



The Ristic-Balakrishnan odd log-logistic family of distributions: Properties and Applications

Hamid Esmaeili¹, Fazlollah Lak^{1,*}, Emrah Altun²

¹*Department of Statistics, Persian Gulf University, Bushehr, Iran.*

²*Department of Statistics, Bartin University, Bartin, Turkey*

Abstract We introduce and study general mathematical properties of a new generator of continuous distributions with two extra parameters called the *Ristic-Balakrishnan odd log-logistic family of distributions*. We present some special models. We investigate the asymptotes. 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, generating functions and order statistics, which hold for any baseline model, are determined. Further, we discuss the estimation of the model parameters by maximum likelihood and we studied a simulation study based on maximum likelihood estimation. A regression model based on the proposed model was introduced. We illustrate the potentiality of the family utilizing three applications to real data.

Keywords Estimation, Ristic-Balakrishnan distributions, Generated family, Maximum likelihood, Mean deviation, Moment, Quantile function

AMS 2010 subject classifications 60E05, 62F10, 62G05

DOI: 10.19139/soic-2310-5070-715

1. Introduction

The statistics literature is filled with hundreds of continuous univariate distributions: see Johnson *et al.* [13], [14]. Adding parameters to a well-known distribution is widely used to obtain a more flexible new family of distributions. The different characteristics of the data sets have been made to define new generated families of distributions to control skewness and kurtosis through the tail weights and provide much more flexibility in modeling skewed data, including the two-piece approach introduced by Hansen [12] and the generators pioneered by Eugene *et al.* [6], Cordeiro and de Castro [5] and Alexander *et al.* [2]. Many articles apply these techniques to induce skewness into well-known symmetric distributions such as the symmetric Student *t*; see, Aas and Haff [1], for a review.

We study several general mathematical properties of the *gamma generated* (“Gamma-G” for short) family of distributions. This family was motivated by a pioneered work by Ristic and Balakrishnan [23]. It is also important to mention that the results presented in this paper follow similar lines of the results developed by Nadarajah *et al.* [19], although their model is completely different from that one discussed in this paper.

The proposed family can extend several common distributions such as normal, Weibull and Gumbel distributions by adding two extra generator parameters. Indeed, for any baseline *G* distribution, we can define the associated *RBOLL-G* (“RBOLL-G”) distribution. The main characteristics of the GE family, such as moments and generating and quantile functions, have tractable mathematical properties. The role of the generator parameters has been investigated and is related to the skewness and kurtosis of the generated distribution. The family studied here can

*Correspondence to: Fazlollah Lak (Email: fazlollahlak@gmail.com). Department of Statistics, Persian Gulf University, Bushehr, Iran.

be considered a special subfamily of that one proposed recently by Ristic and Balakrishnan [23] (“RB-G” from now on). The Gamma-G family can be constructed as follows. Let $G(x)$ be any continuous distribution defined on a finite or an infinite interval. The RB family of distributions is defined from the cumulative distribution function (cdf) (for $\alpha > 0$)

$$F_{RB-G}(x) = \frac{1}{\Gamma(\beta)} \int_0^{-\log[1-G(x;\boldsymbol{\tau})]} t^{\beta-1} e^{-t} dt, \quad x \in \mathbb{R}, \quad (1)$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ denotes the gamma function, $G(x; \boldsymbol{\tau})$ ($x \in \mathbb{R}$), denote the baseline cdf and $\boldsymbol{\tau}$ denotes the parameters in the parent $G(\cdot)$.

Gleatn and Lynch [7], introduced odd log-logistic family (OLL-G) with cdf

$$F_{OLL-G}(x; \alpha, \boldsymbol{\tau}) = \frac{G(x; \boldsymbol{\tau})^\alpha}{G(x; \boldsymbol{\tau})^\alpha + \bar{G}(x; \boldsymbol{\tau})^\alpha}, \quad (2)$$

where $\alpha > 0$ and $\bar{G}(x; \boldsymbol{\tau}) = 1 - G(x; \boldsymbol{\tau})$ is the baseline survival function. By compounding RB-G and OLL-G we define the cdf given by

$$\begin{aligned} F(x; \alpha, \beta, \boldsymbol{\tau}) &= 1 - \frac{1}{\Gamma(\beta)} \int_0^{-\log\left[\frac{G(x;\boldsymbol{\tau})^\alpha}{G(x;\boldsymbol{\tau})^\alpha + \bar{G}(x;\boldsymbol{\tau})^\alpha}\right]} t^{\beta-1} e^{-t} dt \\ &= 1 - \frac{1}{\Gamma(\beta)} \gamma\left\{\beta, -\log\left[\frac{G(x;\boldsymbol{\tau})^\alpha}{G(x;\boldsymbol{\tau})^\alpha + \bar{G}(x;\boldsymbol{\tau})^\alpha}\right]\right\}. \end{aligned} \quad (3)$$

where $\gamma(\beta, z) = \int_0^z t^{\beta-1} e^{-t} dt$ denotes the incomplete gamma function. The model (3) is called the Ristic-Balakrishnan Odd Log-logistic (RBOLL-G for short) distribution with $\alpha > 0$ and $\beta > 0$.

Let $g(x; \boldsymbol{\tau}) = dG(x; \boldsymbol{\tau})/dx$ be the corresponding baseline probability density function (pdf). The density function corresponding to (3) becomes

$$f(x; \alpha, \beta, \boldsymbol{\tau}) = \frac{\alpha g(x; \boldsymbol{\tau}) G(x; \boldsymbol{\tau})^{\alpha-1} \bar{G}(x; \boldsymbol{\tau})^{\alpha-1}}{\Gamma(\beta)[G(x; \boldsymbol{\tau})^\alpha + \bar{G}(x; \boldsymbol{\tau})^\alpha]^2} \left\{-\log\left[\frac{G(x; \boldsymbol{\tau})^\alpha}{G(x; \boldsymbol{\tau})^\alpha + \bar{G}(x; \boldsymbol{\tau})^\alpha}\right]\right\}^{\beta-1}. \quad (4)$$

The hazard rate function (hrf) of X is given by

$$h(x; \alpha, \beta, \boldsymbol{\tau}) = \frac{\alpha g(x; \boldsymbol{\tau}) G(x; \boldsymbol{\tau})^{\alpha-1} \bar{G}(x; \boldsymbol{\tau})^{\alpha-1}}{[G(x; \boldsymbol{\tau})^\alpha + \bar{G}(x; \boldsymbol{\tau})^\alpha]^2} \times \frac{\left\{-\log\left[\frac{G(x; \boldsymbol{\tau})^\alpha}{G(x; \boldsymbol{\tau})^\alpha + \bar{G}(x; \boldsymbol{\tau})^\alpha}\right]\right\}^{\beta-1}}{\gamma\left\{\beta, -\log\left[\frac{G(x; \boldsymbol{\tau})^\alpha}{G(x; \boldsymbol{\tau})^\alpha + \bar{G}(x; \boldsymbol{\tau})^\alpha}\right]\right\}} \quad (5)$$

The RBOLL-G distribution has the same parameters of the G distribution plus two additional parameters α and β . From now on, a random variable X with density function (4) is denoted by $X \sim \text{RBOLL-G}(\alpha, \beta, \boldsymbol{\tau})$. For $\alpha = \beta = 1$ the RBOLL-G distribution reduces to the baseline G distribution, for $\alpha = 1$ we obtain RB-G distributions and for $\beta = 1$ we obtain a Odd log logistic (OLL-G) distributions.

Each new RBOLL-G distribution can be obtained from a specified G distribution. From the statistical modeling point of view, the RBOLL-G distribution has two important aspects. First, the proposed model has more parameters than the baseline distribution and the additional parameters α and β of the generated model have clear interpretations. In fact, the RBOLL-G family of distributions is easily simulated by inverting (3) as follows: if V has $U(0, 1)$ distribution, $Q_G(\cdot)$ denote the quantile function of baseline G and $Q_{\Gamma(\beta, 1)}(\cdot)$ denote the quantile function of $\Gamma(\beta, 1)$ random variable, then solution of the nonlinear equation

$$X_V = Q_G \left\{ \frac{e^{-\frac{1}{\alpha} Q_{\Gamma(\beta, 1)}(1-V)}}{e^{-\frac{1}{\alpha} Q_{\Gamma(\beta, 1)}(1-V)} + [1 - e^{-Q_{\Gamma(\beta, 1)}(1-V)}]^\frac{1}{\alpha}} \right\}. \quad (6)$$

The parameters α and β have a clear interpretation. Following the key idea of Zografos and Balakrishnan [25] and Ristic and Balakrishnan [23], we can also interpret (4) in this way: if $X_{L(1)}, X_{L(2)}, \dots, X_{L(n)}$ are lower record values from a sequence of independent random variables with common pdf

$$g(x; \alpha, \tau) = \frac{\alpha g(x; \tau) \{G(x; \tau)[1 - G(x; \tau)]\}^{\alpha-1}}{\{G(x; \tau)^\alpha + \bar{G}(x; \tau)^\alpha\}^2},$$

then the pdf of the n th lower record value has the pdf (4).

