699, 1.124, 1.981, 2.661, 3.779, 1.248, 2.010, 2.688, 3.924, 1.281, 2.038, 2.823, 4.035, 1.281, 2.085, 2.890, 4.121, 1.303, 2.089, 2.902, 4.167, 1.432, 2.097, 2.934, 4.240, 1.480, 2.135, 2.962, 4.255, 1.505, 2.154, 2.964, 4.278, 1.506, 2.190, 3.000, 4.305, 1.568, 2.194, 3.103, 4.376, 1.615, 2.223, 3.114, 4.449, 1.619, 2.224, 3.117, 4.485, 1.652, 2.229, 3.166, 4.570, 1.652, 2.300, 3.344, 4.602, 1.757, 2.324, 3.376, 4.663.

The MLEs of the parameters and the standard errors (SEs) in parentheses and the goodness-of-fit statistics for D1 are listed in Table 3. We note that the OLLLN model outperforms all the fitted competitive models under these statistics.

The histogram of the data D1 and fitted densities are displayed in Fig. 6. We note that the fitted OLLLN distribution best captures the empirical histogram. Based on the equations given in Section 4, we give some measures based on the moments of the OLLLN distribution. The expected value and the variance of the failure times of windshield devices are: $E(X) = 2.58$ and $Var(X) = 1.24$., respectively. Also, the skewness

Table 3 The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D1

Model	μ	σ	α	β	$\hat{\ell}_n$	W^*	A^*	K-S	p-value
OLLLN	2.887 (0.212)	0.532 (0.180)	0.379 (0.192)	-1.639 (1.947)	126.178	0.035	0.331	0.053	0.963
Normal	2.557 (0.121)	1.112 (0.086)			128.119	0.092	0.608	0.092	0.445
EN	1.832 (2.366)	1.337 (0.708)	1.939 (3.878)		128.064	0.075	0.522	0.085	0.557
BN	0.808 (7.151)	2.443 (8.162)	7.113 (48.582)	2.469 (14.619)	128.085	0.074	0.520	0.084	0.562
GN	2.805 (1.059)	0.541 (0.265)	0.290 (0.382)	0.197 (0.216)	127.757	0.058	0.438	0.075	0.7105
KN	1.654 (1.067)	0.748 (0.539)	0.920 (1.017)	0.320 (0.524)	127.848	0.063	0.469	0.079	0.642
LN	3.172 (0.562)	1.079 (0.094)		-7.091 (16.134)	127.570	0.048	0.391	0.066	0.828
OLL-N	2.626 (0.126)	0.603 (0.218)	0.452 (0.232)		127.062	0.076	0.523	0.095	0.408



and kurtosis measures are given by $Skewness(X) = 0.23$ and $Kurtosis(X) = 2.42$, thus indicating that the tail on the right side is longer and then it is a platykurtic distribution.

8.2 Application 2

The second data set D2 consists of lifetimes of 43 blood cancer patients (in days) from one of the Health Hospitals in Saudi Arabia (Abouammoh and Abdulghani 1994). These data are: 115, 181, 255, 418, 441, 461, 516, 739, 743, 789, 807, 865, 924, 983, 1025, 1062, 1063, 1165, 1191, 1222, 1222, 1251, 1277, 1290, 1357, 1369, 1408, 1455, 1478, 1519, 1578, 1578, 1599, 1603, 1605, 1696, 1735, 1799, 1815, 1852, 1899, 1925, 1965.

The MLEs of the parameters and SEs (in parentheses) and the goodness-of-fit statistics for D2 are listed in Table 4. We conclude that the KN model has the smallest $\hat{\ell}_n$ and the OLLLN has the third smallest $\hat{\ell}_n$, the BN model has the smallest W^* and the OLLLN and KN models have the second smallest W^* s. The OLLLN and BN models have the smallest A^* s. The OLLLN and GN models have the smallest K-S statistics. The GN model has the largest p -value and the OLLLN and KN models have the second largest p -values. Therefore, we can conclude that the OLLLN model has either the best fit or is very close to the best fit with respect to the current criteria. The histogram of D2 and the fitted densities are displayed in Fig. 7.

8.3 Application 3

The third data set D3 includes the lower discharge of at least seven consecutive days and return period (time) of ten years of the Cuiabá River, Cuiabá, Mato Grosso, Brazil. These data have also been studied by Cordeiro et al. (2012). The MLEs of the parameters and SEs (in parentheses) and the goodness-of-fit statistics for D3 are listed in Table 5. We note that the OLLLN model outperforms all other fitted models.

The fitted densities for the models listed in Table 5 are displayed in Fig. 8. We verify that the fitted OLLLN distribution best captures the histogram of these data.

In summary, we conclude that the OLLLN distribution outperforms all the fitted competitive models under the selected criterion for D1, D2 and D3. For all three data sets, we verify that the fitted OLLLN distribution best captures the three histograms, especially for the third data set, which indicates the outstanding performance of this distribution.

