

# The Generalized Odd Gamma-G Family of Distributions: Properties and Applications

## Abstract

Recently, new continuous distributions have been proposed to apply in statistical analysis. In this paper, the Generalized Odd Gamma-G distribution is introduced. In particular, G has been considered as the Uniform distribution and some statistical properties such as quantile function, asymptotics, moments, entropy and order statistics have been calculated. The fitness capability of this model has been investigated by fitting this model and others based on real data sets. The parameters of this model are estimated by the maximum likelihood estimation method with simulated real data in order to test validity of maximum likelihood estimators.

*Keywords:* Generalized Odd Gamma-G; Maximum likelihood; Moment; Entropy.  
*MSC 2010:* 60E05; 62E10; 62N05

## 1 Introduction

The classic statistical distributions which have essential limitations and problems in data modeling, has led statistical researcher to make of the new flexible distributions. The new distributions are often made through the classic distributions and give the required flexibility to the classic distributions. The most important distributions among them are Marshall-Olkin generated (MO-G) by Marshall and Olkin (1997), Kumaraswamy-G (Kw-G) by Cordeiro and de Castro (2011), McDonald-G (Mc-G) by Alexander et al. (2012), Weibull-G by Bourguignon et al. (2014), exponentiated half-logistic by Cordeiro et al. (2014a), transformer (T-X) by Alzaatreh et al. (2013), and Lomax generator by Cordeiro et al. (2014b), Beta Marshall-OLkin family by AUTHORC, et al. (2015), the type I half-logistic family by Cordeiro et al. (2016) .

Based on T-X idea by Alzaatreh et al. (2013), by the following definition, the Generalized Odd Gamma-G distribution (GOGa-G) would be made

$$F(x; \alpha, \beta, \xi) = \int_0^{\frac{G(x; \xi)^\beta}{1-G(x; \xi)^\beta}} \frac{t^{\alpha-1} e^{-t}}{\Gamma(\alpha)} dt = \frac{\gamma\left(\alpha, \frac{G(x; \xi)^\beta}{1-G(x; \xi)^\beta}\right)}{\Gamma(\alpha)} \quad (1)$$

where  $\alpha, \beta > 0$  are two additional shape parameters,  $\xi$  is the parameter for baseline  $G$  and  $\gamma(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} dt$  denote the incomplete gamma function.

In this case, the probability density function (pdf) of the GOGa-G distribution will be as follows:

$$f(x; \alpha, \beta, \xi) = \frac{\beta g(x; \xi) G(x; \xi)^{\alpha\beta-1}}{\Gamma(\alpha) [1 - G(x; \xi)^\beta]^{\alpha+1}} e^{\frac{-G(x; \xi)^\beta}{1-G(x; \xi)^\beta}} \quad (2)$$

where  $g(x; \xi)$  is the pdf of the  $G(x; \xi)$  distribution. From now on, the random variable  $X$  with

pdf (2) is shown with  $X \sim \text{GOGa-G}(\alpha, \beta, \xi)$ . According to (1) and (2) hrt of  $X$  is as follows:

$$\tau(x; \alpha, \beta, \xi) = \frac{\beta g(x; \xi) G(x; \xi)^{\alpha\beta-1} e^{\frac{-G(x; \xi)^\beta}{1-G(x; \xi)^\beta}}}{\left[1 - G(x; \xi)^\beta\right]^{\alpha+1} \left[\Gamma(\alpha) - \gamma\left(\alpha, \frac{G(x; \xi)^\beta}{1-G(x; \xi)^\beta}\right)\right]} \quad (3)$$

An interpretation of the GoGa-G family (1) can be given as follows:

Let  $T$  be a random variable describing a stochastic system by the cdf  $G(x)^\beta$  (for  $\beta > 0$ ). If the random variable  $X$  represents the odds ratio, the risk that the system following the lifetime  $T$  will be not working at time  $x$  is given by  $\frac{G(x)^\beta}{1-G(x)^\beta}$ . If we are interested in modeling the randomness of the odds ratio by the Gamma pdf  $r(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1} e^{-t}$  (for  $t > 0$ ), the cdf of  $X$  is given by

$$Pr(X \leq x) = R\left(\frac{G(x)^\beta}{1 - G(x)^\beta}\right),$$

which is exactly the cdf (1) of the new family.

Theorem 1 provides some relations of the GOGa family with other distributions.

**Theorem 1.** Let  $X \sim \text{GOGa-G}(\alpha, \beta, \xi)$  and  $Y = \frac{G(X; \xi)^\beta}{1 - G(X; \xi)^\beta}$ , then  $Y \sim \Gamma(\alpha, 1)$ .

**Proof:**

$$\begin{aligned} F_Y(y) &= p\left(\frac{G(x)^\beta}{1 - G(x)^\beta} \leq y\right) = P(X \leq G^{-1}\left[\left(\frac{y}{1+y}\right)^{\frac{1}{\beta}}\right]) = \int_0^{\frac{y}{1+y}} \frac{t^{\alpha-1} e^{-t}}{\Gamma(\alpha)} dt \\ &= \int_0^y \frac{t^{\alpha-1} e^{-t}}{\Gamma(\alpha)} dt, \quad \text{so } Y \sim \Gamma(\alpha, 1). \end{aligned}$$

The basic motivations for using the GOGa family in practice are the following:

(i) to make the kurtosis more flexible compared to the baseline model; (ii) to produce a skewness for symmetrical distributions; (iii) to construct heavy-tailed distributions that are not longer-tailed for modeling real data; (iv) to generate distributions with symmetric, left-skewed, right-skewed and reversed-J shaped; (v) to define special models with all types of the hrf; (vi) to provide consistently better fits than other generated models under the same baseline distribution.

In the following, the paper would be like this: In Section 2, a special distribution is introduced by selecting  $G$ . In Section 3, the features of the GOGa- model will be assessed using quantile function, asymptotics, functions expansion, quantile power series, moments, entropy and order statistics. In Section 4, MLE calculation method and in Section 5, estimability of the model additional parameters will be discussed using simulation. In Section 6, the proposed model is fitted based on two real data sets and compared to other famous models.

## 2 Special Models

### 2.1 The Generalized Odd Gamma-Uniform (GOGa-U)

Different distributions family can be reached by selecting different  $G$ s in equation (2). Here as Torabi and Montazeri Hedesh (2012),  $G$  has been considered as uniform distribution. In this

case, by letting  $\xi = (a, b)$  equation (2) will changed as follows:

$$f(x; \alpha, \beta, a, b) = \frac{\frac{\beta}{b-a} \left(\frac{x-a}{b-a}\right)^{\alpha\beta-1} e^{\frac{-(x-a)^\beta}{(b-a)^\beta - (x-a)^\beta}}}{\Gamma(\alpha) \left[1 - \left(\frac{x-a}{b-a}\right)^\beta\right]^{\alpha+1}} \quad (4)$$

where  $a, b \in \mathbb{R}$  and  $a < b$ . If  $X$  be a random variable with density function (4), then it will be displayed by  $\text{GOGa-U}(\alpha, \beta, a, b)$ . In Figure 1 some density and hazard functions for  $\text{GOGa-U}$  have been drawn.