The aim of this paper is to derive several mathematical properties of (3) and (4) in the most simple, explicit and general forms. We obtain general expressions for asymptotic properties of (3), (4) and (5), ordinary and incomplete moments, moment generating function (mgf), probability weighted moments, mean deviations and general properties of the order statistics. The rest of the paper is organized as follows. In Section 2, we present some new distributions. A range of mathematical properties of the RBOLL-G model (4) is derived in Sections 3. Estimation of the model parameters by maximum likelihood and the Fisher information matrix are presented in Section 4. In section 5, we studied a regression model based on proposed model. In section 6, three applications to real data illustrate the performance of the new family. Finally, some conclusions and future work are noted in Section 7.

2. Special RBOLL-G distributions

The RBOLL-G family of density functions (4) can be used to extend several well-known distributions to provide more flexibility in tails which can be applied in many scientific fields such as biology, actuarial and engineering. When the cdf, $G(x; \tau)$, and pdf, $g(x; \tau)$, of the baseline distribution have simple mathematical forms, the density in (4) will be msot tractable.

2.1. Ristic-Balakrishnan odd log-logistic Weibull (RBOLL-W) distribution

If $G(x; \tau)$ is the Weibull cdf with scale parameter $\gamma > 0$ and shape parameter $\lambda > 0$, where $\tau = (\lambda, \gamma)^T$, say $G(x; \tau) = 1 - \exp\{-(x/\lambda)^\gamma\}$, the RBOLL-W density function (for $x > 0$) reduces to

$$f_{\text{RBOLL-W}}(x) = \frac{\alpha \gamma \lambda^{-\gamma} x^{\gamma-1} \exp[-(x/\lambda)^\gamma] \{1 - \exp[-(x/\lambda)^\gamma]\}^{\alpha-1} \exp[-(\alpha-1)(x/\lambda)^\gamma]}{\Gamma(\beta) \{ \{1 - \exp[-(x/\lambda)^\gamma]\}^\alpha + \exp[-\alpha(x/\lambda)^\gamma] \}^2} \times \left\{ \gamma \left(\beta, -\log \left[\frac{\exp[-\alpha(x/\lambda)^\gamma]}{\{1 - \exp[-(x/\lambda)^\gamma]\}^\alpha + \exp[-\alpha(x/\lambda)^\gamma]} \right] \right) \right\}^{\beta-1}. \tag{7}$$

Some plots of pdf and hrf related to 7 are given in Figure 1.

As shown in figure 1, the pdf of RBOLL-W, can be unimodal, bimodal or almost symmetric. The hazard rate function of RBOLL-W can be constant, decreasing, increasng or bathtub shape. It is important in reliability theory.

2.2. Ristic-Balakrishnan odd log-logistic normal (RBOLL-N) distribution

The RBOLL-N distribution is defined from (4) by taking $G(x; \tau) = \Phi(\frac{x-\mu}{\sigma})$ and $g(x; \tau) = \frac{1}{\sigma} \phi(\frac{x-\mu}{\sigma})$ to be the cdf and pdf of the normal $N(\mu, \sigma^2)$ distribution, where $\tau = (\mu, \sigma)^T$. Its density function becomes

$$f_{\text{RBOLL-N}}(x) = \frac{\alpha \phi(z) \Phi^{\alpha-1}(z) [1 - \Phi(z)]^{\alpha-1}}{\sigma \Gamma(\beta) \{ \Phi^\alpha(z) + [1 - \Phi(z)]^\alpha \}^2} \left\{ \gamma \left(\beta, -\log \left[\frac{[1 - \Phi(z)]^\alpha}{\Phi^\alpha(z) + [1 - \Phi(z)]^\alpha} \right] \right) \right\}^{\beta-1}, \tag{8}$$

where $z = \frac{x-\mu}{\sigma}$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, α and β are shape and scale parameters, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively. A random variable with density (8) is denoted by $X \sim \text{GE-N}(\alpha, \beta, \mu, \sigma^2)$. For $\mu = 0$ and $\sigma = 1$, we obtain the RBOLL-standard normal (RBOLL-SN) distribution.

Some plots of pdf of 8 are given in in figure 2. . As shown in Figure 2, RBOLL-N can be symmetric, skew and bimodal pdf. It provide more flexibility than other well-known Normal extensions.

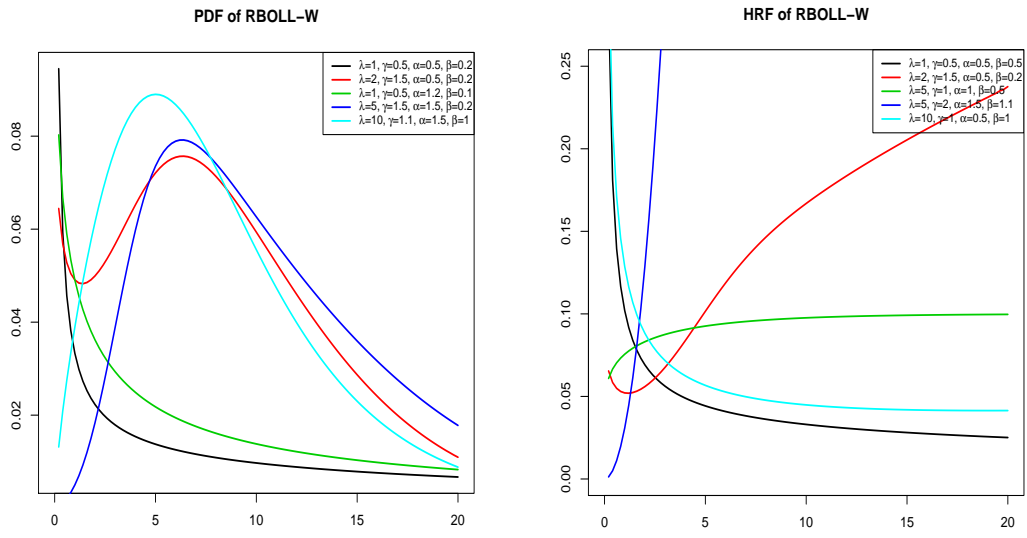


Figure 1. Plots of pdf and hrf of RBOLL-W for some selected parameters.

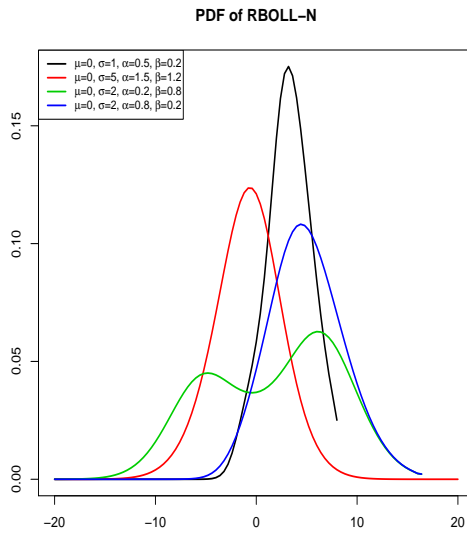


Figure 2. Plots of RBOLL-N for some selected parameters.

3. Main Properties

In this section, we study some general properties of RBOLL-G family.

3.1. Asymptotics

Proposition 1

Let $\delta = \min\{x|G(x) > 0\}$ the asymptotics of equations (3), (4) and (5) as $x \rightarrow \delta$ are given by

$$\begin{aligned} F(x) &\sim \frac{1}{\Gamma(\beta + 1)} \{-\alpha \log [\bar{G}(x)]\}^\beta \quad \text{as } x \rightarrow \delta, \\ f(x) &\sim \frac{\alpha}{\Gamma(\beta)} g(x) G(x)^{\alpha\beta-1} \quad \text{as } x \rightarrow \delta, \\ h(x) &\sim \frac{\alpha}{\Gamma(\beta)} g(x) G(x)^{\alpha\beta-1} \quad \text{as } x \rightarrow \delta. \end{aligned}$$

Proposition 2

The asymptotics of equations (3), (4) and (5) as $x \rightarrow \infty$ are given by

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

3.2. Expansions for pdf and cdf

Using the exponentiated distributions, we derive the expansions of (4) and (3) which is required to obtain the statistical properties of the RBOLL-G family.

