Journal of Statistical Modelling: Theory and Applications
Vol. 1, No. 1, 2020, pp. 29-52
Yazd University Press 2020

JSMTA

Research Paper

The exponentiated odd log-logistic family of distributions: properties and applications

Morad Alizadeh, Saeid Tahmasebi *, Hossein Haghbin
[1] Department of Statistics, Persian Gulf University, Bushehr, Iran

Received: September 1, 2018/ Revised: December 20, 2018, 2018/ Accepted: December 28, 2018

Abstract: Based on the generalized log-logistic family (Gleaton and Lynch (2006)) of distributions, we propose a new family of continuous distributions with two extra shape parameters called the exponentiated odd log-logistic family. It extends the class of exponentiated distributions, odd log-logistic family (Gleaton and Lynch (2006)) and any continuous distribution by adding two shape parameters. Some special cases of this family are discussed. We investigate the shapes of the density and hazard rate functions. The proposed family has also tractable properties such as various explicit expressions for the ordinary and incomplete moments, quantile and generating functions, probability weighted moments, Bonferroni and Lorenz curves, Shannon and Rényi entropies, extreme values and order statistics, which hold for any baseline model. The model parameters are estimated by maximum likelihood and the usefulness of the new family is illustrated by means of three real data sets.

Keywords: Generated family; Maximum likelihood; Moment; Odd log-logistic distribution; Probability weighted moment; Quantile function; Rényi entropy.
Mathematics Subject Classification (2010): 60E05, 62H10.

1 Introduction

Numerous classical distributions have been extensively used over the past decades for modeling data in several areas such as engineering, actuarial, environmental and medical sciences, biological studies, demography, economics, finance and insurance. However, in many applied areas such as lifetime analysis, finance and insurance, there is a clear need for extended forms of these distributions. For that reason, several methods for generating new families have been investigated to extend well-known distributions and at the same time provide great flexibility in modelling data in practice. The use of new generators of continuous distributions from classical ones has become more

*Corresponding author: `tahmasebi@pgu.ac.ir`

common in the last ten years or so. Some examples are the *beta-generated* (Eugene et al. (2002)), *gamma-generated* (Zografos and Balakrishnan (2009)) and *generalized Kumaraswamy* (Cordeiro and de Castro (2011)) families.

Based on a baseline cumulative distribution function (cdf) $G(x;\boldsymbol{\xi})$ depending on a parameter vector $\boldsymbol{\xi}$, survival function $\bar{G}(x;\boldsymbol{\xi}) = 1-G(x;\boldsymbol{\xi})$ and probability density function (pdf) $g(x;\boldsymbol{\xi})$, we define the *exponentiated odd log-logistic family* ("EOLL-G", for short) (for $x \in \mathbb{R}$) following the same approach of "odds of death" by Cooray (2006) for constructing the odd Weibull distribution. The exponentiated odd log-logistic family is a continuous univariate parametric family of distributions for modelling continuous univariate data that can be in any interval of the real line. Therefore, the new family is motivated to analyze continuous univariate data that have any type of support.

Let T be a random variable describing a stochastic system having a baseline G distribution. The odds x that the system will not be working at time t is $G(t)/\bar{G}(t)$. We are interested in modelling the random variable X of this odds using the two-parameter Dagum model. We can write

$$\Pr(X \le x) = \Pi_X(x) = F_X\left(\frac{G(t)}{\bar{G}(t)}\right),$$

and then by replacing x in the Dagum cdf (Dagum (1975)) by the ration $G(x;\boldsymbol{\xi})/\bar{G}(x;\boldsymbol{\xi})$, the cdf of the new family follows as

$$F(x;a,b,\boldsymbol{\xi}) = \int_0^{\frac{G(x;\boldsymbol{\xi})}{\bar{G}(x;\boldsymbol{\xi})}} \frac{a\,b\,x^{a\,b-1}}{(1+x^a)^{b+1}}dt = \frac{G(x;\boldsymbol{\xi})^{ab}}{\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a\right\}^b}, \tag{1}$$

where $a > 0$ and $b > 0$ are two additional shape parameters to the vector $\boldsymbol{\xi}$ of parameters in G. Clearly, if we take $a = b = 1$, equation (1) leads to the baseline G. For $a = 1$, we obtain the exponentiated-G ("exp-G") sub-family with power parameter b. This sub-family is also known as the proportional reversed hazard rate model. If $G(x;\boldsymbol{\xi}) = x/(1+x)$, it reduces to the Dagum distribution.

By differentiating (1), we obtain the density function of X as

$$f(x;a,b,\boldsymbol{\xi}) = \frac{a\,b\,g(x;\boldsymbol{\xi})\,G(x;\boldsymbol{\xi})^{ab-1}\,\bar{G}(x;\boldsymbol{\xi})^{a-1}}{\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a\right\}^{b+1}}. \tag{2}$$

For each continuous G distribution (henceforth "G" denotes the baseline distribution), we can associate the EOLL-G distribution with two extra parameters α and β defined by the pdf (2). Hereafter, a random variable X with pdf (2) is denoted by $X \sim$EOLL-G$(\alpha,\beta,\boldsymbol{\xi})$. The two additional parameters induced by the EOLL-G generator are sought as a manner to furnish a more flexible distribution. For a given G, it has a wide variety of shapes and it is able to model comfortable bathtub-shaped failure rate data. Further, the new family can easily be adapted for discriminating between the G and the EOLL-G distributions. When b is an integer, we consider a system formed by b independent components following the Odd-G family (Gleaton and Lynch (2006)) given by

$$H(x;a,\boldsymbol{\xi}) = \frac{G(x;\boldsymbol{\xi})^a}{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a}.$$

Suppose the system fails if all of the b components fails and let X denote the lifetime of the entire system. Then, the cdf of X is

$$F(x;a,b,\boldsymbol{\xi}) = H(x;a,a,\boldsymbol{\xi})^b = \frac{G(x;\boldsymbol{\xi})^{ab}}{\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a\right\}^b},$$

which is the proposed generator.

Also, for $b=1$, we have $a = \log\left[\frac{F(x)}{\bar{F}(x)}\right] / \log\left[\frac{G(x)}{\bar{G}(x)}\right]$ which revals that the parameter a is the log odd ratio between the proposed model and baseline cdf G(.).

The hazard rate function (hrf) of X is obtained from equations (1) and (2) as

$$h(x;a,b,\boldsymbol{\xi}) = \frac{abg(x;\boldsymbol{\xi})G(x;\boldsymbol{\xi})^{a\,b-1}[\bar{G}(x;\boldsymbol{\xi})]^{a-1}}{\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a\right\}\left\{\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a\right\}^b - G(x;\boldsymbol{\xi})^{ab}\right\}}. \tag{3}$$

We can easily simulate data from this family. If $U \sim U(0,1)$ then

$$Q_G\left\{\frac{u^{\frac{1}{ab}}}{u^{\frac{1}{ab}} + (1-u^{\frac{1}{b}})^{\frac{1}{a}}}\right\} \sim EOLL - G(a,b;\boldsymbol{\xi}), \tag{4}$$

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

The remaining of the paper is organized as follows. In Section 2, two special cases of the EOLL-G family are given. In Section 3, the shapes of the density and hazard rate functions are described analytically. A useful expansion for the EOLL-G density family is obtained in Section 4. In Section 5, we derive a power series for the EOLL-G qf. In Section 6, the quantile measures of EOLL-G family are obtained. General explicit expressions and some special cases for the EOLL-G moments are provided in Section 7. The incomplete moments are investigated in Section 8. In Section 9, we derive its generating function. In Section 10, we obtain the mean deviations. Section 11 provides expressions for the Rényi and Shannon entropies. The order statistics and their moments are determined in Section 12. Estimation of the model parameters by maximum likelihood is performed in Section 13. A simulation study is presented in Section 14. Applications to three real data sets illustrate the potentiality of the new family in Section 15. The paper is concluded in Section 16.

2 Special cases

This section introduces some of the many distributions which can arise as special models within the EOLL-G family of distributions. We consider only two parents: normal and Weibull distributions.

2.1 EOLL-Normal (EOLLN)

The EOLLN distribution is defined from (2) by taking $G(x;\boldsymbol{\xi}) = \Phi(\frac{x-\mu}{\sigma})$ and $g(x;\boldsymbol{\xi}) = \phi(\frac{x-\mu}{\sigma})$ where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, and $\boldsymbol{\xi} = (\mu, \sigma^2)$.

