Reparametrized Exponentiated-G Distribution Family: Properties and Applications

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

 \bullet The main objective of this paper is to introduce the reparametrized exponentiated-G distribution family and outline its general mathematical properties. In contrast to the exponentiated-G distribution, this reparametrized family is characterized by a parameter η_{τ} , which offers an interpretation based on the τ th quantile of the distribution. To illustrate the derivation of submodels within the reparametrized exponentiated-G distribution family, we employed the reparameterized Weibull model as baseline, resulting in the reparameterized exponentiated Weibull model. Through Monte Carlo simulations implemented using the R language, we conducted a numerical evaluation of the performance of maximum likelihood estimators and their associated asymptotic confidence intervals. Furthermore, we applied the proposed distribution family to a real dataset to provide additional illustrations.

Keywords:

 Baseline distribution; Maximum likelihood estimator; Monte Carlo simulation; Quantile; Reparametrized exponentiated-G distribution.

AMS Subject Classification:

 $62J20, 62G08.$

 $\overline{\boxtimes}$ Corresponding author

1. INTRODUCTION

Let Y be a random variable with an arbitrary (continuous) cumulative distribution function (CDF), denoted as $G(y)$. If this CDF is raised to the exponent $\alpha(\alpha > 0)$, denoted as $G^{\alpha}(y)$, then we obtain a exponentiated-G distribution with parameter α [\(Lehmann,](#page-18-0) [1953\)](#page-18-0). To simplify notation, we will refer to this distribution as α -G. The probability density function (PDF) of the α -G distribution is given by

$$
f(y) = \alpha g(y) G^{\alpha - 1}(y),
$$

where, $g(y) = dG(y)/dy$ corresponds the PDF of $G(\cdot)$, called the baseline distribution. It is important to observe that when the parameter α takes on the value 1, we obtain, as a particular case, the baseline model of the α -G distribution.

Several probabilistic models within the α -G family of distributions have been proposed in the literature. For instance, the exponentiated exponential (EE) distribution, introduced in [Gupta and Kundu](#page-18-1) [\(2001\)](#page-18-1), extends the exponential distribution and finds application in survival data analysis within medical and engineering domains. Similarly, the exponentiated Weibull (EW) distribution, proposed in [Mudholkar and Srivastava](#page-18-2) [\(1993\)](#page-18-2), extends the Weibull distribution and offers versatility in representing failure rate functions. Note that, the EE distribution becomes a special case of the EW model when the shape parameter κ equals 1. Additionally, the exponentiated Fréchet (EF) [\(Nadarajah and Kotz,](#page-18-3) [2003\)](#page-18-3), exponentiated Rayleigh (ER)[\(Kundu and Raqab,](#page-18-4) [2005\)](#page-18-4), exponentiated Birnbaum-Saunders (EBS) [\(Cordeiro](#page-17-0) [et al.,](#page-17-0) [2013\)](#page-17-0), exponentiated Lomax (EL) [\(Abdul-Moniem and Abdel-Hameed,](#page-17-1) [2012\)](#page-17-1), exponentiated Gompertz (EG) [\(A. El-Gohary,](#page-17-2) [2013\)](#page-17-2), exponentiated Gumbel (EGU) [\(Nadarajah,](#page-18-5) [2006\)](#page-18-5), exponentiated log-logistic (ELL) [\(Rosaiah et al.,](#page-18-6) [2006\)](#page-18-6) and exponentiated generalized Birnbaum-Saunders (EGBS) [\(Gallardo and Santos-Neto,](#page-18-7) [2021\)](#page-18-7) distribution have been introduced to generalize their respective baseline distributions and explore their mathematical properties. These models find applications in various fields such as reliability analysis, survival studies, and statistical modeling. Furthermore, studies have investigated their theoretical properties and applied them to real-world datasets.

In this paper, we introduce the reparametrized α -G distribution family and outline its general mathematical properties. This family proves valuable for modeling distribution quantiles, as it is indexed by a parameter η_{τ} , representing the quantile of interest. Our proposal will generalize a wide range of existing distribution families, thereby providing significant flexibility for modeling data.