Table 4 The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D2

Model	μ	σ	α	β	$\hat{\ell}_n$	W^*	A^*	K-S	p -value
OLLLLN	930.893 (151.780)	215.288 (93.152)	0.323 (0.224)	0.854 (0.202)	325.684	0.018	0.157	0.057	0.997
Normal	1191.628 (76.357)	500.709 (53.993)			328.303	0.068	0.489	0.083	0.901
EN	2083.910 (81.898)	52.5006 (14.640)	0.005 (0.003)		325.847	0.020	0.174	0.076	0.9485
BN	2153.003 (313.548)	173.89 (512.195)	0.0496 (0.268)	4.139 (6.681)	325.680	0.017	0.157	0.059	0.996
GN	2138.76 (0.036)	251.154 (0.036)	0.101 (0.016)	2.607 (1.319)	326.237	0.020	0.183	0.057	0.998
KN	1825.282 (288.596)	71.445 (65.481)	0.0097 (0.019)	0.529 (0.319)	325.4645	0.018	0.162	0.058	0.997
LN	748.777 (275.414)	472.416 (71.311)		0.974 (0.070)	326.628	0.026	0.221	0.064	0.991
Oll-N	1131.115 (75.566)	223.172 (99.640)	0.338 (0.231)		327.292	0.058	0.415	0.125	0.473

8.4 Application 4: OLLL-G long-term survival models

These data consist of $n = 493$ lifetimes (t_i in months) of patients diagnosed with breast cancer. The steps to construct these data can be found in Gendoo et al. (2015). In many applications there is qualitative information about the hazard shape, which can help for selecting a particular polyhazard model. In this context, a device called the total time on test (TTT) plot is useful. The TTT plot is obtained by plotting $G(r/n) = [(\sum_{i=1}^r T_{i:n}) + (n - r)T_{r:n}] / (\sum_{i=1}^n T_{i:n})$, where $r = 1, \dots, n$, and $T_{i:n}$ (for $i = 1, \dots, n$) are the order statistics of the sample, against r/n . It is a straight diagonal for constant hazards leading to an exponential model. It is convex for decreasing hazards and concave for increasing hazards leading to a single-Weibull model. It is first convex and then concave if the hazard is bathtub-shaped leading to a bi-Weibull model. It is first concave and

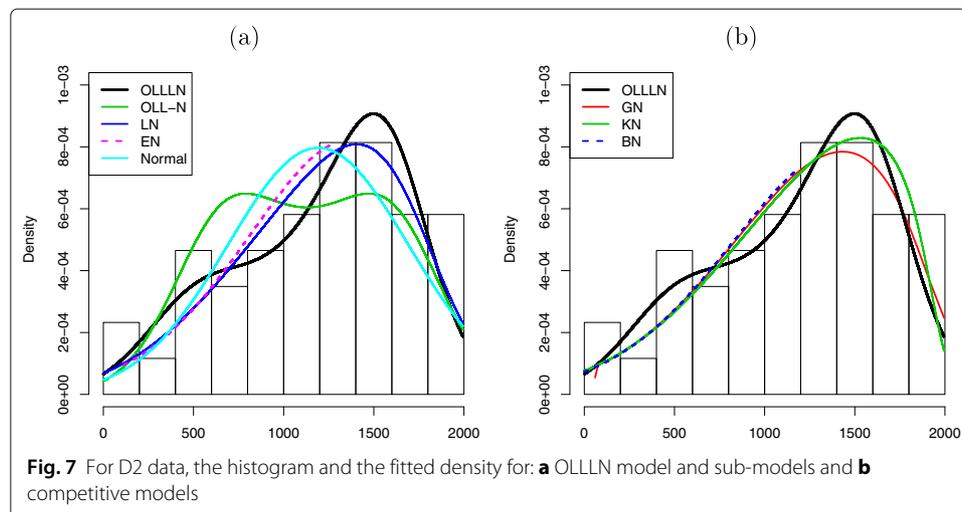


Fig. 7 For D2 data, the histogram and the fitted density for: **a** OLLLLN model and sub-models and **b** competitive models

Table 5 The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D3

Model	μ	σ	α	β	$\hat{\ell}_n$	W^*	A^*	K-S	p -value
OLLLN	102.050 (5.611)	7.879 (5.275)	0.100 (0.119)	0.521 (0.381)	188.193	0.024	0.206	0.076	0.9705
Normal	110.214 (6.144)	37.873 (4.344)			192.021	0.109	0.671	0.144	0.378
EN	169.145 (26.432)	13.261 (13.067)	0.062 (0.142)		191.360	0.074	0.491	0.113	0.674
BN	172.325 (22.216)	18.326 (14.485)	0.109 (0.160)	1.785 (2.635)	191.273	0.068	0.465	0.105	0.759
GN	172.393 (20.263)	17.696 (15.038)	0.102 (0.159)	1.340 (2.564)	191.263	0.068	0.465	0.106	0.752
KN	129.188 (30.610)	13.163 (5.980)	0.082 (0.104)	0.361 (0.297)	191.071	0.065	0.450	0.111	0.695
LN	62.065 (23.996)	33.056 (6.044)		0.997 (0.011)	190.819	0.059	0.417	0.109	0.717
Oll-N	105.546 (4.163)	7.830 (4.921)	0.098 (0.110)		188.655	0.034	0.271	0.110	0.7025