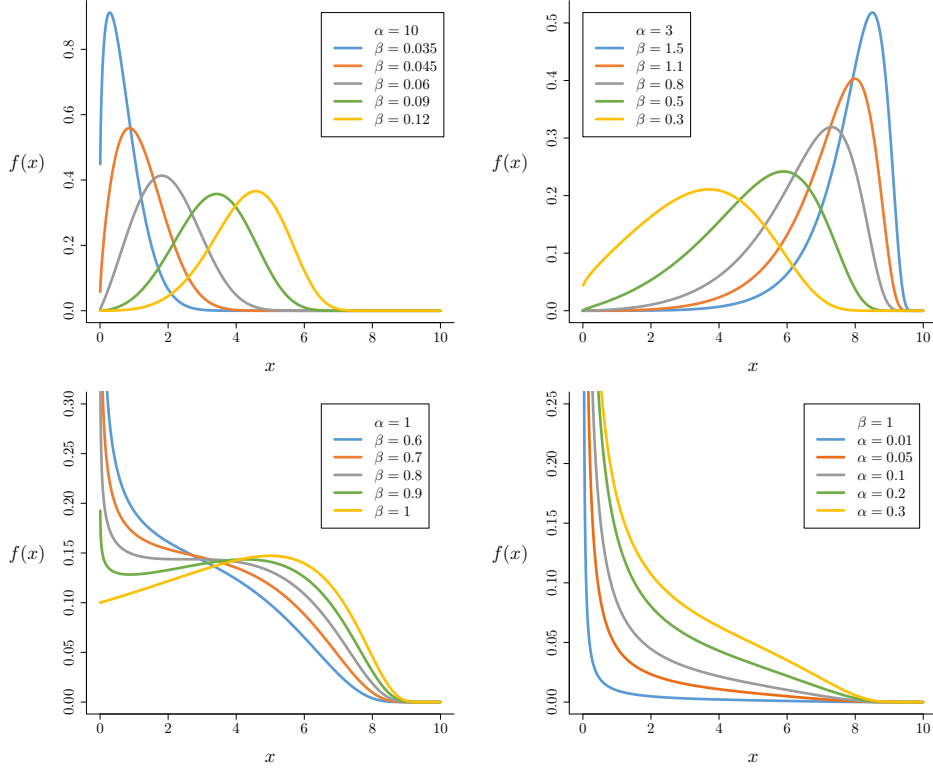


Figure 1: The sample curves of density function of  $\text{GOGa-U}(\alpha, \beta, 0, 10)$ .

One can see in the curves of Figure 1 that the different states of density function including symmetric density function (approximately), mild and high skewed (right and left) and bimodal (in the right bottom curve, one mode is in point zero) have been produced. In Figure 2 one can see some curves of the hazard function of the  $\text{GOGa-U}$  distribution for some paramentreters. According to Figure 2 you see that the U shape hazard functions are producible by  $\text{GOGa-U}$ .

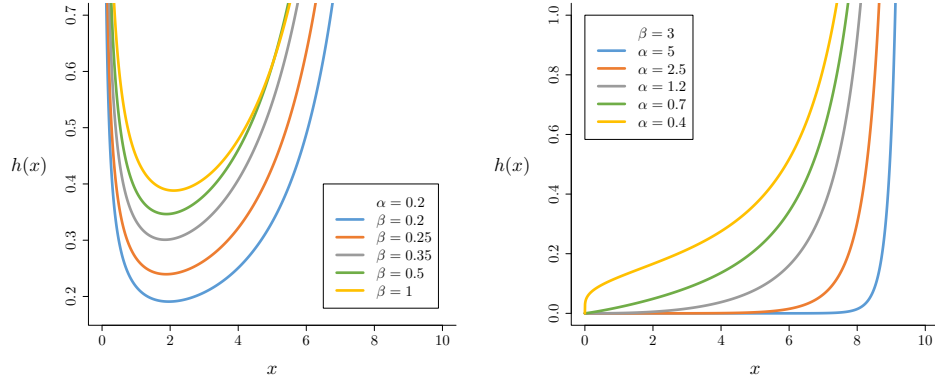


Figure 2: The sample curves of hazard function of  $\text{GOGa-U}(\alpha, \beta, 0, 10)$ .

## 2.2 The Generalized Odd Gamma-Weibull (GOGa-W)

In GOGa-G, suppose G is as follows Weibull distribution function:

$$G(x; \alpha, \beta, \lambda, k) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}, \quad x \geq 0$$

In this case, by letting  $\xi = (\lambda, k)$  equation (2) will be changed as follows

$$f(x; \alpha, \beta, \lambda, k) = \frac{\beta k \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]^{\alpha\beta-1} \frac{-\left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]^\beta}{1 - \left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]^\beta}}{\lambda \Gamma(\alpha) \left\{1 - \left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]^\beta\right\}^{\alpha+1}} e^{-\left(\frac{x}{\lambda}\right)^k}, \quad x \geq 0 \quad (5)$$

where  $\alpha, \beta, \lambda, k > 0$ . If  $X$  be a random variable with density function (5), then it will be displayed by  $\text{GOGa-W}(\alpha, \beta, \lambda, k)$ . In Figure 3 some pdfs for GOGa-W have been drawn.

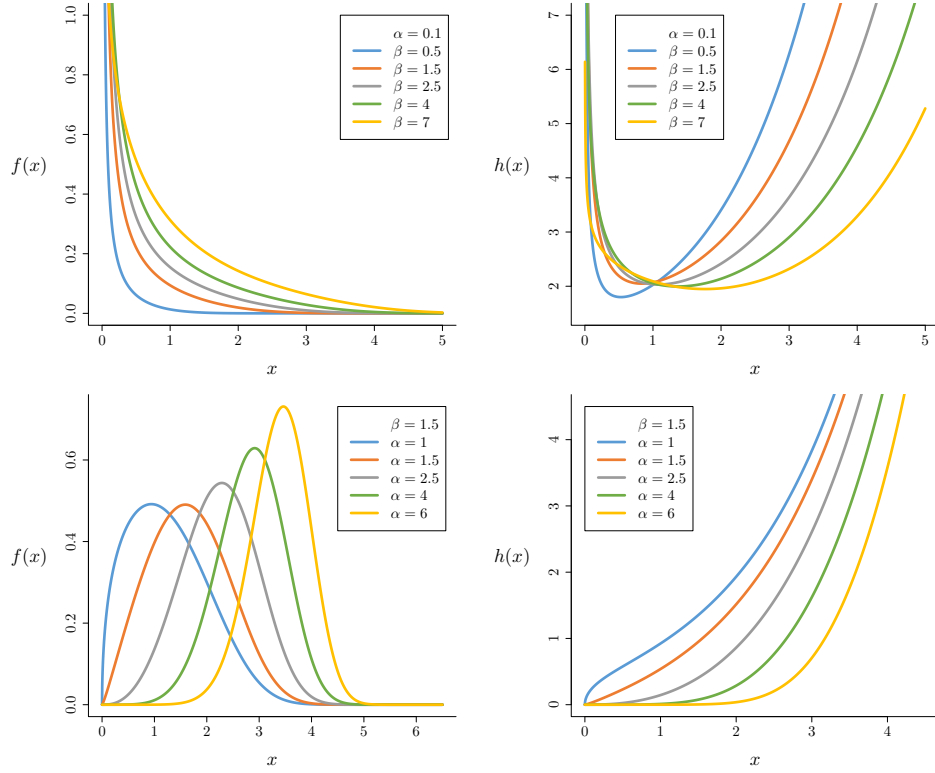


Figure 3: The sample curves of density and hazard function of  $\text{GOGa-W}(\alpha, \beta, 1, 1.5)$ .

### 3 Main Features

#### 3.1 Quantile Function

By considering (1) quantile function (qf)  $X$  is obtained as follows: If  $V \sim \Gamma(\alpha, 1)$  then the solution of nonlinear equation  $x_v = Q_G \left[ \left( \frac{V}{1+V} \right)^{\frac{1}{\beta}} \right]$  has cdf (1).

#### 3.2 Asymptotics

**Proposition 1.** Let  $a = \inf \{x | f(x) > 0\}$ , then the asymptotic of equation (1), (2) and (3) when  $x \rightarrow a$  are given by

$$F(x) \sim \frac{G(x)^{\alpha\beta}}{\alpha\Gamma(\alpha)}$$

$$f(x) \sim \frac{\beta g(x) G(x)^{\alpha\beta-1}}{\Gamma(\alpha)}$$

$$\tau(x) \sim \frac{\beta g(x) G(x)^{\alpha\beta-1}}{\Gamma(\alpha)}$$

**Proposition 2.** The asymptotic of equation (1), (2) and (3) when  $x \rightarrow +\infty$  are given by

$$\overline{F}(x) \sim 1 - \frac{\gamma\left(\alpha, \frac{1}{\beta G(x)}\right)}{\Gamma(\alpha)}$$

$$f(x) \sim \frac{g(x)}{\beta^\alpha \Gamma(\alpha) \overline{G}(x)^{\alpha+1}} e^{\frac{-1}{\beta \overline{G}(x)}}$$

$$\tau(x) \sim \frac{g(x) e^{\frac{-1}{\beta \overline{G}(x)}}}{\beta^\alpha \left[ \Gamma(\alpha) - \gamma \left( \alpha, \frac{1}{\beta \overline{G}(x)} \right) \right] \overline{G}(x)^{\alpha+1}}$$

### 3.3 Expansion for Pdf and Cdf

Using generalized binomial and taylor expansion one can obtain

$$\begin{aligned} f(x) &= \frac{\beta g(x) G(x)^{\alpha\beta-1}}{\Gamma(\alpha) [1 - G(x)^\beta]^{\alpha+1}} \sum_{i=0}^{\infty} \frac{(-1)^i \left( \frac{G(x)^\beta}{1 - G(x)^\beta} \right)^i}{i!} \\ &= \frac{\beta g(x)}{\Gamma(\alpha)} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^i}{i!} \binom{-\alpha - i - 1}{j} G(x)^{\beta(\alpha+i+j)-1} \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_{i,j} h_{\beta(\alpha+i+j)}(x) \end{aligned} \quad (6)$$

where

$$w_{i,j} = \frac{(-1)^i \binom{-\alpha - i - 1}{j}}{i! [\alpha + i + j] \Gamma(\alpha)}$$

and  $h_\beta(x) = \beta g(x) G(x)^{\beta-1}$ , denote the pdf of exp-G distribution with power parameter  $\beta$ .