For a baseline cdf $G(x)$, if the random variable follows a exponentiated- G distribution with parameter $a > 0$, say $X \sim \text{exp-}G(a)$, if its pdf and cdf are

$$h_a(x) = aG(x)g(x)^{a-1} \tag{9}$$

and

$$H_a(x) = G(x)^a, \tag{10}$$

respectively. The properties of exponentiated distributions have been studied by many authors in recent years, see Mudholkar and Srivastava [17] for exponentiated Weibull, Gupta *et al.* [9] for exponentiated Pareto, Gupta and Kundu [10] for exponentiated exponential, Nadarajah [20] for exponentiated Gumbel, Kakde and Shirke [15] for exponentiated lognormal, and Nadarajah and Gupta [21] for exponentiated gamma distributions. Also, we can refer to recent articles by Mozafari *et al.* [16] and Alizadeh *et al.* [3]. The binomial coefficient generalized to real arguments is given by $\binom{x}{y} = \Gamma(x + 1)/[\Gamma(y + 1)\Gamma(x - y + 1)]$.

First using Taylor expansion we have

$$F(x) = 1 - \frac{1}{\Gamma(\beta)} \sum_{i=0}^{\infty} \frac{(-1)^i}{i!(\beta + i)} \left\{ -\log \left[\frac{G(x)^\alpha}{G(x)^\alpha + \bar{G}(x)^\alpha} \right] \right\}^{\beta+i}$$

For any real parameter $a > 0$, the following formula holds ([http:// functions.wolfram.com/ ElementaryFunctions/ Log/ 06/ 01/ 04/ 03/](http://functions.wolfram.com/ElementaryFunctions/Log/06/01/04/03/))

$$\begin{aligned} \left\{ -\log \left[1 - \frac{\bar{G}(x)^\alpha}{G(x)^\alpha + \bar{G}(x)^\alpha} \right] \right\}^{\beta+i} &= (\beta + i) \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{j+k} \binom{k - \beta - i}{k} \binom{k}{j} p_{j,k}}{(\beta + i - j)} \\ &\times \left[\frac{\bar{G}(x)^\alpha}{G(x)^\alpha + \bar{G}(x)^\alpha} \right]^{\beta+i+k}, \end{aligned} \tag{11}$$

where the constants $p_{j,k}$ can be calculated recursively by

$$p_{j,k} = k^{-1} \sum_{m=1}^k [k - m(j+1)] c_m p_{j,k-m} \quad (12)$$

for $k = 1, 2, \dots, p_{j,0} = 1$ and $c_k = (-1)^{k+1} (k+1)^{-1}$.

Also

$$\left[\frac{\bar{G}(x)^\alpha}{G(x)^\alpha + \bar{G}(x)^\alpha} \right]^{\beta+i+k} = \frac{\sum_{r=0}^{\infty} \lambda_r G(x)^r}{\sum_{r=0}^{\infty} \rho_r G(x)^r} = \sum_{r=0}^{\infty} a_r G(x)^r$$

where

$$\lambda_r = (-1)^r \binom{\alpha(\beta+i+k)}{r} \quad (13)$$

$$\rho_r = h_r(\alpha, \beta+i+k) \quad (14)$$

$$a_r = a_r(\alpha, \beta, i, k) = \frac{1}{\rho_0} \left[\rho_r - \frac{1}{\rho_0} \sum_{s=1}^r \rho_s a_{r-s} \right], \text{ for } r \geq 1 \quad (15)$$

, $a_0 = \frac{\lambda_0}{\rho_0}$ and $h_r(\alpha, \beta+i+k)$ is defined in Appendix. and then (3) can be expressed as

$$F(x) = \sum_{r=0}^{\infty} b_r H_r(x), \quad (16)$$

where

$$b_0 = 1 - \frac{1}{\Gamma(\beta)} \sum_{i,k=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k} \binom{k-\beta-i}{k} \binom{k}{j} p_{j,k} a_0(\alpha, \beta, i, k)}{i!(\beta+i-j)},$$

$$b_r = \frac{1}{\Gamma(\beta)} \sum_{i,k=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k+1} \binom{k-\beta-i}{k} \binom{k}{j} p_{j,k} a_r(\alpha, \beta, i, k)}{i!(\beta+i-j)},$$

for $r \geq 1$ and $H_r(x)$ denotes the cdf of the exp- $G(r)$ distribution. The corresponding (4) can be expressed as

$$f(x) = \sum_{r=0}^{\infty} b_{r+1} h_{r+1}(x), \quad (17)$$

where $h_{r+1}(x)$ denotes the pdf of the exp- $G(r+1)$ distribution. So, several properties of the gamma- G distribution can be obtained by knowing those of the exp- G distribution, see, for example, Mudholkar *et al.* [18], Gupta and Kundu [11] and Nadarajah and Kotz [22], among others.

3.3. Moments

From now on, we 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 (4). The moments of the RBOLL- G distribution can be obtained from

the (r, k) th probability weighted moments (PWMs) of Y defined by

$$\tau_{s,k} = E[X^s G(X)^k] = \int_{-\infty}^{\infty} x^s G(x)^k g(x) dx. \quad (18)$$

In fact, we have

$$\mu'_s = E(X^s) = \sum_{r=0}^{\infty} (r+1) b_{r+1} \tau_{s,k}. \quad (19)$$

where b_r is defined by equation (16). Thus, the moments of any RBOLL-G distribution can be expressed as an infinite linear combination of the baseline PWMs.

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

$$\tau_{s,k} = \int_0^1 Q_G(u)^s u^k du. \quad (20)$$

The PWMs for various distributions can be determined by using equations (18) and (20). The following special cases were already published by Cordeiro and Nadarajah [5].

3.4. Incomplete moments

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

$$m_n(y) = \sum_{r=0}^{\infty} b_{r+1} \int_0^{G(y; \xi)} Q_G(u)^n u^r du. \quad (21)$$

The integral in (21) can be computed at least numerically for most baseline distributions.

3.5. Generating function

In this section, we provide two formulae for the moment generating function (mgf) $M(s) = E(e^{sX})$ of a random variable X with the RBOLL-G distribution. A first formula for $M(s)$ comes from equation (17) as

$$M(t) = \sum_{r=0}^{\infty} b_{r+1} M_{r+1}(t), \quad (22)$$

where $M_{r+1}(t)$ is the generating function of the exp-G distribution with power parameter $t + 1$. 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

$$M(t) = \sum_{r=0}^{\infty} (r+1) b_{r+1} \rho_r(s), \quad (23)$$

where the quantity $\rho_r(s) = \int_{-\infty}^{\infty} \exp(tx) G(x)^r g(x) dx$ follows from the baseline qf as

$$\rho_r(s) = \int_0^1 \exp[s Q_G(u)] u^r du. \quad (24)$$

3.6. 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

$$\delta_1(Y) = 2\mu'_1 F(\mu'_1) - 2m_1(\mu'_1) \quad \text{and} \quad \delta_2(Y) = \mu'_1 - 2m_1(M), \quad (25)$$

respectively, where $M = Q_G \left\{ \frac{e^{-\frac{1}{\alpha} Q_{\Gamma(\beta,1)}(0.5)}}{e^{-\frac{1}{\alpha} Q_{\Gamma(\beta,1)}(0.5)} + [1 - e^{-Q_{\Gamma(\beta,1)}(0.5)}]^{-\frac{1}{\alpha}}} \right\}$ is the median of Y , $\mu'_1 = E(Y)$ comes from equation (19), $F(\mu'_1)$ is easily calculated from equation (3) 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 (17) as

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

where

$$J_{r+1}(z) = \int_{-\infty}^z x h_{r+1}(x) dx. \quad (27)$$

Equation (27) is the basic quantity to compute the mean deviations for the RBOLL-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

$$\delta_1(Y) = 2\mu'_1 F(\mu'_1) - 2 \sum_{r=0}^{\infty} b_{r+1} J_{r+1}(\mu'_1) \quad \text{and} \quad \delta_2(Y) = \mu'_1 - 2 \sum_{r=0}^{\infty} b_{r+1} J_{r+1}(M).$$

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

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

where $T_r(z)$ is given by

$$T_r(z) = \int_0^{G(z)} Q_G(u) u^r du.$$

3.7. Order statistics

Let X_1, \dots, X_n be a random sample from the RBOLL-G family. Denote the random variables in the ascending order as $X_{1:n} \leq \dots \leq X_{n:n}$. The pdf of $X_{i:n}$ is given by

$$\begin{aligned} f_{i:n}(x) &= K f(x) F^{i-1}(x) \{1 - F(x)\}^{n-i} = K \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} f(x) F(x)^{j+i-1} \\ &= \sum_{r,k=0}^{\infty} \sum_{j=0}^{n-i} m_{j,r,k} h_{r+k+1}(x), \end{aligned} \quad (28)$$

where $K = n! / [(i-1)!(n-i)!]$, $h_{r+k+1}(x)$ denotes the exp-G density function with parameter $r+k+1$ and

$$m_{j,r,k} = \frac{(-1)^j n!}{(i-1)!(n-i-j)! j!} \frac{(r+1) b_{r+1} f_{j+i-1,k}}{[r+k+1]}, \quad (29)$$

where b_k is defined by (16). Here, the quantities $f_{j+i-1,k}$ are obtained recursively by $f_{j+i-1,0} = b_0^{j+i-1}$ and (for $k \geq 1$)

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

Thus, one can easily obtain ordinary and incomplete moments and generating function of order statistics for any given G.