then convex if the hazard is bimodal-shaped leading to a log-logistic model. For multimodal hazards, the TTT plot contains several concave and convex regions. The TTT plot in Fig. 10a indicates an increasing-decreasing-increasing hrf. So, the OLLW distribution would be a good option to model these data.

Next, we compare the results by fitting the OLLWcr model and some of its sub-models such as: the odd log-logistic Weibull cure rate (OLL-Wcr) model (OLLNcr distribution when $\beta \rightarrow 1$), the logarithmic Weibull cure rate (LWcr) model (OLLNcr distribution when $\alpha = 1$) and Weibull cure rate model (OLLNcr distribution when $\alpha = 1$ and $\beta \rightarrow 1$). Table 6 provides the MLEs (and the corresponding SEs in parentheses) of the model parameters and the values of the AIC and BIC statistics. The results indicate that the OLLWcr model has the lowest values of these statistics among those values of the

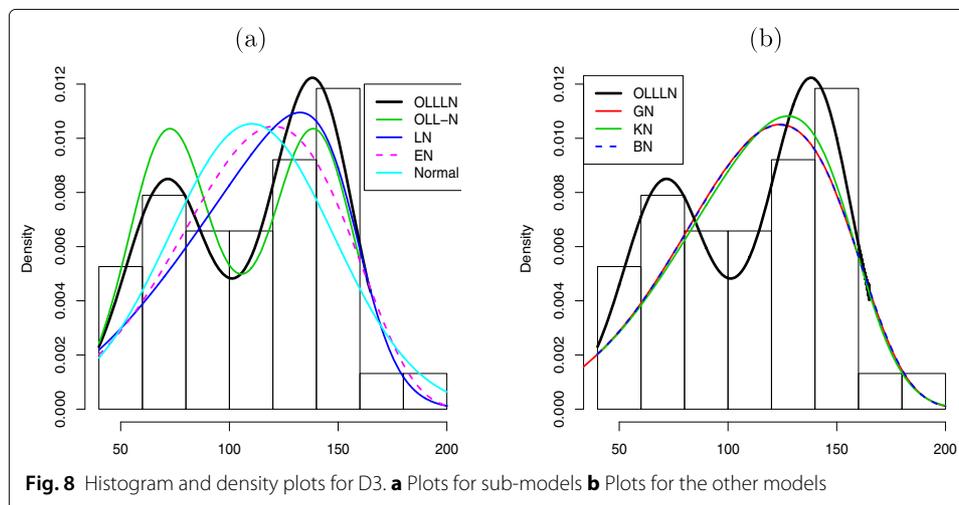


Table 6 MLEs, SEs in parentheses and goodness-of-fit statistics for breast cancer data

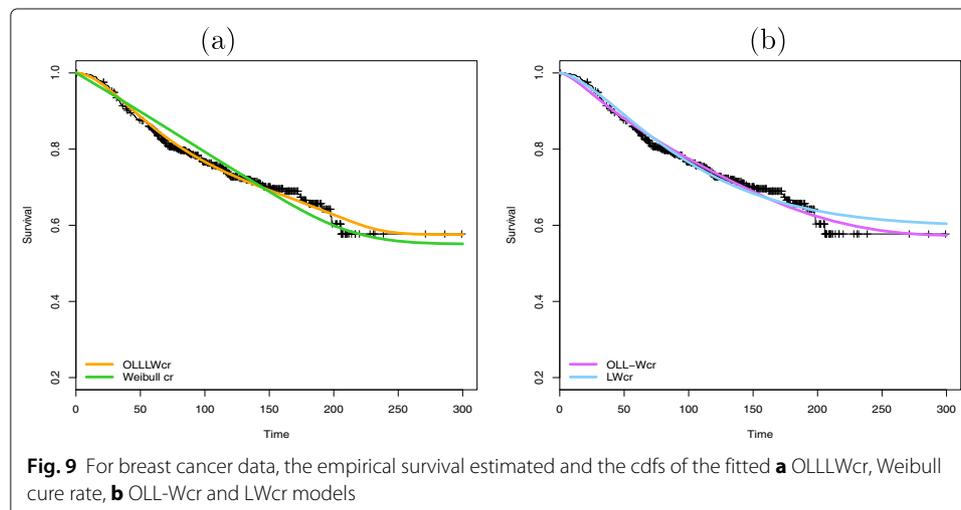
Model	μ	σ	α	β	ρ	AIC	BIC
OLLLWcr	5.933 (0.079)	0.006 (0.001)	0.317 (0.084)	-9.131 (4.319)	0.575 (0.061)	1878.15	1899.15
OLL-Wcr	0.778 (0.085)	0.036 (0.009)	1.788 (0.426)	- -	0.569 (0.059)	1886.37	1903.17
LWcr	1.581 (0.188)	0.007 (0.001)	- -	-1.334 (1.712)	0.597 (0.044)	1885.27	1902.07
Weibull cure rate	1.027 (0.082)	0.040 (0.008)	- -	- -	0.550 (0.064)	1900.03	1912.63