This article is structured as follows: Section [2](#page-2-0) introduces the PDF and CDF of the reparametrized α -G distribution family, along with its submodels derived by changing the baseline distribution $G(y)$. We also discuss the survival function $S(y)$, failure rate function $h(y)$, and cumulative failure rate function $H(y)$ for this family. Moreover, we outline the parameter estimation process using maximum likelihood and present general mathematical properties applicable to various submodels by selecting different base distributions G. Section [3](#page-6-0) reviews the Weibull model and its key mathematical properties, including a reparameterization based on the median. In Section [4,](#page-6-1) we introduce a specific case of the reparametrized α -G distribution family, known as the reparameterized exponentiated Weibull model (REW), and detail the parameter estimation procedure. Section [5](#page-8-0) presents Monte Carlo simulations

to assess the performance of maximum likelihood estimators for the REW model and calculate descriptive statistics such as variance, coefficient of skewness, coefficient of kurtosis, percentage relative bias in absolute value, and mean squared error. We also evaluate asymptotic confidence intervals numerically. Section [6](#page-13-0) demonstrates the application of the theory to real-world data. Finally, Section [7](#page-15-0) offers concluding remarks.

2. REPARAMETRIZED α -G DISTRIBUTION FAMILY

If the α -G distribution is indexed by its median, η , and for any quantile τ ($0 < \tau < 1$), we can rewrite the parameter α as follows [\(Gallardo and Santos-Neto,](#page-18-7) [2021\)](#page-18-7):

$$
G^{\alpha}(\eta) = \tau
$$

$$
\left(\frac{1}{2}\right)^{\alpha} = \tau
$$

$$
\log\left(\frac{1}{2}\right)^{\alpha} = \log(\tau)
$$

$$
-\alpha \log(2) = \log(\tau)
$$

$$
\alpha = -\frac{\log(\tau)}{\log(2)}.
$$

where $G^{\alpha}(\eta)$ represents the CDF of the α -G distribution, and $\alpha_{\tau} = \alpha$ becomes a known fixed value depending on the quantile of interest. Now, the parameter $\eta = \eta_{\tau}$ represents the τ th quantile of this distribution family.

The PDF and CDF of the reparametrized α_{τ} -G distribution family are given respectively by:

(2.1)
$$
f(y) = \alpha_{\tau} g(y) G^{\alpha_{\tau}-1}(y),
$$

and

$$
(2.2) \t\t F(y) = G^{\alpha_{\tau}}(y).
$$

Additionally, if the baseline distribution is not indexed by the median, we need to reparameterize the baseline distribution for this to occur. Note that the reparametrized α_{τ} -G distribution family includes all submodels of the family of α -G distributions that are indexed by the median or have the median in closed form.

Furthermore, the survival function (SF), the failure rate function (FRT), and the cumulative failure rate function (CFR) for the reparametrized α_{τ} -G distribution family are given respectively by:

$$
S(y) = 1 - F(y) = 1 - G^{\alpha_{\tau}}(y),
$$

$$
h(y) = \alpha_{\tau} g(y) G^{\alpha_{\tau}-1}(y) [1 - G^{\alpha_{\tau}}(y)]^{-1},
$$

and

$$
H(y) = -\log[1 - G^{\alpha_{\tau}}(y)].
$$

Table [1](#page-5-0) presents some submodels of the reparametrized α_{τ} -G distribution family. These submodels are derived from equations [\(2.1\)](#page-2-1) and [\(2.2\)](#page-2-2) by substituting $G(y)$ and $g(y)$, which are the CDF and PDF, respectively, of the baseline distribution. For instance, the REW model is derived using the reparameterized Weibull baseline distribution.