By integrating from equation (6) with respect to  $x$ , we have

$$F(x) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_{i,j} H_{\beta(\alpha+i+j)}(x) \quad (7)$$

where  $H_\beta(x) = G(x)^\beta$ .

By considering  $G(x) = 1 - [1 - G(x)]$  and binomial expansion we have:

$$\begin{aligned} G(x)^{\beta(\alpha+i+j)} &= \sum_{l=0}^{\infty} (-1)^l \binom{\beta(\alpha+i+j)}{l} [1 - G(x)]^l \\ &= \sum_{l=0}^{\infty} \sum_{k=0}^l (-1)^{l+k} \binom{\beta(\alpha+i+j)}{l} \binom{l}{k} G(x)^k \\ &= \sum_{k=0}^{\infty} \sum_{l=k}^{\infty} (-1)^{l+k} \binom{\beta(\alpha+i+j)}{l} \binom{l}{k} G(x)^k \end{aligned}$$

In this case, regarding to (7) cdf extends as follows

$$F(x) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=k}^{\infty} w_{i,j} (-1)^{l+k} \binom{\beta(\alpha+i+j)}{l} \binom{l}{k} G(x)^k$$

then

$$F(x) = \sum_{k=0}^{\infty} b_k G(x)^k \quad (8)$$

where

$$b_k = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{l=k}^{\infty} w_{i,j} (-1)^{l+k} \binom{\beta(\alpha + i + j)}{l} \binom{l}{k} \quad (9)$$

and finally regarding to (8) for cdf we also have

$$f(x) = \sum_{k=0}^{\infty} b_{k+1} h_{k+1}(x)$$

### 3.4 Moments

The  $r$ th ordinary moment of  $X$  is given by

$$\mu'_r = E(X^r) = \int_{-\infty}^{+\infty} x^r f(x) dx$$

Using (1), we obtain the following:

$$\mu'_r = \sum_{k=0}^{\infty} b_{k+1} E(Y_{k+1}^r) \quad (10)$$

Hereafter,  $Y_{k+1}$  denotes the Exp-G distribution with power parameter  $(k+1)$ . Setting  $r = 1$  in (10), We have the mean of  $X$ . The last integration can be computed numerically for most parent distributions. The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships. The  $n$ th central moment of  $X$ , say  $M_n$ , follows as

$$M_n = E(X - \mu)^n = \sum_{h=0}^n (-1)^h \binom{n}{h} (\mu'_1)^n \mu'_{n-h}$$

The cumulants  $(\kappa_n)$  of  $X$  follow recursively from

$$\kappa_n = \mu'_n - \sum_{r=0}^{n-1} \binom{n-1}{r-1} \kappa_r \mu'_{n-r}$$

where  $\kappa_1 = \mu'_1$ ,  $\kappa_2 = \mu'_2 - \mu'^2_1$ ,  $\kappa_3 = \mu'_3 - 3\mu'_2\mu'_1 + \mu'^3_1$ , etc. The skewness and kurtosis measures also can be calculates from the ordinary moment using well-known relationships. The moment generating function (mgf) of  $X$ , say  $M_X(t) = E(e^{tX})$ , is given by

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu'_r = \sum_{k,r=0}^{\infty} \frac{t^r b_{k+1}}{r!} E(Y_{k+1}^r)$$

### 3.5 Incomplete Moments

The main application of the first incomplete moment refers to Bonferroniand Lorenz curves. These curves are very useful in economics, reliability, demography, insurance and medicine. The answers to many important questions in economics require more than just knowing the mean

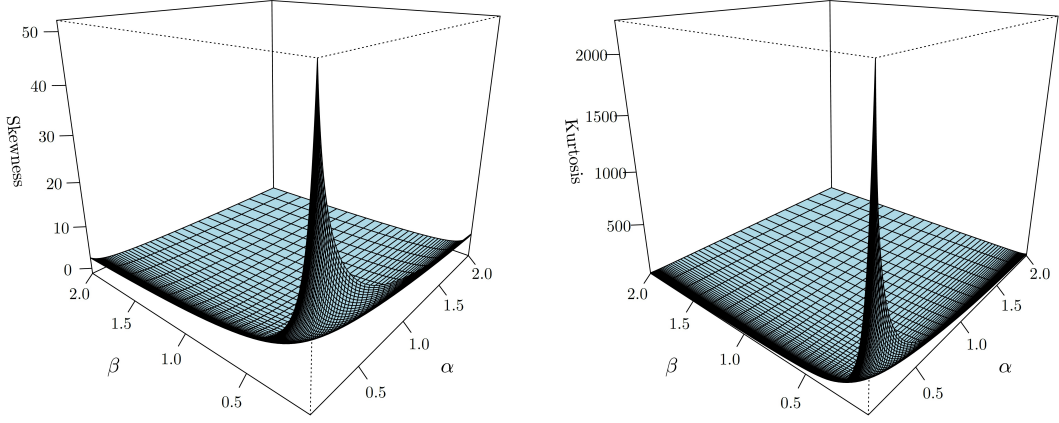


Figure 4: Skewness and Kurtosis for GOGa-U.

of the distribution, its shape as well. This is obvious both in the study of econometrics and in areas as well. The  $s$ th incomplete moments, say  $\varphi_s(t)$ , is given by

$$\varphi_s(t) = \int_{-\infty}^t x^s f(x) dx$$

Using equation(8), we obtain

$$\varphi_s(t) = \sum_{k=0}^{\infty} b_{k+1} \int_{-\infty}^t x^s h_{k+1}(x) dx \quad (11)$$

The first incomplete of the GOGa-G family,  $\varphi_1(t)$ , can be obtained by setting  $s = 1$  in (11). Another application of the first incomplete moment is related to meanresidual life and mean waiting tie given by  $m_1(t) = [1 - \varphi_1(t)] / R(t) - t$  and  $M_1(t) = t - [\varphi_1(t) / F(t)]$ , respectively.

### 3.6 Entropy

Entropy is an index for measuring variation or uncertainty of a random variable. The measure of entropy, Rényi (1961), is defined as follows

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

for  $\gamma > 0$  and  $\gamma \neq 1$ . The Shannon entropy measure is also defined by  $E \{-\log [f(x)]\}$  that is a special state of the Rényi entropy when  $\gamma \uparrow 1$ .



$$\begin{aligned}
f(x)^\gamma &= \left[ \frac{\beta g G^{\alpha\beta-1} e^{\frac{-G^\beta}{1-G^\beta}}}{\Gamma(\alpha)[1-G^\beta]^{\alpha+1}} \right]^\gamma \\
&= \frac{\beta^\gamma g^\gamma G^{\gamma(\alpha\beta-1)} e^{\frac{-G^\beta}{1-G^\beta}}}{[\Gamma(\alpha)]^\gamma [1-G^\beta]^{\gamma(\alpha+1)}} \\
&= \frac{\beta^\gamma}{[\Gamma(\alpha)]^\gamma} \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \gamma^i \frac{G^{\gamma(\alpha\beta-1)+\beta i} e^{\frac{-G^\beta}{1-G^\beta}}}{[1-G^\beta]^{\gamma(\alpha+1)+i}} g^\gamma \\
&= \frac{\beta^\gamma}{[\Gamma(\alpha)]^\gamma} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{i+j}}{i!} \binom{-\gamma(\alpha+1)-i}{j} \gamma^i g^\gamma G^{\gamma(\alpha\beta-1)+\beta(i+j)} \\
\Rightarrow I_R(\gamma) &= \frac{1}{1-\gamma} \log \left[ \int_{-\infty}^{+\infty} f^\gamma(x) dx \right] \\
&= \frac{\gamma}{1-\gamma} \log \left[ \frac{\beta}{\Gamma(\alpha)} \right] + \frac{1}{1-\gamma} \log \left[ \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_{i,j} I(\gamma, \alpha, \beta, i, j) \right]
\end{aligned}$$

where  $v_{i,j} = \frac{(-1)^{i+j} \gamma^i}{i!} \binom{-\gamma(\alpha+1)-i}{j}$  and  $I(\gamma, \alpha, \beta, i, j) = \int_{-\infty}^{+\infty} g(x)^\gamma G(x)^{\gamma(\alpha\beta-1)+\beta(i+j)} dx$ .