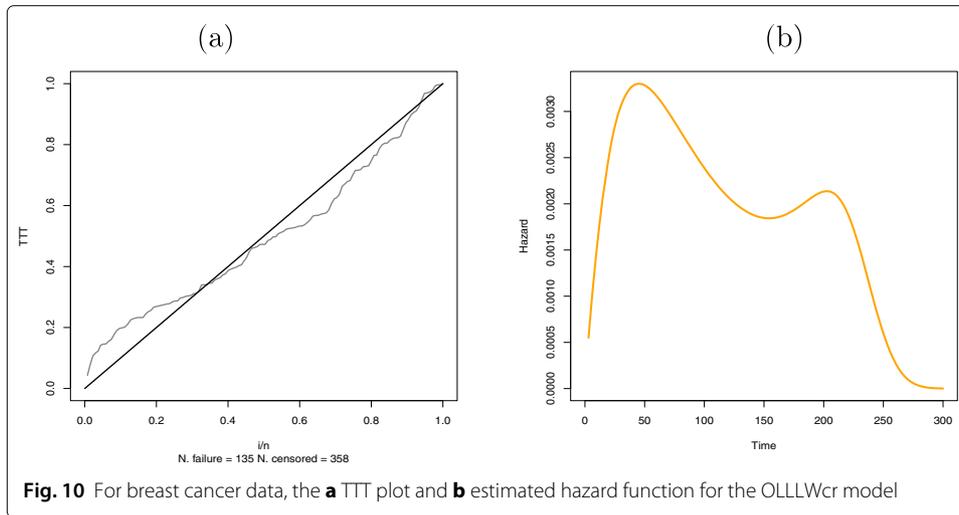
fitted models, and therefore it could be chosen as the best model. On the other hand, the proportion of cured individuals obtained by the Kaplan-Meier estimator is 0.577. Thus, we can conclude based on the figures in Table 6 that the OLLLNcr model gives a more accurate estimate for the proportion of cured individuals.

The adequacy of the fitted models can also be noted in Fig. 9, which presents the empirical and estimated survival functions. Based on these plots, we can conclude that the OLLLWcr model provides a good fit for the breast cancer data. In addition, the empirical scaled TTT transform can be used to identify the shape of the hazard function. The fitted hazard function for the OLLLWcr model is displayed in Fig. 10b, which we observe bimodal shapes, thus indicating a good fit.

9 Conclusions

We study some mathematical properties of the odd log-logistic logarithmic-G family of distributions with two extra shape parameters $\alpha > 0$ and $\beta \in (0, 1)$. We provide some special models, a very useful linear representation for the density function in terms of exponentiated densities, explicit expressions for the moments, generating function, entropies and order statistics. The model parameters are estimated by the method of maximum likelihood. We perform a simulation study to verify the adequacy of the estimators. We also introduce a long-term survival model based on the new family. The importance





of the proposed models is illustrated by means of four real life data sets. The new models provide consistently better fits than other competitive models for the current data.

Appendix 1: Asymptotes and shapes

Corollary 1 Let $a = \inf\{x|F(x) > 0\}$. The asymptotics of Eqs. (2), (3) and (4) when $x \rightarrow a$ are given by

$$\begin{aligned}
 F(x) &\sim \frac{\beta G(x)^\alpha}{-\log(1-\beta)} \text{ as } x \rightarrow a, \\
 f(x) &\sim \frac{\alpha\beta g(x) G(x)^{\alpha-1}}{-\log(1-\beta)} \text{ as } x \rightarrow a, \\
 h(x) &\sim \frac{\alpha\beta g(x) G(x)^{\alpha-1}}{-\log(1-\beta)} \text{ as } x \rightarrow a.
 \end{aligned}$$

Corollary 2 The asymptotics of Eqs. (2), (3) and (4) when $x \rightarrow \infty$ are given by

$$\begin{aligned}
 1 - F(x) &\sim 1 - \frac{\log[1 - \beta G(x)^\alpha]}{\log(1-\beta)} \text{ as } x \rightarrow \infty, \\
 f(x) &\sim \frac{\alpha\beta g(x) G(x)^{-1}}{-(1-\beta)\log(1-\beta)} \text{ as } x \rightarrow \infty, \\
 h(x) &\sim \frac{\alpha\beta g(x) G(x)^{-1}}{-(1-\beta)\log\left[\frac{1-\beta}{1-\beta G(x)^\alpha}\right]} \text{ as } x \rightarrow \infty.
 \end{aligned}$$