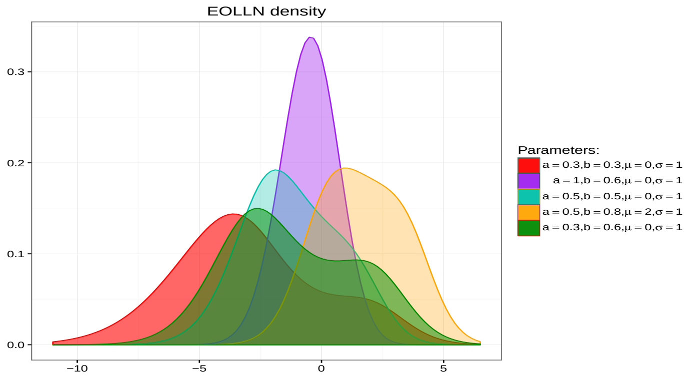

Figure 1: Density function of EOLL-N distribution

The EOLLN pdf is given by

$$f(x;a,b,\mu,\sigma) \;=\; \frac{ab\phi(\frac{x-\mu}{\sigma})[\Phi(\frac{x-\mu}{\sigma})]^{ab-1}[1-\Phi(\frac{x-\mu}{\sigma})]^{a-1}}{\sigma\left\{\Phi(\frac{x-\mu}{\sigma})^{a}+[1-\Phi(\frac{x-\mu}{\sigma})]^{a}\right\}^{b+1}}, \tag{5}$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter and $\sigma > 0$ is a scale parameter.

A random variable with density (5) is denoted by $X \sim \mathrm{EOLLN}(a,b,\mu,\sigma^2)$.

Plots of the EOLLN density function for selected shape parameter values are displayed in Figure 1. These plots indicate that decreasing of the a and b values causes a flattening of the pdf curves.

2.2 EOLL-Weibull (EOLLW)

The pdf and cdf of the Weibull distribution with scale parameter α and shape parameter β is given by

$$g(x;\alpha,\beta)=\alpha\beta x^{\beta-1}\mathrm{e}^{-\alpha \mathrm{x}^{\beta}} \qquad \text{and} \qquad \mathrm{G}(\mathrm{x};\alpha,\beta)=1-\mathrm{e}^{-\alpha\mathrm{x}^{\beta}}.$$

Applying these expressions in (2) gives the EOLLW density function

$$f(x;\alpha,\beta,a,b)=\frac{ab\alpha\beta x^{\beta-1}\mathrm{e}^{-\mathrm{ab}\alpha\mathrm{x}^{\beta}}[1-\mathrm{e}^{-\alpha\mathrm{x}^{\beta}}]^{\mathrm{a}-1}}{\left\{\left[1-\mathrm{e}^{-\alpha\mathrm{x}^{\beta}}\right]^{a}+\mathrm{e}^{-\mathrm{a}\,\alpha\mathrm{x}^{\beta}}\right\}^{b+1}}. \tag{6}$$

Plots of the density function of EOLL-Weibull distribution for selected parameter values are displayed in Figure 2 and the hazard function displayed in Figure 3.

3 Asymptotics and shapes

Let X be a nonnegative random variable. Then, the asymptotics of equations (1), (2) and (3) as $x \to 0$ are given by

$$F(x) \sim G(x)^{a\,b} \;\text{ as }\; \mathrm{x}\to 0,$$

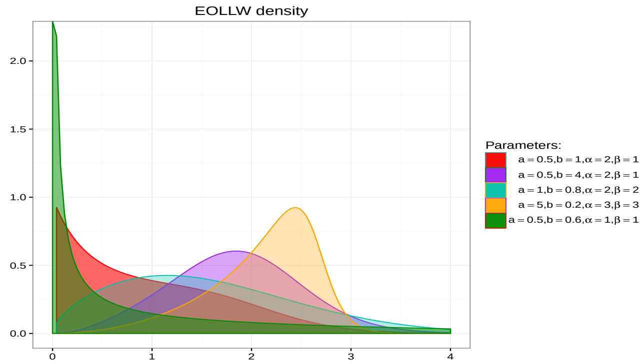

Figure 2: Plot of the density function of EOLL-Weibull distribution

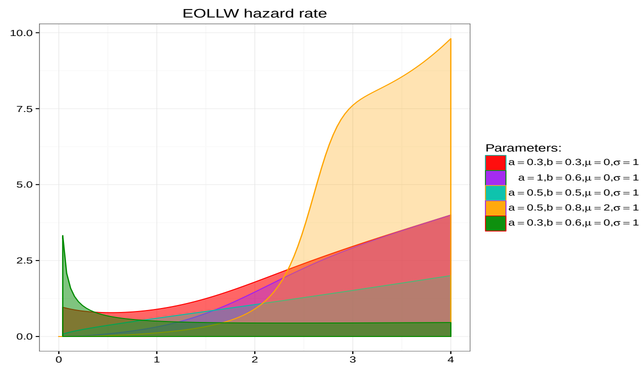

Figure 3: Plot of the hazard function of EOLL-Weibull distribution

$$f(x) \sim abg(x)G(x)^{a\,b-1} \quad \text{as} \;\; \text{x} \to 0,$$
$$h(x) \sim \frac{abg(x)G(x)^{a\,b-1}}{1-G(x)^{a\,b}} \quad \text{as} \;\; \text{x} \to 0.$$

The asymptotics of equations (1), (2) and (3) as $x \to \infty$ are given by

$$1-F(x) \sim b\,\bar{G}(x)^{a} \quad \text{as} \;\; \text{x} \to \infty,$$
$$f(x) \sim abg(x)\bar{G}(x)^{a-1} \quad \text{as} \;\; \text{x} \to \infty,$$
$$h(x) \sim \frac{a\,g(x)}{\bar{G}(x)} \quad \text{as} \;\; \text{x} \to \infty.$$

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

$$\frac{g'(x)}{g(x)} + (ab-1)\frac{g(x)}{G(x)} + (1-a)\frac{g(x)}{\bar{G}(x)} - a(b+1)g(x)\frac{G(x)^{a-1}-\bar{G}(x)^{a-1}}{G(x)^{a}+\bar{G}(x)^{a}} = 0. \qquad (7)$$

There may be more than one root to (7). Let $\lambda(x) = \frac{d^2 \log[f(x)]}{dx^2}$. We have

$$\lambda(x) \;\; = \;\; \frac{g'(x)g(x)-[g'(x)]^2}{g(x)^2} + (ab-1)\frac{g(x)G(x)-g(x)^2}{G(x)^2}$$

$$
\begin{aligned}
&+ \quad (1-a)\frac{g'(x)\bar{G}(x)+g(x)^2}{\bar{G}(x)^2} \\
&- \quad a(b+1)g'(x)\frac{G(x)^{a-1}-\bar{G}(x)^{a-1}}{G(x)^a+\bar{G}(x)^a} \\
&- \quad a(a-1)(b+1)g(x)^2\,\frac{G(x)^{a-2}+\bar{G}(x)^{a-2}}{G(x)^a+\bar{G}(x]^a} \\
&- \quad (b+1)\left\{ag(x)\frac{G(x)^{a-1}-\bar{G}(x)^{a-1}}{G(x)^a+\bar{G}(x)^a}\right\}^2.
\end{aligned}
$$

If $x = x_0$ is a root of (7) then it corresponds to a local maximum if $\lambda(x) > 0$ for all $x < x_0$ and $\lambda(x) < 0$ for all $x > x_0$. It corresponds to a local minimum if $\lambda(x) < 0$ for all $x < x_0$ and $\lambda(x) > 0$ for all $x > x_0$. It refers to a point of inflection if either $\lambda(x) > 0$ for all $x \neq x_0$ or $\lambda(x) < 0$ for all $x \neq x_0$.