In Figure 5 one can see some curves of the entropy function of the GOGa-U distribution for some parameters.

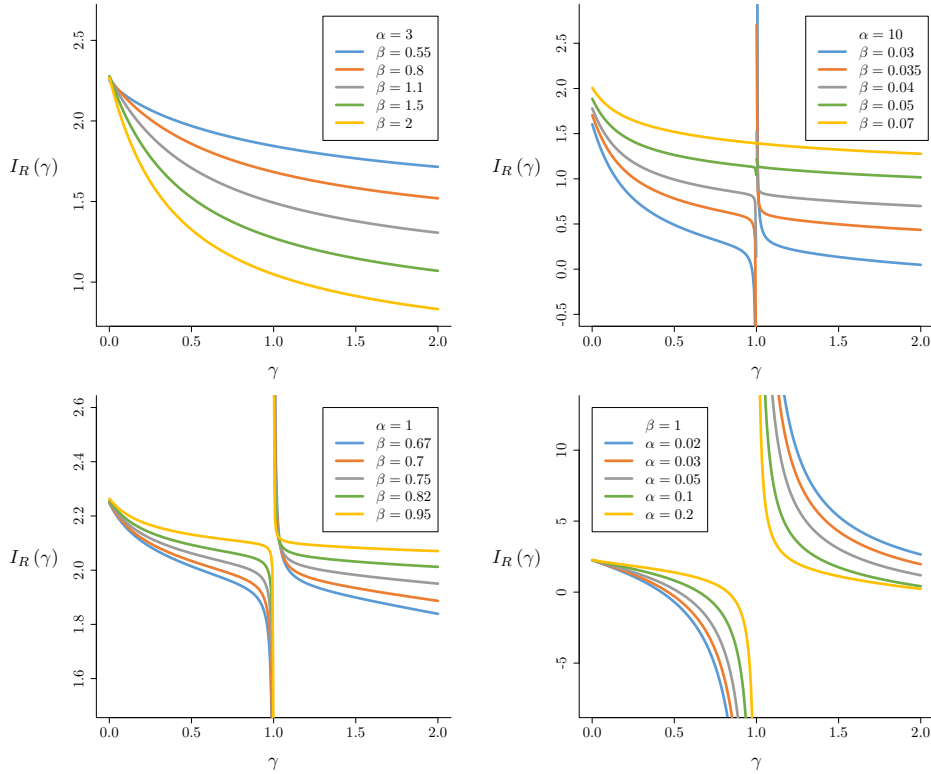


Figure 5: Curves of the GOGa-U Entropy function for some parameter values.

### 3.7 Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose  $X_1, \dots, X_n$  is a random sample from any GOGa-G distribution. Let  $X_{i:n}$  denote the  $i$ th order statistic. The pdf of  $X_{i:n}$  can be expressed as

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

where  $c = \frac{1}{B(i, n-i+1)}$ .

We use the result 0.314 of Gradshteyn and Ryzhik (2000) for a power series raised to a positive integer  $n$  (for  $n \geq 1$ )

$$\left( \sum_{i=0}^{\infty} a_i u^i \right)^n = \sum_{i=0}^{\infty} c_{n,i} u^i, \quad (12)$$

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

$$c_{n,i} = (i a_0)^{-1} \sum_{m=1}^i [m(n+1) - i] a_m c_{n,i-m}. \quad (13)$$

By using equations (9), (12), (13), We can demonstrate that the density function of the  $i$ th order statistic of any GOGa-G distribution can be expressed as follows:

$$f_{i:n}(x) = \sum_{r,k=0}^{\infty} m_{r,k} h_{r+k+1}(x), \quad (14)$$

where  $h_{r+k+1}(x)$  denotes the exp-G density function with parameter  $r + k + 1$ ,

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

$b_r$  is given by equation (9) and the quantities  $f_{j+i-1,k}$  can be determined given that  $f_{j+i-1,0} = b_0^{j+i-1}$  and recursively for  $k \geq 1$

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

We can obtain the ordinary and incomplete moments, generating function and mean deviations of the GOGa-G order statistics from equation (14) and some properties of the exp-G model.

## 4 The Maximum Likelihood Estimator

The MLE is one of the most common point estimators. This estimator is very applicable in confidence intervals and hypothesis testing. By MLE, various statistics is built for assessing the goodness-of-fit in a model, such as: the maximum log-likelihood ( $\hat{\ell}_{max}$ ), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Anderson-Darling ( $A^*$ ) and Cramér-von

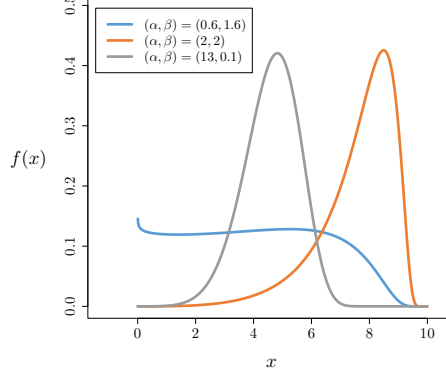


Figure 6: Three density functions for simulation study.

Mises ( $W^*$ ), described by Chen and Balakrishnan (1995). The lower values of these statistics indicate that the model have better fitting. We use these statistics in section 5.

To calculating the MLE, let  $x_1, x, \dots, x_n$  are observations from pdf (2). In this case, by letting  $\theta = (\alpha, \beta, \xi)$  we have

$$\begin{aligned} \ell_n(\theta) = & n \ln(\beta) + \sum_{i=0}^n \ln(g(x_i; \xi)) + (\alpha\beta - 1) \sum_{i=0}^n \ln(G(x_i; \xi)) \\ & - \sum_{i=0}^n \frac{G(x_i; \xi)^\beta}{1 - G(x_i; \xi)^\beta} - n \ln(\Gamma(\alpha)) - (\alpha + 1) \sum_{i=0}^n \ln(1 - G(x_i; \xi)^\beta) \end{aligned}$$

By numerically solving the following equations, the maximum likelihood estimators can be obtained.

$$\begin{cases} \frac{\partial \ell_n(\theta)}{\partial \alpha} = \beta \sum_{i=0}^n \ln G(x_i) - n \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=0}^n \ln(1 - G(x_i)^\beta) = 0 \\ \frac{\partial \ell_n(\theta)}{\partial \beta} = \frac{n}{\beta} + \alpha \sum_{i=0}^n \ln G(x_i) - \sum_{i=0}^n \frac{G(x_i)^\beta \ln G(x_i)}{(1 - G(x_i)^\beta)^2} + (\alpha + 1) \sum_{i=0}^n \frac{\ln G(x_i) G(x_i)^\beta}{(1 - G(x_i)^\beta)} = 0 \\ \frac{\partial \ell_n(\theta)}{\partial \xi} = \sum_{i=0}^n \frac{g(x_i; \xi)}{g(x_i)} + (\alpha\beta - 1) \sum_{i=0}^n \frac{G_i(\xi)}{G(x_i)} - \sum_{i=0}^n \frac{\beta G_i(\xi) G(x_i)^{\beta-1}}{(1 - G(x_i)^\beta)^2} + (\alpha + 1) \sum_{i=0}^n \frac{\beta G_i(\xi) G(x_i)^{\beta-1}}{1 - G(x_i)^\beta} = 0 \end{cases}$$

where  $g_i(\xi) = \frac{\partial g(x_i; \xi)}{\partial \xi}$  and  $G_i(\xi) = \frac{\partial G(x_i; \xi)}{\partial \xi}$

## 5 Simulation Study

In this section, the Maximum likelihood estimators for additional parameters  $\alpha$  and  $\beta$  in pdf (4) for three different states, has been assessed by simulating:  $(\alpha, \beta) = (0.6, 1.6)$ ,  $(\alpha, \beta) = (2, 2)$  and  $(\alpha, \beta) = (13, 0.1)$ . In each three case, the uniform distribution parameters in (4) are  $(a, b) = (0, 10)$ . The density functions for one of the three states, has been indicated in Figure 6. One can see three different states of GOGa-U density functions, means skewed to the left, right and the symmetric.