2.1. Maximum Likelihood Method

Let Y_1, \ldots, Y_n denote a random sample of size n from the reparametrized α_{τ} -G distribution family, and let θ represent the parameter vector of dimension q, where $\eta_{\tau} \in \theta$. The likelihood function for this family can be expressed as:

$$
L(\boldsymbol{\theta}) = \prod_{i=1}^n f(y_i; \boldsymbol{\theta}) = \prod_{i=1}^n \alpha_{\tau} g(y_i; \boldsymbol{\theta}) G^{\alpha_{\tau}-1}(y_i; \boldsymbol{\theta})
$$

$$
= (\alpha_{\tau})^n \prod_{i=1}^n g(y_i; \boldsymbol{\theta}) \prod_{i=1}^n G^{\alpha_{\tau}-1}(y_i; \boldsymbol{\theta}).
$$

The logarithm of the likelihood function, $\ell(\theta) = \log [L(\theta)]$, is given by:

$$
\ell(\boldsymbol{\theta}) = n \log(\alpha_{\tau}) + \sum_{i=1}^{n} \log \left[g(y_i; \boldsymbol{\theta}) \right] + (\alpha_{\tau} - 1) \sum_{i=1}^{n} \log \left[G(y_i; \boldsymbol{\theta}) \right].
$$

The score vector is defined by:

$$
U(\boldsymbol{\theta}) = \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^{\top},
$$

and its elements, for $j = 1, \ldots, q$, are:

$$
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j} = \sum_{i=1}^n \frac{1}{g(y_i; \boldsymbol{\theta})} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \theta_j} + (\alpha_{\tau} - 1) \sum_{i=1}^n \frac{1}{G(y_i; \boldsymbol{\theta})} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \theta_j}.
$$

The maximum likelihood estimators (MLEs) of θ are solutions to the system of equations $U(\theta) = 0$ However, in this case, the system of equations does not have a closed-form solution. Therefore, the MLEs must be obtained using some numerical method. Here, we employ the quasi-Newton method, specifically BFGS, as described in [Press et al.](#page-18-8) [\(1992\)](#page-18-8). In the R programming language [\(Team,](#page-18-9) [2020\)](#page-18-9), we can utilize the BFGS method through the optim() function.

The corresponding Fisher information matrix is a $q \times q$ matrix denoted by $I(\theta) =$ $\mathbb{E}(U(\theta)U(\theta)^{\top})$. If certain regularity conditions [\(Cox and Hinkley,](#page-18-10) [1974\)](#page-18-10) for the likelihood function are satisfied, then $\mathbb{E}(U(\theta)) = \mathbf{0}$ and $I(\theta) = -\mathbb{E}(\partial^2 \ell(\theta)/\partial \theta \partial \theta^\top)$. Therefore, its elements can be expressed as:

$$
\mathbb{E}\left(-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right) = n \mathbb{E}\left(\frac{1}{g^2(Y;\boldsymbol{\theta})} \frac{\partial g(Y;\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial g(Y;\boldsymbol{\theta})}{\partial \theta_j}\right) + n(\alpha_{\tau} - 1) \mathbb{E}\left(\frac{1}{G^2(Y;\boldsymbol{\theta})} \frac{\partial G(Y;\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial G(Y;\boldsymbol{\theta})}{\partial \theta_j}\right) - n \mathbb{E}\left(\frac{1}{g(Y;\boldsymbol{\theta})} \frac{\partial^2 g(Y;\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right) - n \mathbb{E}\left(\frac{1}{G(Y;\boldsymbol{\theta})} \frac{\partial^2 G(Y;\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right).
$$

According to [Lehmann and Casella](#page-18-11) [\(1998\)](#page-18-11), when certain regularity conditions are satisfied, the MLEs $\hat{\theta}$ are asymptotically distributed as a q-variate normal distribution. In other words: √

$$
\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\stackrel{a}{\sim}N_q\left(\boldsymbol{0},\boldsymbol{I}^{-1}(\boldsymbol{\theta})\right),
$$

where $\stackrel{a}{\sim}$ denotes asymptotically distributed. This result can be utilized to construct asymptotic confidence intervals and perform hypothesis tests for the parameters of the reparametrized α_{τ} -G distribution family. Additionally, when it is difficult to obtain the Fisher information matrix $I(\theta)$, we can substitute $I(\theta)$ with the observed Fisher information matrix, $-\ddot{L}(\theta)$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, where $\ddot{L} = (\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top)$.