4. Estimation and Simulation

The maximum likelihood estimation (MLE) method is preferred to obtain the unknown parameters of the RBOLL-G family for any specific baseline distribution.

4.1. Estimation

Let x_1, \dots, x_n be observed values from the RBOLL-G distribution with parameters α, β and τ . Let $\Theta = (\alpha, \beta, \tau)^\top$ be the $r \times 1$ parameter vector. The total log-likelihood function for Θ is given by

$$\begin{aligned} \ell_n &= \ell_n(\Theta) = n \log(\alpha) - n \log[\Gamma(\beta)] + \sum_{i=1}^n \log[g(x_i; \tau)] \\ &+ (\alpha - 1) \sum_{i=1}^n \log[G(x_i; \tau)\bar{G}(x_i; \tau)] - 2 \sum_{i=1}^n \log[G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha] \\ &+ (\beta - 1) \sum_{i=1}^n \log \left\{ -\log \left[\frac{G(x_i; \tau)^\alpha}{G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha} \right] \right\} \end{aligned} \tag{30}$$

The log-likelihood function can be maximized either directly by using the SAS (PROC NLMIXED) or the Ox program (sub-routine MaxBFGS) or by solving the nonlinear likelihood equations obtained by differentiating (30). The components of the score function $U_n(\Theta) = (\partial \ell_n / \partial \alpha, \partial \ell_n / \partial \beta, \partial \ell_n / \partial \tau)^\top$ are

$$\begin{aligned} \frac{\partial \ell_n}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \log[G(x_i; \tau)\bar{G}(x_i; \tau)] - 2 \sum_{i=1}^n \frac{G(x_i; \tau)^\alpha \log[G(x_i; \tau)] + \bar{G}(x_i; \tau)^\alpha \log[\bar{G}(x_i; \tau)]}{G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha} \\ &- (\beta - 1) \sum_{i=1}^n \frac{G(x_i; \tau)^\alpha \log[G(x_i; \tau)]}{[G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha] \log \left[1 - \frac{G(x_i; \tau)^\alpha}{G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha} \right]} \\ \frac{\partial \ell_n}{\partial \beta} &= -n \psi(\beta) + \sum_{i=1}^n \log \left\{ -\log \left[\frac{G(x_i; \tau)^\alpha}{G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha} \right] \right\} \\ \frac{\partial \ell_n}{\partial \tau} &= \sum_{i=1}^n \frac{\dot{g}(x_i; \tau)]_\tau}{g(x_i; \tau)} + (\alpha - 1) \sum_{i=1}^n \frac{[\dot{G}(x_i; \tau)]_\tau}{G(x_i; \tau)} + (1 - \alpha) \sum_{i=1}^n \frac{G^{(\tau)}(x_i; \tau)}{\bar{G}(x_i; \tau)} \\ &- \alpha \sum_{i=1}^n [\dot{G}(x_i; \tau)]_\tau \frac{G(x_i; \tau)^{\alpha-1} - \bar{G}(x_i; \tau)^{\alpha-1}}{G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha} \\ &- \alpha(\beta - 1) \sum_{i=1}^n G^{(\tau)}(x_i; \tau) \frac{G(x_i; \tau)^{\alpha-1} \bar{G}(x_i; \tau)^{\alpha-1}}{[G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha]^2 \log \left[\frac{G(x_i; \tau)^\alpha}{G(x_i; \tau)^\alpha + \bar{G}(x_i; \tau)^\alpha} \right]} \end{aligned}$$

where $g^{(\tau)}(\cdot)$ means the derivative of the function g with respect to τ .

where

$$[\dot{g}(x_i; \boldsymbol{\tau})]_{\boldsymbol{\tau}} = \frac{dg(x_i; \boldsymbol{\tau})}{d\boldsymbol{\tau}}, \quad [\dot{G}(x_i; \boldsymbol{\tau})]_{\boldsymbol{\tau}} = \frac{dG(x_i; \boldsymbol{\tau})}{d\boldsymbol{\tau}},$$

the functions $g(\cdot)$ and $G(\cdot)$ are defined in Section 1 and $\psi(\cdot)$ is the digamma function.

The MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by solving the nonlinear likelihood equations $U_{\alpha}(\boldsymbol{\theta}) = 0$, $U_{\beta}(\boldsymbol{\theta}) = 0$ and $U_{\boldsymbol{\tau}}(\boldsymbol{\theta}) = 0$. These equations cannot be solved analytically and statistical software can be used to solve them numerically. We can use iterative techniques such as a Newton-Raphson type algorithm to obtain the estimate $\hat{\boldsymbol{\theta}}$. We employ the numerical procedure NLMixed in SAS.

For interval estimation of $(\alpha, \beta, \boldsymbol{\tau})$ and hypothesis tests on these parameters we obtain the observed information matrix since the expected information matrix is very complicated and requires numerical integration. The $(p+2) \times (p+2)$ observed information matrix $J(\boldsymbol{\theta})$, where p is the dimension of the vector $\boldsymbol{\tau}$, becomes

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

whose elements can compute easily.

Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $N_{p+2}(\mathbf{0}, I(\boldsymbol{\theta})^{-1})$, where $I(\boldsymbol{\theta})$ is the expected information matrix. The multivariate normal $N_{p+2}(\mathbf{0}, J(\hat{\boldsymbol{\theta}})^{-1})$ distribution, where $I(\boldsymbol{\theta})$ is replaced by $J(\hat{\boldsymbol{\theta}})$, i.e., the observed information matrix evaluated at $\hat{\boldsymbol{\theta}}$, can be used to construct approximate confidence intervals for the individual parameters.

We can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain likelihood ratio (LR) statistics for testing some sub-models of the Ga-G distribution. Tests of the hypotheses of the type $H_0 : \boldsymbol{\psi} = \boldsymbol{\psi}_0$ versus $H : \boldsymbol{\psi} \neq \boldsymbol{\psi}_0$, where $\boldsymbol{\psi}$ is a subset of parameters of $\boldsymbol{\theta}$, can be performed through LR statistics. For example, we may use the LR statistic to check if the fit using the Ga-G distribution is statistically "superior" to a fit using the G distribution for a given data set.

4.2. Simulation

This section contains the simulation results of the RBOLL-W distribution for MLE method. The finite sample performance of the MLE method for the parameters of the RBOLL-W distribution is discussed via simulation study. The simulation results are interpreted based on the results of estimated bias mean square error (MSE), estimated average length (AL) and coverage probability (CP). The simulation replication is determined as $N = 1,000$ and sample size is $n = 50, 55, \dots, 1000$. The true parameter values are $\alpha = 2, \beta = 3, \lambda = 5, \gamma = 7$. The MLEs and corresponding standard errors of the parameters are obtained for each generated sample and denoted as $(\hat{\alpha}_i, \hat{\beta}_i, \hat{\lambda}_i, \hat{\gamma}_i)$, $(s_{\hat{\alpha}_i}, s_{\hat{\beta}_i}, s_{\hat{\lambda}_i}, s_{\hat{\gamma}_i})$, respectively, where $i = 1, \dots, N$. The required formulas of the bias, MSE, AL and CP are given below

$$\widehat{Bias}_{\epsilon}(n) = \frac{1}{N} \sum_{i=1}^N (\hat{\epsilon}_i - \epsilon), \quad \widehat{MSE}_{\epsilon}(n) = \frac{1}{N} \sum_{i=1}^N (\hat{\epsilon}_i - \epsilon)^2,$$

$$CP_{\epsilon}(n) = \frac{1}{N} \sum_{i=1}^N I(\hat{\epsilon}_i - 1.95996s_{\hat{\epsilon}_i}, \hat{\epsilon}_i + 1.95996s_{\hat{\epsilon}_i}), \quad AL_{\epsilon}(n) = \frac{3.919928}{N} \sum_{i=1}^N s_{\hat{\epsilon}_i}.$$

where $\epsilon = \alpha, \beta, \lambda, \gamma$.

The results of the above metrics are computed and displayed in Figure 3. We expect the estimated biases and MSEs are near the zero, CP are near 95% and AL treats a decreasing function of the sample size. The figures confirm our expectations. The estimated biases and MSEs approach the zero when the sample size increases. The CP are near the nominal value and AL decreases when the sample size increases. The similar results can be obtained for different parameter vectors and also different baseline distribution of the RBOLL-G family.