To verify the validity of the maximum likelihood estimator, Mean Square Error of the Estimate (MSE), Coverage Probability (CP) and Coverage Length (CL) have been used. For example, as described in Section 3.1, for  $(\alpha, \beta) = (0.6, 1.6)$ ,  $N = 10000$  times have been simulated samples of  $n = 30, 40, \dots, 500$  of GOGa-U(0.6, 1.6, 0, 10). To estimate the numerical value of the maximum likelihood, the *optim* function (in the *stat* package) and L-BFGS-B method

in R software has been used. If  $\theta = (\alpha, \beta)$ , for any simulation by  $n$  volume and  $i = 1, 2, \dots, N$ , the maximum likelihood estimates are obtained as  $\hat{\theta}_i = (\hat{\alpha}_i, \hat{\beta}_i)$ . The standard deviation of estimations, which is obtained through the information matrix is shown by  $s_{\hat{\theta}_i} = (s_{\hat{\alpha}_i}, s_{\hat{\beta}_i})$ . In this case, the MLE, Bias, MSE, CP and CL are calculated by the following formula

$$MLE_{\hat{\theta}}(n) = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_i$$

$$Bias_{\hat{\theta}}(n) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)$$

$$MSE_{\hat{\theta}}(n) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2$$

$$CP_{\hat{\theta}}(n) = \frac{1}{N} \sum_{i=1}^N I(\hat{\theta}_i - 1.96s_{\hat{\theta}_i}, \hat{\theta}_i + 1.96s_{\hat{\theta}_i})$$

$$CL_{\hat{\theta}}(n) = \frac{3.92}{N} \sum_{i=1}^N s_{\hat{\theta}_i}$$

In Figures 7 represent the Biases, MSEs, CPs and CLs plots for  $(\alpha, \beta) = (0.6, 1.6)$ . As expected, the biases and MSE of estimated parameters converges to zero while  $n$  growing. The CPs plots should converge to 0.95 and CLs plots should be descending they are correct in Figures 7. Plots of parameters vector  $(\alpha, \beta) = (2, 2)$  and  $(\alpha, \beta) = (13, 0.1)$  have the same position that one can see in Appendix A.

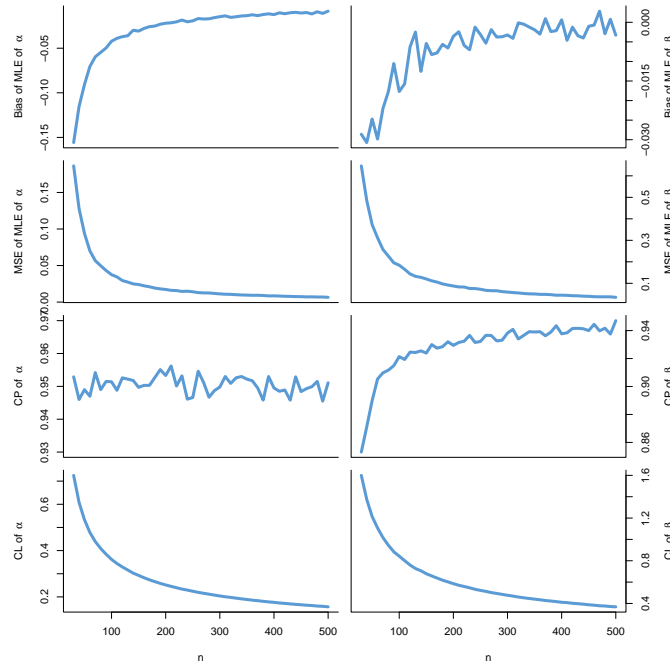


Figure 7: Biases, MSEs, CPs and CLs of  $\hat{\alpha}, \hat{\beta}$  versus  $n$  when  $(\alpha, \beta) = (0.6, 1.6)$ .

## 6 Applications

In this section, fitting of GOGa-U and some famous models to the two real data sets has been assessed. The Akaike information criterion (AIC), Bayesian information criterion (BIC), Anderson-Darling (  $A^*$  ) and Cramér-von Mises (  $W^*$  ) , KolmogorovSmirnov (K.S) and the P-Value of K.S test , have been chosen to comparison of the models. The distributions: Beta Exponential (BE) (Nadarajah and Kotz (2006)), Beta Generalized Exponential (BGE) (Barreto-Souza et al. (2010)), Beta Generalized Half-Normal, (BGHN) (Pescima et al. (2010)), Beta Pareto (BP) (Akinsete et al. (2008)), Exponentiated Pareto (EP) (Kus (2007)), Generalized Half-Normal (GHN) (Cooray and Ananda (2008) ), Gamma-Uniform (GU) (Torabi and Montazeri Hedesh (2012)), Kumaraswamy Gumbel (KwGu) (Cordeiro et al. (2012)) and Weibull-G{ $E$ } (Alzaatreh et al. (2015)) have been selected for comparison. The parameters of models have been estimated by the MLE method.

We also analyzed the hazard rates of these two data sets. In order to identify the shapes of data, we consider the graphical method based on total time on test (TTT) transformed, introduced by Barlow and Campo (1975). The empirical illustration of TTT-transform is given by Aarset (1987). The TTT plot is obtained by plotting  $F(r/n) = \frac{[\sum_{i=1}^r T_{i:n} + (n-r)T_{r:n}]}{[\sum_{i=1}^n T_{i:n}]}$  versus  $r/n$  ( $r = 1, 2, \dots, n$ ).  $T_{i:n}$ , for  $i = 1, 2, \dots, n$ , are the order statistics of the sample.

The first data set presents upside-down bathtub shaped hazard function while the second data sets present decreasing shaped hazard function. From figure 9, the TTT-plot for the data set AG negative shows that hazard function  $\tau(x)$  is first concave and then convex, giving an indication of bathtub- upside-down shape, while in figures 10, TTT-plot for the data set Sun of Skin Folds shows that the hazard rate function is concave . Hence, the GOGa-U family could be in principle an appropriate model for fitting these data sets.

### 6.1 The Myelogenous Leukemia Data for AG Positive

This sub-section is related to study of AG data which presented by Feigl and Zelen (1965) that include 16 observations. Observed survival times (weeks) for AG positive were identified by the presence of Auer rods and significant granulative of the leukemic cells in the bone marrow at diagnosis. For the AG negative patients these factors were absent. The data set is: 56, 65, 17, 17, 16, 22, 3, 4, 2, 3, 8, 4, 3, 30, 4, 43.

the Tables 1 and 2 display a summary of the fitted information criteria and MLEs for this data with different models , respectively. Models have been sorted from the lowest to the highest value of AIC. As you see, the GOGa-U is selected as the best model with all the criteria. Note that P-Value for GOGa-U is also more than all other distributions. The histogram of the AG negative data and the plots of fitted pdf are displayed in Figure 8.

