STATISTICA, anno LXXX, n. 1, 2020

THE ODDS GENERALIZED GAMMA-G FAMILY OF DISTRIBUTIONS: PROPERTIES, REGRESSIONS AND APPLICATIONS

M. Arslan Nasir
Department of Statistics, Govt S.E College Bahawalpur, Bahawalpur, Punjab 63100, Pakistan
Muhammad H. Tahir
Department of Statistics, The Islamia University Bahawalpur, Punjab, Pakistan
Christophe Chesneau [1]
LMNO, University of Caen Normandie, Caen 14032, France
Farrukh Jamal
Department of Statistics, Govt. S.A Postgraduate College Dera Nawab Sahib, Bahawalpur, Punjab 63360, Pakistan
M. Akbar Ali Shah
Department of Statistics, The Islamia University Bahawalpur, Punjab, Pakistan

1. INTRODUCTION

The deep analysis of data from complex phenomena is often limited by the use of classical (probability) distributions. This motivates the developments of new distributions/ models having the ability to capture the fine features hidden behind the data. In this regard, the most common approach is to use "generators of distributions" aiming to provide more flexible properties to a well-known baseline distribution. Among the most popular generators, there are the beta generator (see Eugene *et al.*, 2002; Jones, 2004), the transmuted generator (see Shaw and Buckley, 2007), the Kumaraswamy generator (see Cordeiro and de Castro, 2011), the McDonald generator (see Alexander *et al.*, 2012), the Kummer beta generator (see Pescim *et al.*, 2012), the gamma generator (see Zografos and Balakrishnan, 2009; Ristíc and Balakrishnan, 2012; Torabi and Montazari, 2012), the log-gamma generator (see Amini *et al.*, 2014), the logistic generator (see Torabi and Montazari, 2014), the beta extended Weibull generator (see Cordeiro *et al.*, 2012), the transformed-transformer (T-X) generator (see Alzaatreh *et al.*, 2013), the exponentiated

[1] Corresponding Author. E-mail: christophe.chesneau@unicaen.fr

T-X generator (see Alzaghal *et al.*, 2013), the Weibull generator (see Alzaatreh *et al.*, 2013; Bourguignon *et al.*, 2014), the exponentiated half-logistic generator (see Cordeiro *et al.*, 2014), the sine generator (see Kumar *et al.*, 2015), the odds Burr III generator (see Jamal *et al.*, 2017), the cosine-sine generator (see Chesneau *et al.*, 2018) and the generalized odds gamma generator (see Hosseini *et al.*, 2018).

In particular, it is demonstrated in Brito *et al.* (2017) that the combination of gamma generator with the odds transformation generates very flexible distributions, with great advantages in data analysis (providing more flexible kurtosis in comparison to the baseline distribution, producing a skewness for symmetrical distributions, generating distributions with symmetric, left-skewed, right-skewed and reversed-J shaped ...). It was recently generalized by Hosseini *et al.* (2018) by the consideration of a generalized odds function, with practical benefits. The aim of this paper is to propose an interesting alternative to the generator introduced by Brito *et al.* (2017) and Hosseini *et al.* (2018). It is constructed from the generalized gamma distribution and the (standard) odds transformation. It allows us to define a new family of distributions described below. Let us recall that the probability density function (pdf) of generalized gamma distribution is specified by

$$f(x;\alpha,\beta,\delta)=\frac{\alpha\delta^{\alpha\beta}}{\Gamma(\beta)}x^{\alpha\beta-1}\exp[-(\delta x)^{\alpha}],\qquad x>0,$$

where $\alpha>0$ and $\beta>0$ are shape parameters and $\delta>0$ is a scale parameter. If we take $\delta=1$, then $f(x;\alpha,\beta,\delta)$ is reduced to the pdf $r(x;\alpha,\beta)$ given by

$$r(x;\alpha,\beta)=\frac{\alpha}{\Gamma(\beta)}x^{\alpha\beta-1}\exp(-x^{\alpha}),\qquad x>0,$$

where $\Gamma(\beta)=\int_0^{+\infty}w^{\beta-1}\exp(-w)dw$ is the gamma function. Let us now consider a cumulative distribution function (cdf) of a baseline distribution denoted by $G(x;\xi)$, where ξ represents the related parameter vector. Then, we propose to use a generalized gamma generator with the standard odds transformation defined by $W[G(x;\xi)]=G(x;\xi)/[1-G(x;\xi)]$. This yields the following cdf:

$$F(x;\alpha,\beta,\xi)=\int_{-\infty}^{W[G(x;\xi)]}r(t;\alpha,\beta)dt=\frac{\gamma\left[\beta,\left(\frac{G(x;\xi)}{1-G(x;\xi)}\right)^{\alpha}\right]}{\Gamma(\beta)},\qquad x\in\mathbb{R},\qquad (1)$$

where $\gamma(\beta,x)=\int_0^{x}w^{\beta-1}\exp(-w)dw$ is the (lower) incomplete gamma function. By (almost everywhere) differentiation with respect to x, the corresponding pdf is given as

$$f(x;\alpha,\beta,\xi)=\frac{\alpha}{\Gamma(\beta)}\frac{g(x;\xi)G^{\alpha\beta-1}(x;\xi)}{(1-G(x;\xi))^{\alpha\beta+1}}\exp\left\{-\left[\frac{G(x;\xi)}{1-G(x;\xi)}\right]^{\alpha}\right\},\qquad x\in\mathbb{R}.\qquad (2)$$

The corresponding hazard rate function (hrf) is obtained as

$$h(x;\alpha,\beta,\xi) = \frac{f(x;\alpha,\beta,\xi)}{1-F(x;\alpha,\beta,\xi)}$$
$$= \frac{\alpha g(x;\xi) G^{\alpha\beta-1}(x;\xi)}{(1-G(x;\xi))^{\alpha\beta+1}\left\{\Gamma(\beta)-\gamma\left[\beta,\left(\frac{G(x;\xi)}{1-G(x;\xi)}\right)^{\alpha}\right]\right\}} \exp\left\{-\left[\frac{G(x;\xi)}{1-G(x;\xi)}\right]^{\alpha}\right\},$$
$$x \in \mathbb{R}. \quad (3)$$

For the purpose of the paper, the generator characterized by the cdf (1) is called the odds generalized gamma generator and the corresponding family of distributions is called odds generalized gamma G (GG-G) family of distributions. To the best of our knowledge, it is new in the literature, even if connections exist with the generalized odds gamma-G family introduced by Hosseini *et al.* (2018) (the families coincide by taking $\alpha = 1$ and, in the definition of the generalized odds gamma-G family, $\beta = 1$). In this study, we show how our family is complementary on several aspects, and can be superior in terms of goodness-of-fit in comparison to the generalized odds gamma-G family.

The rest of this paper is planned as follows. In Section 2, three special models are given, with plots of their pdfs and hrfs illustrating their flexibility. The one using the Fréchet distribution as baseline will be at the heart of our applied study. The main mathematical properties of the GG-G family are studied in Section 3, including shapes and asymptotes of the pdf and hrf, mixture representations of the pdf and cdf in terms of functions of the baseline distribution, explicit expressions for the rth moment, rth incomplete moment, moment generating function, mean deviations, Rényi entropy, reliability parameter and the pdf of the ith order statistic. The estimation of the related parameters by the maximum likelihood method is discussed in Section 4. In Section 5, a regression model is proposed, with a simulation study. Section 6 is devoted to the residual analysis. Applications to real life data sets are performed in Section 7.

2. Special sub distributions

In this section, we study three special sub distributions of the GG-G family, namely odds generalized gamma Fréchet (GGFr), odds generalized gamma Weibull (GGW) and odds generalized gamma Lomax (GGLx) distributions.

2.1. *Odds generalized gamma Fréchet distribution*

Let us consider the Fréchet distribution as baseline distribution, i.e., with pdf $g(x;a,b) = (a^b b/x^{b+1})\exp\left[-(a/x)^b\right]$ and cdf $G(x;a,b) = \exp\left[-(a/x)^b\right]$, where $a,b,x > 0$. Then, the cdf of the GGFr distribution is given by

$$F(x;\alpha,\beta,a,b) = \frac{\gamma\left(\beta,\left[\exp\left[(a/x)^b\right]-1\right]^{-\alpha}\right)}{\Gamma(\beta)}, \qquad x > 0.$$

The pdf can be expressed as

$$f(x;\alpha,\beta,a,b)=$$
$$\frac{\alpha a^b b\exp\left[-\alpha\beta(a/x)^b\right]}{\Gamma(\beta)x^{b+1}\left\{1-\exp\left[-(a/x)^b\right]\right\}^{\alpha\beta+1}}\exp\left\{-\left[\exp\left[\left(\frac{a}{x}\right)^b\right]-1\right]^{-\alpha}\right\},\qquad x>0.$$

Also, the hrf is given as

$$h(x;\alpha,\beta,a,b)=$$
$$\frac{\alpha a^b b\exp\left[-\alpha\beta(a/x)^b\right]}{x^{b+1}\left\{1-\exp\left[-(a/x)^b\right]\right\}^{\alpha\beta+1}\left[\Gamma(\beta)-\gamma\left(\beta,\left[\exp\left[(a/x)^b\right]-1\right]^{-\alpha}\right)\right]}\times$$
$$\exp\left\{-\left[\exp\left[\left(\frac{a}{x}\right)^b\right]-1\right]^{-\alpha}\right\},\qquad x>0.$$

2.2. *Odds generalized gamma Weibull distribution*

Now, let us consider the gamma Weibull distribution as baseline distribution, i.e., with pdf $g(x;a,b)=abx^{b-1}\exp\left(-ax^b\right)$ and cdf $G(x;a,b)=1-\exp\left(-ax^b\right)$, where $a,b,x>0$. Then, the cdf of the GGW distribution is defined by

$$F(x;\alpha,\beta,a,b)=\frac{\gamma\left(\beta,\left[\exp\left(ax^b\right)-1\right]^{\alpha}\right)}{\Gamma(\beta)},\qquad x>0.$$

The pdf is given by

$$f(x;\alpha,\beta,a,b)=$$
$$\frac{\alpha}{\Gamma(\beta)}abx^{b-1}\left[1-\exp\left(-ax^b\right)\right]^{\alpha\beta-1}\exp\left\{a\alpha\beta x^b-\left[\exp\left(ax^b\right)-1\right]^{\alpha}\right\},\qquad x>0.$$

Also, the hrf is given as

$$h(x;\alpha,\beta,a,b)=$$
$$\frac{\alpha abx^{b-1}\left[1-\exp\left(-ax^b\right)\right]^{\alpha\beta-1}}{\Gamma(\beta)-\gamma\left(\beta,\left[\exp\left(ax^b\right)-1\right]^{\alpha}\right)}\exp\left\{a\alpha\beta x^b-\left[\exp\left(ax^b\right)-1\right]^{\alpha}\right\},\qquad x>0.$$

2.3. *Odds generalized gamma Lomax distribution*

Finally, let us chose the Lomax distribution as baseline distribution, i.e., with pdf $g(x;a,b)=(a/b)(1+x/b)^{-a-1}$ and cdf $G(x;a,b)=1-(1+x/b)^{-a}$, where $a,b,x>0$. Then, the cdf of the GGLx distribution is expressed as

$$F(x;\alpha,\beta,a,b)=\frac{\gamma(\beta,[(1+x/b)^{a}-1]^{\alpha})}{\Gamma(\beta)},\qquad x>0.$$