The asymptotic confidence intervals with $(1 - \gamma) \times 100\%$ confidence level for the parameter θ_i are given by:

$$
\hat{\theta}_j \pm z_{(1-\gamma/2)}\sqrt{v_j},
$$

where $z_{(1-\gamma/2)}$ represents the $(1 - \gamma/2)$ th quantile of the standard normal distribution. The asymptotic variance v_j of $\hat{\theta}_j$ is the jth diagonal element of the matrix $I(\theta)^{-1}$ evaluated at $\hat{\theta}$.

3. BASELINE DISTRIBUTION FOR THE REW DISTRIBUTION

In this section, we introduce the baseline distribution for the REW distribution. The Weibull distribution, proposed by [Weibull](#page-18-12) [\(1939\)](#page-18-12), is a probabilistic model characterized by shape parameter ($\kappa > 0$) and scale parameter ($\beta > 0$). Its PDF is given by:

$$
g(y; \kappa, \beta) = \frac{\kappa}{\beta^{\kappa}} y^{(\kappa - 1)} \exp\left[-\left(\frac{y}{\beta}\right)^{\kappa}\right], \text{ for } y > 0.
$$

The cumulative distribution function is defined as:

$$
G(y; \kappa, \beta) = \begin{cases} 0, & \text{if } y \le 0; \\ 1 - \exp[-(y/\beta)^{\kappa}], & \text{if } y > 0. \end{cases}
$$

Additionally, its mean and variance are:

$$
\mathbb{E}(Y) = \beta \Gamma\left(\frac{1}{\kappa} + 1\right) \quad \text{and} \quad Var(Y) = \beta^2 \left[\Gamma\left(1 + \frac{2}{\kappa}\right) - \Gamma^2\left(\frac{1}{\kappa} + 1\right)\right],
$$

respectively. The median of the Weibull distribution is $\eta = \beta \log(2) \cdot 1/\kappa$.

Reparameterizing the Weibull distribution in terms of the median, we use a new parameterization where $\beta = \eta/[\log(2)]^{\frac{1}{\kappa}}$. Thus, the PDF of the reparameterized Weibull distribution is:

(3.1)
$$
g(y; \kappa, \eta) = \frac{\kappa \log(2) y^{(\kappa - 1)} 2^{-\left(\frac{y}{\eta}\right)^{\kappa}}}{\eta^{\kappa}},
$$

and its CDF is:

(3.2)
$$
G(y; \kappa, \eta) = \begin{cases} 0, & \text{if } y \le 0; \\ 1 - 2^{-\left(\frac{y}{\eta}\right)^{\kappa}}, & \text{if } y > 0. \end{cases}
$$

4. REW distribution

The PDF and CDF of the REW distribution are obtained by applying equations [\(3.1\)](#page-6-2) and [\(3.2\)](#page-6-3) to equations [\(2.1\)](#page-2-1) and [\(2.2\)](#page-2-2), respectively. Consequently,

$$
f(y; \kappa, \eta_\tau, \alpha_\tau) = \frac{\log(2) \alpha_\tau \kappa y^{\kappa - 1} 2^{-\frac{y^{\kappa}}{\eta_\tau^{\kappa}}}}{\eta_\tau^{\kappa}} \left(1 - 2^{-\frac{y^{\kappa}}{\eta_\tau^{\kappa}}}\right)^{\alpha_{\tau} - 1},
$$

and

$$
F(y; \kappa, \eta_{\tau}, \alpha_{\tau}) = \begin{cases} 0, & \text{if } y \leq 0; \\ \left(1 - 2^{-\frac{y^{\kappa}}{\eta_{\tau}^{\kappa}}}\right)^{\alpha_{\tau}}, & \text{if } y > 0, \end{cases}
$$

respectively.