5. Log-RBOLL-W regression model

The parametric survival regression models are widely used to analyze the lifetimes of individuals with associated covariates. Here, we introduce a new parametric survival regression model based on the RBOLL-W density. Let the random variable X follows a RBOLL-W density function with four parameters $\alpha > 0$, $\beta > 0$, $\lambda > 0$ and $\gamma > 0$, given in (7). Considering

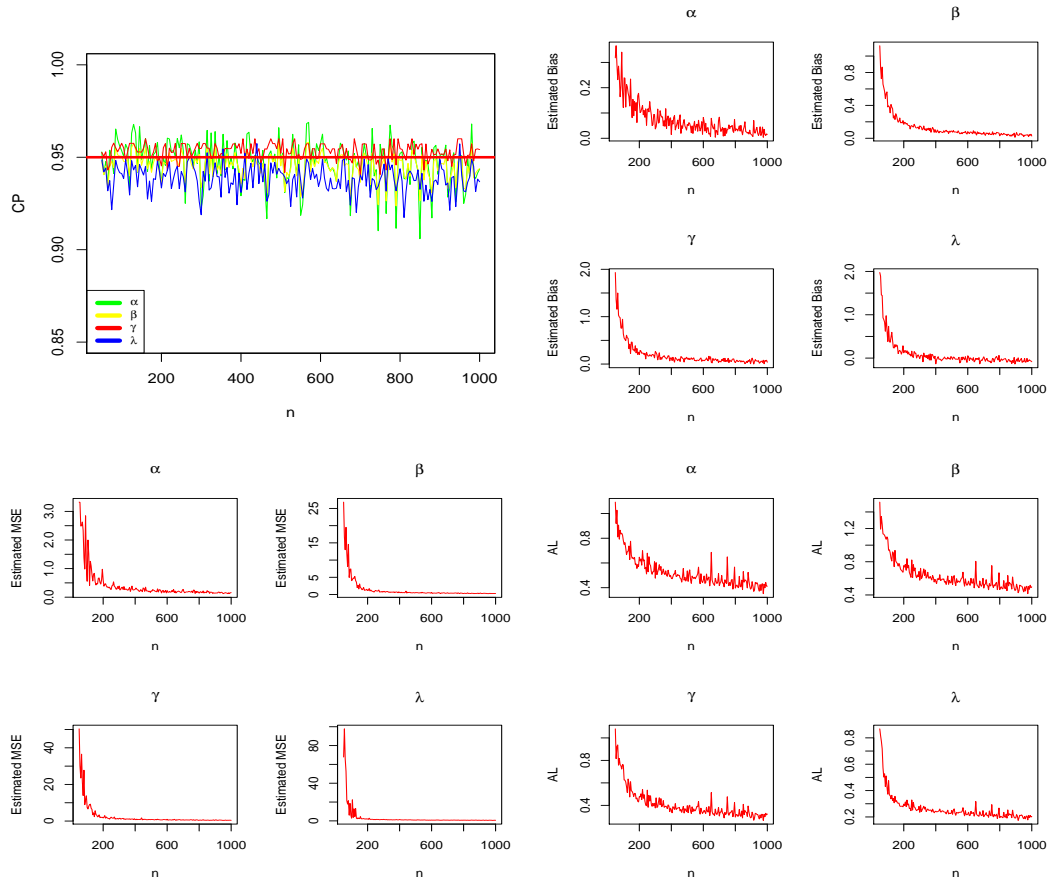


Figure 3. Estimated CPs, biases, MSEs and ALs for the selected parameters.

$Y = \log(X)$ transformation with $\lambda = 1/\sigma$ and $\gamma = \exp(\mu)$ re-parametrizations, we have (for $y \in \mathfrak{R}$)

$$f(y) = \frac{\frac{\alpha}{\sigma \Gamma(\beta)} \exp\left[\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right] \{1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)]\}^{\alpha-1} (2 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)])^{\alpha-1}}{\left[\{1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)]\}^\alpha + (2 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)])^\alpha\right]^2} \times \left\{ -\log \left[\frac{\{1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)]\}^\alpha}{\{1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)]\}^\alpha + (2 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)])^\alpha} \right] \right\}^{\beta-1}, \quad (31)$$

where $\mu \in \mathfrak{R}$ is the location parameter and $\sigma > 0$ is the scale parameter. The parameters $\alpha > 0$ and $\beta > 0$ are the shape parameter. The density in (31) is denoted as $Y \sim \text{LRBOLLW}(\alpha, \beta, \sigma, \mu)$. The survival function of the LRBOLLW density is

$$S(y) = \frac{1}{\Gamma(\beta)} \gamma \left(\beta, -\log \left[\frac{\{1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)]\}^\alpha}{\{1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)]\}^\alpha + (2 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)])^\alpha} \right] \right) \quad (32)$$

Let $Z = (Y - \mu)/\sigma$ be a standardized random variable. Its density is

$$f(z) = \frac{\frac{\alpha}{\Gamma(\beta)} \exp[z - \exp(z)] \{1 - \exp[-\exp(z)]\}^{\alpha-1} (2 - \exp[-\exp(z)])^{\alpha-1}}{\left[\{1 - \exp[-\exp(z)]\}^\alpha + (2 - \exp[-\exp(z)])^\alpha\right]^2} \times \left\{ -\log \left[\frac{\{1 - \exp[-\exp(z)]\}^\alpha}{\{1 - \exp[-\exp(z)]\}^\alpha + (2 - \exp[-\exp(z)])^\alpha} \right] \right\}^{\beta-1}. \quad (33)$$

Now, we propose a new location-scale regression model by linking the covariates to location of the LRBOLLW distribution by means of identity link function. Let y_i represents the response variable and $\mathbf{v}_i^\top = (v_{i1}, \dots, v_{ip})$ is the explanatory variable vector. The location-scale regression model is given by

$$y_i = \mathbf{v}_i^\top \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \dots, n, \quad (34)$$

where the random error z_i has density function (33), $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$, $\sigma > 0$, $\alpha > 0$ and $\beta > 0$ are unknown parameters. Let the response variable is defined as $y_i = \min\{\log(x_i), \log(c_i)\}$ where $\log(x_i)$ is the log-lifetime and $\log(c_i)$ is the log-censoring times. We define two sets F and C to represent the log-lifetime and log-censoring times, respectively. Under these specifications, the log-likelihood function for the parameter vector $\boldsymbol{\tau} = (\alpha, \beta, \sigma, \boldsymbol{\beta}^\top)^\top$ from model (34) has the form

$$\begin{aligned} \ell(\boldsymbol{\tau}) = & r \log\left(\frac{\alpha}{\sigma \Gamma(\beta)}\right) + c \log\left(\frac{1}{\Gamma(\beta)}\right) + \sum_{i \in F} (z_i - u_i) + (\alpha - 1) \sum_{i \in F} \log\{1 - \exp[-u_i]\} \\ & + (\alpha - 1) \sum_{i \in F} \log\{2 - \exp[-u_i]\} \\ & - 2 \sum_{i \in F} \log\left[\{1 - \exp[-u_i]\}^\alpha + \{2 - \exp[-u_i]\}^\alpha\right] \\ & + (\beta - 1) \sum_{i \in F} \log\left\{-\log\left[\frac{\{1 - \exp[-u_i]\}^\alpha}{\{1 - \exp[-u_i]\}^\alpha + \{2 - \exp[-u_i]\}^\alpha}\right]\right\} \\ & + \sum_{i \in C} \log\left(\gamma\left(\beta, -\log\left[\frac{\{1 - \exp[-u_i]\}^\alpha}{\{1 - \exp[-u_i]\}^\alpha + \{2 - \exp[-u_i]\}^\alpha}\right]\right)\right) \end{aligned} \quad (35)$$

where $\boldsymbol{\tau} = (\alpha, \beta, \sigma, \boldsymbol{\beta}^\top)^\top$ is the unknown parameter vector and $u_i = \exp(z_i)$, $z_i = (y_i - \mathbf{v}_i^\top \boldsymbol{\beta})/\sigma$ and r represents the number of uncensored observations and c is the number of censored observations. The MLE of the unknown parameter vector $\hat{\boldsymbol{\tau}}$ is obtained by direct maximization of the log-likelihood function, given in (35). The asymptotic confidence intervals of the parameters are constructed by using the observed information matrix evaluated at the MLE of $\boldsymbol{\tau}$. Note that the under standard regularity conditions of the MLE method, the asymptotic distribution of $(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})$ is $p + 3$ -variate normal distribution, denoted as $N_{p+3}(0, K(\boldsymbol{\tau})^{-1})$ where $K(\boldsymbol{\tau})$ is the expected information matrix which can be approximated by the inverse of observed information matrix.

6. Applications

This section is devoted to demonstrate the empirical importance of the RBOLL-G distributions. So, we compare the data modeling ability of the RBOLL-G distributions with other competitive distributions based on three data sets. The model selection criteria are used to compare the fitted models and verify which model gives the best fit to the data. These are estimated log-likelihood values, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) obtained for all fitted models to choose the best model. The smallest values of these statistics show the best fitted model on the data.