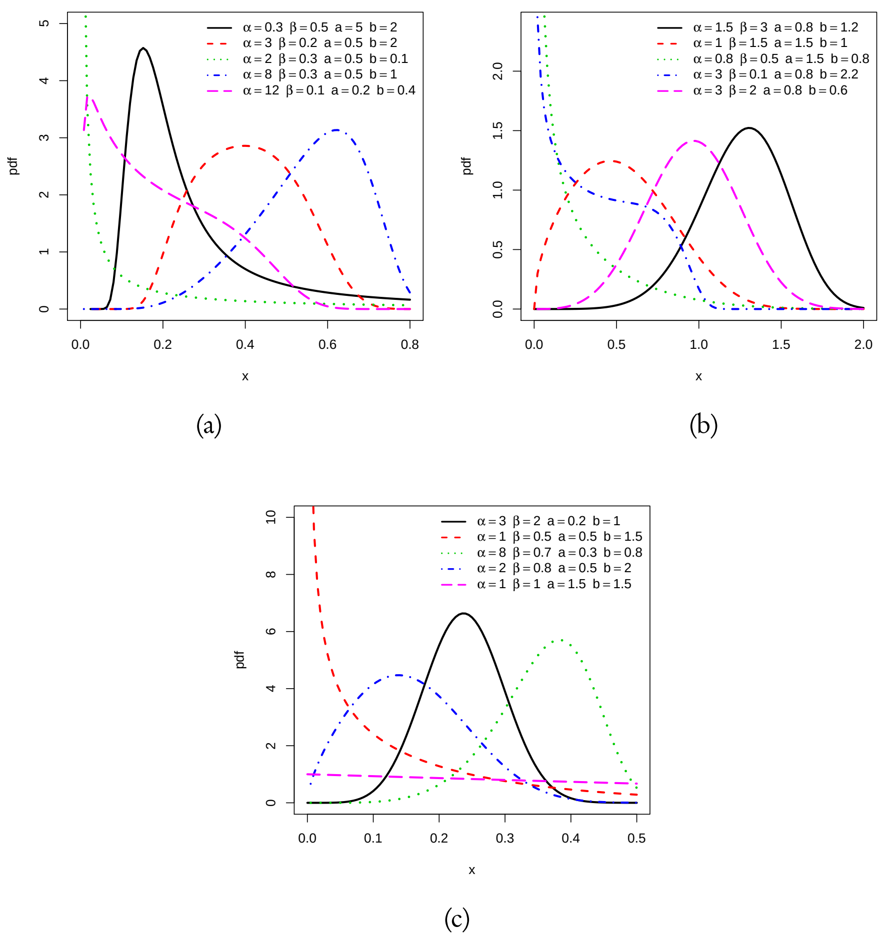


Figure 1 – Plots of pdfs of the GGFr (a), GGW (b) and GGLx (c) distributions.

The pdf is given as

$$f(x;\alpha,\beta,a,b)=\frac{\alpha}{\Gamma(\beta)}\frac{a}{b}\left(1+\frac{x}{b}\right)^{a\alpha\beta-1}\left[1-\left(1+\frac{x}{b}\right)^{-a}\right]^{\alpha\beta-1}\exp\left\{-\left[\left(1+\frac{x}{b}\right)^{a}-1\right]^{\alpha}\right\},$$
$$x>0.$$

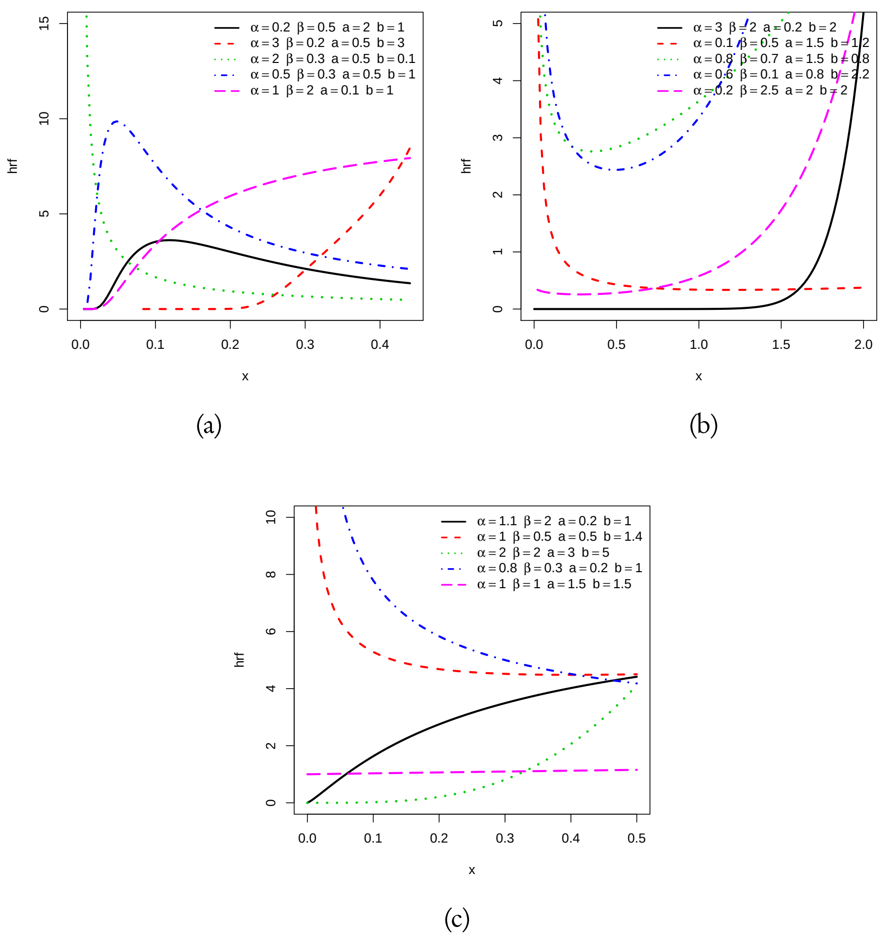


Figure 2 – Plots of hrfs of the GGFr (a), GGW (b) and GGLx (c) distributions.

Also, the hrf can be expressed as

$$h(x;\alpha,\beta,a,b)=$$
$$\frac{\alpha a(1+x/b)^{a\alpha\beta-1}\left[1-(1+x/b)^{-a}\right]^{\alpha\beta-1}}{b\left[\Gamma(\beta)-\gamma(\beta,[(1+x/b)^{a}-1]^{\alpha})\right]}\exp\left\{-\left[\left(1+\frac{x}{b}\right)^{a}-1\right]^{\alpha}\right\},\qquad x>0.$$

Figure 1 illustrates the pdfs of the GGFr, GGW and GGLx distributions for selected values of the parameters. Also, Figure 2 illustrates the hrfs of the GGFr, GGW and GGLx distributions for selected values of the parameters. Various forms of shapes are observed, showing the great flexibility of these special sub distributions, and, a fortiori, of the overall GG-G family.

3. Mathematical properties

This section is devoted to some notable mathematical properties of the GG-G family of distributions.

3.1. *Characterization*

Let $G^{-1}(x;\xi)$ be the inverse function of $G(x;\xi)$ (i.e., the quantile function of the baseline distribution), and V a random variable having the gamma distribution with parameters 1 and β. Then, the random variable $X=G^{-1}\left\{V^{1/\alpha}/(1+V^{1/\alpha});\xi\right\}$ has the cdf of the GG-G family.

3.2. *Quantile function*

Let $\gamma^{-1}(\beta,x)$ be the inverse function of $\gamma(\beta,x)$. Then, after some developments, the quantile function of the GG-G family can be expressed as

$$Q(y,\xi)=G^{-1}\left(\frac{[\gamma^{-1}(\beta,y\Gamma(\beta))]^{1/\alpha}}{1+[\gamma^{-1}(\beta,y\Gamma(\beta))]^{1/\alpha}},\xi\right),\quad y\in(0,1).$$

Among others, several descriptive parameters of the GG-G family can be defined, such as the median defined as $M_{ed}=Q(0.5,\xi)$.

3.3. Shape analysis of the crucial functions

The shapes of the pdf and the hrf of the GG-G family can be described analytically. The critical points of the pdf $f(x;\alpha,\beta,\xi)$ given by (2) are the roots of the following equation:

$$\frac{g^x(x;\xi)}{g(x;\xi)}+(\alpha\beta-1)\frac{g(x;\xi)}{G(x;\xi)}+(\alpha\beta+1)\frac{g(x;\xi)}{1-G(x;\xi)}$$
$$-\alpha\left\{\frac{G(x;\xi)}{1-G(x;\xi)}\right\}^{\alpha-1}\frac{g(x;\xi)}{\{1-G(x;\xi)\}^2}=0, \tag{4}$$

where $g^x(x;\xi)=\partial g(x;\xi)/\partial x$. The critical points of $h(x;\alpha,\beta,\xi)$ expressed by (3) are obtained from the following equation:

$$\frac{g^x(x;\xi)}{g(x;\xi)}+(\alpha\beta-1)\frac{g(x;\xi)}{G(x;\xi)}+(\alpha\beta+1)\frac{g(x;\xi)}{1-G(x;\xi)}$$
$$-\alpha\left\{\frac{G(x;\xi)}{1-G(x;\xi)}\right\}^{\alpha-1}\frac{g(x;\xi)}{\{1-G(x;\xi)\}^2}$$
$$+\frac{\alpha g(x;\xi)G^{\alpha\beta-1}(x;\xi)}{(1-G(x;\xi))^{\alpha\beta+1}\left\{\Gamma(\beta)-\gamma\left[\beta,\left(\frac{G(x;\xi)}{1-G(x;\xi)}\right)^{\alpha}\right]\right\}}\exp\left\{-\left[\frac{G(x;\xi)}{1-G(x;\xi)}\right]^{\alpha}\right\}=0. \tag{5}$$

By using most of the symbolic computation software platforms, we can examine the equations (4) and (5) to determine the local maxima and minima and inflexion points.

3.4. Mixture representation

Here, we plan to give the mixture representations of the pdf and cdf of the GG-G family in terms of functions of the exp-G family, which will be useful for the derivation of further properties.

First of all, let us recall that the exponential series expansion is formulated as, for any $x\in\mathbb{R}$,

$$\exp(-ax)=\sum_{i=0}^{+\infty}\frac{(-1)^i a^i}{i!}x^i. \tag{6}$$

On the other side, the generalized binomial series expansion is given by, for any x such that $|x|<1$,

$$(1-x)^b=\sum_{j=0}^{+\infty}\binom{b}{j}(-1)^j x^j, \tag{7}$$

where $\binom{b}{j}=b(b-1)\ldots(b-j+1)/(j!)$. These two formulas will be useful in our mathematical developments. Using the series (6) and (7), the pdf $f(x;\alpha,\beta,\xi)$ given by (2) becomes

$$f(x;\alpha,\beta,\xi)=\sum_{i,j=0}^{+\infty}u_{i,j}\,g(x;\xi)G^{\alpha(\beta+i)+j-1}(x;\xi),$$

with $u_{i,j}=\frac{\alpha}{\Gamma(\beta)}\binom{-\alpha(\beta+i)-1}{j}\frac{(-1)^{i+j}}{i!}$. Now, by writing

$$G^{\alpha(\beta+i)+j-1}(x;\xi)=[1-(1-G(x;\xi))]^{\alpha(\beta+i)+j-1},$$

we get

$$f(x;\alpha,\beta,\xi)=\sum_{i,j,m=0}^{+\infty}\sum_{\ell=m}^{+\infty}w_{i,j,\ell,m}\,g(x;\xi)G^{m}(x;\xi),$$

where $w_{i,j,\ell,m}=u_{i,j}\binom{\alpha(\beta+i)+j-1}{\ell}\binom{\ell}{m}(-1)^{\ell+m}$. Rewriting the above expression, we arrive at

$$f(x;\alpha,\beta,\xi)=\sum_{m=0}^{+\infty}a_m h_{m+1}(x;\xi), \tag{8}$$

where

$$a_m=\frac{\alpha}{\Gamma(\beta)}\frac{(-1)^m}{m+1}\sum_{i,j=0}^{+\infty}\sum_{\ell=m}^{+\infty}\binom{-\alpha(\beta+i)-1}{j}\binom{\alpha(\beta+i)+j-1}{\ell}\binom{\ell}{m}\frac{(-1)^{i+j+\ell}}{i!} \tag{9}$$

and $h_{m+1}(x;\xi)=(m+1)g(x;\xi)G^m(x;\xi)$.

Moreover, by integrating this equation with respect to x, the cdf of the GG-G family can be expressed as

$$F(x;\alpha,\beta,\xi)=\sum_{m=0}^{+\infty}a_m H_{m+1}(x;\xi), \tag{10}$$

where $H_{m+1}(x;\xi)=G^{m+1}(x;\xi)$. The expression in (8) and (10) are the infinite mixtures representation of the pdf of the GG-G family in terms of functions (pdf and cdf) of the exp-G family.