Figure [1](#page-7-0) displays the density curves of the REW distribution for various values of the parameter $\kappa \in \{1.5, 2.5, 5\}$ and $\tau = \{0.25, 0.50, 0.75\}$, with $\eta_{\tau} = 1$. Note that the parameter κ controls the level of kurtosis of the distribution. The higher the value of κ , the flatter the curve becomes. Conversely, τ appears to influence the degree of asymmetry of the distribution. Increasing the value of τ enhances the asymmetry of the curve.

Figure 1: REW densities for different values of κ and τ with $\eta_{\tau} = 1$.

4.1. Maximum Likelihood Method

Let Y_1, \ldots, Y_n be a random sample of size n from the REW distribution with parameter vector $\boldsymbol{\theta} = (\kappa, \eta_{\tau}, \alpha_{\tau})^{\top}$, where α_{τ} is known. The likelihood function can be written as:

$$
L(\theta) = \frac{\left[\log(2)\right]^n (\alpha_\tau)^n \kappa^n (\prod_{i=1}^n y_i)^{k-1} 2^{-\frac{\sum_{i=1}^n y_i^k}{\eta_\tau^k}}}{\eta_\tau^{nk}} \prod_{i=1}^n \left(1 - 2^{-\frac{y_i^k}{\eta_\tau^k}}\right)^{\alpha_\tau - 1}
$$

The logarithm of the likelihood function $\ell(\kappa, \eta_\tau, \alpha_\tau) = \log [L(\kappa, \eta_\tau, \alpha_\tau)]$ is given by:

$$
\ell(\kappa, \eta_{\tau}, \alpha_{\tau} \mid y_i) = n \log \left[\log(2) \right] + n \log \left(\alpha_{\tau} \right) + n \log(\kappa) + (\kappa - 1) \sum_{i=1}^n \log \left(y_i \right)
$$

$$
- \frac{1}{\eta_{\tau}^{\kappa}} \sum_{i=1}^n y_i^{\kappa} \log(2) - n \kappa \log(\eta_{\tau}) + (\alpha_{\tau} - 1) \sum_{i=1}^n \log \left(1 - 2^{-\frac{y_i^{\kappa}}{\eta_{\tau}^{\kappa}}} \right).
$$

5. SIMULATION

In this Section, we present the analyses and numerical results of the Monte Carlo simulations. The simulation codes were written using the R programming language in its version 3.5.1 [\(Team,](#page-18-9) [2020\)](#page-18-9) available for free at <www.r-project.org/>. Additionally, the graphs were carried out using the R software. The R language was created by Ross Ihaka and Robert Gentleman. In the Monte Carlo simulations, the numerical method BFGS [\(Press](#page-18-8) [et al.,](#page-18-8) [1992\)](#page-18-8) and the optim function were used. According to [Cordeiro et al.](#page-18-13) [\(2011\)](#page-18-13) the initial values for the parameters can be obtained from the fit model baseline . Therefore, the initial values used for the shape parameter κ of the REW model were given by the estimates of the method of moments of the Weibull baseline distribution and for the parameter η_{τ} the initial values were given by the quantile of the model. The results are displayed in tables and figures. The measures used to evaluate the behavior of the point estimates obtained by the maximum likelihood method for the REW distribution were: CS_{κ} and $CS_{\eta_{\tau}}$ are coefficient of skewness for κ e η_{τ} , respectivelly, CK_{κ} e $CK_{\eta_{\tau}}$ coefficient of skewness of kurtosis κ e η_{τ} , respectivelly, percentual relative bias in absolute value for κ and τ (RB_{κ} e RB_{η_{τ}} respectivelly. In addition, $(\text{VAR}_{\kappa} \text{ and } \text{VAR}_{\eta_{\tau}} \text{ are variances for } \kappa \in \eta_{\tau},$ respectivelly). Finally, mean squared error (MSE_{κ} e $MSE_{\eta_{\tau}}$. In the simulation, 5000 Monte Carlo replicas were carried out, the sample sizes considered were $n = 10, 30, 50,$ and 100. In this work, we consider the following values for the shape parameter $\kappa \in \{1.5, 2.5, 5\}$ $\kappa \in \{1.5, 2.5, 5\}$ $\kappa \in \{1.5, 2.5, 5\}$. In, Figure 1 we have some values of shape parameter κ can be indicate degree of kurtosis. Therefore, The values of κ was fixed for some different degrees kurtosis, strong asymmetry, moderate and weak. Without loss of generality, the parameter η_{τ} was considered fixed, with $\eta_{\tau} = 1$ in all Monte Carlo experiments. We also consider $\tau \in \{0.25, 0.50, 0.75\}$ to study the behavior of the parameters for different values of τ . Furthermore, we present the numerical results of the simulations for the asymptotic confidence intervals (CI).

