Chilean Journal of Statistics
Volume 14, Number 1 June 2023, Pages 1–25
DOI: 10.32372/chjs.14-01-01

Distributions Theory
Research Paper

An extended Rayleigh model: Properties, regression and COVID-19 application

Gauss M. Cordeiro[1], Gabriela M. Rodrigues[2], Edwin M. M. Ortega[2], Luís H. de Santana[3] and Roberto Vila[4,*]

[1]Department of Statistics, Federal University of Pernambuco, Pernambuco, Brazil
[2]Department of Exact Sciences, University of São Paulo, São Paulo, Brazil
[3]Department of Technology, State University of Maringá, Umuarama, Brazil
[4]Department of Statistics, University of Brasília, Brasília, Brazil

(Received: 17 August 2022 · Accepted in final form: 19 May 2023)

Abstract

We define a four-parameter extended Rayleigh distribution which is a quite flexible model to analyze positive data, and includes as special models the odd log-logistic generalized Rayleigh, exponentiated generalized Rayleigh and generalized Rayleigh distributions. We obtain a linear representation and some of its mathematical properties. We construct a regression from the new distribution with two systematic components for censored data. The estimation is done by maximum likelihood. The utility of the new models is proved in two real applications.

Keywords: Censored data · Linear representation · Simulation · Voltage data · Odd log-logistic.

Mathematics Subject Classification: Primary 62F40 · Secondary 62P30.

1. Introduction

The generalized Rayleigh (GR) distribution is a very important model that has been widely used in recent decades to analyze data in survival analysis, reliability, industrial engineering to represent manufacturing and delivery times, extreme value theory, forecasting wind speed, wireless communications, and insurance to predict the size of reinsurance claims.

Some surveys based on the GR distribution are discussed by several authors: Cordeiro et al. (2011) introduced the beta GR distribution which has some known distributions as sub-models, Gomes et al. (2014) presented the Kumaraswamy GR distribution for analysing lifetime data, Naqash et al. (2016) studied a Bayesian estimation of the shape parameter of the two-parameter GR distribution using single and double priors, and more recently, Barranco-Chamorro et al. (2021) presented a GR family based on the modified Slash distribution, and Shen et al. (2022) introduced an extended form of the GR distribution to model the Reddit advertising and breast cancer data sets. Note that all these cited papers fail to model bimodal data.

*Corresponding author. Email: rovig161@gmail.com.

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

The need for extended forms of the GR distribution arises in many applied areas, for example, to try to model bimodal and asymmetric data without using mixing of distributions. Cordeiro et al. (2017) introduced the generalized odd log-logistic-G (GOLL-G) family which presents different ways to model bimodal, asymmetric and unimodal data. Based on this paper, the first objective is to propose a new generalized odd log-logistic generalized Rayleigh (GOLLGR) distribution that has as special cases the little known odd log-logistic generalized Rayleigh (OLLGR) and exponentiated generalized Rayleigh (EGR) distributions. Thus, we will have several distributions to model data, and are able to discriminate between these distributions. Further, we present several mathematical properties of the new distribution which can be used by statistics users in future research.

We also know that in many practical applications, the lifetimes are affected by explanatory variables, and thus parametric regression models are widely used to estimate univariate survival functions for censored data.

Thus, the second objective of this article is to propose a regression model based on the GOLLGR distribution under two systematic components and the presence of censored data. Several simulations for different scenarios are presented to study the behavior of the parameter estimators.

We consider two data sets for applications. The first set concerns failure times and operating times for a sample of devices from a field tracking study system. The second application considers patients infected with the COVID-19 virus confirmed by the RT-PCR test method in the city of Campinas in the state of São Paulo (Brazil).

The paper is structured as follows. Section 2 defines the GOLLGR distribution, addresses some asymptotes and quantile function (QF), gives a linear representation for the family density, reports the moments and generating function, and addresses maximum likelihood estimation. A new regression is constructed in Section 3 as well as the residual analysis. Some simulations are carried out in Section 4. Two lifetime data sets in Section 5 show the utility of the new models. Some conclusions are reported in Section 6.

2. The GOLLGR model, properties and estimation

2.1 Context

The GR distribution has been employed in many areas (Johnson et al., 1994). Several of its extensions have been published so far.

The cumulative distribution function (CDF) of the GR distribution (Vodă, 1976) is

$$G_{\mathrm{GR}}(x;\delta,\theta) = \gamma_1(\delta+1,\theta x^2), \quad x>0, \tag{2.1}$$

where $\delta > -1$ and $\theta > 0$, $\Gamma(p) = \int_0^\infty w^{p-1}\,\mathrm{e}^{-w}\mathrm{d}w$ and $\gamma_1(p,x) = \int_0^x w^{p-1}\,\mathrm{e}^{-w}\mathrm{d}w/\Gamma(p)$ are the gamma function and lower incomplete gamma function ratio, respectively.

We can write

$$\frac{\partial}{\partial\theta}\gamma_1(\delta+1,\theta x^2) = \frac{\theta^{\delta}x^{2(\delta+1)}\mathrm{e}^{-\theta x^2}}{\Gamma(\delta+1)} \tag{2.2}$$

and

$$\frac{\partial}{\partial\delta}\gamma_1(\delta+1,\theta x^2) = H_1(\delta+1,\theta x^2) - \gamma_1(\delta+1,\theta x^2)\,\psi(\delta+1)\,, \tag{2.3}$$

where

$$H_1(\delta+1,\theta x^2) = \frac{1}{\Gamma(\delta+1)}\int_0^{\theta x^2} t^{\delta}\,\mathrm{e}^{-t}\log(t)\;dt$$

and

$$\psi(\alpha) = \frac{1}{\Gamma(\alpha)}\int_0^{\infty} t^{\alpha-1}\mathrm{e}^{-t}\log(t)\,dt$$

is the digamma function.

The GR distribution includes well-known sub-models. The classical Rayleigh distribution follows when $\delta = 0$ and $\theta = 1/\lambda^2$. If $\delta = 1/2$ and $\theta = 1/(2\lambda^2)$, it gives the Maxwell distribution. The chi-square refers to $\theta = 1/(2\tau^2)$, $\tau > 0$, and $\delta = (n/2)-1$, $n \in \mathrm{N}$, and the half-normal to $\delta = -1/2$ and $\theta = 2/\sigma^2$.

The probability density function (PDF) corresponding to Equation (2.1) is

$$g_{\mathrm{GR}}(x;\delta,\theta) = \frac{2\theta^{\delta+1}}{\Gamma(\delta+1)}\,x^{2\delta+1}\,\mathrm{e}^{-\theta x^2}. \tag{2.4}$$

Let $Z \sim \mathrm{GR}(\delta,\theta)$ be a random variable with density given in Equation (2.4). The moments of Z are easily obtained from the integral given in Section 3.478 of Gradshteyn and Ryzhik (2000): $\int_0^\infty x^{\nu-1}\mathrm{e}^{-\mu x^p}\mathrm{d}x = \Gamma(\nu/p)(p\mu^{\nu/p})$, where $p,\nu,\mu > 0$. Indeed, the sth ordinary moment of Z (for a positive real number s) is

$$\mu_s'(\delta,\theta) = \frac{\Gamma(\frac{s}{2}+\delta+1)}{\theta^{s/2}\,\Gamma(\delta+1)}. \tag{2.5}$$

Further, generalized distributions can be used for model discrimination. Eugene et al. (2002) pioneered the generator function based on the beta distribution, Cordeiro et al. (2011) presented the generating function from the Kumaraswamy distribution, Alexander et al. (2012) defined the generating function based on the McDonald distribution, Nadarajah et al. (2015a) proposed the Zografos-Balakrishnan-G family of distributions, Nadarajah et al. (2015b) studied the exponentiated G geometric family of distributions, Pescim et al. (2012) proposed the generating function based on the Kummer beta distribution, Cordeiro et al. (2017) proposed the generalized odd log-logistic family of distributions, Cordeiro et al. (2018) defined the odd Lomax generator of distributions, and recently Cordeiro et al. (2020) defined the extended beta family of distributions to generalize the beta generator pioneered by Eugene et al. (2002).

The CDF of the generalized odd log-logistic-G (GOLL-G) family follows from Cordeiro et al. (2017)

$$F(x;\alpha,\beta,\boldsymbol{\xi}) = \frac{G(x;\boldsymbol{\xi})^{\alpha\beta}}{G(x;\boldsymbol{\xi})^{\alpha\beta}+[1-G(x;\boldsymbol{\xi})^{\beta}]^{\alpha}}, \tag{2.6}$$

where $\alpha > 0$ and $\beta > 0$ are two extra parameters.

The generalized log-logistic (Gleaton and Lynch, 2006) and proportional reversed hazard rate (Gupta and Gupta, 2007) are special models of Equation (2.6). Further, for $\alpha = \beta = 1$, we obtain the GR distribution. Prataviera et al. (2018) discussed the generalized odd log-logistic flexible Weibull distribution.

If $g(x;\boldsymbol{\xi})$ is the parent density, the PDF corresponding to Equation (2.6) can be written as

$$f(x;\alpha,\beta,\boldsymbol{\xi})=\frac{\alpha\beta g(x;\boldsymbol{\xi})\,G(x;\boldsymbol{\xi})^{\alpha\beta-1}[1-G(x;\boldsymbol{\xi})^{\beta}]^{\alpha-1}}{\left\{G(x;\boldsymbol{\xi})^{\alpha\beta}+[1-G(x;\boldsymbol{\xi})^{\beta}]^{\alpha}\right\}^{2}}\,. \tag{2.7}$$

The GOLLGR density is defined (for $x>0$) by inserting Equations (2.1) and (2.4) in Equation (2.7)

$$f(x;\alpha,\beta,\delta,\theta)=\frac{2\alpha\beta\theta^{\delta+1}\,x^{2\delta+1}\,\mathrm{e}^{-\theta x^{2}}\,\gamma_1\left(\delta+1,\theta x^{2}\right)^{\alpha\beta-1}\left[1-\gamma_1\left(\delta+1,\theta x^{2}\right)^{\beta}\right]^{\alpha-1}}{\Gamma(\delta+1)\left\{\gamma_1\left(\delta+1,\theta x^{2}\right)^{\alpha\beta}+\left[1-\gamma_1\left(\delta+1,\theta x^{2}\right)^{\beta}\right]^{\alpha}\right\}^{2}}, \tag{2.8}$$

and its hazard rate function (HRF) is

$$h(x)=\frac{2\alpha\beta\theta^{\delta+1}\,x^{2\delta+1}\,\mathrm{e}^{-\theta x^{2}}\,\gamma_1\left(\delta+1,\theta x^{2}\right)^{\alpha\beta-1}}{\Gamma(\delta+1)\left[1-\gamma_1\left(\delta+1,\theta x^{2}\right)^{\beta}\right]\left\{\gamma_1\left(\delta+1,\theta x^{2}\right)^{\alpha\beta}+\left[1-\gamma_1\left(\delta+1,\theta x^{2}\right)^{\beta}\right]^{\alpha}\right\}}\,.$$