The shapes of the density and hazard rate functions can be described analytically. The critical points of the OLLL-G density function are the roots of the equation:

$$\begin{aligned}
 \frac{g'(x)}{g(x)} + (\alpha - 1) \frac{g(x)}{G(x)} + (1 - \alpha) \frac{g(x)}{\bar{G}(x)} - \alpha g(x) \frac{G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{G(x)^\alpha + \bar{G}(x)^\alpha} \\
 - \alpha g(x) \frac{(1-\beta)G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{(1-\beta)G(x)^\alpha + \bar{G}(x)^\alpha} = 0.
 \end{aligned} \tag{18}$$

There may be more than one root to (18). Let $\lambda(x) = \frac{d^2 \log[f(x)]}{dx^2}$. We have

$$\begin{aligned} \lambda(x) = & \frac{g''(x)g(x) - g'(x)^2}{g(x)^2} + (\alpha - 1) \frac{g'(x)G(x) - g(x)^2}{G(x)^2} + (1 - \alpha) \frac{g'(x)\bar{G}(x) + g(x)^2}{\bar{G}(x)^2} \\ & - \alpha g'(x) \frac{G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{G(x)^\alpha + \bar{G}(x)^\alpha} - \alpha(\alpha - 1) g(x)^2 \frac{G(x)^{\alpha-2} + \bar{G}(x)^{\alpha-2}}{G(x)^\alpha + \bar{G}(x)^\alpha} \\ & + \left\{ \alpha g(x) \frac{G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{G(x)^\alpha + \bar{G}(x)^\alpha} \right\}^2 - \alpha g'(x) \frac{(1 - \beta)G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{(1 - \beta)G(x)^\alpha + \bar{G}(x)^\alpha} \\ & - \alpha(\alpha - 1) g(x)^2 \frac{(1 - \beta)G(x)^{\alpha-2} + \bar{G}(x)^{\alpha-2}}{(1 - \beta)G(x)^\alpha + \bar{G}(x)^\alpha} \\ & + \left\{ \alpha g(x) \frac{(1 - \beta)G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{(1 - \beta)G(x)^\alpha + \bar{G}(x)^\alpha} \right\}^2. \end{aligned}$$

If $x = x_0$ is a root of (18) then it corresponds to a local maximum (minimum) if $\lambda(x_0) < 0$ ($\lambda(x_0) > 0$). It refers to a point of inflexion if $\lambda(x_0) = 0$.

The critical point of $h(x)$ are obtained from the equation

$$\begin{aligned} & \frac{g'(x)}{g(x)} + (\alpha - 1) \frac{g(x)}{G(x)} + (1 - \alpha) \frac{g(x)}{\bar{G}(x)} - \alpha g(x) \frac{G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{G(x)^\alpha + \bar{G}(x)^\alpha} - \\ & \alpha g(x) \frac{(1 - \beta)G(x)^{\alpha-1} - \bar{G}(x)^{\alpha-1}}{(1 - \beta)G(x)^\alpha + \bar{G}(x)^\alpha} - \\ & \frac{\alpha \beta g(x) G(x)^{\alpha-1} \bar{G}(x)^{\alpha-1}}{[G(x)^\alpha + \bar{G}(x)^\alpha] [(1 - \beta)G(x)^\alpha + \bar{G}(x)^\alpha] \left\{ \log \left[\frac{1 - \beta}{1 - \frac{\beta G(x)^\alpha}{G(x)^\alpha + \bar{G}(x)^\alpha}} \right] \right\}} = 0. \end{aligned} \tag{19}$$

There may be more than one root to (19).

Appendix 2: Useful power series

The power series derived in this appendix are required for the proofs of the linear representations in Section 3. All power series given below are convergent for $|u| \leq 1$. In Sections 3 and 5.1, they can be applied for the support of X since the quantity $\beta G(x, \xi)^\alpha / [G(x, \xi)^\alpha + \bar{G}(x, \xi)^\alpha]$ does belong to the interval $(0, 1)$ when $\beta \in (0, 1)$.

First, for $a > 0$, we have the generalized binomial expansion

$$(1 - u)^a = \sum_{j=0}^{\infty} (-1)^j \binom{a}{j} u^j, \tag{20}$$

which holds for $|u| \leq 1$.