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

$$
\begin{aligned}
&\frac{g'(x)}{g(x)}+(ab-1)\frac{g(x)}{G(x)}+(1-a)\frac{g(x)}{1-G(x)}-ag(x)\frac{G(x)^{a-1}-[\bar{G}(x)]^{a-1}}{G(x)^a+\bar{G}(x)^a} \\
&-abg(x)\frac{\left\{\left\{G(x)^a+\bar{G}(x)^a\right\}^{b-1}[G(x)^{a-1}-\bar{G}(x)^{a-1}]-G(x)^{ab-1}\right\}}{\left\{G(x)^a+\bar{G}(x)^a\right\}^b-G(x)^{ab}}=0. \qquad (8)
\end{aligned}
$$

There may be more than one root to (8). Let $\tau(x) = d^2\log[h(x)]/dx^2$. We have

$$
\begin{aligned}
\tau(x) \quad &= \quad \frac{g'(x)g(x)-[g'(x)]^2}{g(x)^2}+(ab-1)\frac{g(x)G(x)-g(x)^2}{G(x)^2} \\
&+ \quad (1-a)\frac{g'(x)\bar{G}(x)+g(x)^2}{\bar{G}(x)^2} \\
&- \quad ag'(x)\frac{G(x)^{a-1}-\bar{G}(x)^{a-1}}{G(x)^a+\bar{G}(x)^a} \\
&- \quad a(a-1)g(x)^2\,\frac{G(x)^{a-2}+\bar{G}(x]^{a-2}}{G(x)^a+\bar{G}(x)^a} \\
&- \quad \left\{ag(x)\frac{G(x)^{a-1}-\bar{G}(x)^{a-1}}{G(x)^a+\bar{G}(x)^a}\right\}^2 \\
&- \quad abg'(x)\frac{\left\{\left\{G(x)^a+\bar{G}(x)^a\right\}^{b-1}[G(x)^{a-1}-\bar{G}(x)^{a-1}]-G(x)^{ab-1}\right\}}{\left\{G(x)^a+\bar{G}(x)^a\right\}^b-G(x)^{ab}} \\
&- \quad ab(a-1)g(x)\frac{\left\{\left\{G(x)^a+\bar{G}(x)^a\right\}^{b-1}[G(x)^{a-2}+\bar{G}(x)^{a-2}]\right\}}{\left\{G(x)^a+\bar{G}(x)^a\right\}^b-G(x)^{ab}} \\
&- \quad ab(b-1)g(x)\frac{\left\{\left\{G(x)^a+\bar{G}(x)^a\right\}^{b-2}[G(x)^{a-1}-\bar{G}(x)^{a-1}]^2-G(x)^{ab-2}\right\}}{\left\{G(x)^a+\bar{G}(x]^a\right\}^b-G(x)^{ab}}
\end{aligned}
$$

$$-\quad \left\{ abg(x) \frac{\left\{ \left\{ G(x)^a + \bar{G}(x)^a \right\}^{b-1} \left[G(x)^{a-1} - \bar{G}(x)^{a-1} \right] - G(x)^{ab-1} \right\}}{\left\{ G(x)^a + \bar{G}(x)^a \right\}^b - G(x)^{ab}} \right\}^2 .$$

If $x = x_0$ is a root of (8) then it refers to a local maximum if $\tau(x) > 0$ for all $x < x_0$ and $\tau(x) < 0$ for all $x > x_0$. It corresponds to a local minimum if $\tau(x) < 0$ for all $x < x_0$ and $\tau(x) > 0$ for all $x > x_0$. It gives an inflection point if either $\tau(x) > 0$ for all $x \neq x_0$ or $\tau(x) < 0$ for all $x \neq x_0$.

4 A useful expansion

For an arbitrary baseline cdf $G(x)$, a random variable Z has the exp-G distribution with power parameter $c > 0$, say $Z \sim$exp-G(c), if its pdf and cdf are given by

$$h_c(x) = c\, G(x)^{c-1}\, g(x) \quad \text{and} \quad H_c(x) = G(x)^c,$$

respectively. Some structural properties of the exp-G distributions are explored by Mudholkar et al. (1995), Gupta and Kundu (2001) and Nadarajah and Kotz (2006), among others.

We can demonstrate that the cdf (1) of X admits the expansion

$$F(x) = \frac{G(x)^{a\, b}}{\left[G(x)^a + \bar{G}(x)^a \right]^b} = \frac{\sum_{k=0}^{\infty} \alpha_k\, G(x)^k}{\sum_{k=0}^{\infty} \beta_k\, G(x)^k} = \sum_{k=0}^{\infty} b_k\, G(x)^k, \tag{9}$$

where $b_0 = \alpha_0 / \beta_0$,

$$\alpha_k = s_k(a\, b),\ \beta_k = h_k(a, b) \ \ \textit{and} \ \ b_k = \frac{1}{\beta_0} \left[\alpha_k - \sum_{r=1}^{k} \beta_r b_{k-r} \right], \text{for k} \geq 1, \tag{10}$$

and $s_k(a\, b)$ and $h_k(a, b)$ are defined by equations (34) and (37) (see Appendix) and $H_k(x) = G(x)^k$ denotes the exp-G cumulative distribution with power parameter k. So, the density function of X can be expressed as an infinite linear combination of exp-G density functions

$$f(x; a, b, \boldsymbol{\xi}) = \sum_{k=0}^{\infty} b_{k+1}\, h_{k+1}(x; \boldsymbol{\xi}), \tag{11}$$

where $h_{k+1}(x; \boldsymbol{\xi}) = (k+1)\, g(x; \boldsymbol{\xi})\, G(x; \boldsymbol{\xi})^k$ denotes the exp-G density function with power parameter $k+1$. Hereafter, a random variable with the density function $h_{k+1}(x; \boldsymbol{\xi})$ is denoted by $Y_{k+1} \sim$ exp-G$(k+1)$. Equation (11) reveals that the EOLL-G density function is a linear combination of exp-G densities. Thus, some mathematical properties of the new model can be obtained directly from those properties of the exp-G distribution. For example, the ordinary and incomplete moments and moment generating function (mgf) of X can be obtained from those quantities of the exp-G distribution.

5 Quantile power series

In this section, we derive a power series expansion for the qf $x = Q(u) = F^{-1}(u)$ of X by expanding (12). If the G qf, say $Q_G(u) = G^{-1}(u)$, does not have a closed-form expression, it can usually be expressed in terms of a power series

$$Q_G(u) = \sum_{i=0}^{\infty} a_i\, u^i, \tag{12}$$

where the coefficients a_i are suitably chosen real numbers which depend on the parameters of the G distribution. For several important distributions, such as the normal, Student t, gamma and beta distributions, $Q_G(u)$ does not have explicit expressions but it can be expanded as in equation (12). As a simple example, for the normal $N(0,1)$ distribution, $a_i = 0$ for $i = 0, 2, 4, \ldots$ and $a_1 = 1$, $a_3 = 1/6$, $a_5 = 7/120$ and $a_7 = 127/7560, \ldots$

We use throughout the paper a result of Gradshteyn and Ryzhik (2000) for a power series raised to a positive integer n (for $n \geq 1$)

$$Q_G(u)^n = \left(\sum_{i=0}^{\infty} a_i\, u^i\right)^n = \sum_{i=0}^{\infty} c_{n,i}\, u^i, \tag{13}$$

where the coefficients $c_{n,i}$ (for $i = 1, 2, \ldots$) are easily obtained from the recurrence equation (with $c_{n,0} = a_0^n$)

$$c_{n,i} = (i\, a_0)^{-1} \sum_{m=1}^{i} [m(n+1) - i]\, a_m\, c_{n,i-m}. \tag{14}$$

Clearly, $c_{n,i}$ can be determined from $c_{n,0}, \ldots, c_{n,i-1}$ and then from the quantities $a_0, \ldots, a_i$.

Next, we derive an expansion for the argument of $Q_G(\cdot)$ in (4)

$$A = \frac{u^{\frac{1}{ab}}}{u^{\frac{1}{ab}} + (1 - u^{\frac{1}{b}})^{\frac{1}{a}}}.$$

Using the the generalized binomial expansion since $u \in (0,1)$, we can write

$$A = \frac{\sum_{k=0}^{\infty} a_k^*\, u^k}{\sum_{k=0}^{\infty} b_k^*\, u^k} = \sum_{k=0}^{\infty} \delta_k u^k,$$

where $a_k^* = s_k(a/b)$, $b_k^* = \sum_{i=0}^{\infty} (-1)^i s_k(i/b) \binom{\frac{1}{a}}{i}$, and the coefficient δ_k (for$k \geq 0$) is determined from the recurrence equation

$$\delta_k = \frac{1}{b_0^*}\left[a_k^* - \frac{1}{b_k^*} \sum_{r=1}^{k} b_r^*\, \delta_{k-r}\right].$$

Also $s_i(1/ab)$ and $s_k(i/b)$ are given by equation (34). Then, the qf of X can be expressed from (4) as

$$Q(u) = Q_G\left(\sum_{k=0}^{\infty} \delta_k\, u^k\right). \tag{15}$$

For any baseline G distribution, we can combine (12) with (15) to obtain

$$Q(u) = Q_G\left(\sum_{m=0}^{\infty} \delta_m\, u^m\right) = \sum_{i=0}^{\infty} a_i \left(\sum_{m=0}^{\infty} \delta_m\, u^m\right)^i,$$

and then using (13) and (14), we have

$$Q(u) = \sum_{m=0}^{\infty} e_m\, u^m, \tag{16}$$

where $e_m = \sum_{i=0}^{\infty} a_i\, d_{i,m}$, $d_{i,0} = \delta_0^i$ and, for $m > 1$,

$$d_{i,m} = (m\,\delta_0)^{-1} \sum_{n=1}^{m} [n(i+1) - m]\, \delta_n\, d_{i,m-n}.$$

Equation (16) is the main result of this section since it allows to obtain various mathematical quantities for the EOLL-G family as shown in the next sections.