In the context of the GGFr distribution: We can express $f(x;\alpha,\beta,\xi)$ as (8) with $h_{m+1}(x;\xi)=(m+1)(a^b b/x^{b+1})\exp\left[-(m+1)(a/x)^b\right]$, $\xi=(a,b)$, $x,a,b>0$.

3.5. *Moments and moment generating function*

Now, we give the explicit expression for the rth moment, rth incomplete moment, moment generating function and mean deviation about the mean.
The rth moment of the GG-G family can be obtained by using following formula

$$\mu'_r=\int_{-\infty}^{+\infty}x^r f(x;\alpha,\beta,\xi)dx,$$

where $f(x;\alpha,\beta,\xi)$ is given by (2). Assuming that the sum and integral are interchangeable, it follows from the infinite mixture representation expressed in (8) that

$$\mu'_r=\sum_{m=0}^{+\infty}a_m\tau_m^r,$$

where

$$\tau_m^r=\int_{-\infty}^{+\infty}x^r h_{m+1}(x;\xi)dx. \tag{11}$$

In the context of the GGFr distribution: One can show that, for $r<b$, we have $\tau_m^r=(m+1)^{r/b}a^r\Gamma(1-r/b)$.

Similarly, the rth incomplete moment of the GG-G family can be expressed as

$$T^r(x;\alpha,\beta,\xi)=\sum_{m=0}^{+\infty}a_m\Delta_m^r(x;\xi),$$

where

$$\Delta_m^r(x;\xi)=\int_{-\infty}^{x}t^r h_{m+1}(t;\xi)dt. \tag{12}$$

In the context of the GGFr distribution: One can show that, for $r<b$, we have $\Delta_m^r(x;\xi)=(m+1)^{r/b}a^r\Gamma\left(1-r/b,(m+1)a^b x^{-b}\right)$, $\xi=(a,b)$, $a,b>0$, where $\Gamma(s,x)$ denotes the upper incomplete gamma function defined by $\Gamma(s,x)=\int_x^{+\infty}t^{s-1}e^{-t}dt$, $s,x>0$.

The mean deviation about the mean of the GG-G family is given by

$$D_1=2\mu'_1F(\mu'_1;\alpha,\beta,\xi)-2T^1(\mu'_1;\alpha,\beta,\xi)$$

and the mean deviation about the median of the GG-G family is given as

$$D_2=\mu'_1F(\mu'_1;\alpha,\beta,\xi)-2T^1(M_{ed};\alpha,\beta,\xi).$$

The moment generating function of the GG-G family is

$$M_0(t;\alpha,\beta,\xi)=\sum_{m=0}^{+\infty}a_mM_x(t;\xi),$$

where

$$M_x(t;\xi)=\int_{-\infty}^{+\infty}\exp(tx)h_{m+1}(x;\xi)dx. \tag{13}$$

Note that the integrals in (11), (12) and (13) only depend on the choice of baseline distribution.

3.6. *Reliability parameter*

In the context of reliability, the stress-strength model defines the life of an element which has a random strength X_1 that is subjected to an accidental stress X_2. The component fails at the instant that the stress applied to it exceeds the strength, and the component will work suitably whenever $X_1 > X_2$. Hence, by modeling X_1 and X_2 as random variables, the probability $R = P(X_2 < X_1)$ is a measure of components reliability. It has many applications, especially in the area of reliability and engineering. In what follows, we derive the reliability R when X_1 and X_2 are independent and belongs to the GG-G family with pdf $f(x;\alpha_1,\beta_1,\xi)$ for X_1 and with cdf $F(x;\alpha_2,\beta_2,\xi)$ for X_2 (note that they have the same baseline parameter(s) ξ). From (1) and (2), we can write

$$R = \int_{-\infty}^{+\infty} f(x;\alpha_1,\beta_1,\xi)F(x;\alpha_2,\beta_2,\xi)dx.$$

Using the mixture representations of the pdf and cdf provided in (8) and (10), respectively, we have

$$R = \sum_{m=0}^{+\infty}\sum_{k=0}^{+\infty} a_m^{(1)} a_k^{(2)} \int_{-\infty}^{+\infty} h_{m+1}(x;\xi)H_{k+1}(x;\xi)dx,$$

where $a_m^{(1)}$ and $a_k^{(2)}$ are defined by (9) with (α_1,β_1) and (α_2,β_2) instead of (α,β), respectively.

Since $h_{m+1}(x;\xi)H_{k+1}(x;\xi) = (m+1)g(x;\xi)G^{m+k+1}(x;\xi)$, we have

$$\begin{aligned} R &= \sum_{m=0}^{+\infty}\sum_{k=0}^{+\infty} a_m^{(1)} a_k^{(2)} (m+1)\int_{-\infty}^{+\infty} g(x;\xi)G^{m+k+1}(x;\xi)dx \\ &= \sum_{m=0}^{+\infty}\sum_{k=0}^{+\infty} a_m^{(1)} a_k^{(2)} \frac{m+1}{m+k+2}. \end{aligned}$$

Let us observe that, under the above setting, R does not depend on the choice of baseline distribution.

3.7. *Entropies*

Let $\delta > 0$ with $\delta \neq 1$. Then, the Réyni entropy of the GG-G family is defined by

$$I_\delta = \frac{1}{1-\delta}\log\left[\int_{-\infty}^{+\infty} f^{\delta}(x;\alpha,\beta,\xi)dx\right]. \tag{14}$$

Using (6) and (7), we have

$$f^{\delta}(x;\alpha,\beta,\xi) = \left(\frac{\alpha}{\Gamma(\beta)}\right)^{\delta} \sum_{i,j=0}^{+\infty} u_{i,j,\delta}\, g^{\delta}(x;\xi)G^{j+\delta(\alpha\beta-1)+\alpha i}(x;\xi),$$

where $u_{i,j,\delta} = \binom{-\alpha i - \delta(\alpha\beta+1)}{j}\frac{(-1)^{i+1}\delta^i}{i!}$. Now (14) becomes

$$I_\delta = \frac{1}{1-\delta}\Big\{\delta\log\alpha - \delta\log\Gamma(\beta) + \log\Big[\sum_{i,j=0}^{+\infty} u_{i,j,\delta}\int_{-\infty}^{+\infty} g^\delta(x;\xi)G^{j+\delta(\alpha\beta-1)+\alpha i}(x;\xi)dx\Big]\Big\}.$$

We observe that I_δ depends only for any choice of baseline distribution.

The δ-entropy (or Tsallis entropy) is defined by

$$H_\delta = \frac{1}{\delta-1}\Big[1 - \int_{-\infty}^{+\infty} f^\delta(x;\alpha,\beta,\xi)dx\Big].$$

So, we have

$$H_\delta = \frac{1}{\delta-1}\left[1 - \left(\frac{\alpha}{\Gamma(\beta)}\right)^\delta \sum_{i,j=0}^{+\infty} u_{i,j,\delta}\int_{-\infty}^{+\infty} g^\delta(x;\xi)G^{j+\delta(\alpha\beta-1)+\alpha i}(x;\xi)dx\right].$$

Finally, the Shannon entropy of the GG-G family is defined by

$S = -\int_{-\infty}^{+\infty}\log[f(x;\alpha,\beta,\xi)]f(x;\alpha,\beta,\xi)dx$. It is in fact a particular case of the Rényi entropy, obtained when δ tends to 1^+.

3.8. *Order statistics*

Let $X_1,\ldots,X_n$ be independent and identically distributed random variables with common pdf given by (2). Then, by applying the well-know theory on order statistics, the pdf of the ith order statistic is specified as

$$f_{i:n}(x;\alpha,\beta,\xi) = \frac{n!}{(i-1)!(n-i)!}f(x;\alpha,\beta,\xi)F^{i-1}(x;\alpha,\beta,\xi)[1-F(x;\alpha,\beta,\xi)]^{n-i}.$$

Using the series expansion in (7), we have

$$f_{i:n}(x;\alpha,\beta,\xi) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^j f(x;\alpha,\beta,\xi)F^{j+i-1}(x;\alpha,\beta,\xi). \tag{15}$$

By the infinite mixture representation in (8) and (10), an alternative expression is

$$f_{i:n}(x;\alpha,\beta,\xi) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^j\sum_{m=0}^{+\infty}a_m h_{m+1}(x;\xi)\left[\sum_{k=0}^{+\infty}a_k H_{k+1}(x;\xi)\right]^{j+i-1}. \tag{16}$$

Using a general result of power series raised to a positive power (see Gradshteyn and Ryzhik, 2000), we get

$$\left(\sum_{k=0}^{+\infty} c_k x^k\right)^n = \sum_{k=0}^{+\infty} d_k x^k,$$

where $d_0 = c_0^n$ and $d_m = [1/(mc_0)]\sum_{k=1}^{m}(k(n+1)-m)c_k d_{m-k}$ for $m \geq 1$. Therefore,

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

with $d_0 = a_0^{j+i-1}$ and $d_m = [1/(ma_0)]\sum_{k=1}^{m}(k(j+i)-m)a_k d_{m-k}$ for $m \geq 1$. Now, (16) becomes

$$f_{i:n}(x;\alpha,\beta,\xi) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^j \sum_{m=0}^{+\infty}\sum_{k=0}^{+\infty} d_k a_m (m+1) g(x;\xi) G^{m+j+i+k-1}(x;\xi).$$

From this expression, several mathematical properties can be obtained, as moments, moment generating function, various entropies....

In the context of the GGFr distribution: One can observe that

$$\begin{aligned} g(x;\xi)G^{m+j+i+k-1}(x;\xi) &= \frac{a^b b}{x^{b+1}}\exp\left[-(m+j+i+k)\left(\frac{a}{x}\right)^b\right] \\ &= \frac{1}{m+j+i+k} u_{m+j+i+k}(x), \end{aligned}$$

where $u_{m+j+i+k}(x)$ denotes the pdf of the Fréchet distribution with parameters $(m+j+i+k)^{1/b}a$ and b. So, the pdf of the ith order statistic of the GGFr distribution can be expressed as a linear combination of Fréchet pdfs.

4. Estimation of the GG-G family parameters

Let $x_1, \ldots, x_2$ be independent observations of a random variable with the pdf given by (2) and vector of parameters $\Theta = [\alpha, \beta, \xi]^T$. The log-likelihood function for Θ is expressed as

$$\begin{aligned} \ell(\Theta) = n\log\alpha - n\log\Gamma(\beta) + \sum_{i=1}^{n}\log g(x_i,\xi) + (\alpha\beta-1)\sum_{i=1}^{n}\log G(x_i,\xi) \\ -(\alpha\beta+1)\sum_{i=1}^{n}\log(1-G(x_i;\xi)) - \sum_{i=1}^{n}\left(\frac{G(x_i,\xi)}{1-G(x_i;\xi)}\right)^{\alpha}. \end{aligned}$$

The components of score vector $\Theta = [\alpha, \beta, \xi]^T$ are given by

$$U_\alpha = \frac{n}{\alpha} + \beta \sum_{i=1}^{n} \log G(x_i, \xi) - \beta \sum_{i=1}^{n} \log(1 - G(x_i; \xi))$$
$$- \sum_{i=1}^{n} \left(\frac{G(x_i, \xi)}{1 - G(x_i; \xi)} \right)^{\alpha} \log \left(\frac{G(x_i, \xi)}{1 - G(x_i; \xi)} \right),$$
$$U_\beta = -n \frac{d}{d\beta} \log \Gamma(\beta) + \alpha \sum_{i=1}^{n} \log G(x_i, \xi) - \alpha \sum_{i=1}^{n} \log(1 - G(x_i; \xi)),$$
$$U_\xi = \sum_{i=1}^{n} \frac{g^{\xi}(x_i, \xi)}{g(x_i, \xi)} + (\alpha\beta - 1) \sum_{i=1}^{n} \frac{G^{\xi}(x_i, \xi)}{G(x_i, \xi)} + (\alpha\beta + 1) \sum_{i=1}^{n} \frac{G^{\xi}(x_i, \xi)}{1 - G(x_i, \xi)}$$
$$- \alpha \sum_{i=1}^{n} \left[\frac{G(x_i, \xi)}{1 - G(x_i, \xi)} \right]^{\alpha - 1} \frac{G^{\xi}(x_i, \xi)}{\{1 - G(x_i, \xi)\}^2},$$

where $G^{\xi}(x_i, \xi) = \partial G(x_i, \xi)/\partial \xi$ and $g^{\xi}(x_i, \xi) = \partial g(x_i, \xi)/\partial \xi$. Since these equations are nonlinear according to the parameters and of complex nature, they can not be solved analytically but can be solved numerically by any software like R-language or Mathematica.