In Table [2](#page-10-0) we can observe that when the sample size increases, the variance, mean squared error, and percentage relative bias in absolute value decrease in all κ e τ values. Furthermore, it can be seen that for fixed values of κ and n when the value of τ increases the percentage relative bias in absolute value, the variance and the mean squared error for κ also increase. Therefore, the estimates of the κ parameter performed better for smaller values of τ . In addition, for fixed values of κ and n, the η_{τ} estimator performed better for values of $\tau = 0.5$. Also, it is noted that, for fixed values of τ and n, when the value of the shape parameter κ increases, the variance and mean squared error also increase and the estimator $\hat{\kappa}$ showed similar behavior when compared to estimates of relative percentage bias in absolute value. Furthermore, we can observe that, for small sample sizes $n = 10$, the estimator $\hat{\kappa}$ presented largekurtosis. Additionally, for $n = 10$, $\kappa = 5$ and $\tau \in \{0.25, 0.50, 0.75\}$ the estimator $\hat{\kappa}$ presented large variance and mean squared error, Finally, we can observe that when sample size increase , the coefficient of skewness and coefficient of kurtosis close to zero and 3 respectivelly.

According to Table [3,](#page-11-0) we can observe that all confidence intervals contain true value parameter of κ and η_{τ} . In addition, all confidence intervals (CI) were calculated with nominal coverage level $(1 - \gamma)$ equal to 0.95. Furthermore, note that the mean lengths of the confidence intervals $(MLCI)$ for the parameter κ , when $\kappa = 1.5$ and varying values of $\tau \in \{0.25, 0.50, 0.75\}$ presented similar MLCI for fixed sample sizes. For example, $\kappa = 1.5$,

 $n = 100$ and $\tau \in \{0.25, 0.50, 0.75\}$ the *MLCI* were equal to 0.428, 0.451 and 0.496, respectively. Similarly, the same behavior can be seen in the case where $\kappa = 2.5$ and varying the values of $\tau \in \{0.25, 0.50, 0.75\}$, that is, the CI presented close to MLCI for fixed sample sizes. It is also noted that, with $\kappa = 5$ and varying the values of $\tau \in \{0.25, 0.50, 0.75\}$ the CI showed similar MLCI for fixed sample sizes. Furthermore, we can observe that for sample sizes and fixed τ as the value of the parameter κ increases the MLCI also increases. For example, for $\kappa = 1.5$, $\tau = 0.25$ and $n = 100$ the MLCI was equal to 0.428. On the other hand, for $\kappa = 5$, $\tau = 0.25$ and fixed $n = 100$ the MLCI was equal to 1.428. This indicates that the smaller the value of the κ parameter, the better the interval estimate to the κ parameter. Additionally, considering fixed κ values and fixed sample sizes as the τ value increases the MLCI to the parameter η_{τ} also increases. This shows that the smaller the value of τ , the better the interval estimate for the parameter η_{τ} . It is also noticed that, for all confidence intervals, with the increase in the sample size the MLCI decreases, that is, as the sample size grows the interval estimates for the parameters improve.