### 6.2 The Sun of Skin Folds Data

The second data set which contains 202 observation can be seen in Weisberg (2005) that have been used in Alzaatreh (2015) (article not yet published). These data are the sum of skin folds in 202 athletes collected at the Australian Institute of Sports and are as follows:

28.0, 98, 89.0, 68.9, 69.9, 109.0, 52.3, 52.8, 46.7, 82.7, 42.3, 109.1, 96.8, 98.3, 103.6, 110.2, 98.1, 57.0, 43.1, 71.1, 29.7, 96.3, 102.8, 80.3, 122.1, 71.3, 200.8, 80.6, 65.3, 78.0, 65.9, 38.9, 56.5, 104.6, 74.9, 90.4, 54.6, 131.9, 68.3, 52.0, 40.8, 34.3, 44.8, 105.7, 126.4, 83.0, 106.9, 88.2, 33.8, 47.6, 42.7, 41.5, 34.6, 30.9, 100.7, 80.3, 91.0, 156.6, 95.4, 43.5, 61.9, 35.2, 50.9, 31.8, 44.0, 56.8, 75.2, 76.2, 101.1, 47.5, 46.2, 38.2, 49.2, 49.6, 34.5, 37.5, 75.9, 87.2, 52.6, 126.4, 55.6, 73.9, 43.5, 61.8,

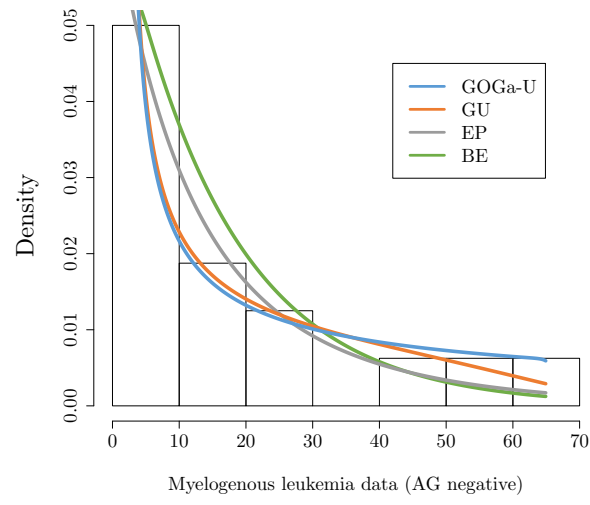


Figure 8: Histogram and estimated pdfs for the AG negative data.

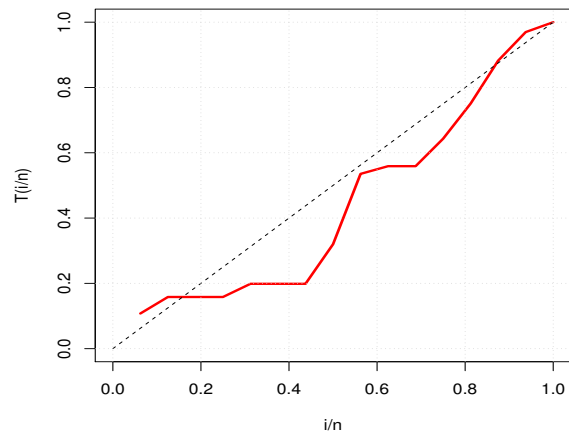


Figure 9: TTT Plot for the AG negative data.

Table 1: Information criteria for the AG negative data.

| Model             | AIC    | BIC    | $W^*$ | $A^*$ | K.S  | P-Value |
|-------------------|--------|--------|-------|-------|------|---------|
| GOGa-U            | 121.29 | 124.38 | 0.07  | 0.46  | 0.18 | 0.687   |
| G-U               | 122.68 | 125.77 | 0.06  | 0.39  | 0.18 | 0.678   |
| EP                | 129.09 | 130.64 | 0.1   | 0.65  | 0.21 | 0.475   |
| BE                | 129.66 | 131.98 | 0.1   | 0.72  | 0.3  | 0.105   |
| GHN               | 130.21 | 131.76 | 0.11  | 0.65  | 0.22 | 0.422   |
| BP                | 131.41 | 134.5  | 0.11  | 0.66  | 0.22 | 0.404   |
| Weibull-G $\{E\}$ | 131.55 | 134.64 | 0.11  | 0.68  | 0.22 | 0.441   |
| BGHN              | 131.83 | 134.93 | 0.11  | 0.67  | 0.23 | 0.356   |
| BGE               | 132.55 | 135.64 | 0.1   | 0.67  | 0.23 | 0.343   |
| KwGu              | 134.22 | 137.31 | 0.1   | 0.65  | 0.3  | 0.123   |

Table 2: MLEs for the the AG negative data.

| Model             | Parameters  |
|-------------------|---|
| GOGa-U            | $(\hat{\alpha}, \hat{\beta}, \hat{a}, \hat{b}) = (0.01, 51.13, 1.99, 66.67)$<br>$(s_{\hat{\alpha}}, s_{\hat{\beta}}, s_{\hat{a}}, s_{\hat{b}}) = (0.01, 62.04, 0.01, 2.63)$             |
| G-U               | $(\hat{\alpha}, \hat{\beta}, \hat{a}, \hat{b}) = (0.40, 0.81, 1.99, 98.91)$<br>$(s_{\hat{\alpha}}, s_{\hat{\beta}}, s_{\hat{a}}, s_{\hat{b}}) = (0.15, 1.20, 0.01, 53.41)$              |
| EP                | $(\hat{\lambda}, \hat{\beta}) = (1.01, 0.04)$<br>$(s_{\hat{\lambda}}, s_{\hat{\beta}}) = (1.88, 0.02)$  |
| BE                | $(\hat{a}, \hat{b}, \hat{\lambda}) = (8.24, 0.04, 1.54)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\lambda}}) = (40.43, 0.08, 2.85)$   |
| GHN               | $(\hat{\alpha}, \hat{\theta}) = (0.74, 22.79)$<br>$(s_{\hat{\alpha}}, s_{\hat{\theta}}) = (0.15, 6.04)$   |
| BP                | $(\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{k}) = (98.66, 3.01, 0.01, 0.53)$<br>$(s_{\hat{\alpha}}, s_{\hat{\beta}}, s_{\hat{\theta}}, s_{\hat{k}}) = (593.87, 17.51, 0.02, 1.80)$  |
| Weibull-G $\{E\}$ | $(\hat{c}, \hat{\gamma}, \hat{\alpha}, \hat{\beta}) = (0.48, 3.09, 5.02, 1.40)$<br>$(s_{\hat{c}}, s_{\hat{\gamma}}, s_{\hat{\alpha}}, s_{\hat{\beta}}) = (0.09, 1.47, 0.50, 0.01)$      |
| BGHN              | $(\hat{a}, \hat{b}, \hat{\alpha}, \hat{\theta}) = (0.03, 76.12, 508.34, 270.67)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\alpha}}, s_{\hat{\theta}}) = (0.04, 4235.56, 1349.84, 471.83)$ |
| BGE               | $(\hat{a}, \hat{b}, \hat{\lambda}, \hat{\alpha}) = (14.23, 6.84, 0.00, 0.13)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\lambda}}, s_{\hat{\alpha}}) = (33.74, 4.52, 0.00, 0.27)$          |
| KwGu              | $(\hat{a}, \hat{b}, \hat{\mu}, \hat{\sigma}) = (0.01, 0.11, 10.51, 1.93)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\mu}}, s_{\hat{\sigma}}) = (0.01, 0.03, 0.01, 0.02)$                   |

88.9, 31.0, 37.6, 52.8, 97.9, 111.1, 114.0, 62.9, 36.8, 56.8, 46.5, 48.3, 32.6, 31.7, 47.8, 75.1, 110.7, 70.0, 52.5, 67, 41.6, 34.8, 61.8, 31.5, 36.6, 76.0, 65.1, 74.7, 77.0, 62.6, 41.1, 58.9, 60.2, 43.0, 32.6, 48, 61.2, 171.1, 113.5, 148.9, 49.9, 59.4, 44.5, 48.1, 61.1, 31.0, 41.9, 75.6, 76.8, 99.8, 80.1, 57.9, 48.4, 41.8, 44.5, 43.8, 33.7, 30.9, 43.3, 117.8, 80.3, 156.6, 109.6, 50.0, 33.7, 54.0, 54.2, 30.3, 52.8, 49.5, 90.2, 109.5, 115.9, 98.5, 54.6, 50.9, 44.7, 41.8, 38.0, 43.2, 70.0, 97.2, 123.6, 181.7, 136.3, 42.3, 40.5, 64.9, 34.1, 55.7, 113.5, 75.7, 99.9, 91.2, 71.6, 103.6, 46.1, 51.2, 43.8, 30.5, 37.5, 96.9, 57.7, 125.9, 49.0, 143.5, 102.8, 46.3, 54.4, 58.3, 34.0, 112.5, 49.3, 67.2, 56.5, 47.6, 60.4, 34.9

the Tables 3 and 4 display a summary of the fitted information criteria and MLEs for this data with different models, respectively. Models have been sorted from the lowest to the highest

value of AIC. As you see, the GOGa-U is selected as the best model with all the criteria. Here P-Value for GOGa-U is also more than all other distributions. The histogram of the sun of skin folds data and the plots of fitted pdf are displayed in Figure 10.