5. Log-generalized gamma Fréchet

In many applied areas, the lifetimes are affected by explanatory variables such as the cholesterol level, blood sugar, gender and many other explanatory variables. Parametric survival models to estimate the survival functions for censored data are widely used. For instance, recently Lanjoni *et al.* (2016) defined the extended Burr XII regression model and Prataviera *et al.* (2018) proposed the heteroscedastic odds log-logistic generalized gamma regression model for censored data. Thus, by using the same approach of these papers, a distribution obtained from the log-generalized gamma Fréchet (LGGFr) distribution will be expressed in the form of the class of location-scale models with two additional parameters to the shape. In this way, we propose a model of regression location, scale and shape.

Let X be a random variable following the Fréchet distribution with parameters 1 and 1. Then, the random variable $Y = \sigma \log(X) + \mu$ has the following cdf and pdf:

$$G(y; \mu, \sigma) = \exp \left\{ -\exp \left\{ -\left(\frac{y - \mu}{\sigma} \right) \right\} \right\}$$

and

$$g(y; \mu, \sigma) = \frac{1}{\sigma} \exp \left\{ -\left(\frac{y - \mu}{\sigma} \right) - \exp \left\{ -\left(\frac{y - \mu}{\sigma} \right) \right\} \right\},$$

respectively, and the pdf specified in (2) becomes

$$f(y;\alpha,\beta,\mu,\sigma)=$$
$$\frac{\alpha}{\Gamma(\beta)}\frac{g(y;\mu,\sigma)G^{\alpha\beta-1}(y;\mu,\sigma)}{(1-G(y;\mu,\sigma))^{\alpha\beta+1}}\exp\left\{-\left[\frac{G(y;\mu,\sigma)}{1-G(y;\mu,\sigma)}\right]^{\alpha}\right\}, \tag{17}$$

where $y,\mu\in\mathbb{R}$ and $\sigma>0$, μ is the location parameter, σ is the scale parameter and α and β are shape parameters. Thus, if $X\sim \mathrm{GGFr}(1,1,\alpha,\beta)$ we set $Y=\sigma\log(X)+\mu\sim \mathrm{LGGFr}(\mu,\sigma,\alpha,\beta)$. Also, if $Y\sim\mathrm{LGGFr}(\mu,\sigma,\alpha,\beta)$, by putting

$$G(z)=\exp\{-\exp\{-z\}\},\quad g(z)=\exp\{-z-\exp\{-z\}\}.$$

the pdf of the random variable $Z=(Y-\mu)/\sigma$ becomes

$$f(z;\alpha,\beta)=\frac{\alpha}{\Gamma(\beta)}\frac{g(z)G^{\alpha\beta-1}(z)}{(1-G(z))^{\alpha\beta+1}}\exp\left\{-\left[\frac{G(z)}{1-G(z)}\right]^{\alpha}\right\}. \tag{18}$$

In this case, we write $Z\sim\mathrm{LGGFr}(0,1,\alpha,\beta)$.

In order to introduce a regression structure in the class of models (18), we assume that the parameters μ_i, σ_i, α_i and β_i vary across observations through the following structure:

$$y_i=\mu_i+\sigma_i z_i,\qquad i=1,\ldots,n, \tag{19}$$

where the random error z_i has the pdf given by (18), μ_i and σ_i are parameterized as

$$\mu_i=\mu_i(\theta_1),\qquad \sigma_i=\sigma_i(\theta_2),\qquad \alpha_i=\alpha_i(\theta_3),\qquad \beta_i=\beta_i(\theta_4),$$

where $\theta_1=(\theta_{11},\ldots,\theta_{1p_1})^T$, $\theta_2=(\theta_{21},\ldots,\theta_{2p_2})^T$, $\theta_3=(\theta_{31},\ldots,\theta_{3p_3})^T$ and $\theta_4=(\theta_{41},\ldots,\theta_{4p_4})^T$. The usual systematic component for the location parameter is $\mu_i=\mathbf{x}_i^T\theta_1$, where $\mathbf{x}_i^T=(x_{i1},\ldots,x_{ip_1})$ is a vector of known explanatory variables, i.e., $\boldsymbol{\mu}=\mathbf{X}\theta_1$, and $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)^T$, $\mathbf{X}=(\mathbf{x}_1,\ldots,\mathbf{x}_n)^T$ is a specified $n\times p_1$ matrix of full rank with $p_1<n$. Analogously, we consider the systematic component $g(\sigma_i)=\eta_i=\mathbf{v}_i^T\theta_2$ for the dispersion parameter, where $g(\cdot)$ is the dispersion link function, and $\mathbf{v}_i^T=(v_{i1},\ldots,v_{ip_2})$ is a vector of known explanatory variables. We have $g(\boldsymbol{\sigma})=\boldsymbol{\eta}=\mathbf{V}\theta_2$, where $\boldsymbol{\sigma}=(\sigma_1,\ldots,\sigma_n)^T$, $\boldsymbol{\eta}=(\eta_1,\ldots,\eta_n)^T$ and $\mathbf{V}=(\mathbf{v}_1,\ldots,\mathbf{v}_n)^T$ is a specified $n\times p_2$ matrix of full rank with $p_2<n$. For α_i and β_i, we consider the systematic component analogous. We have $g(\boldsymbol{\alpha})=\boldsymbol{\delta}=\mathbf{W}\theta_3$, where $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_n)^T$, $\boldsymbol{\delta}=(\delta_1,\ldots,\delta_n)^T$ and $\mathbf{W}=(\mathbf{w}_1,\ldots,\mathbf{w}_n)^T$ is a specified $n\times p_3$ matrix of full rank with $p_3<n$ and $g(\beta)=\boldsymbol{\lambda}=\mathbf{S}\theta_2$, where $\beta=(\beta_1,\ldots,\beta_n)^T$, $\boldsymbol{\lambda}=(\lambda_1,\ldots,\lambda_n)^T$ and $\mathbf{S}=(\mathbf{s}_1,\ldots,\mathbf{s}_n)^T$ is a specified $n\times p_4$ matrix of full rank with $p_4<n$. It is assumed that θ_1, θ_2, θ_3 and θ_4 are functionally independent and that $g(\cdot)$ is a known one-to-one continuously twice differentiable function.

5.1. Maximum likelihood estimation

Consider a sample $(y_1, \mathbf{x}_1, \mathbf{v}_1, \mathbf{w}_1, \mathbf{s}_1), \cdots, (y_n, \mathbf{x}_n, \mathbf{v}_n, \mathbf{w}_n, \mathbf{s}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(x_i), \log(c_i)\}$. Here, the x_i's are the failure times and the c_i's are the censored times. We assume non-informative censoring such that the observed lifetimes and censoring times are independent. Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. Then, we can apply conventional likelihood estimation techniques to estimate the model parameters. The log-likelihood function for the vector of parameters $\boldsymbol{\phi} = (\beta_1^T, \beta_2^T, \beta_3^T, \beta_4^T)^T$ from model (19) has the form $l(\boldsymbol{\phi}) = \sum_{i \in F} l_i(\boldsymbol{\phi}) + \sum_{i \in C} l_i^{(c)}(\boldsymbol{\phi})$, where $l_i(\boldsymbol{\phi}) = \log[f(y_i)]$, $l_i^{(c)}(\boldsymbol{\phi}) = \log[S(y_i)]$, $f(y_i)$ is the pdf specified by (17) and $S(y_i)$ is the corresponding survival function, respectively. The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\phi}}$ of the vector of model parameters can be computed by maximizing the log-likelihood $l(\boldsymbol{\phi})$.

The asymptotic distribution of $(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi})$ is, under standard regularity conditions, multivariate normal $N_{p_1+p_2+p_3+p_4}(0, K(\boldsymbol{\phi})^{-1})$, where $K(\boldsymbol{\phi})$ is the total expected information matrix. The asymptotic covariance matrix $K(\boldsymbol{\phi})^{-1}$ of $\widehat{\boldsymbol{\phi}}$ can be approximated by the inverse of the $(p_1 + p_2 + p_3 + p_4) \times (p_1 + p_2 + p_3 + p_4)$ observed information matrix $-\ddot{\mathbf{L}}(\boldsymbol{\phi})$. The elements of $-\ddot{\mathbf{L}}(\boldsymbol{\phi})$ can be evaluated numerically. The approximate multivariate normal distribution $N_{p_1+p_2+p_3+p_4}(0, -\ddot{\mathbf{L}}(\boldsymbol{\phi})^{-1})$ for $\widehat{\boldsymbol{\phi}}$ can be used in the classical way to construct approximate confidence intervals for the components of $\boldsymbol{\phi}$.

5.2. Simulation study

We now perform a Monte Carlo simulation study to assess the finite sample behavior of the MLEs. The results are obtained from 1000 Monte Carlo simulations performed using the R software. In each replication, a random sample of size n is drawn from the LGGFr$(\theta_1, \theta_2, \theta_3, \theta_4)$ model and the parameters are estimated by maximum likelihood. The log-lifetimes denoted by $\log(x_1), \cdots, \log(x_n)$ are generated from the LGGFr regression model (19), where $\mu_i = \theta_{10} + \theta_{11} x_i$, $\sigma_i = \exp(\theta_{20} + \theta_{21} x_i)$, $\alpha_i = \exp(\theta_{30} + \theta_{31} x_i)$, $\beta_i = \exp(\theta_{40} + \theta_{41} x_i)$ and x_i is generated from a normal distribution $N(0, 0.65)$. Thus, we consider for the simulations with sample sizes $n = 100$, $n = 350$ and $n = 850$, and censoring percentages approximately equal to 0%, 10% and 20%. The values considered for the parameters are $\theta_{10} = 5.6350$, $\theta_{11} = 0.4948$, $\theta_{20} = 0.1688$, $\theta_{21} = 0.5219$, $\theta_{30} = -0.5805$, $\theta_{31} = 0.0313$, $\theta_{40} = -2.6156$ and $\theta_{41} = 3.2250$. The survival times are generated considering the random censoring mechanism as follows.

- Generate $x_i \sim$ Normal(1,0.65).
- Generate $c \sim$ Uniform$(0, \tau)$, where τ denotes the proportion of censored observations.

- Generate $z \sim \text{LGGFr}(0, 1, \alpha_i, \alpha_i)$, the values from the pdf given by (18).
- Write $y^* = \mu_i x_i + \sigma_i z$.
- Set $y = \min(y^*, c)$.
- Create a vector x of dimension n which receives 1's if ($y^* <= c$) and zero otherwise.