In Figures [2](#page-12-0)[-4](#page-14-0) we illustrate the histograms of the 5000 maximum likelihood estimates for the REW model parameters. The line segments indicate the length of the confidence interval. It can be seen in Figures [2-](#page-12-0)[4](#page-14-0) that, as the sample size increases regardless of the values of κ and τ , the empirical distribution becomes more symmetrical around the parameter κ . To evaluate the interval estimates, the probability of empirical coverage of the asymptotic intervals denoted in Table [3](#page-11-0) by CP, was calculated, to the empirical coverage probability on the left denoted by LIE and the empirical coverage probability on the right denoted by LSD, with nominal coverage level $(1 - \gamma)$ equal to 0.95 for all sample sizes. According to Table [3,](#page-11-0) it is noted that for all sample sizes, the coverage probability for the shape parameter κ is close to 0.95, so the interval estimate for κ is accurate, that is, with a small coverage error. The observed coverage probabilities for the η_{τ} parameter were better for larger sample sizes. Therefore, as the sample size increases, the coverage probabilities observed for parameter η_{τ} approach the nominal coverage probability of 0.95. Furthermore, when the sample size increase, the confidence intervals tend to balanced of coverage probability. In addition the histograms tend to normal distribution.

Table 2: Maximum likelihood estimates and performance measures esti- $\kappa \in \{1.5, 2.5, 5\}, \eta_{\tau} = 1$ fixed, κmated by the BFGS method for $\tau \in \{0.25, 0.50, 0.75\}.$ and Table 2:

τ

Table 3: Interval estimates and empirical coverage probabilities (CP) of the asymptotic CI and empirical CP of the left tails (LIE) and right (LSD) for κ $\kappa \in \{1.5, 2.5, 5\}, \eta_{\tau} = 1 \text{ fixed}, \text{ and}$
" τ $\tau \in \{0.25, 0.50, 0.75\}$, with confidence level 1 $\gamma = 0.95.$

Figure 2: Empirical distribution of $\hat{\kappa}$ with $\tau = 0.25$ and $\eta_{\tau} = 1$ fixed.

Figure 3: Empirical distribution of $\hat{\kappa}$ with $\tau = 0.5$ and $\eta_{\tau} = 1$ fixed.

Figure 4: Empirical distribution of $\hat{\kappa}$ with $\tau = 0.75$ and $\eta_{\tau} = 1$ fixed.

6. APPLICATION

In this application, we utilized the dataset concerning the breaking stress of carbon fibers (in GBa) [\(Nichols and Padgett,](#page-18-14) [2006\)](#page-18-14). The dataset comprises 100 measurements and is available in the AdequacyModel [\(Marinho et al.,](#page-18-15) [2019\)](#page-18-15) package of R. To access it, simply execute the command data(carbone).

In Table [4,](#page-15-1) we provide several descriptive statistics for this dataset. These measures were obtained using the basicStats() function from the fbasic package in R. It can be observed that the mean breaking stress of carbon fibers is 2.62. The distribution of the data shows a slight positive skewness. Moreover, the distribution exhibits a mesokurtic shape. To fit the data, we used the goodness.fit() function from the AdequacyModel package.

Mean	ω	\mathbf{r} Median	\mathcal{Q}_3	Variance	\sim Skewness	$ -$ Kurtosis	Minimum	Maximum	\boldsymbol{n}
റ ഭറ 4.V4	1.84	70 $4 \cdot 10$	ഹ $\boldsymbol{\sigma}$. 44	ഹ T.OO	$\rm 0.36$	0.04	$0.39\,$	$5.56\,$	100

Table 4: Descriptive measures.

The maximum likelihood estimates and their standard errors are presented in Table [5.](#page-16-0) The dataset was fitted considering the distributions: REE, REW, REBS, and RELL. Additionally, we considered the following values of τ : 0.25, 0.5, and 0.75.