The formulae derived throughout the paper can be easily handled in most symbolic computation software platforms such as Maple, Mathematica and Matlab. These platforms have currently the ability to deal with analytic expressions of formidable size and complexity. The infinity limit in these sums can be substituted by a large positive integer such as 20 or 30 for most practical purposes

6 Quantile measure

The effects of the extra shape parameters a and b on the skewness and kurtosis of X can be considered based on quantile measures. The Bowley skewness based on quartiles, is given by

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

The Moors kurtosis based on quantiles, is given by

$$M = \frac{Q\left(\frac{3}{8}\right) - Q\left(\frac{1}{8}\right) + Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)}.$$

These measures are less sensitive to outliers and they exist even for distributions without moments. In Figures 4 and 5, we plot the measures B and M for the EOLL-N(0,1,a, b) discussed in Section 2. These plots reveal how both measures B and M vary with the shape parameters.

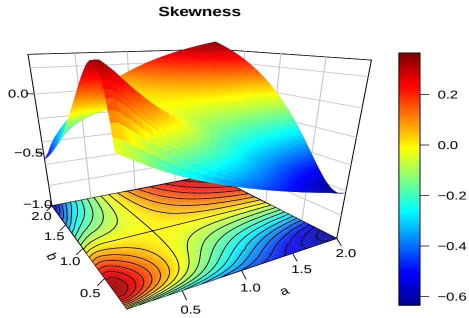

Figure 4: Plot of the Bowly skewness function of EOLL-N(0,1,a, b) distribution

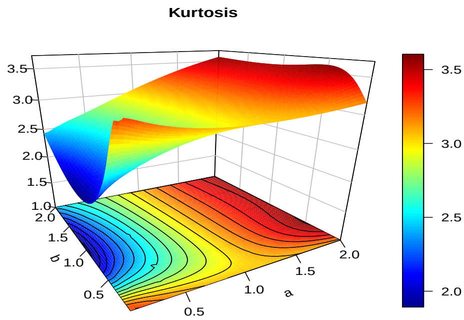

Figure 5: Plot of the Moors kurtosis function of EOLL-N(0,1,a, b) distribution

7 Moments

Hereafter, we shall assume that $G(x)$ is the baseline cdf of a random variable Y and that $F(x)$ is the cdf of the random variable X having density function (2). The moments of the EOLL-G distribution can be obtained from the (r, k)th probability weighted moments (PWMs) of Y defined by

$$\tau_{r,k} = \mathrm{E}[X^r\, G(X)^k] = \int_{-\infty}^{\infty} x^r\, G(x)^k\, g(x) dx. \tag{17}$$

In fact, we have

$$\mu'_k = \mathrm{E}(X^r) = \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, \tau_{r,k}. \tag{18}$$

where b_k is defined by equation (10). Thus, the moments of any EOLL-G distribution can be expressed as an infinite linear combination of the baseline PWMs.

A second formula for $\tau_{r,k}$ can be written in terms of $Q_G(u) = G^{-1}(u)$. Setting $G(x) = u$, we obtain

$$\tau_{r,k} = \int_0^1 Q_G(u)^r\, u^k du. \tag{19}$$

The PWMs for various distributions can be determined by using equations (17) and (19). The following special cases were already published by Cordeiro and Nadarajah (2011).

8 Incomplete moments

The nth incomplete moment of X is defined as $m_n(y) = \int_{-\infty}^{y} x^r f(x)dx$. Here, we propose two methods to determine the incomplete moments of the new family. First, the nth incomplete moment of X can be expressed as

$$m_n(y) = \sum_{k=0}^{\infty} b_{k+1} \int_0^{G(y;\boldsymbol{\xi})} Q_G(u)^n \, u^k \, du. \tag{20}$$

The integral in (20) can be computed at least numerically for most baseline distributions. A second method to obtain the incomplete moments of X follows from (20) using equations (13) and (14). We obtain

$$m_n(y) = \sum_{k,m=0}^{\infty} \frac{(k+1)\, b_{k+1}\, c_{n,m}}{m+k+1} G(y;\boldsymbol{\xi})^{m+k+1}, \tag{21}$$

where the coefficients $c_{n,m}$ are given by (14). The Bonferroni and Lorenz curves and the Bonferroni and Gini indices have many applications in economics, reliability, demography, insurance and medicine. The Bonferroni and Lorenz curves are given by

$$\begin{aligned} B(p) &= \frac{1}{p\mu} \int_0^q x f(x)dx = \frac{1}{p\mu} \int_0^p F^{-1}(x)dx \\ L(p) &= \frac{1}{\mu} \int_0^q x f(x)dx = \frac{1}{\mu} \int_0^p F^{-1}(x)dx \end{aligned}$$

where $q = F^{-1}(p)$. The Bonferroni and Gini indices are defined by

$$B = 1 - \int_0^1 B(p)dp \text{ and } L = 1 - 2\int_0^1 L(p)dp.$$

After some simple calculation, the Bonferroni and Gini indices of the EOLL-G distribution are given by

$$\begin{aligned} B &= 1 - \frac{1}{\mu} \sum_{k,m=0}^{\infty} \frac{(k+1)\, b_{k+1}\, c_{1,m}}{(m+k+1)p} \int_0^1 G(q;\boldsymbol{\xi})^{m+k+1} dp, \\ L &= 1 - \frac{2}{\mu} \sum_{k,m=0}^{\infty} \frac{(k+1)\, b_{k+1}\, c_{1,m}}{(m+k+1)} \int_0^1 G(q;\boldsymbol{\xi})^{m+k+1} dp \end{aligned}$$

where b_{k+1} and $c_{1,m}$ are determined from equation(10) and (14), respectively.

9 Generating function

In this section, we provide two formulae for the moment generating function (mgf) $M(s) = E(\mathrm{e}^{\mathrm{s\,X}})$ of a random variable X with the EOLL-G distribution. A first formula for $M(s)$ comes from equation (11) as

$$M(s) = \sum_{k=0}^{\infty} b_{k+1}\, M_{k+1}(s), \tag{22}$$

where $M_{k+1}(s)$ is the generating function of the exp-G distribution with power power parameter $k+1$. Hence, $M(s)$ can be determined from the exp-G generating function.

A second formula for $M(s)$ can be derived from equation (22) as

$$M(s) = \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, \rho_k(s), \tag{23}$$

where the quantity $\rho_k(s) = \int_{-\infty}^{\infty} \exp(sx)\, G(x)^k\, g(x) dx$ follows from the baseline qf as

$$\rho_k(s) = \int_0^1 \exp\left[s\, Q_G(u)\right] u^k du. \tag{24}$$

10 Mean deviations

The mean deviations about the mean ($\delta_1(Y) = E(|Y - \mu_1'|)$) and about the median ($\delta_2(Y) = E(|Y - M|)$) of Y can be expressed as

$$\delta_1(Y) = 2\mu_1'\, F\left(\mu_1'\right) - 2m_1\left(\mu_1'\right) \qquad \text{and} \qquad \delta_2(Y) = \mu_1' - 2m_1(M), \tag{25}$$

respectively, where $M = Q_G\left[\frac{0.5^{\frac{1}{ab}}}{0.5^{\frac{1}{ab}} + (1 - 0.5^{\frac{1}{b}})^{\frac{1}{a}}}\right]$ is the median of Y, $\mu_1' = \mathrm{E}(Y)$ comes from equation (18), $F(\mu_1')$ is easily calculated from equation (9) and $m_1(z) = \int_{-\infty}^{z} x\, f(x) dx$ is the first incomplete moment.