Table 3: Information criteria for the Sun of Skin Folds data.

| Model             | AIC     | BIC     | $W^*$ | $A^*$ | K.S  | P-Value |
|-------------------|---------|---------|-------|-------|------|---------|
| GOGa-U            | 1897.11 | 1910.34 | 0.09  | 0.55  | 0.05 | 0.731   |
| G-U               | 1897.15 | 1910.38 | 0.08  | 0.47  | 0.05 | 0.668   |
| KwGu              | 1906.25 | 1919.48 | 0.11  | 5.05  | 0.06 | 0.393   |
| BP                | 1915.68 | 1928.91 | 0.18  | 1.57  | 0.07 | 0.228   |
| Weibull-G $\{E\}$ | 1916.04 | 1929.27 | 0.2   | 2.4   | 0.07 | 0.34    |
| BGE               | 1920.58 | 1933.81 | 0.26  | 0.76  | 0.08 | 0.179   |
| BGHN              | 1925.11 | 1938.34 | 0.32  | 1.21  | 0.08 | 0.135   |
| BE                | 1930.2  | 1940.13 | 0.4   | 1.25  | 0.09 | 0.063   |
| GHN               | 1978.34 | 1984.96 | 0.86  | 2.36  | 0.13 | 0.002   |
| EP                | 2119.1  | 2125.71 | 0.41  | 1.89  | 0.35 | 0       |

Table 4: MLEs for the Sun of Skin Folds data.

| Model             | Parameters   |
|-------------------|--|
| GOGa-U            | $(\hat{\alpha}, \hat{\beta}, \hat{a}, \hat{b}) = (8.95, 0.04, 27.99, 650.04)$<br>$(s_{\hat{\alpha}}, s_{\hat{\beta}}, s_{\hat{a}}, s_{\hat{b}}) = (1.96, 0.01, 0.02, 169.55)$          |
| G-U               | $(\hat{\alpha}, \hat{\beta}, \hat{a}, \hat{b}) = (1.27, 0.07, 27.88, 579.87)$<br>$(s_{\hat{\alpha}}, s_{\hat{\beta}}, s_{\hat{a}}, s_{\hat{b}}) = (0.17, 0.05, 0.25, 334.16)$          |
| KwGu              | $(\hat{a}, \hat{b}, \hat{\mu}, \hat{\sigma}) = (0.01, 0.21, 68.93, 7.15)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\mu}}, s_{\hat{\sigma}}) = (0.01, 0.06, 2.22, 1.78)$                  |
| BP                | $(\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{k}) = (102.94, 4.20, 3.31, 1.14)$<br>$(s_{\hat{\alpha}}, s_{\hat{\beta}}, s_{\hat{\theta}}, s_{\hat{k}}) = (183.57, 4.22, 3.43, 0.63)$ |
| Weibull-G $\{E\}$ | $(\hat{c}, \hat{\gamma}, \hat{\alpha}, \hat{\beta}) = (0.01, 0.21, 68.93, 7.15)$<br>$(s_{\hat{c}}, s_{\hat{\gamma}}, s_{\hat{\alpha}}, s_{\hat{\beta}}) = (0.01, 0.06, 2.22, 1.78)$    |
| BGE               | $(\hat{a}, \hat{b}, \hat{\lambda}, \hat{\alpha}) = (1.23, 0.73, 0.05, 8.77)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\lambda}}, s_{\hat{\alpha}}) = (1.98, 0.25, 0.01, 16.52)$          |
| BGHN              | $(\hat{a}, \hat{b}, \hat{\alpha}, \hat{\theta}) = (0.27, 34.18, 43.46, 13.78)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\alpha}}, s_{\hat{\theta}}) = (0.10, 37.32, 31.62, 0.50)$        |
| BE                | $(\hat{a}, \hat{b}, \hat{\lambda}) = (5.34, 5.86, 0.01)$<br>$(s_{\hat{a}}, s_{\hat{b}}, s_{\hat{\lambda}}) = (0.53, 1.62, 0.00)$   |
| GHN               | $(\hat{\alpha}, \hat{\theta}) = (1.65, 86.05)$<br>$(s_{\hat{\alpha}}, s_{\hat{\theta}}) = (0.09, 2.89)$  |
| EP                | $(\hat{\lambda}, \hat{\beta}) = (0.01, 0.01)$<br>$(s_{\hat{\lambda}}, s_{\hat{\beta}}) = (1.65, 86.05)$  |

## 7 Conclusions

In many applied areas there is a clear need for extended forms of the well-known distributions. Generally, the new distributions are more flexible to model real data that present a high



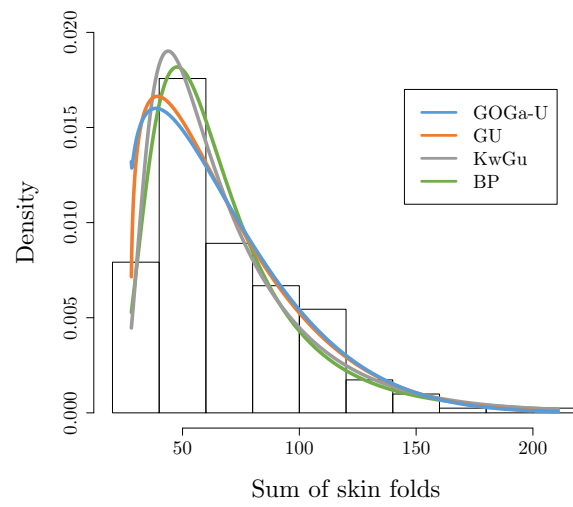


Figure 10: Histogram and estimated pdfs for the Sun of Skin Folds data.

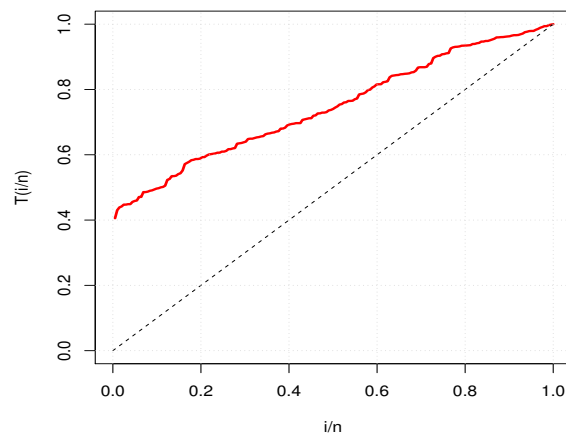


Figure 11: TTT Plot for the Sun of Skin Folds data.

degree of skewness and kurtosis. We propose Generalized Odd Gamma-G (GOGa-G) family of distributions. Many well-known models emerge as special cases of the GOGa-G family by using special parameter values. Some mathematical properties of the new class including explicit expansions for the ordinary and incomplete moments, quantile and generating functions, mean deviations, entropies and order statistics are provided. We survey the theoretical outcomes with numerical computation by using *R* software. The parameters of this model are estimated by the maximum likelihood estimation method. We prove empirically by means of an application to a real data set that special cases of the proposed family can give better fits than other models generated from well-known families.