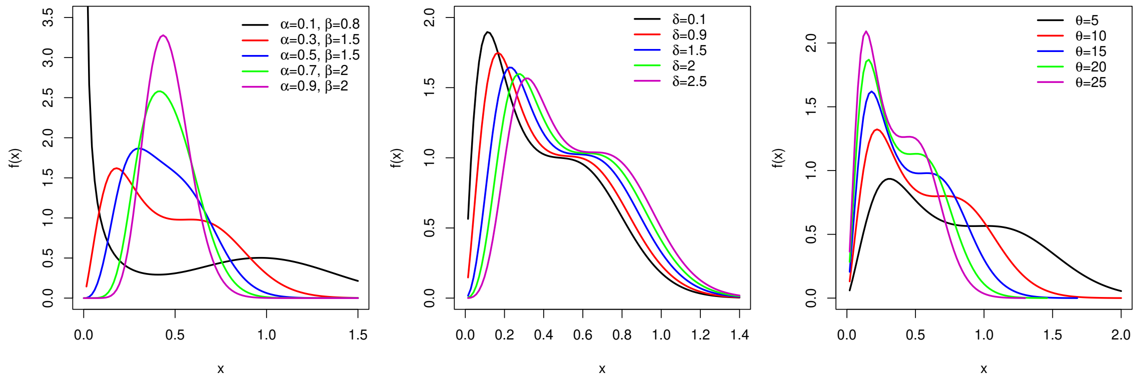
We have $\lim_{x\to\infty}f(x)=0$, and

$$\lim_{x\to0^{+}}f(x)=\begin{cases}\frac{2\sqrt{\theta}}{\sqrt{\pi}}, & \alpha\beta-1=0,\ 2\delta+1=0,\\ 0, & \alpha\beta-1>0,\ 2\delta+1>0,\\ \infty, & \alpha\beta-1>0,\ 2\delta+1<0,\ 0<\delta+1<\frac{2\delta+1}{2(1-\alpha\beta)},\\ \frac{2\alpha\beta\sqrt{\theta}}{\Gamma(\delta+1)}\left[\frac{2(1-\alpha\beta)}{\Gamma(\delta+1)(2\delta+1)}\right]^{\alpha\beta-1}, & \alpha\beta-1>0,\ 2\delta+1<0,\ \delta+1=\frac{2\delta+1}{2(1-\alpha\beta)},\\ 0, & \alpha\beta-1>0,\ 2\delta+1<0,\ \delta+1>\frac{2\delta+1}{2(1-\alpha\beta)},\\ \infty, & \alpha\beta-1<0,\ 2\delta+1>0,\ \delta+1>\frac{2\delta+1}{2(1-\alpha\beta)},\\ \frac{2\alpha\beta\sqrt{\theta}}{\Gamma(\delta+1)}\left[\frac{2(1-\alpha\beta)}{\Gamma(\delta+1)(2\delta+1)}\right]^{\alpha\beta-1}, & \alpha\beta-1<0,\ 2\delta+1>0,\ \delta+1=\frac{2\delta+1}{2(1-\alpha\beta)},\\ 0, & \alpha\beta-1<0,\ 2\delta+1>0,\ 0<\delta+1<\frac{2\delta+1}{2(1-\alpha\beta)},\\ \infty, & \alpha\beta-1\le0,\ 2\delta+1\le0,\ \alpha\beta-1\neq0 \text{ or } 2\delta+1\neq0.\end{cases} \tag{2.9}$$

Further, $\lim_{x\to\infty}h(x)=\infty$ and $\lim_{x\to0^{+}}h(x)=\lim_{x\to0^{+}}f(x)$.

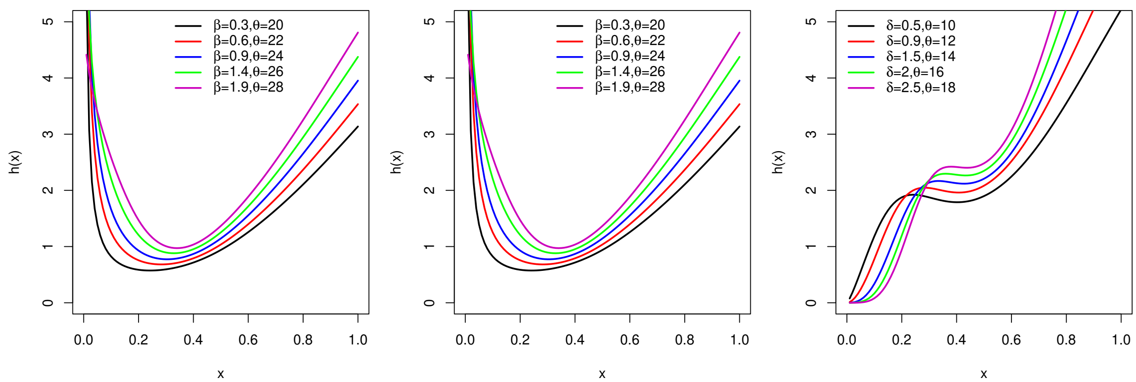
Hereafter, we omit the arguments of the functions, and let $X\sim\mathrm{GOLLGR}(\alpha,\beta,\delta,\theta)$ be a random variable with PDF given in Equation (2.8). The odd log-logistic GR follows when $\beta=1$, and the proportional reversed hazard rate GR when $\alpha=1$. Plots of the PDF and HRF of X are displayed in Figures 1 and 2, respectively. They reveal the bimodality and asymmetry of the PDF, and different shapes of the HRF.

In recent years, bimodal distributions have played an important role in the applied statistical literature; see, for example, Alizadeh et al. (2017); Vila et al. (2020, 2021a,b); Ribeiro-Reis et al. (2020).



(a) (b) (c)

Figure 1. The GOLLGR PDF. (a) $\delta = 1.5$ and $\theta = 15$. (b) $\alpha = 0.3$, $\beta = 2$ and $\theta = 15$. (c) $\alpha = 0.3$, $\beta = 1.5$ and $\delta = 1.5$.



(a) (b) (c)

Figure 2. The GOLLGR HRF. (a) $\alpha = 0.1$ and $\delta = 1.5$. (b) $\alpha = 0.3$ and $\beta = 2.5$. (c) $\alpha = 0.1$.

2.2 Asymptotes and quantile function

The asymptotes below follow from Equations (2.6) and (2.7)

$$F(x) \sim G(x)^{\alpha\beta}, \quad f(x)\,,\, h(x) \sim \alpha\beta g(x) G(x)^{\alpha\beta-1} \quad \text{as} \quad x \to 0^+,$$

and

$$f(x) \sim \alpha\beta\, g(x) \left[1 - G(x)^{\beta}\right]^{\alpha-1}, \quad h(x) \sim \frac{\alpha\beta g(x)}{[1 - G(x)^{\beta}]} \quad \text{as} \quad x \to \infty.$$

For the GOLLGR distribution, we obtain

$$\begin{aligned} F(x) &\sim \gamma_1(\delta+1, \theta x^2)^{\alpha\beta}\,, \\ f(x), h(x) &\sim \frac{2\alpha\beta\theta^{\delta+1}}{\Gamma(\delta+1)}\, x^{2\delta+1}\, \mathrm{e}^{-\theta x^2}\, \gamma_1(\delta+1, \theta x^2)^{\alpha\beta-1} \quad \text{as} \quad x \to 0^+, \end{aligned}$$

and

$$f(x) \sim \frac{2\alpha\beta\theta^{\delta+1}\, x^{2\delta+1}\, \mathrm{e}^{-\theta x^2} \left[1 - \gamma_1\left(\delta+1, \theta x^2\right)^{\beta}\right]^{\alpha-1}}{\Gamma(\delta+1)},$$

$$h(x) \sim \frac{2\alpha\beta\theta^{\delta+1}\, x^{2\delta+1}\, \mathrm{e}^{-\theta x^2}}{\Gamma(\delta+1)\left[1 - \gamma_1\left(\delta+1, \theta x^2\right)^{\beta}\right]} \quad \text{as} \quad x \to \infty.$$

By combining the inverse functions of Equations (2.1) and (2.6), the QF of X can be expressed as

$$x = Q(u) = Q_{\mathrm{GR}}\left(\left[\frac{(\frac{u}{1-u})^{1/\alpha}}{1 + (\frac{u}{1-u})^{1/\alpha}}\right]^{1/\beta}\right), \tag{2.10}$$

where the GR QF $Q_{\mathrm{GR}}(z) = G_{\mathrm{GR}}^{-1}(z; \delta, \theta)$ comes as

$$Q_{\mathrm{GR}}(z) = \left[\theta^{-1}\, \gamma_1^{-1}(\delta+1; z)\right]^{1/2}. \tag{2.11}$$

Here, $\gamma^{-1}\,(\delta+1; z)$ is the gamma QF with shape $\delta+1$ and unity scale. So, the GOLLGR variates follow easily from Equation (2.11).

2.3 Stochastic representation

The PDF of the log-logistic random variable $Y \sim \mathrm{LL}(\nu, \alpha)$ is

$$f_Y(y; \nu, \alpha) = \frac{(\alpha/\nu)(y/\nu)^{\alpha-1}}{\left[1 + (y/\nu)^{\alpha}\right]^2},$$

where $\nu > 0$ and $\alpha > 0$ are scale and shape, respectively.

If $X \sim \mathrm{GOLLGR}(\alpha, \beta, \delta, \theta)$ and $Y \sim \mathrm{LL}(1, \alpha)$, we can write from Equations (2.6) and (2.11)

$$\begin{aligned} F(x; \alpha, \beta, \delta, \theta) &= \mathrm{P}\left(Y \le \frac{G_{\mathrm{GR}}(x; \delta, \theta)^{\beta}}{1 - G_{\mathrm{GR}}(x; \delta, \theta)^{\beta}}\right) \\ &= \mathrm{P}\left(\left[\theta^{-1}\, \gamma_1^{-1}\left(\delta+1; \left(\frac{Y}{1+Y}\right)^{1/\beta}\right)\right]^{1/2} \le x\right), \end{aligned} \tag{2.12}$$

where $G_{\mathrm{GR}}(x; \delta, \theta) = \gamma_1(\delta+1, \theta x^2)$, $x > 0$. So, we prove the following:

Proposition 2.1 The GOLLGR random variable X admits the stochastic representation:

$$X = \left[\theta^{-1}\, \gamma_1^{-1}\left(\delta+1; \left(\frac{Y}{1+Y}\right)^{1/\beta}\right)\right]^{1/2} \quad \text{for} \quad Y \sim \mathrm{LL}(1, \alpha).$$

2.4 Critical points and modality

For brevity, we define

$$T(x) = T(x;\beta,\delta,\theta) = \frac{G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}}{1 - G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}}. \tag{2.13}$$

Since $T(x/k;\beta,\delta,\theta) = T(x;\beta,\delta,\theta/k^2)$, $k > 0$, the next result follows from Equation (2.12):

Proposition 2.2 If $X \sim \mathrm{GOLLGR}(\alpha,\beta,\delta,\theta)$, then $k\,X \sim \mathrm{GOLLGR}(\alpha,\beta,\delta,\theta/k^2)$.