6.1. First application

The first data set represents the failure times for a particular wind-shield model including 85 observations that are classified as failed times of wind-shields. The used data set is given in Appendix C. The estimated parameters and their standard errors, $-\ell$, AIC and BIC values are reported in Table 1. Based on Table 1, since the RBOLL-W distribution has the lowest values of model selection criteria, it provides better results than other competitive models. The accuracy of the fitted RBOLL-W distribution is displayed in Figure 4. Figure 4(a) shows the pdfs of the fitted distributions on the histogram of the first data. Figure 4(b) shows the estimated functions of the RBOLL-W distribution such as fitted pdf, hrf, survival and probability-probability (PP) plot. These figures point out that the RBOLL-W distribution is the best model for the first data set.

The RBOLL-W distribution is compared with its sub-models based on the LR test. The LR test results are given in Table 2. These results show that the RBOLL-W distribution provides better fit than its sub-models since the p-values are less than 0.05. Here, the null hypothesis is that there is no significant difference between two models. Since the p-value is less than 0.05, we reject the null hypothesis in favour of the RBOLL-W distribution.

6.2. Second Application

The second data set contains 32 observations corresponding to the birth weights of newborn babies in ounces. The used data set is given in Appendix C. The estimated parameters and model selection criteria are reported in Table 3. From these results, we conclude that the RBOLL-N distribution is the best model among others.

Table 4 lists the LR test results with its p-values for the second data. Since all computed p-values are less than 0.05, the null hypotheses are rejected in favour of the RBOLL-N distribution. Therefore, it is concluded that the RBOLL-N distribution provides more acceptable results than its sub-models.

The visual comparison of the fitted models are displayed in Figures 5(a) and 5(b). From these figures, we compare the fitted pdfs of the models on the histogram of the data set and conclude that the RBOLL-N distribution has higher compatibility than other models for the second data. Figure 5(b) displays the some fitted functions of the RBOLL-N distribution.

Table 1. MLEs and their SEs of the fitted models and goodness-of-fit statistics for first data set

Models	α	β	λ	γ	$-\ell$	AIC	BIC
W			2.393 0.210	2.868 0.135	131.288	266.576	271.462
G-W		0.998 0.849	2.392 0.442	2.865 1.356	131.288	268.576	275.905
OLL-W	0.604 0.166		3.605 0.834	2.841 0.139	129.978	265.956	273.284
RBOLL-W	0.132 0.054	2.022 0.243	21.758 9.206	4.110 0.161	126.066	260.134	269.904

Table 2. The LR test results for first data set.

	Hypotheses	LR	p -value
RBOLL-W versus OLL-W	$H_0 : \beta = 1$	7.824	0.005
RBOLL-W versus Gamma-W	$H_0 : \alpha = 1$	10.444	0.001
RBOLL-W versus W	$H_0 : \alpha = \beta = 1$	10.444	0.005

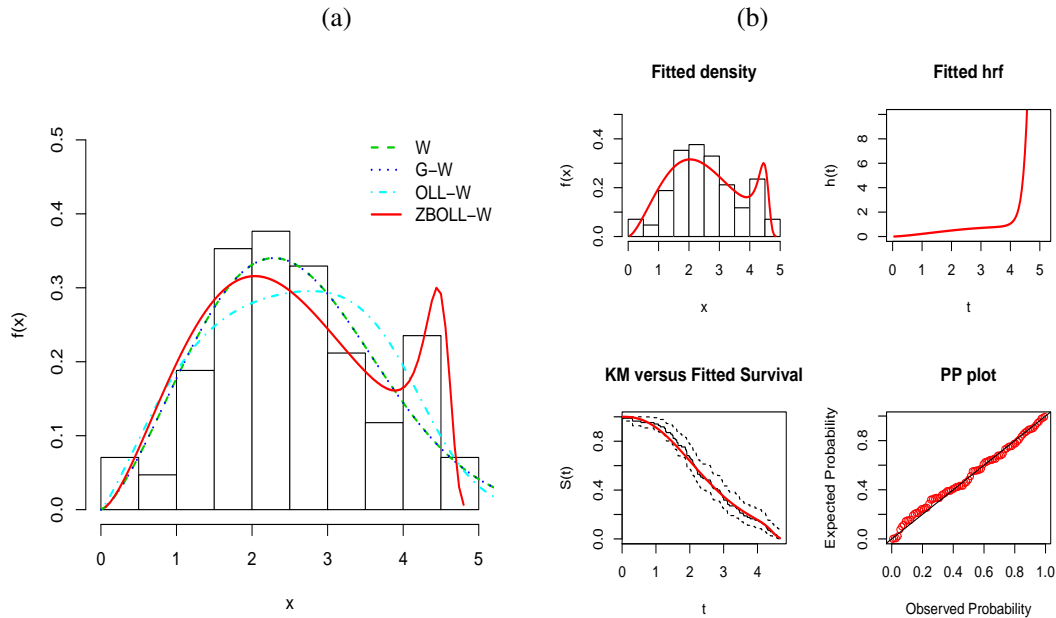


Figure 4. (a) Fitted densities of the models ,(b) fitted functions of the RBOLL-W model for the first data set

6.3. Stanford heart transplant data

In third application, we demonstrate the importance of the LRBOLLW regression model applying it to the real data set which has been recently analyzed by Brito et al. [4] by means of the the Log-Topp-Leone odd log-logistic-Weibull (Log-TLOLL-W) regression model. The Stanford heart transplant data set is used to demonstrate and compare the performance of the LRBOLLW regression model with log-TLOLL-W, log-Weibull and log-OLL-W regression models. The data set can be found in a R package called *p3state.msm*. The sample size is $n = 103$ and censoring rate is 27%. The goal of the study is to

Table 3. MLEs and their SEs of the fitted models and goodness-of-fit statistics for second data set

Models	α	β	μ	σ	$-\ell$	AIC	BIC
N			111.75 3.165	17.907 2.238	137.733	279.467	282.398
G-N	-	4.746 10.729	161.543 100.327	21.947 5.075	137.307	280.614	285.012
OLL-N	0.126 0.018	-	105.776 0.002	4.438 0.002	134.771	275.542	279.939
RBOLL-N	0.164 0.028	0.656 0.118	101.921 0.002	4.911 0.002	132.745	273.490	279.353

Table 4. The LR test results for second data set.

	Hypotheses	LR	p -value
RBOLL-N versus OLL-N	$H_0 : \beta = 1$	4.051	0.044
RBOLL-N versus Gamma-N	$H_0 : \alpha = 1$	9.124	0.002
RBOLL-N versus N	$H_0 : \alpha = \beta = 1$	9.976	0.007

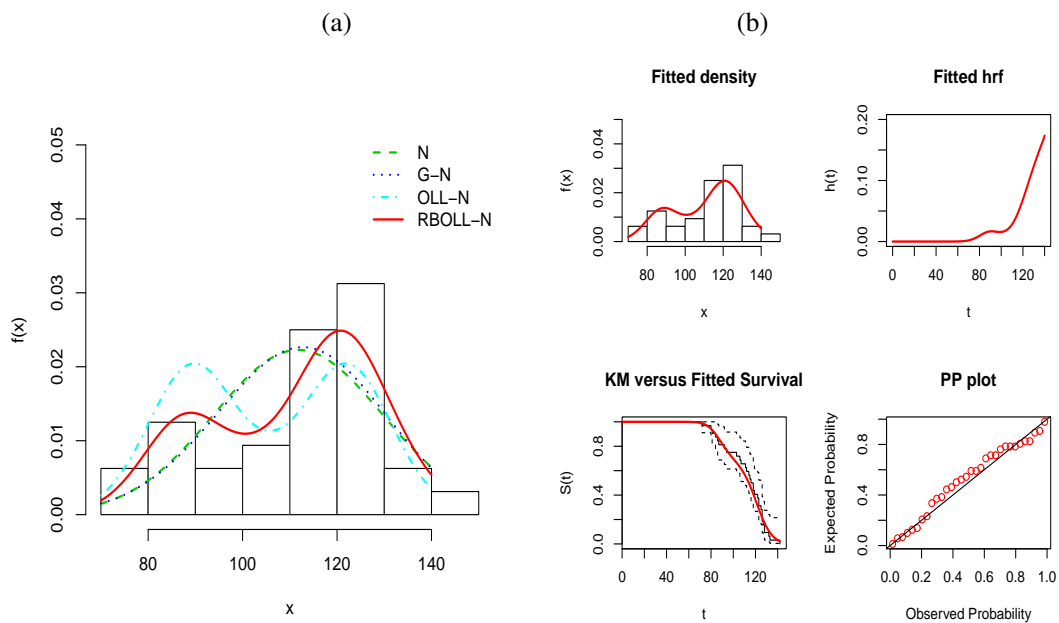


Figure 5. (a) Fitted densities of the models ,(b) fitted functions of the RBOLL-N for the second data set.

analyze the survival times (y_i) of patients with covariates such as x_1 - year of acceptance to the program; x_2 - age of patient (in years); x_3 - previous surgery status ($1 = yes, 0 = no$) and x_4 - transplant indicator ($1 = yes, 0 = no$); c_i .