Second, we obtain an expression for $\left[\frac{u^\alpha}{u^\alpha + (1-u)^\alpha} \right]^m$, where $\alpha > 0$ is a real number, m is a natural number and $|u| \leq 1$. We can write

$$u^\alpha = \sum_{k=0}^{\infty} a_k(\alpha) u^k, \tag{21}$$

where

$$a_k(\alpha) = \sum_{j=k}^{\infty} (-1)^{k+j} \binom{\alpha}{j} \binom{j}{k}.$$

The power series (20) and (21) and the others derived from them converge everywhere. For any real $\alpha > 0$, the power series follows from (20) and (21)

$$u^\alpha + (1 - u)^\alpha = \sum_{k=0}^\infty b_k(\alpha) u^k, \tag{22}$$

where $b_k(\alpha) = a_k(\alpha) + (-1)^k \binom{\alpha}{k}$. Combining (21) and (22), we have (see Gradshteyn and Ryzhik 2000, Section 0.313)

$$\frac{u^\alpha}{u^\alpha + (1 - u)^\alpha} = \frac{\sum_{k=0}^\infty a_k(\alpha) u^k}{\sum_{k=0}^\infty b_k(\alpha) u^k} = \sum_{k=0}^\infty c_k(\alpha) u^k, \tag{23}$$

where $c_0(\alpha) = a_0(\alpha)/b_0(\alpha)$ and the $c_k(\alpha)$'s (for $k \geq 1$) are determined from the recurrence equation

$$c_k(\alpha) = \frac{1}{b_0(\alpha)} \left[a_k(\alpha) - \frac{1}{b_0(\alpha)} \sum_{r=1}^k b_r c_{k-r}(\alpha) \right].$$

Third, based on (23) and using a result of Gradshteyn and Ryzhik (2000 Section 0.314) for a power series raised to a positive integer number, we obtain

$$\left[\frac{u^\alpha}{u^\alpha + (1 - u)^\alpha} \right]^m = \left(\sum_{k=0}^\infty c_k(\alpha) u^k \right)^m = \sum_{k=0}^\infty h_k^*(\alpha, m) u^k, \tag{24}$$

where $h_0^*(\alpha, m) = c_0(\alpha)^m$ and, for $k \geq 1$,

$$h_k^*(\alpha, m) = \frac{1}{kc_0(\alpha)} \sum_{q=1}^k [(m + 1)q - k] c_q(\alpha) h_{k-q}^*(\alpha, m).$$

Fourth, we obtain a power series for $\frac{u^{-\gamma}}{[u^\alpha + (1 - u)^\alpha]^w}$, which is applied in Section 5.1, where w and α are positive real numbers, $\gamma > 0$ and $0 < u < 1$. We have

$$\frac{u^{-\gamma}}{[u^\alpha + (1 - u)^\alpha]^w} = u^{-\alpha w - \gamma} \left[\frac{u^\alpha}{u^\alpha + (1 - u)^\alpha} \right]^w. \tag{25}$$

The second term in (25) can be expanded as

$$\left[\frac{u^\alpha}{u^\alpha + (1 - u)^\alpha} \right]^w = \sum_{i=0}^\infty a_i(w) \left[\frac{u^\alpha}{u^\alpha + (1 - u)^\alpha} \right]^i,$$

where $a_i(w) = \sum_{j=i}^\infty (-1)^{i+j} \binom{w}{j} \binom{j}{i}$. Now, from (24), we have

$$\left[\frac{u^\alpha}{u^\alpha + (1 - u)^\alpha} \right]^w = \sum_{k=0}^\infty s_{1,k}^*(\alpha, w) u^k, \tag{26}$$

where

$$s_{1,k}^*(\alpha, w) = \sum_{i=0}^\infty a_i(w) h_k^*(\alpha, i).$$

The first term $u^{-\alpha w - \gamma}$ in (25) can be expanded as

$$u^{-\alpha w - \gamma} = \sum_{k=0}^\infty s_{2,k}^*(\alpha, w, \gamma) u^k, \tag{27}$$

where $s_{2,k}^*(\alpha, w, \gamma) = \sum_{i=k}^{\infty} (-1)^{i+k} \binom{-\alpha w - \gamma}{i} \binom{i}{k}$. Finally, from (25), (26), (27) and using a result of Gradshteyn and Ryzhik (2000 Section 0.316) for multiplication of two power series, we obtain

$$\frac{u^{-\gamma}}{[u^\alpha + (1-u)^\alpha]^w} = \left(\sum_{k=0}^{\infty} s_{1,k}^*(\alpha, w) u^k \right) \left(\sum_{k=0}^{\infty} s_{2,k}^*(\alpha, w, \gamma) u^k \right) = \sum_{k=0}^{\infty} s_k^*(\alpha, w, \gamma) u^k, \quad (28)$$

where $s_k^*(\alpha, w, \gamma) = \sum_{j=0}^k s_{1,j}^*(\alpha, w) s_{2,k-j}^*(\alpha, w, \gamma)$. Equation (28) is the main result to obtain the Rényi entropy in Section 5.1.