TABLE 1

Simulations for the LGGFr regression model for the parameters values $\theta_{10} = 5.6350$, $\theta_{11} = 0.4948$, $\theta_{20} = 0.1688$, $\theta_{21} = 0.5219$, $\theta_{30} = -0.5805$, $\theta_{31} = 0.0313$, $\theta_{40} = -2.6156$ and $\theta_{41} = 3.2250$ for 0% censored.

n	θ	AE	Bias	MSE
	θ_{10}	5.648	0.013	0.119
	θ_{11}	0.484	-0.010	0.064
	θ_{20}	0.121	-0.047	0.015
100	θ_{21}	0.529	0.007	0.007
	θ_{30}	-0.698	-0.117	0.101
	θ_{31}	0.109	0.078	0.142
	θ_{40}	-2.630	-0.014	0.090
	θ_{41}	3.249	0.024	0.111
	θ_{10}	5.658	0.023	0.015
	θ_{11}	0.479	-0.015	0.009
	θ_{20}	0.155	-0.013	0.002
350	θ_{21}	0.528	0.006	0.000
	θ_{30}	-0.631	-0.051	0.019
	θ_{31}	0.072	0.041	0.024
	θ_{40}	-2.610	0.005	0.017
	θ_{41}	3.223	-0.001	0.016
	θ_{10}	5.647	0.012	0.003
	θ_{11}	0.486	-0.008	0.002
	θ_{20}	0.162	-0.006	0.000
850	θ_{21}	0.526	0.004	0.000
	θ_{30}	-0.611	-0.031	0.005
	θ_{31}	0.059	0.027	0.006
	θ_{40}	-2.608	0.007	0.004
	θ_{41}	3.216	-0.008	0.004

TABLE 2

Simulations for the LGGFr regression model for the parameters values $\theta_{10} = 5.6350$, $\theta_{11} = 0.4948$, $\theta_{20} = 0.1688$, $\theta_{21} = 0.5219$, $\theta_{30} = -0.5805$, $\theta_{31} = 0.0313$, $\theta_{40} = -2.6156$ *and* $\theta_{41} = 3.2250$ *for 10% censored.*

n	θ	AE	Bias	MSE
	θ_{10}	5.661	0.026	0.142
	θ_{11}	0.468	-0.026	0.068
	θ_{20}	0.120	-0.048	0.016
100	θ_{21}	0.536	0.014	0.008
	θ_{30}	-0.681	-0.101	0.103
	θ_{31}	0.086	0.054	0.178
	θ_{40}	-2.666	-0.050	0.115
	θ_{41}	3.295	0.070	0.148
	θ_{10}	5.662	0.027	0.017
	θ_{11}	0.477	-0.017	0.010
	θ_{20}	0.155	-0.013	0.002
350	θ_{21}	0.528	0.006	0.000
	θ_{30}	-0.612	-0.031	0.018
	θ_{31}	0.049	0.018	0.023
	θ_{40}	-2.632	-0.016	0.020
	θ_{41}	3.243	0.018	0.019
	θ_{10}	5.652	0.017	0.004
	θ_{11}	0.484	-0.010	0.002
	θ_{20}	0.160	-0.008	0.000
850	θ_{21}	0.526	0.004	0.000
	θ_{30}	-0.601	-0.021	0.005
	θ_{31}	0.047	0.016	0.006
	θ_{40}	-2.625	-0.010	0.005
	θ_{41}	3.231	0.006	0.005

TABLE 3
Simulations for the LGGFr regression model for the parameters values $\theta_{10} = 5.6350$, $\theta_{11} = 0.4948$, $\theta_{20} = 0.1688$, $\theta_{21} = 0.5219$, $\theta_{30} = -0.5805$, $\theta_{31} = 0.0313$, $\theta_{40} = -2.6156$ and $\theta_{41} = 3.2250$ for 20% censored.

n	θ	AE	Bias	MSE
100	θ_{10}	5.704	0.069	0.126
	θ_{11}	0.443	-0.050	0.066
	θ_{20}	0.125	-0.043	0.015
	θ_{21}	0.538	0.016	0.009
	θ_{30}	-0.660	-0.079	0.111
	θ_{31}	0.029	-0.001	0.187
	θ_{40}	-2.700	-0.084	0.155
	θ_{41}	3.355	-0.084	0.181
350	θ_{10}	5.674	0.039	0.022
	θ_{11}	0.470	-0.024	0.012
	θ_{20}	0.155	-0.013	0.002
	θ_{21}	0.530	0.008	0.001
	θ_{30}	-0.607	-0.027	0.020
	θ_{31}	0.036	0.005	0.027
	θ_{40}	-2.647	-0.032	0.025
	θ_{41}	3.264	0.039	0.026
850	θ_{10}	5.657	0.022	0.005
	θ_{11}	0.481	-0.013	0.003
	θ_{20}	0.159	-0.008	0.001
	θ_{21}	0.527	0.005	0.000
	θ_{30}	-0.596	-0.015	0.005
	θ_{31}	0.037	0.006	0.008
	θ_{40}	-2.634	-0.018	0.008
	θ_{41}	3.242	0.017	0.008

We fit the LGGFr regression model (19) to each generated data set, where $\mu_i = \theta_{10}+\theta_{11}x_i$, $\sigma_i = \exp(\theta_{20}+\theta_{21}x_i)$, $\alpha_i = \exp(\theta_{30}+\theta_{31}x_i)$ and $\beta_i = \exp(\theta_{40}+\theta_{41}x_i)$. From the simulations reported in Tables 1, 2 and 3, we can verify that the mean squared errors (MSEs) and bias of the MLEs of θ_{10}, θ_{11}, θ_{20}, θ_{21}, θ_{30}, θ_{31}, θ_{40} and θ_{41} decay toward zero when the sample size increases, as usually expected under first-order asymptotic theory. The mean estimates of the parameters tend to be closer to the true parameter values when the sample size n increases. This fact supports that the asymptotic normal distribution provides an adequate approximation to the finite sample distribution of the estimates.

6. Residual analysis

The objective of the analysis of the residuals is to verify the adequacy of the model for a given data set, which includes the investigation of intrinsic characteristics in the data. In order to check some of these characteristics, for example, outliers, several approaches have been proposed by Cox and Snell (1968); Cook and Weisberg (1982); Ortega *et al.* (2008); Silva *et al.* (2011). In the context of survival analysis, the deviance residuals have been more widely used because they take into account the information of censored times (Silva *et al.*, 2011). Thus, the plot of the deviance residuals versus the observed times provides a way to test the adequacy of the fitted model and to detect atypical observations.

If the model is appropriate, the martingale and modified deviance residuals must present a random behavior around zero. The plots of the residuals, martingale or modified deviance residuals versus the adjusted values provide a simple way to verify the adequacy of the model and to detect outliers. Atkinson (1985) suggested the construction of envelopes to enable better interpretation of the normal probability plot of the residuals. These envelopes are simulated confidence bands that contain the residuals, such that if the model is well-fitted, the majority of points will be within these bands and randomly distributed.

We perform a simulation study to assess the accuracy of the MLEs of the parameters in the LGGFr regression model with censored data, and also to investigate the behavior of the empirical distribution of the martingale and deviance residuals. For the simulation study, we generate independent observations $z_1, \cdots, z_n$ from the LGGFr distribution defined by (18).

So, 1000 samples are generated for each scenario presented in Subsection 5.2 as well as the algorithm for generating the survival times considering censored.

In Figures 3, 4, and 5, we display the plots of the residuals versus the expected values of the order statistics of the standard normal distribution for different sample sizes. These plots are known as the normal probability plots and serve to assess the departure from the normality assumption of the residuals (Weisberg, 2005). Therefore, the following interpretation is obtained from these plots: the empirical distribution of the deviance residuals agrees with the standard normal distribution when the sample size increases.

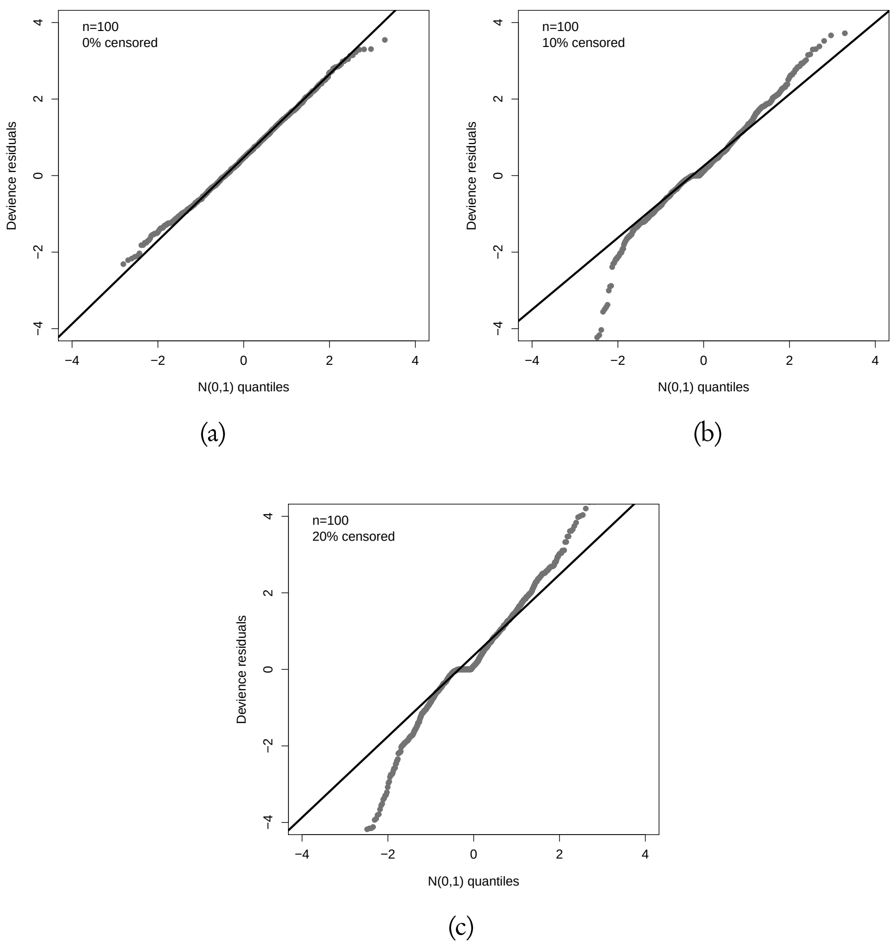
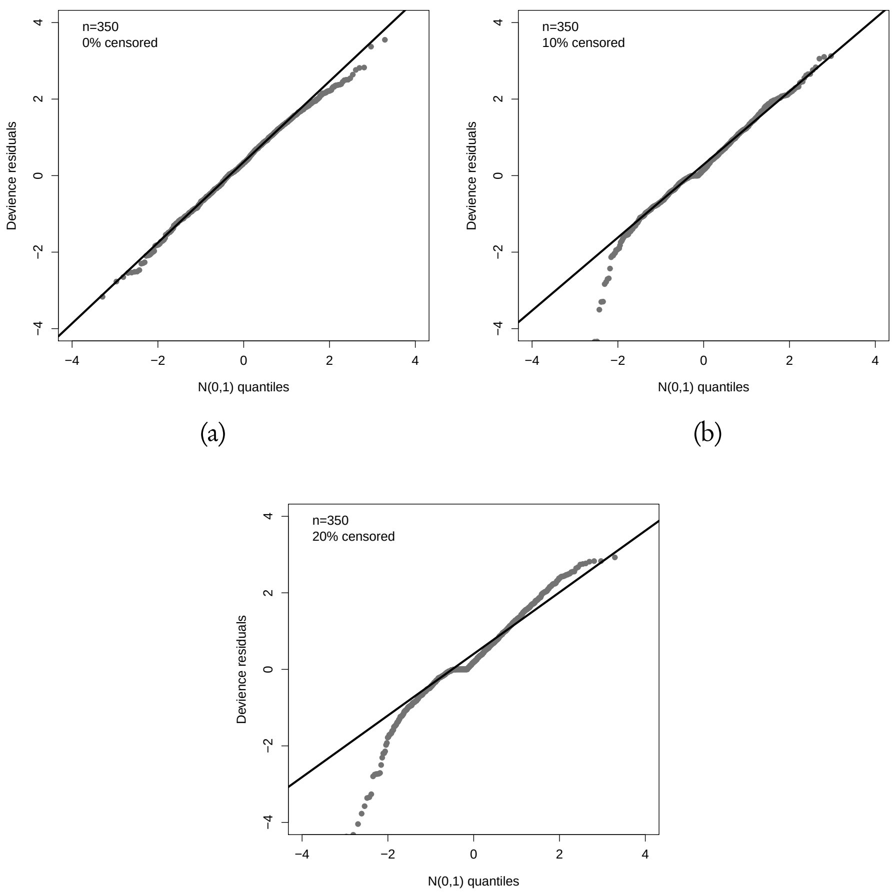


Figure 3 – Normal probability plots for the deviance residuals with sample size $n = 100$ for censored percentages of (a) 0% (b) 10% and (c) 20%.