Proposition 2.3 A critical point of the PDF of X satisfies

$$\frac{y''}{(y')^2} + \frac{(\beta+1)y^{\beta}[y^{\alpha\beta} + (1-y^{\beta})^{\alpha}] - (\alpha\beta+1)y^{\alpha\beta} - 2(1-y^{\beta})^{\alpha}}{y(1-y^{\beta})[y^{\alpha\beta} + (1-y^{\beta})^{\alpha}]} = 0, \tag{2.14}$$

where $y = y(x) = G_{\mathrm{GR}}(x;\delta,\theta) = \gamma_1(\delta+1,\theta x^2)$.

In the remainder of this section, we use Equation (2.14) and the limit in Equation (2.9) to analyze the modality of the PDF of X when $\alpha = 1$.

For $\alpha = 1$, Equation (2.14) reduces to $g_{\mathrm{GR}}(x;\delta,\theta)/G_{\mathrm{GR}}(x;\delta,\theta) = [\delta + \theta x^2 - (1/2)]/x$. Equivalently,

$$G_{\mathrm{GR}}(x;\delta,\theta) = g(x), \tag{2.15}$$

where $g(x) = x g_{\mathrm{GR}}(x;\delta,\theta)$.

A careful analysis shows that, on $(0,\infty)$: (i) $g(x)$ is decreasing/decreasing-increasing-decreasing when $\delta < 1/2$, (ii) $g(x)$ is unimodal when $\delta \geq 1/2$ and (iii) $\lim_{x\to 0^+} g(x) = \lim_{x\to\infty} g(x) = 0$. Moreover, notice that: (a) $g(x)$ has a vertical asymptotic at $x_{\mathrm{va}} = \sqrt{[(1/2)-\delta]/\theta}$ when $\delta < 1/2$, (b) $g(x) < 0$ for all $x < x_{\mathrm{va}}$ and (c) $g(x) > 0$ for all $x > x_{\mathrm{va}}$. Since $G_{\mathrm{GR}}(x;\delta,\theta)$ is increasing on $(0,\infty)$, because this one is a CDF, it is plausible to expect that, by varying the parameters, Equation (2.15) has at most three roots on $(0,\infty)$. So, the PDF of X has at most three critical points on $(0,\infty)$. In the following, we analyze some possible scenarios:

- If the GOLLGR PDF has a single critical point, say x_0, and $\beta > 1$ and $\delta \geq 1/2$, we have by the second limit in Equation (2.9) $\lim_{x\to 0^+} f(x;\alpha,\beta,\delta,\theta) = \lim_{x\to\infty} f(x;\alpha,\beta,\delta,\theta) = 0$. Then, x_0 is really a maximum point, and the PDF of X is unimodal.
- If the GOLLGR PDF has three single critical points, say $x_1 < x_2 < x_3$, and $\beta > 1$ and $\delta \geq 1/2$, again, by the second limit in Equation (2.9), we have $\lim_{x\to 0^+} f(x;\alpha,\beta,\delta,\theta) = \lim_{x\to\infty} f(x;\alpha,\beta,\delta,\theta) = 0$. Hence, x_1 and x_3 are maximum points, and x_2 is a minimum point, and the GOLLGR PDF is bimodal.

In general, as done previously, we can use the limit in Equation (2.9) and the number of critical points of the PDF of X to obtain the result:

Theorem 2.4 The GOLLGR PDF is (i) decreasing, (ii) decreasing-increasing-decreasing or (iii) uni- or bimodal.

2.5 Tail behavior

Here, we prove that, under certain constraints in the parameter space, the distribution of X has thinner tails than an exponential distribution. More precisely, we prove the two results.

Proposition 2.5 For any $\alpha \geq 1$ and any $t > 0$,

$$\lim_{x\to\infty} \frac{\mathrm{e}^{-tx}}{1 - F(x;\alpha,\beta,\delta,\theta)} = \infty. \tag{2.16}$$

Proposition 2.6 For any $0 < \beta \leq 1$ and $\delta > 0$, the limit (for $t > 0$) (2.16) holds.

2.6 Linear Representation

First, the exponentiated-G ("Exp-G") random variable $W \sim \mathrm{Exp}^c\mathrm{G}$ for a continuous CDF $G(x)$ and $c > 0$, has CDF $H_c(x) = G(x)^c$ and PDF $h_c(x) = c g(x) G(x)^{c-1}$. Many Exp-G properties were published in the last three decades.

We obtain after some algebra the power series

$$G(x)^{\alpha\beta} + \left[1 - G(x)^{\beta}\right]^{\alpha} = \sum_{k=0}^{\infty} c_k\, G(x)^k\,, \tag{2.17}$$

where $c_k = c_k(\alpha,\beta) = a_k + \sum_{i=0}^{\infty}\sum_{j=k}^{\infty}(-1)^{i+j+k}\binom{\alpha}{i}\binom{i\beta}{j}\binom{j}{k}$, and $a_k = \sum_{j=k}^{\infty}(-1)^{j+k}\binom{\alpha\beta}{j}\binom{j}{k}$. The expansion given in Equation (2.17) converges for all x such that $0 < G(x) < 1$ due to the generalized binomial theorem.

Equation (2.6) can be rewritten from the ratio of two convergent power series as the convergent expansion

$$F(x) = \sum_{k=0}^{\infty} d_k\, G(x)^k, \tag{2.18}$$

where $d_k = d_k(\alpha,\beta)$'s are found recursively (for $k > 0$, $d_0 = a_0/c_0$) $d_k = c_0^{-1}(a_k + \sum_{r=1}^{k} c_r\, d_{k-r})$.

By differentiating Equation (2.18) and using the "term-wise differentiation of power series theorem", and changing indices $f(x) = \sum_{l=0}^{\infty}(l+1)\, d_{l+1}\, g(x)\, G(x)^l$. For the GR model, we obtain

$$f(x) = \sum_{l=0}^{\infty}(l+1)\, d_{l+1}(\alpha,\beta)\, \frac{2\theta^{\delta+1}}{\Gamma(\delta+1)}\, x^{2\delta+1}\, \mathrm{e}^{-\theta x^2}\, \gamma_1(\delta+1,\theta x^2)^l\,. \tag{2.19}$$

The convergent power series for the incomplete gamma function ratio holds

$$\gamma_1(\delta+1,\theta x^2) = \frac{(\theta x^2)^{\delta+1}}{\Gamma(\delta+1)} \sum_{m=0}^{\infty} \frac{(-1)^m\,(\theta x^2)^m}{m!\,(\alpha+1+m)}\,.$$

Equation 0.314 in Gradshteyn and Ryzhik (2000) gives (for a natural number $l \geq 1$)

$$\left(\sum_{m=0}^{\infty} q_m\, x^m\right)^l = \sum_{m=0}^{\infty} e_m^{(l)}\, x^m\,,$$

where $e_0^{(l)} = q_0^l$, and $e_m^{(l)}$ (for $l \geq 1$) can be found from

$$e_m^{(l)} = \frac{1}{m\, q_0} \sum_{i=1}^{m}[(l+1)\, i - m]\, q_i\, e_{m-i}^{(l)}. \tag{2.20}$$

The next expansion converges because it is a natural power of a convergent power series

$$\gamma_1(\delta+1,\theta x^2)^l = \frac{(\theta x^2)^{l(\delta+1)}}{\Gamma(\delta+1)^l} \sum_{m=0}^{\infty} e_m^{(l)}\, x^{2m}\,, \tag{2.21}$$

where the quantities $e_m^{(l)}$ follow from Equation (2.20) with the constants $q_m = (-1)^m\,\theta^m/[(\delta+1+m)m!]$, for $m = 0, 1, \ldots$.

Further, we set the conditions $e_0^{(0)} = 1$ and $e_m^{(0)} = 0$ for $m \geq 1$. Hence, inserting Equation (2.21) in Equation (2.19) (under these conditions) leads to

$$f(x) = \sum_{l,m=0}^{\infty} \frac{2\,(l+1)\,d_{l+1}(\alpha,\beta)\,e_m^{(l)}}{\Gamma(\delta+1)^{l+1}\theta^m}\,\theta^{[l(\delta+1)+m+\delta]+1} x^{2[l(\delta+1)+m+\delta]+1}\mathrm{e}^{-\theta x^2}$$

and then

$$f(x) = \sum_{l,m=0}^{\infty} w_{l,m}\, g_{\mathrm{GR}}(x;\theta,\delta^*_{l,m})\,, \tag{2.22}$$

where $\delta^*_{l,m} = l(\delta+1)+m+\delta$, and the coefficients are $w_{l,m} = w_{l,m}(\theta,\delta,\alpha,\beta) = (l+1)\,\Gamma(\delta^*_{l,m}+1)\,d_{l+1}(\alpha,\beta)\,e_m^{(l)}/[\theta^m\,\Gamma(\delta+1)^{l+1}]\,\cdot$

Equation (2.22) is useful to obtain some properties of the new distribution from those of the GR model.

2.7 Other Properties

The sth ordinary moment of X comes from Equations (2.5) and (2.22) as

$$\mathrm{E}(X^s) = \sum_{l,m=0}^{\infty} w_{l,m}(\theta,\delta,\alpha,\beta)\,\frac{\Gamma(\frac{s}{2}+\delta^*_{l,m}+1)}{\theta^{s/2}\,\Gamma(\delta^*_{l,m}+1)}.$$

The sth incomplete moment of$X$ follows from Equation (2.22) as

$$\begin{aligned} m_s(x) &= \sum_{l,m=0}^{\infty} \frac{w_{l,m}(\theta,\delta,\alpha,\beta)}{\Gamma(\delta^*_{l,m}+1)} \int_0^x 2\theta^{\delta^*_{l,m}+1}\, t^s\, t^{2\delta^*_{l,m}+1}\,\mathrm{e}^{-\theta t^2}\,\mathrm{d}t \\ &= \sum_{l,m=0}^{\infty} w_{l,m}(\theta,\delta,\alpha,\beta)\frac{\Gamma(\delta^*_{l,m}+\frac{s}{2}+1)}{\Gamma(\delta^*_{l,m}+1)\theta^{\frac{s}{2}}} \int_0^x \frac{2\theta^{\delta^*_{l,m}+\frac{s}{2}+1}}{\Gamma(\delta^*_{l,m}+\frac{s}{2}+1)}\, t^{2(\delta^*_{l,m}+\frac{s}{2})+1}\mathrm{e}^{-\theta t^2}\,\mathrm{d}t \\ &= \sum_{l,m=0}^{\infty} w_{l,m}(\theta,\delta,\alpha,\beta)\frac{\Gamma(\frac{s}{2}+\delta^*_{l,m}+1)}{\theta^{\frac{s}{2}}\,\Gamma(\delta^*_{l,m}+1)}\gamma_1(\delta^*_{l,m}+\frac{s}{2}+1,\theta x^2). \end{aligned}$$

The mean deviations and Bonferroni and Lorenz curves of X are obtained from $m_1(x)$. The generating function (GR) of X can be expressed as

$$M(t) = \int_0^{\infty} \mathrm{e}^{tx} \sum_{l,m=0}^{\infty} w_{l,m}(\theta,\delta,\alpha,\beta)\, g_{\mathrm{GR}}(x;\theta,\delta^*_{l,m})\,\mathrm{d}x,$$

that is,

$$M(t)=\sum_{l,m=0}^{\infty}\frac{2\,w_{l,m}(\theta,\delta,\alpha,\beta)\,\theta^{\delta^*_{l,m}+1}}{\Gamma(\delta^*_{l,m}+1)}\int_0^{\infty}x^{2\delta^*_{l,m}+1}\,\mathrm{e}^{-\theta x^2+tx}\,\mathrm{d}x.$$

From Equation 2.3.15.3 in Prudnikov et al. (1986), we can write

$$\int_0^{\infty}x^{\alpha-1}\mathrm{e}^{-px^2-qx}\mathrm{d}x=\frac{\Gamma(\alpha)}{(2p)^{\frac{\alpha}{2}}}\exp\left(\frac{q^2}{8p}\right)D_{-\alpha}\left(\frac{q}{\sqrt{2p}}\right),$$

where $\mathrm{Re}(\alpha)$, $\mathrm{Re}(p)>0$, and

$$D_p(y)=\frac{\exp(-y^2/4)}{\Gamma(-p)}\int_0^{\infty}\exp\{-(wy+w^2/2)\}w^{-(p+1)}\,\mathrm{d}w.$$

Thus,

$$M(t)=\sum_{l,m=0}^{\infty}\frac{2\,w_{l,m}\,\theta^{\delta^*_{l,m}+1}}{\Gamma(\delta^*_{l,m}+1)}\frac{\Gamma(\tilde{\delta}_{l,m})}{(2\theta)^{\frac{\tilde{\delta}_{l,m}}{2}}}\exp\left(\frac{t^2}{8\theta}\right)D_{-\tilde{\delta}_{l,m}}\left(-\frac{t}{\sqrt{2\theta}}\right),$$

where $\tilde{\delta}_{l,m}=2\,(\delta^*_{l,m}+1)$.