Appendix 3: Properties for a special model

The OLLL-E distribution is defined by inserting $G(x) = 1 - e^{-\lambda x}$ and $g(x) = \lambda e^{-\lambda x}$ in Eq. (3), where $x > 0$ and $\lambda > 0$. Let X be the random variable representing this distribution. We derive some statistical measures of X from the asymptotics in Appendix 1 and the general results in Sections 4 and 5.2

Corollary 3 *The asymptotics of Eqs. (2), (3) and (4) for the OLLL-E distribution when $x \rightarrow 0$ are given by*

$$\begin{aligned} F(x) &\sim \frac{\beta(\lambda x)^\alpha}{-\log(1 - \beta)} \quad \text{as } x \rightarrow 0, \\ f(x) &\sim \frac{\alpha\beta \lambda^\alpha x^{\alpha-1}}{-\log(1 - \beta)} \quad \text{as } x \rightarrow 0, \\ h(x) &\sim \frac{\alpha\beta \lambda^\alpha x^{\alpha-1}}{-\log(1 - \beta)} \quad \text{as } x \rightarrow 0. \end{aligned}$$

Corollary 4 *The asymptotics of Eqs. (2), (3) and (4) for the OLLL-E distribution when $x \rightarrow \infty$ are given by*

$$\begin{aligned} 1 - F(x) &\sim 1 - \frac{\log[1 - \beta(1 - e^{-\lambda x})^\alpha]}{\log(1 - \beta)} \quad \text{as } x \rightarrow \infty, \\ f(x) &\sim \frac{\alpha\beta \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{-1}}{-(1 - \beta) \log(1 - \beta)} \quad \text{as } x \rightarrow \infty, \\ h(x) &\sim \frac{\alpha\beta \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{-1}}{-(1 - \beta) \log\left[\frac{1 - \beta}{1 - \beta G(x)^\alpha}\right]} \quad \text{as } x \rightarrow \infty. \end{aligned}$$

These equations can provide the effects of the parameters on the tails of the OLLL-E distribution.

We provide some statistical measures of X . Its n th ordinary moment follows from (8) as

$$E(X^n) = \lambda n! \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{(-1)^m (k+1) d_{k+1}}{[(m+1)\lambda]^{n+1}} \binom{k}{m},$$

where d_k is defined in Eq. (6).

Further, the n th incomplete moment of the OLLL-E distribution is obtained from (9) as

$$m_n(y) = \lambda \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{(-1)^m (k+1) d_{k+1} \gamma(n+1, y)}{[(m+1)\lambda]^{n+1}} \binom{k}{m},$$

where $\gamma(a, z) = \int_0^z t^{a-1} e^{-t} dt$ denotes the incomplete gamma function.

The mgf of X comes immediately from (10)

$$M(t) = \lambda \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{(-1)^m (k+1) d_{k+1}}{(m+1)\lambda - t} \binom{k}{m}.$$

Finally, the r th ordinary moment of the i th OLLL-E order statistic reduces to

$$E(X_{i:n}^r) = \lambda r! \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{(-1)^m (k+1) s_k}{[(m+1)\lambda]^{r+1}} \binom{k}{m},$$

where s_k is given by (11).

Acknowledgements

The authors are very grateful to the editor Felix Famoye for helpful comments and suggestions.

Authors' contributions

The five authors jointly participated of the study and elaborated the research. Gauss Cordeiro corrected the manuscript throughout. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Statistics, Persian Gulf university, Bushehr, Iran. ²Department of Statistics, University of Mazandaran, Babolsar, Iran. ³Departamento de Ciências Exatas, Universidade de São Paulo, Piracicaba, Brazil. ⁴Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil.