## Appendix A

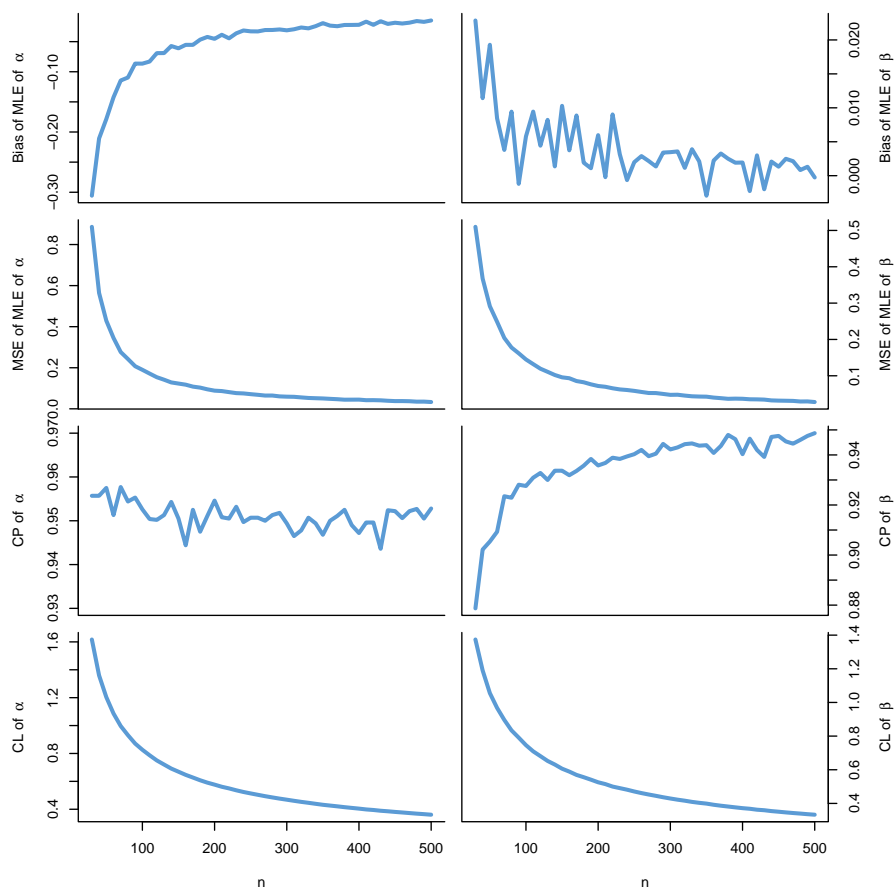


Figure 12: Biases, MSEs, CPs and CLs of  $\hat{\alpha}, \hat{\beta}$  versus  $n$  when  $(\alpha, \beta) = (2, 2)$ .

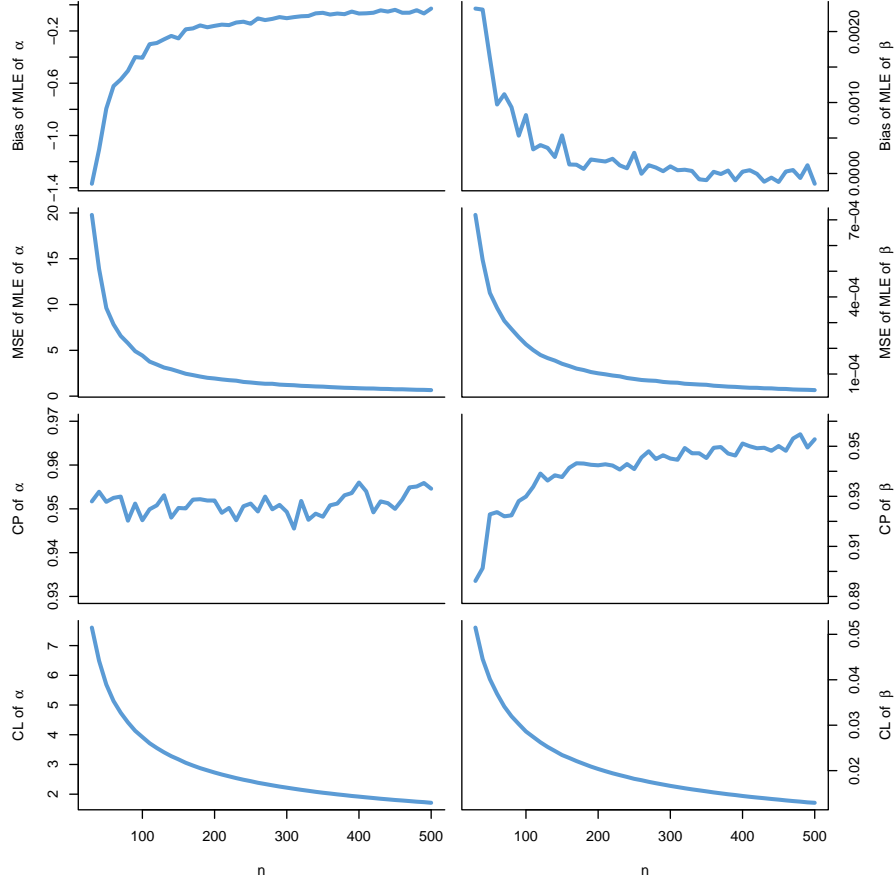


Figure 13: Biases, MSEs, CPs and CLs of  $\hat{\alpha}, \hat{\theta}$  versus  $n$  when  $(\alpha, \beta) = (13, 0.1)$ .

## References

- [1] AUTHORC, et al. (2014a). The Exponentiated half-logistic family of distributions: properties and applications. *Journal of Probability and Statistics*, Article ID 864396, 21 pages,
- [2] AUTHORC, et al. (2015). The beta Marshall-Olkin family of distributions. *Journal of Statistical Distributions and Applications*, **2**(1), 1-18.
- [3] AUTHORC, et al. (2016). The type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation*, **86**(4), 707-728.
- [4] Aarset, M. V. (1987). How to identify bathtub hazard rate. *IEEE Trans. Reliab*, **36**, 106-108.
- [5] Akinsete, A., Famoye, F. & Lee, C. (2008). The beta-Pareto distribution. *A Journal of Theoretical and Applied Statistics*, **42:6**, 547-563.
- [6] Alexander, C., Cordeiro, G.M., Ortega, E.M.M. & Sarabia, J.M. (2012). Generalized beta-generated distributions. *Computational Statistics and Data Analysis*, **56**, 1880-1897.
- [7] Alzaatreh, A., Lee, C. & Famoye, F. (2013). A new method for generating families of distributions. *Metron*, **71**, 63-79.

- [8] Alzaatreh, A., Lee, C. & Famoye, F. (2015). Family of generalized gamma distributions: Properties and applications. *Hacettepe Journal of Mathematics and Statistics* (to appear).
- [9] Barlow, R. E. and Campo, R. A. (1975). Total time on test processes and applications to failure data analysis. In: R.E. Barlow, J.B. Fussel, and N.D. Singpurwalla (eds.), *Reliability and Fault Tree Analysis*, Society for Industrial and Applied Mathematics, Philadelphia, 451-481.
- [10] Barreto-Souza, W., Santos, H.S. A., & Gauss, M.C. (2010). The beta generalized exponential distribution. *Reliability Engineering & System Safety*, **80:2**, 159-172.
- [11] Bourguignon, M., Silva, R.B. & Cordeiro, G.M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science*, **12**, 53-68.
- [12] Cooray, K. & Ananda, M.M.A. (2008). A Generalization of the Half-Normal Distribution with Applications to Lifetime Data. *Communications in Statistics - Theory and Methods*, **37:9**, 1323-1337.
- [13] Cordeiro, G.M. & de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation* **81**: 883-893.
- [14] Cordeiro, G. M., Nadarajah, S., & Ortega, E. M. (2012). The Kumaraswamy Gumbel distribution. *Statistical Methods & Applications*, **21(2)**, 139-168.
- [15] Cordeiro, G.M., Ortega, E.M.M., Popović, B. & Pescim, R. R. (2014b). The Lomax generator of distributions: Properties, minification process and regression model. *Applied Mathematics and Computation*, **247**, 465-486.
- [16] Feigl, P. & Zelen, M. (1965). Estimation of exponential survival probabilities with concomitant information. *Biometrics*, **4**, 826-838.
- [17] Gradshteyn, I.S., Ryzhik, I.M. (2000). Table of integrals, series, and products. Academic Press, San Diego.
- [18] Kus, C. (2007). A new lifetime distribution. *Computational Statistics & Data Analysis*, **51**, 4497-4509.
- [19] Marshall, A.N. & Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika*, **84**, 641-652.
- [20] Nadarajah, S. & Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering & System Safety*, **91:6**, 689-697.
- [21] Pescim, R.R., Demtrio, C.G.B., Cordeiro, G.M., Ortega, E.M.M. & Urbano, M.R. (2010). The beta generalized half-normal distribution. *Computational Statistics & Data Analysis*, **54:4**, 945-957.
- [22] Rényi, A. (1961). On measures of entropy and information. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, **I**, 547-561.
- [23] Torabi, H., Montazeri Hedesh, N., (2012). The gamma-uniform distribution and its applications. *Kybernetika*, **48**, 16-30.
- [24] Weisberg, S. (2005). Applied Linear Regression, Third Edition. *Wiley Series in Probability and Statistics*.