2.8 Estimation

We obtain the maximum likelihood estimate (MLE) of $\boldsymbol{\eta}=(\alpha,\beta,\delta,\theta)^{\top}$ given the data $x_1,x_2,\ldots,x_n$ from the GOLLGR distribution.

The total log-likelihood function for $\boldsymbol{\eta}$ is

$$\begin{aligned}l_n(\boldsymbol{\eta})=&n\,[\log(\alpha\beta)+(\delta+1)\log\theta]+(2\delta+1)\sum_{i=1}^{n}\log x_i-\theta\sum_{i=1}^{n}x_i^2\\&+(\alpha\beta-1)\sum_{i=1}^{n}\log\gamma_1(\delta+1,\theta x_i^2)+(\alpha-1)\sum_{i=1}^{n}\log[1-\gamma_1(\delta+1,\theta x_i^2)^{\beta}]\\&-2\sum_{i=1}^{n}\log\{\gamma_1(\delta+1,\theta x_i^2)^{\alpha\beta}+[1-\gamma_1(\delta+1,\theta x_i^2)^{\beta}]^{\alpha}\}.\end{aligned}\tag{2.23}$$

The maximization of Equation (2.23) can be done using the R software (R Core Team., 2022) or SAS (PROC NLMIXED), among others. The components of the score function

$U_n(\boldsymbol{\eta}) = \left(\frac{\partial l_n(\boldsymbol{\eta})}{\partial \alpha}, \frac{\partial l_n(\boldsymbol{\eta})}{\partial \beta}, \frac{\partial l_n(\boldsymbol{\eta})}{\partial \delta}, \frac{\partial l_n(\boldsymbol{\eta})}{\partial \theta}\right)^\top$ are

$$\begin{aligned}
U_n(\alpha) =& \frac{n}{\alpha} + \beta \sum_{i=1}^{n} \log \gamma_1 \left(\delta+1, \theta x_i^2\right) + \sum_{i=1}^{n} \log \left[1 - \gamma_1 \left(\delta+1, \theta x_i^2\right)^{\beta}\right] \\
& - 2 \sum_{i=1}^{n} \frac{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} \beta \log \gamma_1\left(\delta+1, \theta x_i^2\right)}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \\
& - 2 \sum_{i=1}^{n} \frac{\left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha} \log \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}},
\end{aligned}$$

$$\begin{aligned}
U_n(\beta) =& \frac{n}{\beta} + \alpha \sum_{i=1}^{n} \log \gamma_1 \left(\delta+1, \theta x_i^2\right) - (\alpha-1) \sum_{i=1}^{n} \frac{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta} \log \gamma_1\left(\delta+1, \theta x_i^2\right)}{1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}} \\
& - 2 \sum_{i=1}^{n} \frac{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} \alpha \log \gamma_1\left(\delta+1, \theta x_i^2\right)}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \\
& - 2 \sum_{i=1}^{n} \frac{\alpha\left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha-1} \left(-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right) \log \gamma_1\left(\delta+1, \theta x_i^2\right)}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}},
\end{aligned}$$

$$\begin{aligned}
U_n(\delta) =& n \log \theta + 2 \sum_{i=1}^{n} \log x_i + (\alpha\beta-1) \sum_{i=1}^{n} \frac{1}{\gamma_1\left(\delta+1, \theta x_i^2\right)} \frac{\partial}{\partial \delta} \gamma_1 \left(\delta+1, \theta x_i^2\right) \\
& + (\alpha-1) \sum_{i=1}^{n} \frac{1}{1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}} (-\beta) \gamma_1 \left(\delta+1, \theta x_i^2\right)^{\beta-1} \frac{\partial}{\partial \delta} \gamma_1 \left(\delta+1, \theta x_i^2\right) \\
& - 2 \sum_{i=1}^{n} \frac{\alpha\beta\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \frac{\partial}{\partial \delta} \gamma_1 \left(\delta+1, \theta x_i^2\right) \\
& - 2 \sum_{i=1}^{n} \frac{\alpha\left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha-1} (-\beta)\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \frac{\partial}{\partial \delta} \gamma_1 \left(\delta+1, \theta x_i^2\right),
\end{aligned}$$

$$\begin{aligned}
U_n(\theta) =& \frac{n(\delta+1)}{\theta} - \sum_{i=1}^{n} x_i^2 + (\alpha\beta-1) \sum_{i=1}^{n} \frac{1}{\gamma_1\left(\delta+1, \theta x_i^2\right)} \frac{\partial}{\partial \theta} \gamma_1 \left(\delta+1, \theta x_i^2\right) \\
& + (\alpha-1) \sum_{i=1}^{n} \frac{1}{1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}} (-\beta) \gamma_1 \left(\delta+1, \theta x_i^2\right)^{\beta-1} \frac{\partial}{\partial \theta} \gamma_1 \left(\delta+1, \theta x_i^2\right) \\
& - 2 \sum_{i=1}^{n} \frac{\alpha\beta\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \frac{\partial}{\partial \theta} \gamma_1 \left(\delta+1, \theta x_i^2\right) \\
& - 2 \sum_{i=1}^{n} \frac{\alpha\left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha-1} (-\beta)\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1-\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \frac{\partial}{\partial \theta} \gamma_1 \left(\delta+1, \theta x_i^2\right).
\end{aligned}$$

From Equations (2.2) and (2.3),

$$
\begin{aligned}
U_n(\delta) =& n \log \theta + 2 \sum_{i=1}^n \log x_i + (\alpha\beta - 1) \sum_{i=1}^n \frac{H_1(\delta+1, \theta x_i^2) - \gamma_1(\delta+1, \theta x_i^2)\, \psi(\delta+1)}{\gamma_1\left(\delta+1, \theta x_i^2\right)} \\
& - \beta(\alpha-1) \sum_{i=1}^n \frac{\gamma_1(\delta+1, \theta x_i^2)^{\beta-1} \left[H_1(\delta+1, \theta x_i^2) - \gamma_1(\delta+1, \theta x_i^2)\, \psi(\delta+1)\right]}{1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}} \\
& - 2\alpha\beta \sum_{i=1}^n \frac{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta-1} \left[H_1(\delta+1, \theta x_i^2) - \gamma_1(\delta+1, \theta x_i^2)\, \psi(\delta+1)\right]}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \\
& + 2\alpha\beta \sum_{i=1}^n \frac{\left[1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha-1} \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \\
& \times \left[H_1(\delta+1, \theta x_i^2) - \gamma_1(\delta+1, \theta x_i^2)\, \psi(\delta+1)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
U_n(\theta) =& \frac{n(\delta+1)}{\theta} - \sum_{i=1}^n x_i^2 + (\alpha\beta-1) \sum_{i=1}^n \frac{1}{\gamma_1\left(\delta+1, \theta x_i^2\right)} \frac{\theta^{\delta} x_i^{2(\delta+1)} \mathrm{e}^{-\theta x_i^2}}{\Gamma(\delta+1)} \\
& - \beta(\alpha-1) \sum_{i=1}^n \frac{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta-1}}{1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}} \frac{\theta^{\delta} x_i^{2(\delta+1)} \mathrm{e}^{-\theta x_i^2}}{\Gamma(\delta+1)} \\
& - 2\alpha\beta \sum_{i=1}^n \frac{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \frac{\theta^{\delta} x_i^{2(\delta+1)} \mathrm{e}^{-\theta x_i^2}}{\Gamma(\delta+1)} \\
& + 2\alpha\beta \sum_{i=1}^n \frac{\left[1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha-1} \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta-1}}{\gamma_1\left(\delta+1, \theta x_i^2\right)^{\alpha\beta} + \left[1 - \gamma_1\left(\delta+1, \theta x_i^2\right)^{\beta}\right]^{\alpha}} \frac{\theta^{\delta} x_i^{2(\delta+1)} \mathrm{e}^{-\theta x_i^2}}{\Gamma(\delta+1)}.
\end{aligned}
$$

The elements of the observed information matrix are obtained by differentiating $U_n(\boldsymbol{\eta})$ and computing numerically.

3. The GOLLGR regression

3.1 Context

Recently, some papers on regression models have been published, for example, Hashimoto et al. (2019), Prataviera et al. (2020), Silva et al. (2020) and Vasconcelos et al. (2021). In a similar context, we introduce the regression model based on the GOLLGR distribution.

The systematic components of the GOLLGR regression for X are defined by

$$
\delta_i = \exp(\boldsymbol{v}_i^\top \boldsymbol{\lambda}_2) - 1 \quad \text{and} \quad \theta_i = \exp(\boldsymbol{v}_i^\top \boldsymbol{\lambda}_1), \quad i = 1, 2, \ldots, n, \tag{3.24}
$$

where $\boldsymbol{\lambda}_1 = (\lambda_{11}, \lambda_{12} \ldots, \lambda_{1p})^\top$ and $\boldsymbol{\lambda}_2 = (\lambda_{21}, \lambda_{22}, \ldots, \lambda_{2p})^\top$ are vectors of unknown coefficients, and $\boldsymbol{v}_i^\top = (v_{i1}, v_{i2}, \ldots, v_{ip})$ is a vector of known explanatory variables.

The survival function of X comes from Equation (2.6) as

$$S(x_i|\boldsymbol{v}) = \frac{[1-\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\beta}]^{\alpha}}{\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\alpha\,\beta}+[1-\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\beta}]^{\alpha}}. \tag{3.25}$$

Equation (3.25) opens new possibilities for fitting different types of regressions. The odd log-logistic GR (OLLGR) regression follows when $\beta = 1$, the exponentiated GR (EGR) regression when $\alpha = 1$, and the GR regression when $\beta = \alpha = 1$.