Now, we provide two alternative ways to compute $\delta_1(Y)$ and $\delta_2(Y)$. A general equation for $m_1(z)$ can be derived from equation (11) as

$$m_1(z) = \sum_{k=0}^{\infty} b_{k+1}\, J_{k+1}(z), \tag{26}$$

where

$$J_{k+1}(z) = \int_{-\infty}^{z} x\, h_{k+1}(x) dx.$$

Equation (**??**) is the basic quantity to compute the mean deviations for the EOLL-G distributions. The mean deviations defined in (25) depend only on the first incomplete moment of the Exp-G distributions. So, alternative representations for $\delta_1(Y)$ and $\delta_2(Y)$ are

$$\delta_1(Y) = 2\mu_1'\, F\left(\mu_1'\right) - 2\sum_{k=0}^{\infty} b_{k+1}\, J_{k+1}\left(\mu_1'\right) \quad \text{and} \quad \delta_2(Y) = \mu_1' - 2\sum_{k=0}^{\infty} b_{k+1}\, J_{k+1}(M).$$

A second general formula for $m_1(z)$ can be derived by setting $u = G(x)$ in (26)

$$m_1(z) = \sum_{k=0}^{\infty}(k+1)\, b_{k+1}\, T_k(z),$$

where $T_k(z)$ is given by

$$T_k(z) = \int_0^{G(z)} Q_G(u)\, u^k du.$$

Applications of these equations can be conducted to obtain Bonferroni and Lorenz curves defined for a given probability π by $B(\pi) = m_1(q)/\pi\, \mu_1'$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $q = Q_G([\frac{1-(1-2^{-\alpha})^{1/\lambda}}{1-(1-p\,2^{-\alpha})^{1/\lambda}}])$ is immediately calculated from the parent quantile function.

11 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

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

for $c > 0$ and $c \neq 1$. 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 $c \uparrow 1$. For a random variable X with the EOLL-G distribution, direct calculation yields

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} &= -\log[a\,b] - \mathrm{E}\{\log[g(X;\boldsymbol{\xi})]\} + (1-ab)\,\mathrm{E}\{\log[G(x;\boldsymbol{\xi})]\} \\ &+ (1-a)\,\mathrm{E}\left\{\log\left[\bar{G}(x;\boldsymbol{\xi})\right]\right\} + (b+1)\,\mathrm{E}\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(x;\boldsymbol{\xi})^a\right\} \end{aligned}$$

First, we define and obtain

$$\begin{aligned} A(a_1,a_2,a_3;a) &= \int_0^1 \frac{u^{a_1}(1-u)^{a_2}}{[u^a+(1-u)^a]^{a_3}}du \\ &= \sum_{i=0}^{\infty}(-1)^i\binom{a_3}{i}\int_0^1 \frac{u^{a_1+i}}{[u^a+(1-u)^a]^{a_3}}du \\ &= \sum_{i=0}^{\infty}(-1)^i\binom{a_3}{i}\int_0^1 \frac{\sum_{k=0}^{\infty}\delta_{1,k}u^k}{\sum_{k=0}^{\infty}\delta_{2,k}u^k}du \\ &= \sum_{i=0}^{\infty}(-1)^i\binom{a_3}{i}\int_0^1 \sum_{k=0}^{\infty}\delta_{3,k}u^k \\ &= \sum_{i=0}^{\infty}\frac{(-1)^i\binom{a_3}{i}\delta_{3,k}}{k+1} \qquad (28) \end{aligned}$$

where

$$\delta_{1,k} = s_k(a_1 + i) \ \ , \delta_{2,k} = h_k(a, a_3) \ \ and \ \ \delta_{3,k} = \frac{1}{\delta_{2,0}} \left[\delta_{1,k} - \frac{1}{\delta_{2,0}} \sum_{r=1}^{k} \delta_{2,r} \delta_{3,k-r} \right] \quad (29)$$

After some algebraic manipulations, we obtain

$$\begin{aligned}
\mathrm{E}\left\{\log\left[G(X)\right]\right\} &= a\,b\,\frac{\partial}{\partial t} A(a\,b + t - 1, a - 1, b + 1; a)\Big|_{t=0} \\
\mathrm{E}\left\{\log\left[\bar{G}(X)\right]\right\} &= a\,b\,\frac{\partial}{\partial t} A(a\,b - 1, a + t - 1, b + 1; a)\Big|_{t=0} \\
\mathrm{E}\left\{G(x;\boldsymbol{\xi})^a + \bar{G}(X;\boldsymbol{\xi})^a\right\} &= a\,b\,\frac{\partial}{\partial t} A(a\,b - 1, a - 1, b + 1 - t; a)\Big|_{t=0}
\end{aligned}$$

The simplest formula for the entropy of X is given by

$$\begin{aligned}
\mathrm{E}\left\{-\log[f(X)]\right\} = &-\log[a\,b] - \mathrm{E}\left\{\log\left[g(X;\boldsymbol{\xi})\right]\right\} \\
&+ (1 - a\,b)a\,b\,\frac{\partial}{\partial t} A(a\,b + t - 1, a - 1, b + 1; a)\Big|_{t=0} \\
&+ (1 - a)a\,b\,\frac{\partial}{\partial t} A(a\,b - 1, a + t - 1, b + 1; a)\Big|_{t=0} \\
&+ (b + 1)a\,b\,\frac{\partial}{\partial t} A(a\,b - 1, a - 1, b + 1 - t; a)\Big|_{t=0}
\end{aligned}$$

After some algebraic developments, we obtain an alternative expression for $I_R(c)$

$$I_R(c) = \frac{c}{1-c}\log(a\,b) + \frac{1}{1-c}\log\left[\sum_{i,k=0}^{\infty} w_{i,k}^{*}\,\mathrm{E}_{Y_k}\left(g^{c-1}[G^{-1}(Y)]\right)\right],$$

where $Y_k \sim B(k+1, 1)$,

$$w_{i,k}^{*} = \frac{(-1)^i \binom{c(a-1)}{i} \gamma_{3,k}(a, b, , c, i)}{k+1},$$

and $\gamma_{1,k} = s_k(c(a\,b - 1) + i)$,

$$\gamma_{2,k} = h_k(a, c(b+1)), \ \ \gamma_{3,k} = \frac{1}{\gamma_{2,0}}\left[\gamma_{1,k} - \frac{1}{\gamma_{2,0}}\sum_{r=1}^{k}\gamma_{2,r}\gamma_{3,k-r}\right], \quad (30)$$

where $s_k(c(a\,b - 1) + i)$ and $h_k(a, c(b+1))$ are defined by equation (37).

12 Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1, \ldots, X_n$ is a random sample from the EOLL-G family of distributions. We can write the density of the ith order statistic, say $X_{i:n}$, as

$$f_{i:n}(x) = K f(x) F^{i-1}(x) \{1 - F(x)\}^{n-i} = K \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} f(x) F(x)^{j+i-1},$$

where $K = n!/[(i-1)!\,(n-i)!]$.

Following similar algebraic developments of Nadarajah et al. (2013), we can write the density function of $X_{i:n}$ as

$$f_{i:n}(x) = \sum_{r,k=0}^{\infty} m_{r,k}\, h_{r+k+1}(x), \tag{31}$$

where $h_{r+k}(x)$ denotes the exp-G density function with power parameter $r+k$,

$$m_{r,k} = \frac{n!\,(r+1)\,(i-1)!\,b_{r+1}}{(r+k+1)} \sum_{j=0}^{n-i} \frac{(-1)^j\, f_{j+i-1,k}}{(n-i-j)!\,j!},$$

and b_k is defined in equation (10). Here, the quantities $f_{j+i-1,k}$ are obtained recursively by $f_{j+i-1,0} = b_0^{j+i-1}$ and (for $k \geq 1$)

$$f_{j+i-1,k} = (k\, b_0)^{-1} \sum_{m=1}^{k} [m(j+i) - k]\, b_m\, f_{j+i-1,k-m}.$$

Based on the expansion (31), we can obtain some structural properties (ordinary and incomplete moments, generating function, etc.) for the EOLL-G order statistics from those exp-G properties.