(c)

Figure 4 – Normal probability plots for the deviance residuals with ample size $n = 350$ for censored percentages of (a) 0%, (b) 10% and (c) 20%).

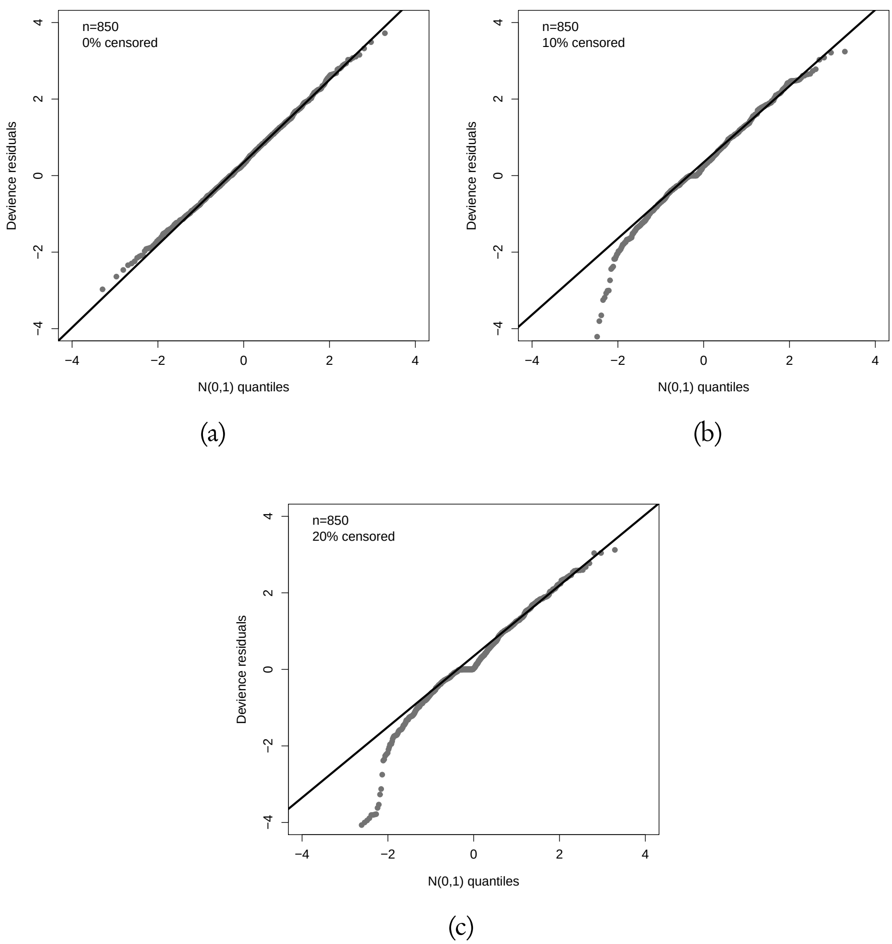


Figure 5 – Normal probability plots for the deviance residuals with sample size $n = 850$ for censored percentages of (0%, 10% and 20%).

7. Data analysis

In this section, several concrete applications of the above methodology are discussed.

7.1. *Analysis of three data sets*

Here, we present three applications to real data to illustrate the potentiality of the GG-G family. To evaluate its performance, we consider its sub model generalized gamma Fréchet (GGFr) and the other competitive models presented in Table 4.

TABLE 4
The competitive models of the GGFr distributions.

Distribution	Author(s)
Generalized odds gamma-Fréchet (GOFr)	(Hosseini *et al.*, 2018)
Log-gamma generated Fréchet (LGFr)	(Amini *et al.*, 2014)
Kumaraswamy Fréchet (KwFr)	(Mead and Abd-Eltawab, 2014)
Beta Fréchet (BFr)	(Nadarajah and Gupta, 2004)
Exponentiated Fréchet (EFr)	(Nadarajah and Kotz, 2003)

We consider the $-\widehat{\ell}$ (where $\widehat{\ell}$ denotes the maximized log-likelihood), AIC (Akaike information criterion), BIC (Bayesian information criterion), CVM (Cramér-Von Mises), AD (Anderson-Darling) and KS (Kolmogorov Smirnov) with its *p*-value (PV) statistics to compare the fitted distributions. The numerical results appearing in this section are obtained using R.

The first data set (Data Set 1) is from Xu *et al.* (2003) and it represents the time to failure (103 h) of turbocharger of one type of engine, recently used by Alzaatreh *et al.* (2015). The time to failure of turbocharger data ($n = 40$) are:
1.6 3.5 4.8 5.4 6.0 6.5 7.0 7.3 7.7 8.0 8.4 2.0 3.9 5.0 5.6 6.1 6.5 7.1 7.3 7.8 8.1 8.4 2.6 4.5 5.1 5.8 6.3 6.7 7.3 7.7 7.9 8.3 8.5 3.0 4.6 5.3 6.0 8.7 8.8 9.0.

The second data set (Data Set 2) was obtained in Proschan (2000) and corresponds to the time of successive failures of the air conditioning system of jet airplanes. These data were also studied by Dahiya and Gurland (1972); Gupta and Kundu (2001); Kus (2007); Andrade *et al.* (2017), among others. The data are:
194, 413, 90, 74, 55, 23, 97, 50, 359,50, 130, 487, 102, 15, 14, 10, 57, 320, 261, 51, 44, 9, 254, 493, 18, 209, 41, 58, 60, 48, 56, 87, 11, 102, 12, 5, 100, 14,29, 37, 186, 29, 104, 7, 4, 72, 270, 283, 7, 57, 33, 100, 61, 502, 220, 120, 141, 22, 603, 35, 98, 54, 181, 65, 49, 12, 239,14, 18, 39, 3, 12, 5, 32, 9, 14, 70, 47, 62, 142, 3, 104, 85, 67, 169, 24, 21, 246, 47, 68, 15, 2, 91, 59, 447, 56, 29, 176, 225,77, 197, 438, 43, 134, 184, 20, 386, 182, 71, 80, 188, 230, 152, 36, 79, 59, 33, 246, 1, 79, 3, 27, 201, 84, 27, 21, 16, 88, 130, 14, 118, 44, 15, 42, 106, 46, 230, 59, 153, 104, 20, 206, 5, 66, 34, 29, 26, 35, 5, 82, 5, 61, 31, 118, 326, 12, 54, 36, 34, 18, 25, 120, 31, 22, 18, 156, 11, 216, 139, 67, 310, 3, 46, 210, 57, 76, 14, 111, 97, 62, 26, 71, 39, 30, 7, 44, 11, 63,23, 22, 23, 14, 18, 13, 34, 62, 11, 191, 14, 16, 18, 130, 90, 163, 208, 1, 24, 70, 16, 101, 52, 208, 95.

The third data set (Data Set 3) represents the annual maximum temperatures at Oxford and Worthing in England for the period of 1901-1980. Chandler and Bate (2007) used the generalized extreme value distribution to model the annual maximum temperatures. Recently study this data by Alzaatreh *et al.* (2015). The annual maximum temperatures data ($n = 80$) are:
75, 92, 87, 86, 85, 95, 84, 87,86 ,82, 77, 89, 79, 83, 79, 85, 89, 84, 84, 82, 86, 81, 84, 84, 87, 89, 80, 86, 85, 84, 89, 80, 87, 84, 85, 82, 86, 87, 86, 89, 90, 90, 91, 81, 85 ,79, 83, 93, 87, 83, 88, 90, 83, 82, 80, 81 ,95 ,89, 86, 89, 87, 92, 89, 87, 87, 83, 89 ,88, 84, 84, 77, 85, 77, 91, 94, 80, 80, 85, 83, 88.

The MLEs and their corresponding standard errors (SEs) (in parentheses) are listed in Tables 5, 6 and 7 for Data Sets 1, 2 and 3, respectively. Tables 8, 9 and 10 provide the values of goodness-of-fit measures for the GGFr model and other fitted models, for Data Sets 1, 2 and 3, respectively.

TABLE 5
MLEs and SEs (in parentheses) for Data Set 1.

Model	Estimate (SE)			
GGFr	39.982	0.055	5.431	0.728
(α,β,a,b)	(35.482)	(0.050)	(1.802)	(0.472)
GOFr	13.077	4.691	0.173	1.174
(α,β,a,b)	(14.145)	(3.416)	(0.850)	(0.363)
LGFr	0.283	135.081	41.751	1.041
(α,β,a,b)	(0.106)	(78.313)	(0.075)	(0.091)
KwFr	5.223	554.611	9.882	0.501
(a,b,λ,α)	(36.851)	(222.077)	(138.846)	(0.092)
BFr	71.150	259.174	74.974	0.168
(a,b,λ,α)	(2.113)	(111.212)	(38.170)	(0.032)
EFr	1.610	1.944	7.929	
(a,b,λ)	(0.167)	(0.203)	(3.314)	

From these tables, we see that the GGFr model has the smallest $-\hat{\ell}$, AIC, BIC, CVM, AD, and KS, and the largest PV, indicating that it gives the best fit for the data in comparison to the other considered models.

Some plots of the fitted pdfs and cdfs are displayed in Figures 6, 7 and 8 for Data Sets 1, 2 and 3, respectively. These figures illustrate the nice fits of the GGFr model.

TABLE 6
MLEs and SEs (in parentheses) for Data Set 2.

Model	Estimate (SE)			
GGFr	1.915	0.687	50.909	0.391
(α,β,a,b)	(0.483)	(0.571)	(46.144)	(0.118)
GOFr	10.211	2.805	0.019	0.430
(α,β,a,b)	(0.483)	(0.571)	(46.144)	(0.118)
LGFr	74.854	165.083	57.952	0.072
(α,β,a,b)	(86.451)	(162.393)	(29.028)	(0.035)
KwFr	7.972	152.567	5.823	0.175
(a,b,λ,α)	(2.571)	(170.432)	(9.582)	(0.036)
BFr	123.682	219.123	69.697	0.056
(a,b,λ,α)	(127.197)	(204.188)	(11.232)	(0.031)
EFr	3.993	0.736	3.968	
(a,b,λ)	(21.510)	(0.034)	(4.465)	

TABLE 7
MLEs and SEs (in parentheses) for Data Set 3.

Model	Estimate (SE)			
GGFr	18.283	14.332	9.771	0.217
(α,β,a,b)	(11.781)	(13.266)	(5.168)	(0.033)
GOFr	40.211	34.297	10.159	3.399
(α,β,a,b)	(16.439)	(20.855)	(11.503)	(0.121)
LGFr	85.658	111.073	69.356	2.303
(α,β,a,b)	(3.121)	(119.737)	(12.213)	(2.261)
KwFr	0.614	35.567	117.481	5.827
(a,b,λ,α)	(6.930)	(16.538)	(28.753)	(2.707)
BFr	123.432	239.655	90.138	1.357
(a,b,λ,α)	(1.234)	(146.649)	(21.323)	(1.666)
EFr	69.090	19.850	38.968	
(a,b,λ)	(38.881)	(1.590)	(8.535)	

TABLE 8
Goodness-of-fit measures for Data Set 1.

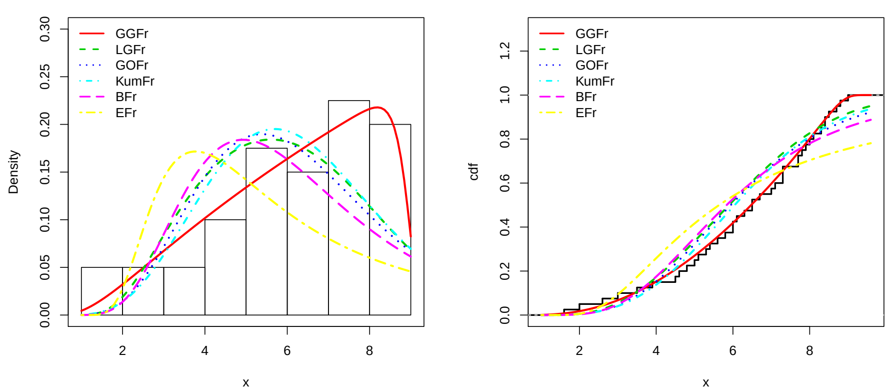
Model	$-\widehat{\ell}$	AIC	BIC	CVM	AD	KS	PV
GGFr	77.686	163.372	170.127	0.014	0.112	0.062	0.997
GOFr	87.878	183.757	190.513	0.220	1.446	0.135	0.456
LGFr	85.802	179.605	186.361	0.165	1.136	0.143	0.383
KwFr	86.150	180.301	187.056	0.171	1.161	0.121	0.592
BFr	91.265	190.531	197.287	0.316	1.981	0.152	0.310
EFr	101.591	209.183	214.250	0.606	3.479	0.243	0.017

TABLE 9
Goodness-of-fit measures for Data Set 2.