Let X_i be the lifetime and C_i be the non-informative censoring time (assuming independent), and $x_i = \min\{X_i, C_i\}$. The total log-likelihood function for $\boldsymbol{\eta} = (\alpha, \beta, \boldsymbol{\lambda}_1^{\top})^{\top}$ from regression given in Equation (3.24) is

$$\begin{aligned} l(\boldsymbol{\eta}) =& r\log\left(\frac{\alpha\,\beta\,2}{\Gamma(\delta_i+1)}\right)+(\delta_i+1)\sum_{i\in F}\log(\theta_i)+(2\,\delta_i+1)\sum_{i\in F}\log(x_i)-\sum_{i\in F}\theta_i x_i^2\\ &+(\alpha\,\beta-1)\sum_{i\in F}\log[\gamma_1(\delta_i+1,\theta_i\,x_i^2)]+(\alpha-1)\sum_{i\in F}\log[1-\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\beta}]\\ &-2\sum_{i\in F}\log\left\{\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\alpha\,\beta}+[1-\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\beta}]^{\alpha}\right\}\\ &+\sum_{i\in C}\log\left(\frac{[1-\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\beta}]^{\alpha}}{\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\alpha\,\beta}+[1-\gamma_1(\delta_i+1,\theta_i\,x_i^2)^{\beta}]^{\alpha}}\right), \end{aligned} \tag{3.26}$$

where F and C are the sets of observed lifetimes and censoring times, r is the number of uncensored observations (failures). The MLE $\widehat{\boldsymbol{\eta}}$ of the vector of unknown parameters can be found by maximizing Equation (3.26). Here, we use the `gamlss` package in `R` (R Core Team., 2022) described in the generalized additive model class for location, scale and shape (Rigby et al., 2005). These models are able to model any of the parameters, location, scale and/or shape, depending on the covariates but without the condition of belonging to any family, such as exponential for example. We choose this package because it basically has two algorithms for maximization: `CG` (Cole et al., 1992) and `RS` (Rigby et al., 2005) which guarantee convergence with simple initial values. We work with the `RS` algorithm. We make available the gamlss framework of this new distribution at https://github.com/gabrielamrodrigues/GOLLGR. From this script, it is possible to extend our regression model with non-parametric covariate effects, random effects, or other additive terms, and also model other shape parameters, if necessary.

3.2 Residual analysis

The quantile residuals (qrs) (Dunn and Smyth, 1996) have been adopted frequently in regression applications to verify possible deviations from the model assumptions. For the proposed regression, they are

$$qr_i = \Phi^{-1}\left(\frac{\gamma_1(\hat{\delta}_i+1,\hat{\theta}_i x_i^2)^{\hat{\alpha}\hat{\beta}}}{\gamma_1(\hat{\delta}_i+1,\hat{\theta}_i x_i^2)^{\hat{\alpha}\hat{\beta}}+[1-\gamma_1(\hat{\delta}_i+1,\hat{\theta}_i x_i^2)^{\hat{\beta}}]^{\hat{\alpha}}}\right), \tag{3.27}$$

where $\Phi^{-1}(\cdot)$ is the standard normal QF.

4. Simulations

In this section, simulation studies are presented by using `gamlss` package in `R` software.

4.1 The GOLLGR distribution

For the first study, we verify the accuracy of the MLEs from 10,000 replicates generated from Equation (2.10) for sample sizes n = 50, 150 and 500 according to the two scenarios below:

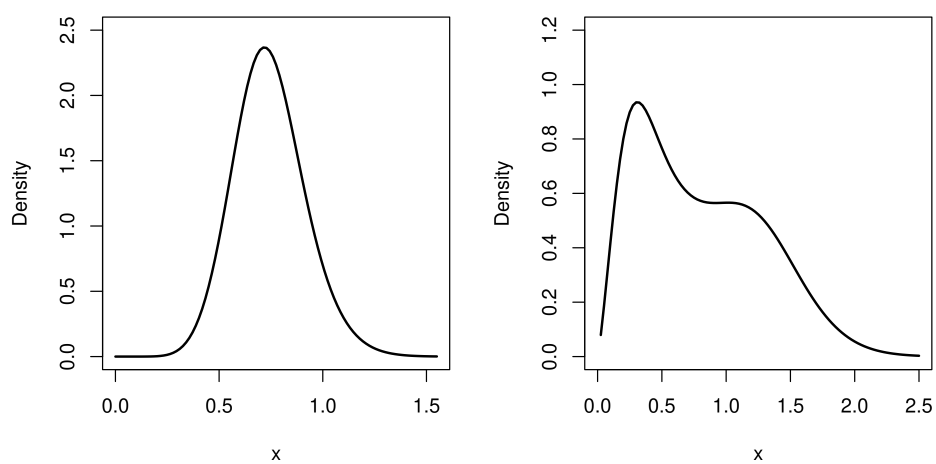
- Scenario 1: Symmetric unimodal density $X \sim \text{GOLLGR}(\alpha = 1.2, \beta = 1.5, \delta = 1.5, \theta = 5)$;
- Scenario 2: Bimodal density with asymmetry $X \sim \text{GOLLGR}(\alpha = 0.3, \beta = 1.5, \delta = 1.5, \theta = 5)$.

PDF plots for each scenario are reported in Figure 3. The average estimates (AEs), biases and mean squared errors (MSEs) of the MLEs are calculated from:

$$\text{AE}(\hat{\boldsymbol{\eta}}) = \frac{1}{r}\sum_{i=1}^{r}\hat{\eta}_i, \quad \text{Bias}(\hat{\boldsymbol{\eta}}) = \frac{1}{r}\sum_{i=1}^{r}(\hat{\eta}_i - \eta_i), \quad \text{MSE}(\hat{\boldsymbol{\eta}}) = \frac{1}{r}\sum_{i=1}^{r}(\hat{\eta}_i - \eta_i)^2, \tag{4.28}$$

where $\hat{\boldsymbol{\eta}}^\top = (\alpha, \beta, \delta, \theta)$ and $r = 10,000$.

Table 1 shows that the AEs converge to the true parameters and the biases and MSEs decay when n increases for both scenarios.



(a) (b)

Figure 3. PDF of X generated for scenarios 1 (a) and 2 (b).

4.2 The GOLLGR regression

The second study examines the accuracy of the MLEs in the proposed regression. Three censoring percentages are taken (approximately) as 0%, 10% and 30% for n = 100, 300 and 500. For each combination, we generate 10,000 replicates. The lifetimes $x_1^*, x_2^*, \ldots, x_n^*$ are generated from the GOLLGR distribution $(\alpha_i, \beta_i, \delta_i, \theta_i)$ and the censoring times $c_1, c_2, \ldots, c_n$

Table 1. Simulation results from the GOLLGR distribution for each scenario (S).

S	$\boldsymbol{\eta}$	True value	$n=50$ AEs	Biases	MSEs	$n=150$ AEs	Biases	MSEs	$n=500$ AEs	Biases	MSEs
1	δ	1.5	1.5050	0.0050	0.0261	1.5010	0.0010	0.0088	1.5001	0.0001	0.0027
	θ	5.0	5.0129	0.0129	0.0116	5.0038	0.0038	0.0034	5.0012	0.0012	0.0012
	α	1.2	1.2289	0.0289	0.0184	1.2097	0.0097	0.0056	1.2030	0.0030	0.0016
	β	1.5	1.5271	0.0271	3.9553	1.5020	0.0020	0.0007	1.5011	0.0011	0.0026
2	δ	1.5	1.5903	0.0903	0.4930	1.5323	0.0323	0.1285	1.5154	0.0154	0.0342
	θ	5.0	5.2360	0.2360	0.8387	5.0782	0.0782	0.2072	5.0213	0.0213	0.0635
	α	0.3	0.3015	0.0015	0.0024	0.3004	0.0004	0.0008	0.3003	0.0003	0.0002
	β	1.5	1.6458	0.1458	0.6741	1.5473	0.0473	0.1338	1.5077	0.0077	0.0285

from a uniform distribution $(0,\nu)$, where ν controls the censoring percentage. We include a covariate $v_{1i} \sim$ Binomial$(1,0.5)$, where v_{1i} is taken in two groups (0 and 1) in the systematic components:

$$\delta_i = \exp(\lambda_{10}+\lambda_{11}v_{1i})-1, \quad \theta_i = \exp(\lambda_{20}+\lambda_{21}v_{1i}), \quad \alpha_i = \exp(\lambda_{30}), \quad \beta_i = \exp(\lambda_{40}). \quad (4.29)$$

The true parameter values are: $\lambda_{10}=0.65$, $\lambda_{11}=2.75$, $\lambda_{20}=0.55$, $\lambda_{21}=1.75$, $\lambda_{30}=0.37$ and $\lambda_{40}=0.61$. The AEs, biases and MSEs of the MLEs are calculated from Equation (4.28), where $\hat{\boldsymbol{\eta}}^\top = (\hat\lambda_{10}, \hat\lambda_{11}, \hat\lambda_{20}, \hat\lambda_{21}, \hat\lambda_{30}, \hat\lambda_{40})$.

The simulation process follows as:

(i) Generate $v_{1i} \sim$ Binomial$(1,0.5)$;

(ii) Calculate δ_i and θ_i from the systematic components given in Equation (4.29);

(iii) Generate $u_i \sim$ U$(0,1)$;

(iv) The previous steps give x_i^*'s from (2.10);

(v) Generate $c_i \sim$ uniform$(0,\nu)$ and determine the survival times $x_i = \min(x_i^*, c_i)$. If $x_i^* < c_i$, then $\gamma_i = 1$; otherwise, $\gamma_i = 0$ (for $i = 1,2,\ldots,n$), where γ_i is the censoring indicator;

(vi) Obtain the MLEs and calculate the AEs, biases and MSEs from Equation (4.28).

The numbers in Table 2 reveal that the AEs converge to the true values, and the biases and MSEs decrease when n increases, thus indicating the consistency of the estimators for these censoring percentages.

Table 2. Simulation results from the GOLLGR regression model.