13 Estimation

Here, we determine the maximum likelihood estimates (MLEs) of the model parameters of the new family from complete samples only. Let $x_1, \ldots, x_n$ be observed values from the EOLL-G distribution with parameters a, b and $\boldsymbol{\xi}$. Let $\Theta = (a, b, \boldsymbol{\xi})^\top$ be the $r \times 1$ parameter vector. The total log-likelihood function for Θ is given by

$$\begin{aligned} \ell_n &= \ell_n(\Theta) = n \log(a\, b) + \sum_{i=1}^{n} \log[g(x_i; \boldsymbol{\xi})] + (ab-1) \sum_{i=1}^{n} \log[G(x_i; \boldsymbol{\xi})] \\ &+ (a-1) \sum_{i=1}^{n} \log[\bar{G}(x_i; \boldsymbol{\xi})] - (b+1) \sum_{i=1}^{n} \log\left\{G(x_i; \boldsymbol{\xi})^a + \bar{G}(x_i; \boldsymbol{\xi})^a\right\}. \end{aligned} \tag{32}$$

The log-likelihood function can be maximized either directly by using the SAS (PROC NLMIXED) or the Ox program (sub-routine MaxBFGS) (see Doornik (1996)) or by solving the nonlinear likelihood equations obtained by differentiating (32). The components of the score function $U_n(\Theta) = (\partial \ell_n/\partial a, \partial \ell_n/\partial b, \partial \ell_n/\partial \boldsymbol{\xi})^\top$ are given by

$$\frac{\partial \ell_n}{\partial a} = \frac{n}{a} + \sum_{i=1}^{n} \log\left\{G(x_i; \boldsymbol{\xi}) \bar{G}(x_i; \boldsymbol{\xi})\right\}$$

$$
\begin{aligned}
&- (b+1)\sum_{i=1}^{n}\frac{G(x_i;\boldsymbol{\xi})^a\log[G(x_i;\boldsymbol{\xi})]+\bar{G}(x_i;\boldsymbol{\xi})^a\log[\bar{G}(x_i;\boldsymbol{\xi})]}{G(x_i;\boldsymbol{\xi})^a+[\bar{G}(x_i;\boldsymbol{\xi})]^a}\\
\frac{\partial\ell_n}{\partial b} &= \frac{n}{b}-\sum_{i=1}^{n}\log\left\{G(x_i;\boldsymbol{\xi})^a+\bar{G}(x_i;\boldsymbol{\xi})^a\right\}+a\sum_{i=1}^{n}\log[G(x_i;\boldsymbol{\xi})],
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial\ell_n}{\partial\boldsymbol{\xi}} &= \sum_{i=1}^{n}\frac{g(x_i;\boldsymbol{\xi})^{(\xi)}}{g(x_i;\boldsymbol{\xi})}+(a-1)\sum_{i=1}^{n}\frac{G(x_i;\boldsymbol{\xi})^{(\xi)}}{G(x_i;\boldsymbol{\xi})}\frac{1-2G(x_i;\boldsymbol{\xi})}{\bar{G}(x_i;\boldsymbol{\xi})}\\
&- a(b+1)\sum_{i=1}^{n}G(x_i;\boldsymbol{\xi})^{(\xi)}\,\frac{G(x_i;\boldsymbol{\xi})^{a-1}-\bar{G}(x_i;\boldsymbol{\xi})]^{a-1}}{G(x_i;\boldsymbol{\xi})^a+\bar{G}(x_i;\boldsymbol{\xi})^a},
\end{aligned}
$$

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

14 Simulation study

In this section, we assess the performance of the MLEs of the EOLL-N distribution parameters with respect to sample size n. The validity of the MLEs is discussed by the measures namely bias, mean square error of the estimate (MSE), coverage lenghts (CL) and coverage probability (CP). We generated $N = 10^4$ parallel samples of size $n = 50, 100, 200, 500, 1000, 2000, 5000$ from EOLL-N distribution with $a = 2, b = 1, \mu = 0, \sigma = 1$, by using inverse transform method. The MLEs of parameters are computed in each generated sample, say $\left(\hat{a_i}, \hat{b_i}, \hat{\mu_i}, \hat{\sigma_i}\right)$ for $i = 1, 2, \cdots, N$. The standard errors of the MLEs are computed by inverting the observed information matrices, say $\left(s_{\hat{a_i}}, s_{\hat{b_i}}, s_{\hat{\mu_i}}, s_{\hat{\sigma_i}}\right)$ for $i = 1, 2, \cdots, N$. The bias and mean square errors are computed by

$$
Bias_\epsilon(n) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\epsilon_i}-\epsilon), \quad MSE_\epsilon(n) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\epsilon_i}-\epsilon)^2,
$$

for $\epsilon = a, b, \mu, \sigma$. The coverage probabilities and coverage lenghts given by

$$
\begin{aligned}
CP_\epsilon(n) &= \frac{1}{N}\sum_{i=1}^{N}I(\hat{\epsilon_i}-1.95996s_{\hat{\epsilon_i}}, \hat{\epsilon_i}+1.95996s_{\hat{\epsilon_i}}),\\
CL_\epsilon(n) &= \frac{3.919928}{N}\sum_{i=1}^{N}s_{\hat{\epsilon_i}}.
\end{aligned}
$$

The obtained numerical results for the above measures are tabulated in Table 1 . As it can be seen from Table 1, the bias is decreases as the sample size n increases, except for the parameter α which is near zero and negligible. The MSEs for each parameter decrease to zero as n increases. This shows the consistency property of the MLEs. The coverage probability is very close to 0.95 and approaches to the nominal value as the sample size increases. Moreover, when the sample size increases the coverage length is decrease for each parameter. Note that the reported results are for only one choice for (a, b, μ, σ). On the other hand, the results were similar for a wide range of other choices for (a, b, μ, σ).

Table 1: Result of simulation from EOLL-N distribution

	Bias				MSE			
n	a	b	μ	σ	a	b	μ	σ
50	-0.0123	0.6309	1.5366	1.0060	0.3294	11.063	46.933	17.784
100	-0.0223	0.7091	1.5693	0.5011	0.1962	11.070	50.315	4.8956
200	-0.0125	0.6532	1.3900	0.2194	0.0992	10.378	45.634	1.4285
500	-0.0062	0.4039	0.8395	0.0573	0.0254	6.6305	27.918	0.1986
1000	-0.0022	0.1409	0.2934	0.0173	0.0078	1.5637	6.7731	0.0412
2000	-0.0005	0.0498	0.1050	0.0058	0.0032	0.3005	1.2943	0.0154
5000	0.0000	0.0193	0.0411	0.0017	0.0011	0.0261	0.1124	0.0050
10000	0.0000	0.0108	0.0231	0.0007	0.0005	0.0119	0.0516	0.0023
	CL				CP			
n	a	b	μ	σ	a	b	μ	σ
50	2.1539	5.2829	11.512	11.095	0.7851	0.6622	0.7505	0.7320
100	1.6482	4.3221	8.9617	6.7426	0.8670	0.7163	0.7255	0.8210
200	1.0896	3.3850	6.9063	3.4553	0.9130	0.7686	0.7547	0.8825
500	0.5356	2.3718	4.9136	1.2803	0.9572	0.8299	0.8368	0.9409
1000	0.3259	1.6344	3.3737	0.6997	0.9565	0.8787	0.8804	0.9535
2000	0.2159	1.0916	2.2622	0.4512	0.9539	0.9046	0.9056	0.9540
5000	0.1321	0.6333	1.3131	0.2744	0.9525	0.9323	0.9339	0.9520
10000	0.0926	0.4230	0.8775	0.1921	0.9578	0.9416	0.9386	0.9552

15 Applications

In this section, we provide four applications to real data. In the first three applications we present some results fitting the special models defined in Section 2.

For the first three applications, the goodness-of-fit statistics including the Cramér-von Mises (W^*) and Anderson-Darling (A^*) test statistics are adopted to compare the fitted models. The smaller the values of A^* and W^*, the better the fit to the data. We also consider the Kolmogrov-Smirnov (K-S) statistic and its corresponding p-value the minus log-likelihood function (-Log(L)) 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 take the odd log-logistic normal (OLLN) distribution and for the purpose of comparison, we fitted the following models to the above data sets:

- The normal distribution.
- The exponentiated normal (EN) distribution.
- The beta normal (BN) distribution with density

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

- The gamma normal (GN) distribution with density

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

- The Kumaraswamy normal (KN) distribution with density

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

- The odd log-logistic normal (OLL-N) distribution (special case of OLLLN distribution when $\beta\to 1$) with density

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

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

15.1 Application 1

Data 1 - First, we consider the data set which represents failure times of a particular windshield device. These data were also studied by Blischke and Murthy (2011) . 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 (SE) in parentheses and the goodness-of-fit statistics for D1 we listed in Table 2. We can note that the OLLLN model outperforms all the fitted competitive models under these statistics. The fitted densities and histogram of the data are displayed in Figure 6. For D1, we note that the fitted OLLLN distribution best captures the empirical histogram.