Model	$-\widehat{\ell}$	AIC	BIC	CVM	AD	KS	PV
GGFr	1174.049	2352.098	2339.543	0.029	0.226	0.034	0.957
GOFr	1175.010	2358.021	2371.466	0.033	0.271	0.035	0.946
LGFr	1180.145	2368.289	2381.735	0.120	0.858	0.058	0.462
KwFr	1174.790	2357.581	2371.026	0.030	0.246	0.038	0.917
BFr	1179.905	2367.809	2381.254	0.115	0.828	0.057	0.487
EFr	1210.316	2426.632	2436.716	0.711	4.575	0.102	0.022

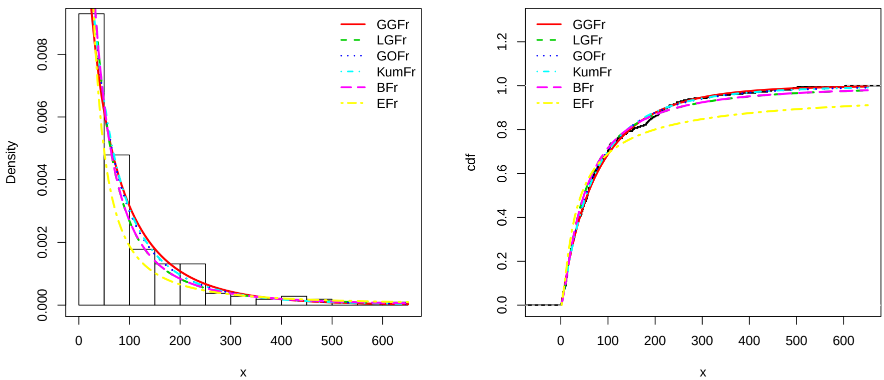
TABLE 10
Goodness-of-fit measures for Data Set 3.

Model	$-\widehat{\ell}$	AIC	BIC	CVM	AD	KS	PV
GGFr	229.103	465.205	474.734	0.048	0.304	0.062	0.899
GOFr	229.055	466.110	475.638	0.050	0.305	0.064	0.897
LGFr	229.477	466.954	476.482	0.066	0.392	0.079	0.685
KwFr	229.037	466.074	475.602	0.049	0.305	0.068	0.819
BFr	229.353	466.707	476.236	0.062	0.369	0.077	0.726
EFr	236.719	479.438	486.584	0.252	1.509	0.134	0.112



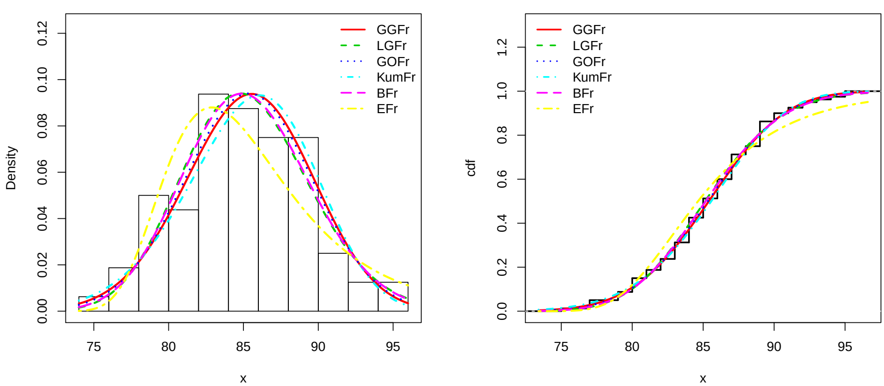
(a) (b)

Figure 6 – Plots of (a) estimated pdfs and (b) estimated cdfs for Data Set 1.



(a) (b)

Figure 7 – Plots of (a) estimated pdfs and (b) estimated cdfs for Data Set 2.



(a) (b)

Figure 8 – Plots of (a) estimated pdfs and (b) estimated cdfs for Data Set 3.

7.2. *Regression model for turbine data*

In this application, we use a real data set available in the book of Lawless (2003, page 262, Table 5.9) to study the LGGFr regression model. This data set presents an experiment designed to compare the performances of high-speed turbine engine bearings made out of five different compounds (McCool, 1979). The model parameters are estimated by the use of the package GAMLSS in R. The experiment tested 10 bearings of each type; the times to fatigue failure are given in units of millions of cycles. The analysis considering the LGGFr regression model is performed with the definition of dummy variables as follows: levels type I ($d_{i1} = 0$, $d_{i2} = 0$, $d_{i3} = 0$ and $d_{i4} = 0$), levels type II ($d_{i1} = 1$, $d_{i2} = 0$, $d_{i3} = 0$ and $d_{i4} = 0$), levels type III ($d_{i1} = 0$, $d_{i2} = 1$, $d_{i3} = 0$ and $d_{i4} = 0$), type IV ($d_{i1} = 0$, $d_{i2} = 0$, $d_{i3} = 1$ and $d_{i4} = 0$) and levels type V ($d_{i1} = 0$, $d_{i2} = 0$, $d_{i3} = 0$ and $d_{i4} = 1$).

The LGGFr regression model for the turbine data can be expressed as

$$y_i = \mu_i + \sigma_i z_i, \quad i = 1, \ldots, 50,$$

where $z_1, \cdots, z_{50}$ are independent random variables with the pdf given by (18).

We consider the three following configurations.

- Configuration I. The first configuration corresponds to the homoscedastic case: we consider the model parameters α, β, μ_i, σ, where $\mu_i = \mu_i = \theta_{10} + \theta_{11} d_{i1} + \theta_{12} d_{i2} + \theta_{13} d_{i3} + \theta_{14} d_{i4}$ and σ is a common variance $\sigma = \sigma_1 = \ldots = \sigma_{50}$.
- Configuration II. The second configuration corresponds to the heteroscedastic case: we consider the parameters α, β, μ_i, σ_i, where $\mu_i = \theta_{10} + \theta_{11} d_{i1} + \theta_{12} d_{i2} + \theta_{13} d_{i3} + \theta_{14} d_{i4}$ and $\sigma_i = \exp(\theta_{20} + \theta_{21} d_{i1} + \theta_{22} d_{i2} + \theta_{23} d_{i3} + \theta_{24} d_{i4})$.
- Configuration III. The third configuration is the more general; we consider the parameters:

$$\begin{aligned}
\mu_i &= \theta_{10} + \theta_{11} d_{i1} + \theta_{12} d_{i2} + \theta_{13} d_{i3} + \theta_{14} d_{i4}, \\
\sigma_i &= \exp(\theta_{20} + \theta_{21} d_{i1} + \theta_{22} d_{i2} + \theta_{23} d_{i3} + \theta_{24} d_{i4}), \\
\alpha_i &= \exp(\theta_{30} + \theta_{31} d_{i1} + \theta_{32} d_{i2} + \theta_{33} d_{i3} + \theta_{34} d_{i4}), \\
\beta_i &= \exp(\theta_{40} + \theta_{41} d_{i1} + \theta_{42} d_{i2} + \theta_{43} d_{i3} + \theta_{44} d_{i4}).
\end{aligned}$$

The MLEs for the LGGFr model parameters are presented in Tables 11 and 12. Thus, when establishing a significance level of 5%, we note that the compounds type level is significant and should be used to model the location, scale and shape.

In order to see if the considered regression model is appropriate, the plot comparing the empirical survival function and estimated survival function for the LGGFr regression model is displayed in Figure 9 under the three presented configurations. We observe that the LGGFr regression model shows a suitable fit. Figure 10 (a) presents the fitted hazard rate functions, (b) the index plot of the deviance residual for the turbine data and (c) the normal probability plot for the deviance component residual with envelopes

TABLE 11
MLEs, SEs and p-values for the LGGFr regression model fitted to the turbine data for Configuration I and Configuration II.

Homoscedastic LGGFr regression				Heteroscedastic LGGFr regression			
Par.	Estimate	SE	p-value	Par.	Estimate	SE	p-value
$\log(\sigma)$	-0.443	0.078	-	$\log(\alpha)$	-0.297	0.119	-
$\log(\alpha)$	0.250	0.103	-	$\log(\beta)$	-0.031	0.114	-
$\log(\beta)$	0.002	0.091	-	θ_{10}	2.310	0.166	0.000
θ_{10}	2.016	0.106	0.000	θ_{11}	-0.466	0.206	0.028
θ_{11}	-0.485	0.175	0.008	θ_{12}	-0.173	0.210	0.414
θ_{12}	-0.162	0.191	0.400	θ_{13}	0.042	0.208	0.840
θ_{13}	0.012	0.153	0.934	θ_{14}	0.366	0.199	0.073
θ_{14}	0.380	0.216	0.086	θ_{20}	-0.384	0.198	0.059
				θ_{21}	-0.309	0.312	0.327
				θ_{22}	-0.270	0.291	0.357
				θ_{23}	-0.214	0.265	0.424
				θ_{24}	-0.396	0.274	0.155

TABLE 12
MLEs, SEs and p-values for the LGGFr regression model fitted to the turbine data for Configuration III.

Location, scale and shape LGGFr regression							
Par.	Estimate	SE	p-value	Par.	Estimate	SE	p-value
θ_{10}	2.298	0.088	0.000	θ_{30}	-3.005	0.319	0.000
θ_{11}	-0.236	0.122	0.060	θ_{31}	2.105	0.423	0.000
θ_{12}	-0.044	0.126	0.727	θ_{32}	0.962	0.465	0.044
θ_{13}	0.352	0.114	0.003	θ_{33}	1.641	0.456	0.000
θ_{14}	0.416	0.122	0.001	θ_{34}	1.420	0.456	0.003
θ_{20}	0.408	0.187	0.034	θ_{40}	3.505	0.289	0.000
θ_{21}	-1.684	0.216	0.000	θ_{41}	-4.453	0.408	0.000
θ_{22}	-0.454	0.331	0.176	θ_{42}	-1.323	0.4216	0.002
θ_{23}	-1.999	0.211	0.000	θ_{43}	-5.283	0.436	0.000
θ_{24}	-0.459	0.334	0.176	θ_{44}	-1.548	0.416	0.000

from the fitted of LGGFr regression model to the turbine data. All these figures shows that the considered LGGFr regression model is appropriate.

From these tables, when establishing a significance level of 5%, we note that the compounds type level is significant and should be used to model the location, scale and shape. In order to see if the considered regression model is appropriate, the plot comparing the empirical survival function and estimated survival function for the LGGFr regression model is displayed in Figure 9 under the three presented configurations.