The fitted regression model is given by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \sigma z_i , \tag{36}$$

where the response variable y_i follows the LRBOLLW distribution.

The results are obtained and summarized in Table 5. Table 5 lists the estimated parameters of the fitted regression models with their corresponding standard errors and model selection criteria, AIC. From the results in Table (5), we conclude that the LRBOLLW regression model provides better fit than the log-TLOLL-W, log-Weibull and log-OLL-W regression models. Additionally, the regression parameters, β_1 and β_2 , are found statistically significant for 5% significance level. According to the estimated regression parameters, we conclude that when the age of patient increases, the survival time of a patient decreases. Also, when the year of acceptance to the program increases, the survival time of a patient increases.

Table 5. MLEs of the parameters to Stanford Heart Transplant Data for Log-Weibull, Log-TLOLL-W, Log-OLL-W and LRBOLLW regression models with corresponding SEs, p-values and $-\ell$ and AIC statistics.

Models	Parameters	α	β	σ	β_0	β_1	β_2	β_3	β_4	$-\ell$	AIC
Log-Weibull	Estimate	-	-	1.478	1.639	0.104	-0.092	1.126	2.544	171.240	354.481
	S.E.	-	-	0.133	6.835	0.096	0.02	0.658	0.378		
	p-value	-	-	-	0.811	0.279	<0.001	0.087	<0.001		
Log-TLOLL-W	Estimate	2.34	24.029	9.68	-0.645	0.074	-0.053	1.676	2.394	164.684	345.368
	S.E.	3.546	3.015	12.526	8.459	0.097	0.02	0.597	0.384		
	p-value	-	-	-	0.939	0.448	0.009	0.005	<0.001		
Log-OLL-W	Estimate	3.529	-	4.434	4.257	0.142	-0.059	0.794	0.084	161.972	337.944
	S.E.	4.398	-	5.47	2.021	0.094	0.018	0.528	0.374		
	p-value	-	-	-	0.035	0.132	0.001	0.132	0.821		
LRBOLLW	Estimate	17.727	1.46	22.796	11.392	0.229	-0.055	0.167	0.397	159.052	334.105
	S.E.	47.554	0.695	61.804	23.002	0.095	0.017	0.485	0.37		
	p-value	-	-	-	0.620	0.015	0.002	0.730	0.283		

Residual Analysis of LRBOLLW model for Stanford heart transplant data set Now, we assess the accuracy of the fitted LRBOLLW regression model with a residual analysis. For this aim, two types of the residuals are used. These are martingale and modified deviance residuals. The martingale residuals for LRBOLLW model is

$$r_{M_i} = \begin{cases} 1 + \log \left\{ \frac{1}{\Gamma(\beta)} \gamma \left(\beta, -\log \left[\frac{\{1 - \exp[-u_i]\}^\alpha}{\{1 - \exp[-u_i]\}^\alpha + (2 - \exp[-u_i])^\alpha} \right] \right) \right\} & \text{if } i \in F, \\ \log \left\{ \frac{1}{\Gamma(\beta)} \gamma \left(\beta, -\log \left[\frac{\{1 - \exp[-u_i]\}^\alpha}{\{1 - \exp[-u_i]\}^\alpha + (2 - \exp[-u_i])^\alpha} \right] \right) \right\} & \text{if } i \in C, \end{cases} \tag{37}$$

The modified deviance residual, proposed by Therneau et al. (1990), is

$$r_{D_i} = \begin{cases} \text{sign}(r_{M_i}) \{ -2 [r_{M_i} + \log(1 - r_{M_i})] \}^{1/2}, & \text{if } i \in F \\ \text{sign}(r_{M_i}) \{ -2r_{M_i} \}^{1/2}, & \text{if } i \in C, \end{cases} \tag{38}$$

where \hat{r}_{M_i} is the martingale residual. When the fitted model is correct, the modified deviance residuals are normally distributed with zero mean and unit variance. The results of the residual analysis are displayed in 6. Figure 6 shows the plot of the modified deviance residuals and its corresponding quantile-quantile (QQ) plot. From these figures, it is clear that there is no possible observation can be evaluated as an outlier. Also, these figures reveal that the LRBOLLW regression model produce accurate fit to the data set.

7. Conclusions

We propose a new class of gamma extended family of distributions, called Ristic-Balakrishnan odd log-logistic G (“RBOLL-G” for short) with two extra generator parameters which can include as special cases several classical continuous distributions. The normal and Weibull distributions are considered as baseline distributions for RBOLL-G family. It is shown that the density of the RBOLL-G family can be expressed as a linear combination of exponentiated-G densities. The statistical properties of the RBOLL-G are derived in detail. More importantly, we introduce a new location-scale regression model based on the Weibull baseline distribution of the RBOLL-G family to analyze the variability of the lifetimes of individuals with covariates. The maximum likelihood estimation method is preferred to obtain unknown parameters of the proposed models. Three real data sets are analyzed to convince the readers in favour of RBOLL-G family.

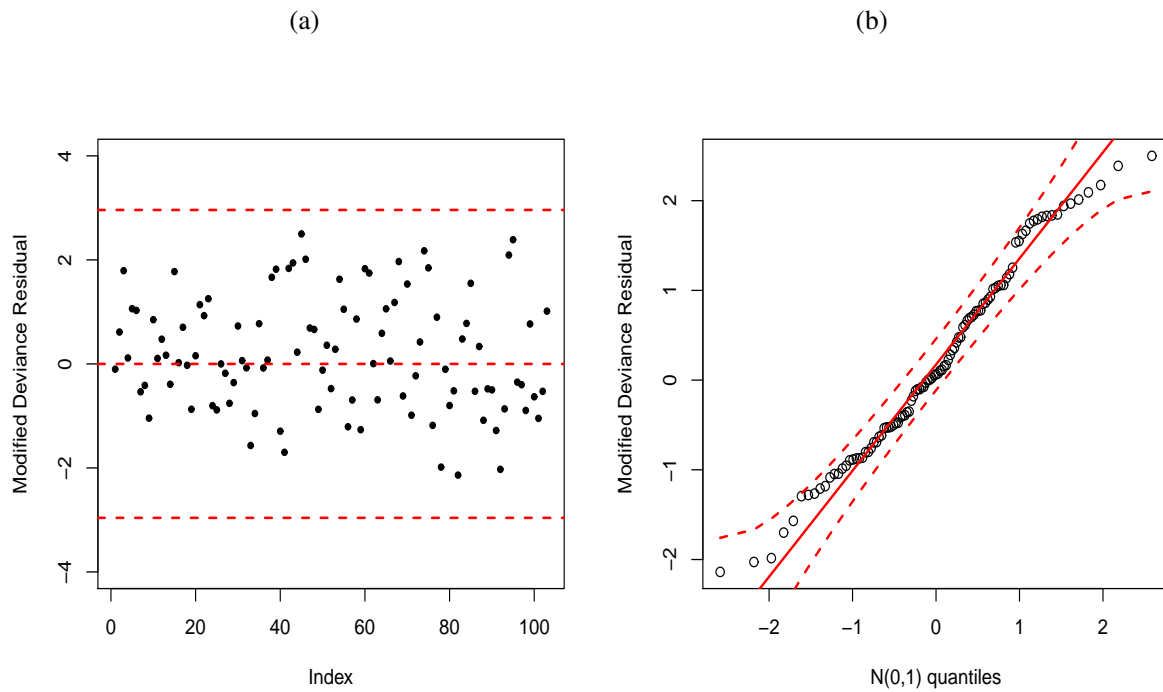


Figure 6. (a) The plot of the modified deviance residual and (b) its QQ plot.

Appendix A

We present four power series for the proof of the linear representation in Section 3. First, for $a > 0$ real non-integer and $|u| < 1$, we have the binomial expansion

$$(1 - u)^a = \sum_{j=0}^{\infty} (-1)^j \binom{a}{j} u^j, \tag{39}$$

where the binomial coefficient is defined for any real.