15.2 Application 2

Data 2- The second data set, referred to as D2, Armitage and Berry (1987) provided the weights in ounces of 32 newborn babies. These data are:

72, 80, 81, 84, 86, 87, 92, 94, 103, 106, 107, 111, 112, 115, 116, 118, 119, 122, 123, 123, 114, 125, 126, 126, 126, 127, 118, 128, 128, 132, 133, 142.

The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D2 are presented in Table 3. We can see that the OLLLN model outperforms all the fitted competitive models under these statistics.

The density and histogram plots are dosplayed in Figure 7. For D2, we note that the fitted OLLLN distribution best captures the empirical histogram.

Table 2: The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D1.

Model	μ	σ	a	b	-log(L)	W^*	A^*	K-S	p-value
EOLLN	3.272	0.351	0.267	0.434	124.591	0.024	0.218	0.053	0.969
	(0.249)	(0.114)	(0.136)	(0.161)					
Normal	2.557	1.112			128.119	0.091	0.607	0.092	0.444
	(0.121)	(0.086)							
EN	1.823	1.339	1.954		128.064	0.074	0.521	0.084	0.560
	(2.342)	(0.701)	(3.864)						
BN	0.808	2.443	7.113	2.469	128.085	0.074	0.519	0.084	0.562
	(7.144)	(8.149)	(48.513)	(14.595)					
GN	2.805	0.541	0.290	0.197	127.757	0.057	0.438	0.074	0.710
	(1.057)	(0.264)	(0.381)	(0.215)					
KN	1.653	0.747	0.918	0.319	127.848	0.063	0.468	0.079	0.641
	(1.063)	(0.534)	(1.013)	(0.518)					
OLL-N	2.626	0.602	0.452		127.062	0.075	0.523	0.095	0.407
	(0.126)	(0.218)	(0.232)						

(a) (b)

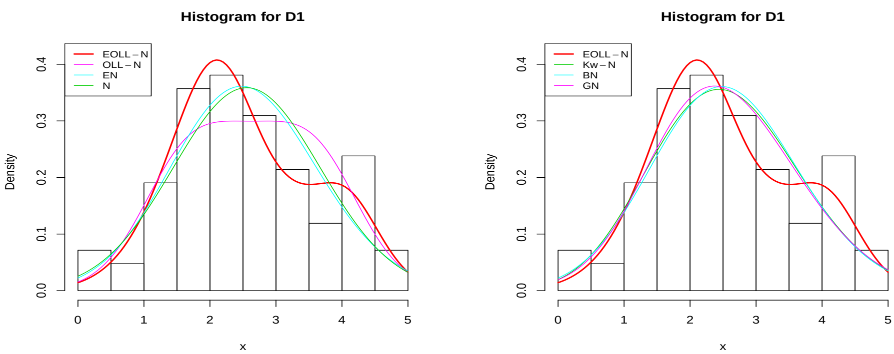

Figure 6: Histogram and density plots for D1. (a) Plots for sub-models (b) Plots for the others models.

Table 3: The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D2.

Model	μ	σ	a	b	-log(L)	W^*	A^*	K-S	p-value
EOLLLN	99.766	5.175	0.129	1.946	132.203	0.048	0.299	0.085	0.974
	(0.002)	(0.002)	(0.017)	(0.344)					
OLL-N	106.197	4.387	0.126		134.786	0.097	0.552	0.216	0.099
	(0.002)	(0.002)	(0.018)						
Normal	111.753	17.907			137.733	0.188	1.083	0.143	0.523
	(3.165)	(2.238)							
EN	142.527	2.704	0.011		135.613	0.101	0.623	0.112	0.816
	(4.026)	(1.780)	(0.015)						
BN	161.943	14.879	0.202	29.070	135.010	0.061	0.425	0.092	0.947
	(48.866)	(14.368)	(0.236)	(127.584)					
GN	145.939	9.060	0.102	3.932	134.915	0.020	0.182	0.056	0.998
	(0.002)	(0.002)	(0.018)	(2.288)					
KN	118.390	3.596	0.0208	0.188	134.282	0.036	0.292	0.097	0.920
	(0.174)	(0.147)	(0.008)	(0.036)					

15.3 Application 3

Data 3 The third data set, denoted by D3, from Nichols and Padgett (2006) on the

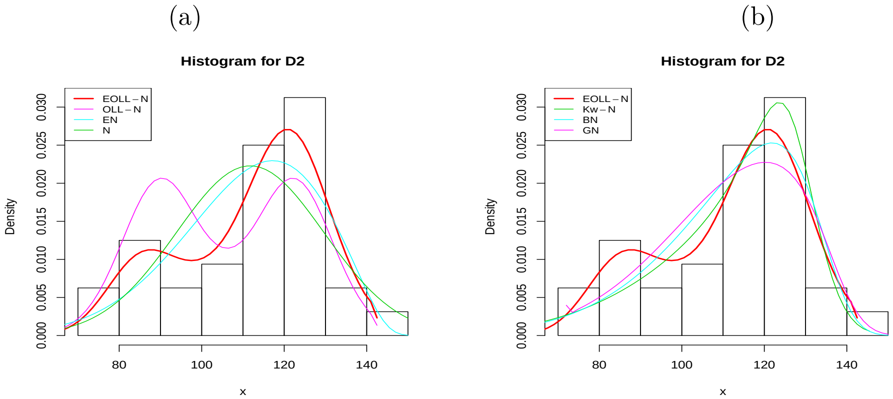

Figure 7: Histogram plots for D2 data set. (a) Plots for sub-models (b) Plots for the others models.

breaking stress of carbon fibers of 50 mm in length.These data are:

3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 3.56, 4.42, 2.41, 3.19, 3.22 ,1.69, 3.28, 3.09, 1.87, 3.15 ,4.90, 1.57, 2.67 ,2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 1.25, 4.38, 1.84, 0.39, 3.68 ,2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.89, 2.88, 2.82, 2.05, 3.65, 3.75 ,2.43, 2.95, 2.97, 3.39 ,2.96, 2.35 ,2.55, 2.59,2.03, 1.61 ,2.12, 3.15, 1.08, 2.56, 1.80, 2.53.

The MLEs of the parameters and SEs in parentheses and the goodness-of-fit statistics for D3 are presented in Table 4. We can see that the OLLLN model outperforms all the fitted competitive models.

Table 4: The MLEs of the parameters and SE in the parentheses and the goodness-of-fit test statistics for D3.

Model	μ	σ	a	b	-log(L)	W^*	A^*	K-S	p-value
EOLLLN	3.075	1.303	1.872	0.652	85.012	0.039	0.266	0.060	0.967
	(0.385)	(1.123)	(1.623)	(0.372)					
Normal	2.759	0.884			85.562	0.073	0.416	0.073	0.866
	(0.108)	(0.077)							
EN	3.061	0.781	0.669		85.513	0.068	0.398	0.073	0.863
	(0.901)	(0.320)	(0.844)						
BN	3.469	1.277	1.336	2.780	85.439	0.062	0.374	0.073	0.870
	(2.879)	(3.626)	(6.073)	(15.936)					
GN	3.348	0.976	0.848	1.664	85.439	0.061	0.374	0.073	0.866
	(3.210)	2.257)	(3.428)	(7.867)					
KN	2.672	1.712	2.699	3.587	85.405	0.060	0.365	0.072	0.881
	(4.622)	(3.073)	(10.916)	(8.805)					
OLL-N	2.772	1.465	1.787		85.272	0.06	0.341	0.062	0.957
	(0.107)	(1.460)	(1.923)						

The density and histogram plots are given in Figure 8. For D3, we see that the fitted OLLLN distribution best captures the empirical histogram.

We note that the OLLLN model outperforms all the fitted competitive models under the selected critrion for D1, D2 and D3. For all three data sets, we note that the fitted OLLLN distribution best captures the empirical histograms, especially for the third data set, which indicates the outstanding performance of the OLLLN distribution.