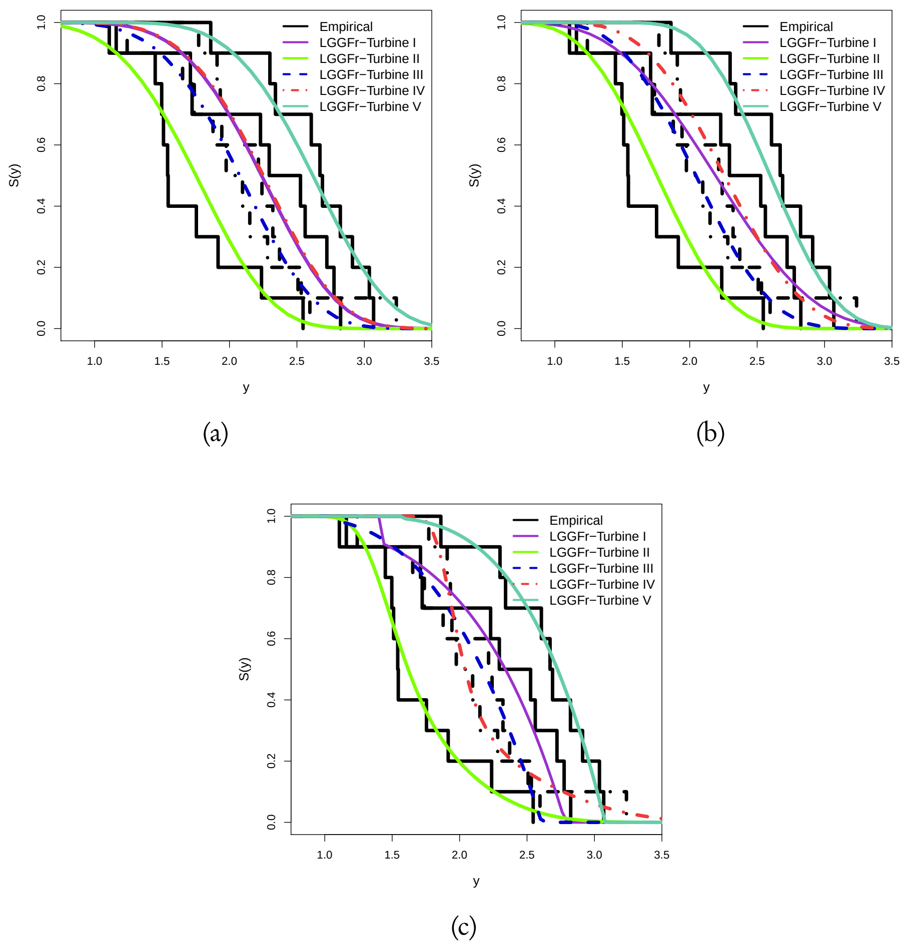
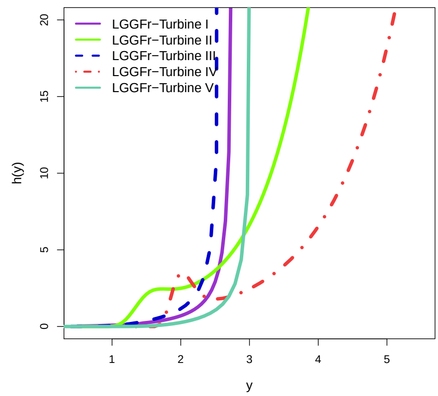
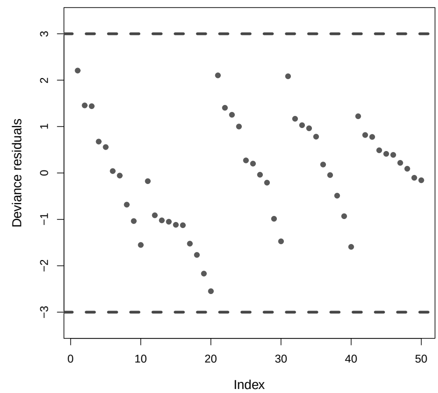


Figure 9 – Estimated survival considering the LGGFr regression model for the (a) Configuration I, (b) Configuration II and (c) Configuration III.

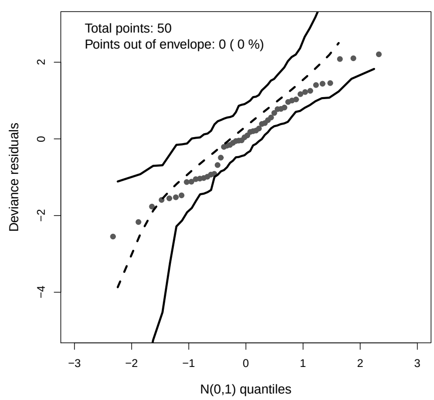
We observe that the LGGFr regression model shows a suitable fit. Figure 10 (a) presents the fitted hazard rate functions, (b) the index plot of the deviance residual for the turbine data and (c) the normal probability plot for the deviance component residual with envelopes from the fitted of LGGFr regression model to the turbine data.



(a)



(b)



(c)

Figure 10 – (a) Fitted hazard rate functions, (b) index plot and (c) normal probability plot for the LGGFr regression model.

All these figures shows that the considered LGGFr regression model is appropriate.

ACKNOWLEDGEMENTS

We are thankful to the two reviewers for constructive comments on the paper.

REFERENCES

C. ALEXANDER, G. CORDEIRO, E. ORTEGA, J. SARABIA (2012). *Generalized beta generated distributions*. Computational Statistics and Data Analysis, 56, pp. 1880–1897.

A. ALZAATREH, C. LEE, F. FAMOYE (2013). *A new method for generating families of continuous distributions*. Metron, 71, pp. 63–79.

A. ALZAATREH, C. LEE, F. FAMOYE (2015). *Family of generalized gamma distributions: Properties and applications*. Hacettepe Journal of Mathematics and Statistic, 45, pp. 869–886.

A. ALZAGHAL, C. LEE, F. FAMOYE (2013). *Exponentiated T-X family of distributions with some applications*. International Journal of Statistics and Probability, 2, pp. 31–49.

M. AMINI, S. MIRMOSTAFAEE, J. AHMADI (2014). *Log-gamma-generated families of distributions*. Statistics, 48, no. 4, pp. 913–932.

T. D. ANDRADE, L. ZEA, S. GOMES-SILVA, G. CORDEIRO (2017). *The gamma generalized Pareto distribution with applications in survival analysis*. International Journal of Statistics and Probability, 6, pp. 141–156.

A. ATKINSON (1985). *Plots, Transformations, and Regression*. Oxford University Press, Oxford.

M. BOURGUIGNON, R. SILVA, G. CORDEIRO (2014). *The Weibull-G family of probability distributions*. Journal of Data Science, 12, pp. 1253–1268.

C. BRITO, F. GOMES-SILVA, L. REGO, W. OLIVEIRA (2017). *A new class of gamma distribution*. Acta Scientiarum: Technology, 39, pp. 79–89.

R. CHANDLER, S. BATE (2007). *Inference for clustered data using the independence log-likelihood*. Biometrika, 94, no. 1, pp. 167–183.

C. CHESNEAU, H. BAKOUCH, T. HUSSAIN (2018). *A new class of probability distributions via cosine and sine functions with applications*. Communications in Statistics - Simulation and Computation, 48, no. 8, pp. 2287–2300.

R. COOK, S. WEISBERG (1982). *Residuals and Influence in Regression*. Chapman and Hall, New York.

G. Cordeiro, M. Alizadeh, E. Ortega (2014). *The exponentiated half-logistic family of distributions: Properties and applications*. Journal of Probability and Statistics, Article ID 864396, 21 pages.

G. Cordeiro, M. de Castro (2011). *A new family of generalized distributions*. Journal of Statistical Computation and Simulation, 81, pp. 883–893.

G. Cordeiro, E. Ortega, G. Silva (2012). *The beta extended Weibull family*. Journal of Probability and Statistical Science, 10, pp. 15–40.

D. Cox, E. Snell (1968). *A general definition of residuals*. Journal of the Royal Statistical Society, Series B, 30, pp. 248–265.

R. Dahiya, J. Gurland (1972). *Goodness of fit tests for the gamma and exponential distributions*. Technometrics, 14, pp. 791–801.

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

I. Gradshteyn, I. Ryzhik (2000). *Table of Integrals, Series and Products*. Academic Press, New York.

R. Gupta, D. Kundu (2001). *Exponentiated exponential family: An alternative to gamma and Weibull distributions*. Biometrical Journal, 43, pp. 117–130.

B. Hosseini, M. Afshari, M. Alizadeh (2018). *The generalized odd gamma-G family of distributions: Properties and applications*. Austrian Journal of Statistics, 47, pp. 69–89.

F. Jamal, M. Nasir, M. Tahir, N. Montazeri (2017). *The odd Burr-III family of distributions*. Journal of Statistics Applications and Probability, 6, pp. 105–122.

M. Jones (2004). *Families of distributions arising from distributions of order statistics*. Test, 13, pp. 1–43.

D. Kumar, U. Singh, S. Singh (2015). *A new distribution using sine function - Its application to bladder cancer patients data*. Journal of Statistics Applications and Probability, 4, no. 3, pp. 417–427.

C. Kus (2007). *A new lifetime distribution*. Computational Statistics and Data Analysis, 51, pp. 4497–4509.

B. Lanjoni, E. Ortega, G. Cordeiro (2016). *Extended Burr XII regression models: Theory and applications*. Journal of Agricultural Biological and Environmental Statistics, 21, pp. 203–224.

J. Lawless (2003). *Statistical Models and Methods for Lifetime Data*. Wiley, New York.

C. McCool (1979). *Distribution of cysticercus bovis in lightly infected young cattle*. Australian Veterinary Journal, 55, pp. 214–216.

M. Mead, A. Abd-Eltawab (2014). *A note on Kumaraswamy Fréchet distribution*. Australian Journal of Basic and Applied Sciences, 8, pp. 294–300.

S. Nadarajah, A. Gupta (2004). *The beta Fréchet distribution*. Far East Journal of Theoretical Statistics, 14, pp. 15–24.

S. Nadarajah, S. Kotz (2003). *The exponentiated Fréchet distribution*. Statistics on the Internet. Http://interstat.statjournals.net/YEAR/2003/articles/0312002.pdf.

E. Ortega, G. Paula, H. Bolfarine (2008). *Deviance residuals in generalized log-gamma regression models with censored observations*. Journal of Statistical Computation and Simulation, 78, pp. 747–764.

R. Pescim, G. Cordeiro, C. Demetrio, E. Ortega, S. Nadarajah (2012). *The new class of Kummer beta generalized distributions*. Statistics and Operations Research Transactions (SORT), 36, pp. 153–180.

F. Prataviera, E. Ortega, G. Cordeiro, R. Pescim, B. Verssani (2018). *A new generalized odd log-logistic flexible Weibull regression model with applications in repairable systems*. Reliability Engineering and System Safety, 176, pp. 13–26.

F. Proschan (2000). *Theoretical explanation of observed decreasing failure rate*. American Statistical Society, reprint in Technometrics, 42, no. 1, pp. 7–11.

M. Ristíc, N. Balakrishnan (2012). *The gamma-exponentiated exponential distribution*. Journal of Statistical Computation and Simulation, 82, pp. 1191–1206.

W. Shaw, I. Buckley (2007). *The alchemy of probability distributions: Beyond Gram-Charlier expansions and a skewkurtotic-normal distribution from a rank transmutation map*. Research report, King's College, London, U.K.

G. Silva, E. Ortega, G. Paula (2011). *Residuals for log-Burr XII regression models in survival analysis*. Journal of Applied and Statistics, 38, pp. 1435–1445.

H. Torabi, N. Montazari (2012). *The gamma-uniform distribution and its application*. Kybernetika, 48, pp. 16–30.

H. Torabi, N. Montazari (2014). *The logistic-uniform distribution and its applications*. Communications in Statistics - Simulation and Computation, 43, pp. 2551–2569.

S. Weisberg (2005). *Applied Linear Regression*. Wiley, Hoboken, 3rd ed.

K. Xu, M. Xie, L. Tang, S. Ho (2003). *Application of neural networks in forecasting engine systems reliability*. Applied Soft Computing, 2, no. 4, pp. 255–268.

K. ZOGRAFOS, N. BALAKRISHNAN (2009). *On families of beta- and generalized gamma-generated distributions and associated inference*. Statistical Methodology, 6, pp. 344–362.

SUMMARY

In this article, a new "odds generalized gamma-G" family of distributions, called the GG-G family of distributions, is introduced. We propose a complete mathematical and statistical study of this family, with a special focus on the Fréchet distribution as baseline distribution. In particular, we provide infinite mixture representations of its probability density function and its cumulative distribution function, the expressions for the Rényi entropy, the reliability parameter and the probability density function of ith order statistic. Then, the statistical properties of the family are explored. Model parameters are estimated by the maximum likelihood method. A regression model is also investigated. A simulation study is performed to check the validity of the obtained estimators. Applications on real data sets are also included, with favorable comparisons to existing distributions in terms of goodness-of-fit.

Keywords: Gamma distribution; Moments; Order statistics; Rényi entropy; Maximum likelihood method; Regression model; Data analysis.