%	$\boldsymbol{\eta}$	$n=100$ AEs	Biases	MSEs	$n=300$ AEs	Biases	MSEs	$n=500$ AEs	Biases	MSEs
0%	λ_{10}	0.6495	-0.0005	0.0095	0.6507	0.0007	0.0032	0.6508	0.0008	0.0019
	λ_{11}	2.7500	-0.0000	0.0120	2.7491	-0.0009	0.0040	2.7490	-0.0010	0.0023
	λ_{20}	0.5495	-0.0005	0.0117	0.5509	0.0009	0.0040	0.5510	0.0010	0.0023
	λ_{21}	1.7502	0.0002	0.0134	1.7489	-0.0011	0.0045	1.7489	-0.0011	0.0027
	λ_{30}	0.3861	0.0161	0.0084	0.3753	0.0053	0.0027	0.3729	0.0029	0.0016
	λ_{40}	0.6123	0.0023	0.0046	0.6110	0.0010	0.0012	0.6110	0.0010	0.0007
10%	λ_{10}	0.6581	0.0081	0.0116	0.6533	0.0033	0.0043	0.6519	0.0019	0.0024
	λ_{11}	2.7419	-0.0081	0.0119	2.7467	-0.0033	0.0043	2.7481	-0.0019	0.0025
	λ_{20}	0.5591	0.0091	0.0104	0.5531	0.0031	0.0038	0.5520	0.0020	0.0022
	λ_{21}	1.7408	-0.0092	0.0106	1.7469	-0.0031	0.0038	1.7480	-0.0020	0.0022
	λ_{30}	0.3755	0.0055	0.0019	0.3711	0.0011	0.0003	0.3707	0.0007	0.0002
	λ_{40}	0.6103	0.0003	0.0001	0.6100	-0.0000	0.0000	0.6100	0.0000	0.0000
30%	λ_{10}	0.6578	0.0078	0.0128	0.6529	0.0029	0.0043	0.6518	0.0018	0.0025
	λ_{11}	2.7422	-0.0078	0.0128	2.7471	-0.0029	0.0043	2.7481	-0.0019	0.0025
	λ_{20}	0.5582	0.0082	0.0114	0.5525	0.0025	0.0038	0.5520	0.0020	0.0022
	λ_{21}	1.7418	-0.0082	0.0116	1.7475	-0.0025	0.0039	1.7481	-0.0019	0.0023
	λ_{30}	0.3705	0.0005	0.0000	0.3701	0.0001	0.0000	0.3700	0.0000	0.0000
	λ_{40}	0.6100	0.0000	0.0000	0.6100	0.0000	0.0000	0.6100	0.0000	0.0000

True values: $\lambda_{10}=0.65$, $\lambda_{11}=2.75$, $\lambda_{20}=0.55$, $\lambda_{21}=1.75$, $\lambda_{30}=0.37$ and $\lambda_{40}=0.61$.

5. Applications

We perform two applications and compare the GOLLGR model with its submodels GR, EGR, OLLGR and with the Maxwell, Rayleigh, Chi-square, Half-normal, Weibull and Inverse gamma models. We calculate the MLEs, and the criteria: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC).

5.1 Application 1: Voltage data

We consider the times of failure and running times for a field-tracking study (Meeker and Escobar, 1998). The failure times and operating times for a sample of devices from a field tracking system study. At a given time, 30 units are determined for each failed unit. The S-mode failures are caused by an accumulation of random damage caused by voltage spikes on the power line during electrical storms, thus resulting in early life. The W Mode failures, caused by normal product wear and tear, started to appear after 100,000 usage cycles.

Table 3 provides a descriptive summary of the failure times which have negative kurtosis. This fact can justify distributions with heavier tails required to be used to model these data. It can be seen that the mean value is smaller than the median and negative skewness, thus indicating a negative asymmetric distribution.

Table 3. Descriptive analysis for voltage data.

Data	n	Mean	Median	Mode	s.d.	Skewness	Kurtosis	Min.	Max.
Times of failure	30	177.03	196.5	300	114.99	-0.2992	-1.6087	2	300

We fit the Rayleigh distribution ($\alpha = \beta = 1$) to find initial values for θ and δ. All computations are done through the `NLMIXED` subroutine in `SAS`. Table 4 lists the MLEs (their standard errors in parentheses), and the previous measures, which reveal that the GOLLG distribution can be chosen as the best model.

The likelihood ratio (LR) statistics in Table 5 indicate that the GOLLGR distribution is the best model among the others. The histogram and the plots of the estimated densities in Figure 4(a), and those of the empirical and estimated survival functions in Figure 4(b) support the previous conclusion.

We note extremum problems (see Figure 4(a)) in the histogram of these data, i.e., there are two peaks (two modes) at the beginning and end of the experiment. Also, from Figure 4(a), we can clearly note that the proposed distribution fits well to these two modes.

5.2 Application 2: COVID-19 data

The second application refers to lifetimes of individuals diagnosed with COVID-19 (Coronavirus Disease 1999) (Galvão and Roncalli, 2021). Since it was declared an international health emergency, many studies have been conducted to obtain information about the clinical, epidemiological and prognostic aspects of the disease; see, for example, Cordeiro et al. (2021a), Cordeiro et al. (2021b) and Marinho et al. (2021).

In Brazil, the epidemiological data are disclosed by the Health Information System (available in: https://opendatasus.saude.gov.br/en/dataset/srag-2021-e-2022. In this analysis, we work with the `gamlss` package of `R`. The codes and dataset used here are available at: https://github.com/gabrielamrodrigues/GOLLGR.

Table 4. Findings for voltage data.

Model	α	β	δ	θ	AIC	CAIC	BIC
GOLLGR	0.0970	0.1663	30.5610	0.0008	353.6	355.2	359.2
	(0.0185)	(0.0517)	(0.0203)	(0.00002)			
OLLGR	0.0246	1	37.0100	0.0010	383.2	384.1	387.4
	(0.0041)		(0.0070)	(0.00006)			
EGR	1	0.0133	35.3427	0.00019	366.5	367.4	370.7
		(0.0014)	(0.0051)	(0.00003)			
GR	1	1	-0.5079	0.000011	368.4	368.8	371.2
			(0.1147)	(5.1E-6)			
	λ		0				
Rayleigh	210.06				380.1	380.2	381.5
	(19.1758)						
	λ		1/2				
Maxwell	121.28				405.2	405.3	406.6
	(9.0394)						
	τ	n					
Chi-square	211.74	0.9841			368.4	368.8	371.2
	(35.5963)	(0.2095)					
	σ		-1/2				
Half-normal	420.11				366.4	366.5	367.8
	(54.2363)						

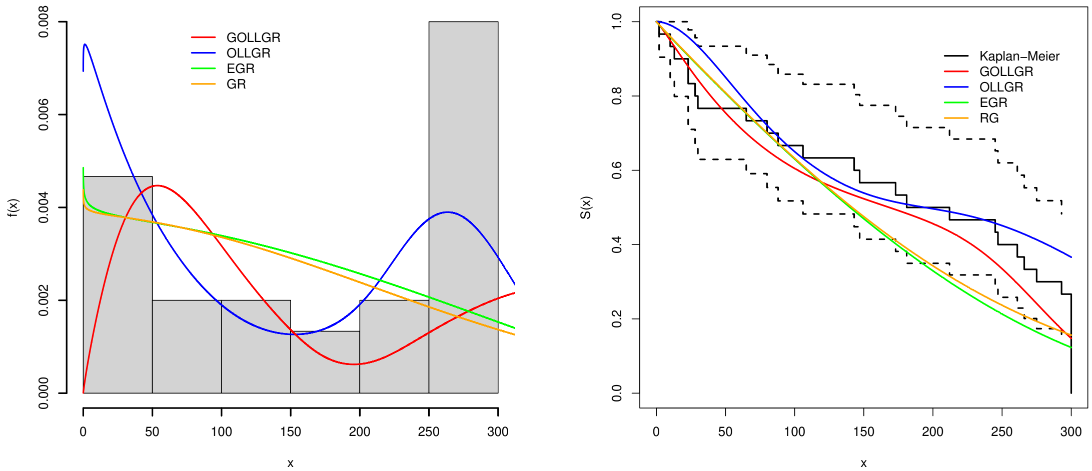
Table 5. LR statistics for voltage data.

Model	Hypotheses	LR statistic	p-value
GOLLGR vs OLLGR	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	23.2	<0.00001
GOLLGR vs EGR	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	6.5	0.0001
GOLLGR vs GR	$H_0 : \beta = \alpha = 1$ vs $H_1 : H_0$ is false	10.4	<0.00001

In this study, 881 patients infected by the virus are considered, confirmed by the RT-PCR test method. The participants consisted of hospitalized patients and outpatients living in the city of Campinas (Brazil) in January and February 2021. The survival consisted of the interval between the first symptoms until the date of death due to COVID-19 (failure). Deaths due to other causes or after the is 73.6%. Equation (3.24) is considered with factors associated with the highest risk of death. The results are compared with the OLLGR, EGR and GR sub-regressions.

The following variables were considered for each patient ($i = 1, 2, \dots, 881$):

- x_i: time until death due to COVID-19 (in days);
- cens$_i$: censoring indicator (0 = censored, 1 = observed lifetime);
- v_{i1} : age (in years);
- v_{i2} : diabetes mellitus (0= no or not reported, 1= yes).



(a) (b)

Figure 4. Estimated densities (a); and Empirical and estimated survival functions (b) for voltage data.

The total number of patients suffering from the comorbidity diabetes was 264 (29.97%), among whom 104 (39.39%) died. In turn, of the 617 patients (70.03%) without the disease or who did not report it, 128 (20.75%) died. Figure 5 presents the Kaplan-Meier survival curve, showing the greater risk of death among patients suffering from diabetes.

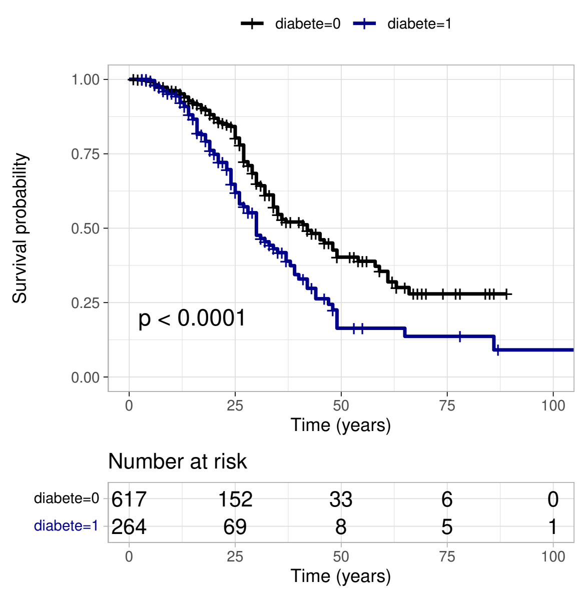


Figure 5. Kaplan-Meier survival curve for the variable diabetes mellitus (1 = yes, 0 = no or not informed).

Table 6 reveals the descriptive statistics according to the diabetes variable. We can verify asymmetry and positive kurtosis for the presence or absence of the disease. Figure 6 displays the histogram of the age variable which indicates a higher frequency of hospitalized patients aged between 55 and 75 years.

Table 6. Descriptive analysis for COVID-19 data.

Diabetes	n	Mean	Median	s.d.	Skewness	Kurtosis	Min.	Max.
0	617	20.49	16.00	14.02	2.08	8.10	1.00	89.00
1	264	21.56	18.00	14.19	2.40	11.82	3.00	108.00

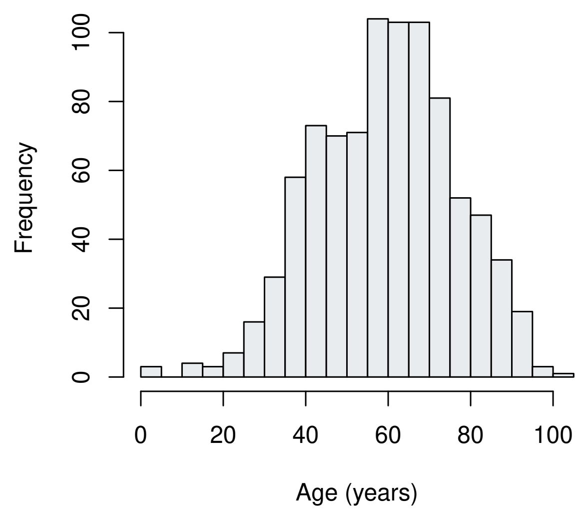


Figure 6. Histogram of the covariate "age".