Received: 26 January 2017 Accepted: 13 June 2017

Published online: 04 July 2017

References

- Abouammoh, AM, Abdulghani, SA: On partial orderings and testing of new better than renewal used classes. *Reliability Eng. Syst. Safety.* **43**, 37–41 (1994)
- Alzaatreh, A, Famoye, F, Lee, C: The gamma-normal distribution: Properties and applications. *Comput. Stat. Data Anal.* **69**, 67–80 (2014)
- Berkson, J, Gage, RP: Survival curve for cancer patients following treatment. *J. Am. Stat. Assoc.* **47**, 501–515 (1952)
- Blishcke, WR, Murthy, DNP: *Reliability: Modeling, Prediction and Optimization*. 1st ed. Wiley, New York (2000)
- Boag, JW: Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J. R. Stat. Soc. Series B.* **11**, 15–53 (1949)
- Bourguignon, M, Silva, RB, Cordeiro, GM: The Weibull-G family of probability distributions. *J. Data Sci.* **12**, 53–68 (2014)
- Braga, AS, Cordeiro, GM, Ortega, GMM, da Cruz, JN: The odd log-logistic normal distribution: Theory and applications in analysis of experiments. *J. Stat. Theory Prac* (2016). doi:10.1080/15598608.2016.1141127
- Chen, G, Balakrishnan, N: A general purpose approximate goodness-of-fit test. *J. Q. Technol.* **27**, 154–161 (1995)
- Cordeiro, GM, Alizadeh, M, Ortega, EMM: The exponentiated half-logistic family of distributions: Properties and applications. *J. Probab. Stat.* **1**, 1–21 (2014a)
- Cordeiro, GM, de Castro, M: A new family of generalized distributions. *J. Stat. Comput. Simul.* **81**, 883–893 (2011)
- Cordeiro, GM, Nadarajah, S: Closed-form expressions for moments of a class of beta generalized distributions. *Braz. J. Prob. Stat.* **25**, 14–33 (2011)
- Cordeiro, GM, Nadarajah, S, Ortega, EMM: The Kumaraswamy Gumbel distribution. *Stat. Methods Appl.* **21**, 139–168 (2012)
- Cordeiro, GM, Ortega, EMM, Bozidar, PV, Pescim, RR: The Lomax generator of distributions: Properties, minification process and regression model. *Appl. Math. Comput.* **247**, 465–486 (2014b)
- Eugene, N, Lee, C, Famoye, F: Beta-normal distribution and its applications. *Commun. Statistics-Theory Methods.* **31**, 497–512 (2002)
- Fachini, JB, Ortega, EMM, Cordeiro, GM: A bivariate regression model with cure fraction. *J. Stat. Comput. Simul.* **84**, 1580–1595 (2014)
- Farewell, VT: The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics.* **38**, 1041–1046 (1982)
- Gendoo, DMA, Ratanasirigulchai, N, Schröder, M, Pare, L, Parker, JS, Prat, A, N Haibe-Kains, B: *genefu: a package for breast cancer gene expression analysis* (2015). Retrieved 2016-03-30, from <https://bioc.ism.ac.jp/packages/devel/bioc/vignettes/genefu/inst/doc/genefu.pdf>, <https://goo.gl/jngJMY>
- Gleaton, JU, Lynch, JD: On the distribution of the breaking strain of a bundle of brittle elastic fibers. *Adv. Appl. Probab.* **36**, 98–115 (2004)
- Gleaton, JU, Lynch, JD: Properties of generalized log-logistic families of lifetime distributions. *J. Probab. Stat. Sci.* **4**, 51–64 (2006)
- Gleaton, JU, Lynch, JD: Extended generalized log-logistic families of lifetime distributions with an application. *J. Probab. Stat. Sci.* **8**, 1–17 (2010)

- Gleaton, JU, Rahman, MM: Asymptotic properties of MLE's for distributions generated from a 2-parameter Weibull distribution by a generalized log-logistic transformation. *J. Probab. Stat. Sci.* **8**, 199–214 (2010)
- Gleaton, JU, Rahman, MM: Asymptotic properties of MLE's for distributions generated from a 2-parameter inverse Gaussian distribution by a generalized log-logistic transformation. *J. Probab. Stat. Sci.* **12**, 85–99 (2014)
- Gradshteyn, IS, Ryzhik, IM: *Table of Integrals, Series, and Products*. 7th ed. Academic Press, San Diego (2000)
- Gupta, RD, Kundu, D: Generalized exponential distributions. *Aust. New Zealand J. Stat.* **42**, 173–188 (1999)
- Hashimoto, EM, Cordeiro, GM, Ortega, EMM: The new Neyman type A beta Weibull model with long-term survivors. *Comput. Stat.* **28**, 933–954 (2013)
- Marshall, AN, Olkin, I: A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika.* **84**, 641–652 (1997)
- Mudholkar, GS, Srivastava, DK: Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Trans. Reliab.* **42**, 299–302 (1993)
- Mudholkar, GS, Srivastava, DK, Freimer, M: The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics.* **37**, 436–445 (1995)
- Murthy, DNP, Xie, M, Jiang, R: *Weibull Models*. 1st ed. Wiley, Hoboken (2004)
- Nadarajah, S: The exponentiated Gumbel distribution with climate application. *Environmetrics.* **17**, 13–23 (2006)
- Nadarajah, S, Kotz, S: The exponentiated type distributions. *Acta Applicandae Mathematicae.* **92**, 97–111 (2006)
- Ortega, EMM, Cordeiro, GM, Kattan, MW: The negative binomial-beta Weibull regression model to predict the cure of prostate cancer. *J. Appl. Stat.* **39**, 1191–1210 (2012)
- Ortega, EMM, Cordeiro, GM, Campelo, AK, Kattan, MW, Cancho, VG: A power series beta-Weibull regression model for predicting breast carcinoma. *Stat. Med.* **34**, 1366–1388 (2015)
- Rényi, A: On measures of entropy and information. *Proc. Fourth Berkeley Symp. Math. Stat. Probab.* **1**, 547–561 (1961)
- R Development Core Team: *R: A Language and Environment for Statistical Computing* (2013)
- Shannon, CE: Prediction and entropy of printed English. *Bell Syst. Tech. J.* **30**, 50–64 (1951)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