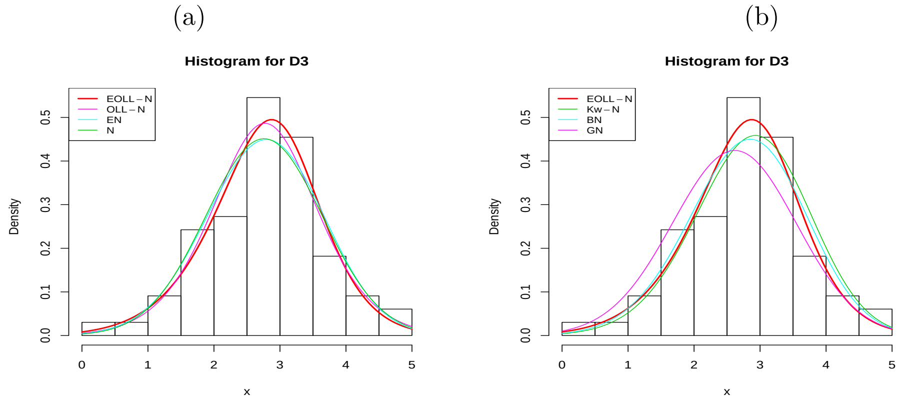

Figure 8: Histogram and density plots for D3. (a) Plots for sub-models (b) Plots for the others models.

16 Conclusions

We study a class of distributions so-called the exponentiated odd log-logistic (EOLL) family to extend several common distributions such as the normal, Weibull, gamma and beta distributions. For each distribution G, we can easily define the associated EOLL-G distribution. The density function of the proposed family can be expressed as a linear combination of exponentiated-G (exp-G) density functions. This mixture representation is important to derive several properties of the proposed family. Some of its characteristics have tractable mathematical properties such as the ordinary and incomplete moments, quantile function and order statistics. The formulae derived are manageable by using modern computer resources with analytic and numerical capabilities. The estimation of the model parameters is approached by the method of maximum likelihood and the observed information matrix is derived. We fit some OLL-G distributions to two real data sets to show the potentiality of the proposed family.

Appendix

We present four power series for the proof of the linear representation in Section 4. First, for $a > 0$ real non-integer and $|u| < 1$, we have the binomial expansion

$$(1-u)^a = \sum_{j=0}^{\infty} (-1)^j \binom{a}{j} u^j, \tag{33}$$

where the binomial coefficient is defined for any real.

Second, the following expansion holds for any $\alpha > 0$ real non-integer

$$G(x)^\alpha = \sum_{r=0}^{\infty} s_r(\alpha)\, G(x)^r, \tag{34}$$

where $s_r(\alpha) = \sum_{j=r}^{\infty} (-1)^{r+j} \binom{\alpha}{j} \binom{j}{r}$.

Third, by expanding z^λ in Taylor series, we have

$$z^\lambda = \sum_{k=0}^{\infty} (\lambda)_k \, (z-1)^k / k! = \sum_{i=0}^{\infty} f_i \, z^i \tag{35}$$

$$f_i = f_i(\lambda) = \sum_{k=i}^{\infty} \frac{(-1)^{k-i}}{k!} \binom{k}{i} (\lambda)_k$$

and $(\lambda)_k = \lambda(\lambda-1)\ldots(\lambda-k+1)$ is the descending factorial.

Fourth, we use throughout an equation of Gradshteyn and Ryzhik (2000), for a power series raised to a positive integer j given by

$$\left(\sum_{j=0}^{\infty} a_j \, v^j\right)^i = \sum_{j=0}^{\infty} c_{i,j} \, v^j, \tag{36}$$

where the coefficients $c_{i,j}$ (for $j = 1, 2, \ldots$) are easily obtained from the recurrence equation (for $j \geq 1$)

$$c_{i,j} = (j a_0)^{-1} \sum_{m=1}^{j} [m(j+1) - j] \, a_m \, c_{i,j-m}$$

and $c_{i,0} = a_0^i$. Hence, the coefficients $c_{i,j}$ can be calculated directly from $c_{i,0}, \ldots, c_{i,j-1}$ and, therefore, from $a_0, \ldots, a_j$. They can be given explicitly in terms of the a_j's, although it is not necessary for programming numerically our expansions in any algebraic or numerical software.

We now obtain an expansion for $[G(x)^a + \bar{G}(x)^a]^c$. We can write from equations (33) and (34)

$$[G(x)^a + \bar{G}(x)^a] = \sum_{j=0}^{\infty} t_j \, G(x)^j,$$

where $t_j = t_j(a) = s_j(a) + (-1)^j \binom{a}{j}$. Then, using (35), we can write

$$[G(x)^a + \bar{G}(x)^a]^c = \sum_{i=0}^{\infty} f_i \left(\sum_{j=0}^{\infty} t_j \, G(x)^j\right)^i,$$

where $f_i = f_i(c)$. Finally, using equations (36) and (14), we obtain

$$[G(x)^a + \bar{G}(x)^a]^c = \sum_{j=0}^{\infty} h_j \, G(x)^j, \tag{37}$$

where $h_j = h_j(a, c) = \sum_{i=0}^{\infty} f_i \, m_{i,j}$ and $m_{i,j} = (j \, t_0)^{-1} \sum_{m=1}^{j} [m(j+1) - j] \, t_m \, m_{i,j-m}$ (for $j \geq 1$) and $m_{i,0} = t_0^i$.

Acknowledgement

The authors are grateful to the editor, associate editor and reviewers for making helpful comments and suggestions.

References

Aarset, M.V. (1987). How to identify bathtub hazard rate. *IEEE Transactions Reliability*, **36**, 106–108.

Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992). A First Course in Order Statistics. John Wiley, New York.

Armitage, P. and Berry, G. (1987). Statistical Methods in Medical Research.*Blackwell Scientific Publications, Oxford.*

Balakrishnan, N. (1985). Order statistics from the half logistic distribution. *Journal of Computational Analysis and Applications*, **20**, 287–309.

Blischke, W. R., and Murthy, D. P. (2011). Reliability: modelling, prediction, and optimization (Vol. 767). John Wiley and Sons.

Cooray, K.(2006). Generalization of the Weibull distribution: the odd Weibull family. *Statistical Modelling*, **6**(3), 265-277.

Cordeiro, G. M., and de Castro, M. (2011). A new family of generalized distributions. Journal of statistical computation and simulation, 81(7), 883-898.

Cordeiro, G. M. and Nadarajah, S. (2011). Closed-form expressions for moments of a class of beta generalized distributions.*Brazilian journal of probability and statistics*, **25**(1), 14-33.

Dagum, C. (1975). A model of income distribution and the conditions of existence of moments of finite order. *of the International Statistical Institute*, **46**, 199-205.

David, H.A. and Nagaraja, H.N. (2003). Order Statistics. John Wiley, Hoboken, New Jersey.

Doornik, J. F. (1996). Object-oriented matrix programming using Ox. *International Thomson Business Press.*

El-Gohary, A., Alshamrani, A. and Al-Otaibi, A. N. (2013). The generalized Gompertz distribu- tion. *Applied Mathematical Modelling*, **37**, 13–24.

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, **31**, 497–512.

Gleaton, J. U., Lynch, J. D. (2006). Properties of generalized log-logistic families of lifetime distributions.*Journal of Probability and Statistical Science*, **4(1)**, 51–64.

Gradshteyn I.S. and Ryzhik I.M. (2000). Table of integrals, series, and products. Academic Press, San Diego.

Gupta, R. D. and Kundu, D. (1999). Generalized Exponential Distributions. *Australian and New Zealand Journal of Statistics*, **41**, 173–188.

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

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

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, **XX**, 299–302.

Nadarajah, S. and Kotz, S. (2006). The exponentiated type distribution. *Acta. Appl. Math.*, **92**, 97–111.

Nadarajah, S., Cordeiro, G.M. and Ortega, E.M.M. (2013). The gamma-G family of distributions: Mathematical properties and applications. Communications in Statistics. Theory and Methods. Accepted.

Nichols, M. D., and Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and reliability engineering international*, f 22(2), 141-151.

Rényi, A. (1961). On measures of entropy and information. HUNGARIAN ACADEMY OF SCIENCES Budapest Hungary.

Shannon, C.E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal*, **30**, 50–64.

Zografos, K. and Balakrishnan N. (2009). On families of beta and generalized gamma-generated distributions and associated inference, *Statistics*, **23**, 344–362.