The covariates are related to the GOLLGR model and its submodels according to the systematic components:

$$\delta_i = \exp(\lambda_{10} + \lambda_{11} v_{1i} + \lambda_{12} v_{2i}) - 1, \quad \theta_i = \exp(\lambda_{20} + \lambda_{21} v_{1i} + \lambda_{22} v_{2i}). \tag{5.30}$$

Note that the Maxwell, Rayleigh and Half-normal models have one positive scale parameter. So, we use the logarithm link function to relate the covariates, just like for δ in (5.30). For the Weibull and Inverse Gamma distributions that have two positive parameters, we adopt the logarithm function to link the covariates to the two parameters.

The statistics in Table 7 support that the GOLLGR regression can be chosen as the best model. Further, the LR statistics in Table 8 indicate that the wider regression yields the best fit. Table 9 reports the MLEs (SEs in parentheses) from the fitted GOLLGR regression.

Table 7. Findings from the fitted regressions to COVID-19 data.

Model	AIC	BIC	CAIC
GOLLGR	2222.54	2260.78	2237.66
OLLGR	2237.73	2271.19	2262.63
EGR	2240.17	2273.63	2265.07
GR	2238.66	2267.34	2273.34
Maxwell	2282.49	2317.18	2299.84
Rayleigh	2236.72	2271.40	2254.06
Half-normal	2342.07	2376.76	2359.41
Weibull	2236.85	2265.54	2271.54
Inverse gamma	2240.60	2269.28	2275.28

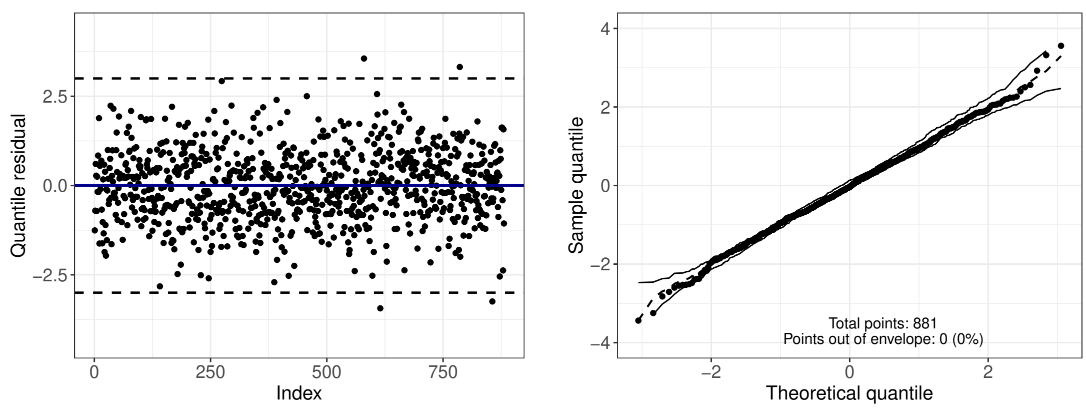
Table 8. LR statistics for COVID-19 data.

Model	Hypotheses	LR statistic	p-value
GOLLGR vs OLLGR	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	17.1915	<0.00001
GOLLGR vs EGR	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	19.6298	<0.00001
GOLLGR vs GR	$H_0 : \beta = \alpha = 1$ vs $H_1 : H_0$ is false	20.1202	<0.00001

Table 9. Results from the fitted GOLLGR regressions to COVID-19 data.

	MLE	SE	p-value
λ_{10}	-0.3599	0.0289	<0.0001
λ_{11}	-0.0028	0.0006	<0.0001
λ_{12}	0.0468	0.0359	0.1920
λ_{20}	-10.1148	0.1176	<0.0001
λ_{21}	0.0480	0.0022	<0.0001
λ_{22}	0.3331	0.0774	<0.0001
$\log(\alpha)$	-0.9748	0.0130	
$\log(\beta)$	2.0127	0.0130	

Figure 7 provides the plots of the quantile residuals (qr_i's) (3.27). The residual index plot (Figure 7(a)) reveals that the qrs have a random behavior and that only four observations are outside the $[-3, 3]$ range. The normal probability plot with simulated envelope for the $qr's$ (Figure 7(b)) indicates that the residuals follow approximately a normal distribution and does not show any points outside the envelope, which support the fitted regression. Thus, there is no evidence against the GOLLGR regression assumptions.



(a) (b)

Figure 7. Index plot of the qr_i's (a), and Normal probability plots with simulated envelope of the qr_i's (b) for COVID-19 data.

We conclude from Table 9 that:

- The covariate age is significant (λ_{11}), thus indicating that older individuals tend to have a progressively shorter period until death due to this coronavirus. Further, the age is significant for the variability of survival times (λ_{21}).
- There is no significant difference between patients with or without diabetes mellitus in terms of time to death due to COVID-19 (λ_{12}). On the other hand, there is a significant difference in relation to the variability of the time until death due to the coronavirus when

comparing patients with or without diabetes mellitus (λ_{22}). So, the patients with diabetes mellitus have greater variability regarding the time of death from COVID-19 compared to patients who do not have diabetes mellitus.

6. Conclusions, limitations, and future research

We introduced the generalized odd log-logistic generalized Rayleigh distribution and studied some of its structural properties. The new distribution generalizes some models studied recently in the literature. It is important for the analysis of asymmetric and bimodal lifetime data. The estimation of parameters is approached by the maximum likelihood method and the observed information matrix is derived. We proposed a new regression model based on the generalized odd log-logistic generalized Rayleigh distribution for censored data. The usefulness of the new models is illustrated by means of two real data sets via classical criterion. Future research using the proposed distribution can be directed to study other estimation methods such as Bayesian and Jackknife. Regarding the generalized odd log-logistic generalized Rayleigh regression model, the diagnostic analysis can be discussed to study its robustness in relation to possible influential points and construct regression models with random effects based on the new distribution. Finally, we can extend the proposed regression model to the multivariate case.

Appendix. Proofs

Proof of Proposition 2.3 By differentiating Equation (2.12) with respect to x, we obtain

$$f(x;\alpha,\beta,\delta,\theta) = f_Y(T(x);1,\alpha)T'(x), \quad Y \sim \mathrm{LL}(1,\alpha). \tag{6.31}$$

Then, the derivative of $f(x;\alpha,\beta,\delta,\theta)$ is

$$f'(x;\alpha,\beta,\delta,\theta) = f'_Y(T(x);1,\alpha)[T'(x)]^2 + f_Y(T(x);1,\alpha)T''(x). \tag{6.32}$$

Since

$$f'_Y(t;1,\alpha) = -f_Y(t;1,\alpha)r[t] \quad \text{with } r[t] = \frac{t^\alpha + \alpha(t^\alpha - 1) + 1}{t(t^\alpha + 1)},$$

Equation (6.32) can be expressed as

$$f'(x;\alpha,\beta,\delta,\theta) = f_Y(T(x);1,\alpha)\{T''(x) - r[T(x)][T'(x)]^2\},$$

where

$$T'(x) = \frac{\beta g_{\mathrm{GR}}(x;\delta,\theta)T(x)}{G_{\mathrm{GR}}(x;\delta,\theta)[1 - G_{\mathrm{GR}}(x;\delta,\theta)^\beta]},$$

$$T''(x) = T'(x)\left\{\frac{g'_{\mathrm{GR}}(x;\delta,\theta)}{g_{\mathrm{GR}}(x;\delta,\theta)} + g_{\mathrm{GR}}(x;\delta,\theta)\frac{(\beta+1)G_{\mathrm{GR}}(x;\delta,\theta)^\beta + \beta - 1}{G_{\mathrm{GR}}(x;\delta,\theta)[1 - G_{\mathrm{GR}}(x;\delta,\theta)^\beta]}\right\},$$

and

$$\frac{g'_{\mathrm{GR}}(x;\delta,\theta)}{g_{\mathrm{GR}}(x;\delta,\theta)} = \frac{2\delta + 2\theta x^2 - 1}{x}.$$

Then,

$$f'(x;\alpha,\beta,\delta,\theta) = f_Y(T(x);1,\alpha)T'(x)\,\frac{g'_{\mathrm{GR}}(x;\delta,\theta)}{g_{\mathrm{GR}}(x;\delta,\theta)}$$

$$+ f_Y(T(x);1,\alpha)T'(x)\,\frac{g_{\mathrm{GR}}(x;\delta,\theta)}{G_{\mathrm{GR}}(x;\delta,\theta)}\,[T(x)+1]\left\{(\beta+1)G_{\mathrm{GR}}(x;\delta,\theta)^{\beta} - \alpha\beta\left[\frac{T(x)^{\alpha}-1}{T(x)^{\alpha}+1}\right] - 1\right\}.$$

Equation (6.31) gives $f_Y(T(x);1,\alpha)T'(x) = f(x;\alpha,\beta,\delta,\theta)$, which is a positive function. Hence, any critical point of the PDF of X satisfies the non-linear equation:

$$\frac{g'_{\mathrm{GR}}(x;\delta,\theta)}{g_{\mathrm{GR}}(x;\delta,\theta)} + \frac{g_{\mathrm{GR}}(x;\delta,\theta)}{G_{\mathrm{GR}}(x;\delta,\theta)}\,[T(x)+1]\left\{(\beta+1)G_{\mathrm{GR}}(x;\delta,\theta)^{\beta} - \alpha\beta\left[\frac{T(x)^{\alpha}-1}{T(x)^{\alpha}+1}\right] - 1\right\} = 0.$$

The previous equation provides the required result.

Proof of Proposition 2.5 If $Y \sim \mathrm{LL}(1,\alpha)$, by Equation (2.12), $F(x;\alpha,\beta,\delta,\theta) = \mathrm{P}(Y \leq T(x))$. Moreover, it is well-known that $\mathrm{P}(Y \leq y) = y^{\alpha}/(1+y^{\alpha})$. Then, from the definition given in Equation (2.13) of T, we have (for any $t > 0$),

$$\frac{\mathrm{e}^{-tx}}{1 - F(x;\alpha,\beta,\delta,\theta)} = \frac{\mathrm{e}^{-tx}}{1 - \mathrm{P}(Y \leq T(x))} = \mathrm{e}^{-tx}[1 + T(x)^{\alpha}]$$

$$\geq \mathrm{e}^{-tx}T(x)^{\alpha} = \frac{\mathrm{e}^{-tx}\,G_{\mathrm{GR}}(x;\delta,\theta)^{\alpha\beta}}{\big[1 - G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}\big]^{\alpha}}. \tag{6.33}$$

The L'Hospital's rule yields

$$\lim_{x\to\infty} \frac{\mathrm{e}^{-tx}}{\big[1 - G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}\big]^{\alpha}} = \lim_{x\to\infty} \frac{t\big[\frac{\mathrm{e}^{-tx}}{g_{\mathrm{GR}}(x;\delta,\theta)}\big]}{\alpha\beta\big[1 - G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}\big]^{\alpha-1}G_{\mathrm{GR}}(x;\delta,\theta)^{\beta-1}}. \tag{6.34}$$

Since (for $\alpha \geq 1$),

$$\lim_{x\to\infty} \frac{\mathrm{e}^{-tx}}{g_{\mathrm{GR}}(x;\delta,\theta)} = \left[\frac{2\theta^{\delta+1}}{\Gamma(\delta+1)}\,x^{2\delta+1}\,\mathrm{e}^{-\theta x^2 + tx}\right]^{-1} = \infty$$

and $\lim_{x\to\infty} G_{\mathrm{GR}}(x;\delta,\theta) = 1$. We have from Equation (6.34)

$$\lim_{x\to\infty} \frac{\mathrm{e}^{-tx}}{\big[1 - G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}\big]^{\alpha}} = \infty.$$

Hence, by taking $x \to \infty$ for both sides of inequality given in Equation (6.33), the limit in Equation (2.16) follows.