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

$$G(x)^\alpha = \sum_{r=0}^{\infty} s_r(\alpha) G(x)^r, \tag{40}$$

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

$$z^\lambda = \sum_{k=0}^{\infty} (\lambda)_k (z - 1)^k / k! = \sum_{i=0}^{\infty} f_i z^i \tag{41}$$

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

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

Fourth, we use throughout an equation of Gradshteyn and Ryzhik ([8], Section 0.314) for a power series raised to a positive integer j given by

$$\left(\sum_{j=0}^{\infty} a_j v^j\right)^i = \sum_{j=0}^{\infty} c_{i,j} v^j, \tag{42}$$

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

$$c_{i,j} = (ja_0)^{-1} \sum_{m=1}^j [m(j+1) - j] a_m c_{i,j-m}$$

and $c_{i,0} = a_0^i$. Hence, the coefficients $c_{i,j}$ can be calculated directly from $c_{i,0}, \dots, c_{i,j-1}$ and, therefore, from a_0, \dots, 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 (39) and (40)

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

where $t_j = t_j(a) = s_j(a) + (-1)^j \binom{a}{j}$. Then, using (41), 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 (42), we obtain

$$[G(x)^a + \bar{G}(x)^a]^c = \sum_{j=0}^{\infty} h_j G(x)^j, \tag{43}$$

where $h_j = h_j(a, c) = \sum_{i=0}^{\infty} f_i m_{i,j}$ and $m_{i,j} = (jt_0)^{-1} \sum_{m=1}^j [m(j+1) - j] t_m m_{i,j-m}$ (for $j \geq 1$) and $m_{i,0} = t_0^i$.

Appendix B

The elements of the observed information matrix $J(\theta)$ for the parameters (α, β, τ) are

$$J_{\alpha\alpha} = -\psi'(\alpha), \quad J_{\alpha\beta} = \frac{n}{\beta}, \quad J_{\alpha\tau} = \sum_{i=1}^n \frac{[\dot{G}(x_i; \tau)]_{\tau}}{G(x_i; \tau)} + \sum_{i=1}^n \frac{[\dot{G}(x_i; \tau)]_{\tau}}{[1 - G(x_i; \tau)]},$$

$$J_{\beta\beta} = -\frac{n\alpha}{\beta^2}, \quad J_{\beta\tau} = -\sum_{i=1}^n \frac{[\dot{G}(x_i; \tau)]_{\tau}}{[1 - G(x_i; \tau)]^2},$$

$$\begin{aligned} J_{\tau\tau} &= (\alpha - 1) \sum_{i=1}^n \left\{ \frac{[\ddot{G}(x_i; \tau)]_{\tau\tau}}{G(x_i; \tau)} - \frac{[\dot{G}(x_i; \tau)]_{\tau}^2}{[G(x_i; \tau)]^2} \right\} \\ &+ (\alpha - 1) \sum_{i=1}^n \left\{ \frac{[\ddot{G}(x_i; \tau)]_{\tau\tau}}{[1 - G(x_i; \tau)]} + \frac{[\dot{G}(x_i; \tau)]_{\tau}^2}{[1 - G(x_i; \tau)]^2} \right\} \\ &- \beta \sum_{i=1}^n \left\{ \frac{[\ddot{G}(x_i; \tau)]_{\tau\tau}}{[1 - G(x_i; \tau)]^2} + \frac{2[\dot{G}(x_i; \tau)]_{\tau}^2}{[1 - G(x_i; \tau)]^3} \right\} + \sum_{i=1}^n \left\{ \frac{[\ddot{g}(x_i; \tau)]_{\tau\tau}}{g(x_i; \tau)} - \frac{[\dot{g}(x_i; \tau)]_{\tau}^2}{[g(x_i; \tau)]^2} \right\}, \end{aligned}$$

where,

$$[\dot{g}(x_i; \tau)]_{\tau} = \frac{dg(x_i; \tau)}{d\tau}, \quad [\dot{G}(x_i; \tau)]_{\tau} = \frac{dG(x_i; \tau)}{d\tau},$$

$$[\ddot{g}(x_i; \tau)]_{\tau\tau} = \frac{d^2g(x_i; \tau)}{d\tau d\tau^T}, \quad [\ddot{G}(x_i; \tau)]_{\tau\tau} = \frac{d^2G(x_i; \tau)}{d\tau d\tau^2},$$

and $g(\cdot)$ and $G(\cdot)$ functions are defined in Section 1.

Appendix C

The first data set: 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, 2.82, 3, 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 second data set: 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.

REFERENCES

1. K. Aas, and I. Haff, *The generalized hyperbolic skew student's t-distribution*, Journal of Financial Econometrics, vol. 4, pp. 275–309, 2006.
2. C. Alexander, G. M. Cordeiro, E. M. M. Ortega, and J. M. Sarabia, *Generalized beta-Generated distributions*, Computational Statistics and Data Analysis, vol. 56, pp. 1880–1897, 2012.
3. M. Alizadeh, M. Emadi and M. Doostparast, *A New Two-Parameter Lifetime Distribution: Properties, Applications and Different Method of Estimations*, Statistics, Optimization and Information Computing, vol. 7(2), pp. 291–310, 2019.
4. E. Brito, G. M. Cordeiro, H. M. Yousof, M. Alizadeh, and G. O. Silva, *The Topp-Leone odd log-logistic family of distributions*, Journal of Statistical Computation and Simulation, vol. 87(15), pp. 3040–3058, 2017.
5. G. M. Cordeiro, and M. de Castro, *A new family of Generalized distributions*, Journal of Statistical Computation and Simulation, vol. 81, pp. 883–898, 2011.
6. N. Eugene, C. Lee, and F. Famoye, *Beta-normal distribution and its applications*, Communications in Statistics-Theory and Methods, vol. 31, pp. 497–512, 2002.
7. J. U. Gleanon, and J. D. Lynch, *Properties of generalized log-logistic families of lifetime distributions*, Journal of Probability and Statistical Science, vol. 4(1), pp. 51–64, 2006.
8. I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integrals, Series, and Products*, sixth edition, Academic Press, San Diego, 2000.
9. R. C. Gupta, R. D. Gupta and P. L. Gupta, *Modeling failure time data by Lehmann alternatives*, Communications in Statistics, Theory and Methods, vol. 27, pp. 887–904, 1998.
10. R. D. Gupta, and D. Kundu, *Generalized exponential distributions*, Australian and New Zealand Journal of Statistics, vol. 41, pp. 173–188, 1999.
11. R. D. Gupta, and D. Kundu, *Exponentiated exponential distribution: an alternative to gamma and Weibull distributions*, Biometrical Journal, vol. 43, pp. 117–130, 2001.
12. B. E. Hansen, *Autoregressive conditional density estimation*, International Economic Review, vol. 35, pp. 705–730, 1994.
13. N. L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions*, Volume 1, 2nd edition. John Wiley and Sons, New York, 1994.
14. N. L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions*, Volume 2, 2nd edition. John Wiley and Sons, New York, 1995.
15. C. S. Kakde, and D. T. Shirke, *On Exponentiated Lognormal distribution*, International Journal of Agricultural and Statistics Sciences, vol. 2, pp. 319–326, 2006.
16. M. Mozafari, M. Afshari, M. Alizadeh, and H. Karamikabir, *The Zografos-Balakrishnan Odd Log-Logistic Generalized Half-Normal Distribution with Mathematical Properties and Simulations*, Statistics, Optimization and Information Computing, vol. 7(1), pp. 211–234, 2019.
17. G. S. Mudholkar, and D. K. Srivastava, *Exponentiated Weibull family for analyzing bathtub failure-real data*, IEEE Transaction on Reliability, vol. 42, pp. 299–302, 1993.
18. G. S. Mudholkar, D. K. Srivastava, and G. D. Kollia, *A generalization of the Weibull distribution with application to the analysis of survival data*, Journal of American Statistical Association, vol. 91, pp. 1575–1583, 1996.
19. S. Nadarajah, G. M. Cordeiro, and E. M. M. Ortega, *The Zografos-Balakrishnan-G family of distributions: Mathematical properties and applications*, Communications in Statistics - Theory and Methods, vol. 44(1), pp. 186–215, (2015).
20. S. Nadarajah, *The exponentiated Gumbel distribution with climate application*, Environmetrics, vol. 17, pp. 13–23, 2005.
21. S. Nadarajah, and A. K. Gupta, *The exponentiated gamma distribution with application to drought data*, Calcutta Statistical Association Bulletin, vol. 59, pp. 29-54, 2007.

22. S. Nadarajah, and S. Kotz, *The exponentiated type distributions*, Acta Applicandae Mathematicae, vol. 92, pp. 97–111, 2006.
23. M. M. Ristic, and N. Balakrishnan, *The gamma exponentiated exponential distribution*, Journal of Statistical Computation and Simulation, pp. 1191–1206, 2012.
24. T. M. Therneau, P. M. Grambsch, and, T.R. Fleming, *Martingale-based residuals for survival models*, Biometrika, vol. 77, 147–160, 1990.
25. K. Zografos, and N. Balakrishnan, *on families of beta- and generalized gamma-generated distributions and associated inference*, Statistical Methodology, vol. 6, pp. 344–362, 2009.