Proof of Proposition 2.6 By inequality in Equation (6.33), it is enough to prove

$$\lim_{x\to\infty} \frac{\mathrm{e}^{-tx}}{\left[1-G_{\mathrm{GR}}(x;\delta,\theta)^{\beta}\right]^{\alpha}} = \lim_{x\to\infty} \frac{1}{\mathrm{e}^{tx}\left[1-\gamma_1(\delta+1,\theta x^2)^{\beta}\right]^{\alpha}} = \infty. \tag{6.35}$$

Indeed, by using the C_p inequality:

$$\forall x, y \geq 0;\ (x+y)^p \leq C_p(x^p+y^p), \quad \text{where } p>0 \text{ and } C_p = \max\{1, 2^{p-1}\};$$

we have (for $0<\beta\leq 1$)

$$1-\gamma_1(\delta+1,\theta x^2)^{\beta} \leq [1-\gamma_1(\delta+1,\theta x^2)]^{\beta} = \Gamma_1(\delta+1,\theta x^2)^{\beta}, \tag{6.36}$$

where $\Gamma_1(p,x) = \Gamma(p)^{-1}\int_x^{\infty} w^{p-1}\,\mathrm{e}^{-w}\mathrm{d}w$ is the upper incomplete gamma function ratio.

By using the inequality of Natalini and Palumbo (2000): for $a>1$, $B>1$ and $x>B(a-1)/(B-1)$,

$$\Gamma(a,x) < B\,x^{a-1}\,\mathrm{e}^{-x};$$

we have (for $x>\sqrt{B\delta/[\theta(B-1)]}$ and $\delta>0$)

$$\Gamma_1(\delta+1,\theta x^2)^{\beta} < B^{\beta}\,\theta^{\beta\delta}\,\Gamma(\delta+1)^{-\beta}x^{2\beta\delta}\,\mathrm{e}^{-\beta\theta x^2}. \tag{6.37}$$

By combining Equations (6.36) and (6.37), we obtain (for any $x>\sqrt{B\delta/[\theta(B-1)]}$)

$$\mathrm{e}^{tx}\left[1-\gamma_1(\delta+1,\theta x^2)^{\beta}\right]^{\alpha} < B^{\alpha\beta}\theta^{\alpha\beta\delta}\Gamma(\delta+1)^{-\alpha\beta}x^{2\alpha\beta\delta}\mathrm{e}^{-\alpha\beta\theta x^2+tx}.$$

Letting $x\to\infty$ in the above inequality, we have $\mathrm{e}^{tx}\left[1-\gamma_1(\delta+1,\theta x^2)^{\beta}\right]^{\alpha}$ tends to zero, proving the limit in Equation (6.35). Thus, we complete the proof.

Author contributions Conceptualization, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; methodology, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; software, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; validation, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; formal analysis, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; investigation, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; data curation, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; writing-original draft preparation, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; writing-review and editing, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; visualization, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.; supervision, G.M.C., G.M.R., E.M.M.O., L.H.S., R.V.. All authors have read and agreed to the published version of the manuscript.

Acknowledgements The authors would like to thank the reviewers for all useful and helpful comments on our manuscript. All remaining errors are our own.

Funding Roberto Vila gratefully acknowledges financial support from CNPq, CAPES, and FAP-DF, Brazil, and Gabriela Rodrigues is grateful for the financial support of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brasil (CAPES) (Finance Code 001).

Conflicts of interest The authors declare no conflict of interest.

References

Alexander, C., Cordeiro, G.M., Ortega, E.M.M., and Sarabia, J.M., 2012. Generalized beta-generated distributions. Computational Statistics and Data Analysis, 56, 1880-1897.

Alizadeh, M., MirMostafaee, S.M.T.K, and Ghosh, I., 2017. A new extension of power Lindley distribution for analyzing bimodal data. Chilean Journal of Statistics, 8, 67–86.

Barranco-Chamorro, I., Gómez, Y.M., Astorga, J.M., and Gomes, H.W., 2021. A generalized Rayleigh family of distributions based on the modified Slash model. Symmetry, 13, 1-18.

Cole, T. J., and Green, P. J. 1992. Smoothing reference centile curves: the LMS method and penalized likelihood. Statistics in medicine, 11, 1305-1319.

Cordeiro, G.M., and de Castro, M., 2011. A new family of generalized distributions. Journal of Statistical Computation and Simulation, 81, 883-898.

Cordeiro, G.M., Cristino, C.T., Hashimoto, E.M., and Ortega, E.M.M., 2011. The beta generalized Rayleigh distribution with applications to lifetime data. Statistical Papers, 54, 133-161.

Cordeiro, G.M., Alizadeh, M., Ozel, G., Hosseini, B., Ortega, E.M.M., and Altun, E., 2017. The generalized odd log-logistic family of distributions: properties, regression models and applications. Journal of Statistical Computation and Simulation, 87, 908-932.

Cordeiro, G.M., Afify, A.Z., Ortega, E.M.M., Suzuki, A.K., and Mead, M.E., 2018. The odd Lomax generator of distributions: Properties, estimation and applications. Journal of Computational and Applied Mathematics, 347, 222-237.

Cordeiro, G.M., Ramires, T.G., Ortega, E.M.M., and Pescim, R.R., 2020. The extended beta generator of distributions: Properties and applications. Studia Scientiarum Mathematicarum Hungarica, 57, 444-464.

Cordeiro, G.M., Figueiredo, D., Silva, L., Ortega, E.M.M., and Prataviera, F., 2021a. Explaining COVID-19 mortality rates in the first wave in Europe. Model Assisted Statistics and Applications, 16, 211-221.

Cordeiro, G.M., de Azevedo Cysneiro, F.J., and Cabral, P.C., 2021b. Estatísticas Básicas e Modelagem de Regressão das taxas de mortalidade por COVID-19 nos Estados Brasileiros Basic Statistics and Regression modeling of COVID-19 mortality rates in Brazilian States. Brazilian Journal of Development, 7, 117735-117749.

Dunn, P., and Smyth, G., 1996. Randomized quantile residuals. Journal of Computational and Graphical Statistics, 5, 236-244.

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

Galvão, M.H.R., and Roncalli, A.G., 2021. Factors associated with increased risk of death from COVID-19: survival analysis based on confirmed cases. Brazilian Journal of Epidemiology, 23, 1-10.

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

Gomes, A.E., da-Silva, C.Q., Cordeiro, G.M., and Ortega, E.M.M., 2014. A new lifetime model: the Kumaraswamy generalized Rayleigh distribution. Journal of Statistical Computation and Simulation, 84, 290-309.

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

Gupta, R.C., and Gupta, R.D., 2007. Proportional reversed hazard rate model and its applications. Journal of Statistical Planning and Inference, 137, 3525-3536.

Hashimoto, E.M., Silva, G.O., Ortega, E.M.M., and Cordeiro, G.M., 2019. Log-Burr XII gamma-Weibull regression model with andom effects and censored data. Journal of Statisyical Theory and Practice, 13, 1-18.

Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994. Continuous univariate distributions. Wiley, New York.

Marinho, P.R.D., Cordeiro, G.M., Coelho, H.F., and Brandão, S.C.S., 2021. Covid-19 in Brazil: A sad scenario. Cytokine & growth factor reviews, 58, 51-54.

Meeker, W.Q., and Escobar, L.A., 1998. Statistical Methods for Reliability Data. Wiley, New York.

Nadarajah, S., Cordeiro, G.M., and Ortega, E.M.M., 2015a. The Zografos-Balakrishnan-G family of distributions: Mathematical properties and applications. Communication in Statistics - Theory and Methods, 44, 186-215.

Nadarajah, S., Cordeiro, G.M., and Ortega, E.M.M., 2015b. The exponentiated G geometric family of distributions. Journal of Statistical Computation and Simulation, 85, 1634-1650.

Naqash, S., Ahmad, S.P., and Ahmed, A., 2016. Bayesian analysis for generalized Rayleigh distribution. Mathematical Theory and Modeling, 6, 85-96.

Natalini, P., and Palumbo, B., 2000. Inequalities for the incomplete gamma function. Mathematical Inequalities & Applications, 3, 69-77.

Pescim, R.R., Cordeiro, G.M., Demétrio, C.G.B., Ortega, E.M.M., and Nadarajah, S., 2012. The new class of Kummer beta generalized distributions. Statistics and Operations Research Transactions, 36, 153-180.

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

Prataviera, F., Loibel, S.M.C., Grego, K.F., Ortega, E.M.M., and Cordeiro, G.M., 2020. Modelling non-proportional hazard for survival data with different systematic components. Environmental and Ecological Statistics, 27, 467-489.

Prudnikov, A.P., Brychkov, Y.A., and Marichev, O.I., 1986. Integrals and series. New York, Gordon and Breach Science Publishers.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2012. R Foundation for Statistical Computing, Vienna, Austria.

Ribeiro-Reis, L.D., Cordeiro, G.M., and de Santana e Silvas, J.J., 2022. The Mc-Donald Chen distribution: A new bimodal distribution with properties and applications. Chilean Journal of Statistics, 13, 91–111.

Rigby, R.A., and Stasinopoulos, D. M., 2005. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society C, 54, 507-554.

Silva, G.O., Cordeiro, G.M., and Ortega, E.M.M., 2020. Surviving and non surviving fraction regression models based on the beta modified Weibull distribution. Model Assisted Statistics and Applications, 15, 111-126.

Shen, Z., Alrumayh, A., Ahmad, Z., Abu-Shanab, R., Al-Mutairi, M., and Aldallal, R., 2022. A new generalized rayleigh distribution with analysis to big data of an online community. Alexandria Engineering Journal, 61, 11523-11535.

Vasconcelos, J.C.S., Cordeiro, G.M., Ortega, E.M.M., and Ribeiro, J.G., 2021. A regression model for extreme events and the presence of bimodality with application in energy generation data. IET Renewable Power Generation, 48, 452-461.

Vila, R., Ferreira, L., Saulo, H., Prataviera, F., and Ortega, E.M.M., 2020. A bimodal gamma distribution: Properties, regression model and applications. Statistics, 54, 469-493.

Vila, R., and Çankaya, M. N., 2021. A Bimodal Weibull Distribution: Properties and Inference. Journal of Applied Statistics, 49, 3044-3062.

Vila, R., Saulo, H., and Roldan, J., 2021. On some properties of the bimodal normal distribution and its bivariate version. Chilean Journal of Statistics, 12, 125–144.

Vodă, V.G., 1976. Inferential procedures on a generalized Rayleigh variate, I. Applications of Mathematics, 21, 395-412.