To analyze which distribution provided the best fit, we considered the following criteria: Akaike information criterion (AIC) [\(Sakamoto et al.,](#page-18-16) [1983\)](#page-18-16), consistent Akaike information criterion (CAIC) [\(Akaike,](#page-17-3) [1973\)](#page-17-3), Bayesian information criterion (BIC) [\(Schwarz,](#page-18-17) [1978\)](#page-18-17), Anderson-Darling (A^{*}) [\(Anderson and Darling,](#page-17-4) [1954\)](#page-17-4), Cramér-Von Mises (W^{*}) (Cramér, [1928;](#page-18-18) [Von Mises,](#page-18-19) [1928\)](#page-18-19), and Kolmogorov-Smirnov (KS) test [\(Kolmogorov,](#page-18-20) [1933\)](#page-18-20) with its associated p -value. Therefore, the model selection process determines the most suitable model among the fitted models based on the lowest values of these statistics. The values of these statistics are presented in Table [6.](#page-16-1)

Overall, the REW distribution showed the best fit according to the criteria used. However, for the 75th quantile the RELL distribution exhibited the lowest values for the statistics used and the highest p -value for the Kolmogorov-Smirnov test. In Figure [5,](#page-17-5) we present the fit of these distributions to the studied dataset.

7. FINAL REMARKS

In this paper, we propose a new class of distributions indexed by a parameter representing the $(\tau \times 100)$ -th quantile. This is highly useful in various fields as there is often interest in modeling quantiles. Additionally, the results presented here generalize several existing works. We conducted Monte Carlo simulation studies to investigate the properties of maximum likelihood estimators and asymptotic confidence intervals for the REW distribution.

Moreover, we noted that the asymptotic confidence intervals for the parameter κ ex-

τ	Models	η_{τ}	κ	β	λ
	REE	1.2429			
		(0.0919)			
0.25	REW	1.8827	1.9429		
		(0.1003)	(0.1471)		
	REBS	1.7280			0.5797
		(0.0872)			(0.0417)
	RELL	1.8357		3.1394	
		(0.0889)		(0.2505)	
	REE	1.8170			
		(0.1817)			
0.50	REW	2.5816	2.7928		
		(0.1085)	(0.2141)		
	REBS	2.3660			0.4621
		(0.1064)			(0.0327)
	RELL	2.4984		4.1178	
		(0.1054)		(0.3441)	
	REE	3.2339			
		(0.4919)			
0.75	REW	3.4316	4.6236		
		(0.1205)	(0.3639)		
	REBS	3.2074			0.3425
		(0.1519)			(0.0246)
	RELL	3.2382		6.5582	
		(0.1191)		(0.5854)	

Table 5: Maximum likelihood estimates of the parameters and estimates of the corresponding standard errors (in parentheses).

τ	Models	AIC	CAIC	BIC	A^*	W^*	KS	p -value
	REE	340.209	340.249	342.814	0.830	0.162	0.210	0.000
0.25	REW	287.420	287.544	292.631	0.503	0.096	0.077	0.585
	REBS	304.785	304.909	309.995	1.670	0.307	0.129	0.070
	RELL	308.544	308.667	313.754	2.161	0.399	0.108	0.191
	REE	394.742	394.782	397.347	0.764	0.149	0.320	0.000
0.50	REW	287.059	287.182	292.269	0.416	0.062	0.060	0.858
	REBS	304.122	304.245	309.332	1.618	0.298	0.130	0.066
	RELL	296.559	296.683	301.769	1.241	0.239	0.090	0.389
	REE	489.636	489.676	492.241	0.750	0.146	0.480	0.000
0.75	REW	294.857	294.981	300.067	0.869	0.109	0.092	0.363
	REBS	302.183	302.307	307.393	1.467	0.271	0.129	0.071
	RELL	286.662	286.786	291.873	0.364	0.066	0.059	0.879

Table 6: AIC, CAIC, BIC, A^* , W^* , and KS (with their p-value) statistics for the dataset on the stress at breakage of carbon fibers (in GBa).

hibited better coverage rates. Finally, we applied the proposed distribution family to a real dataset. In this application, we observed that the REW distribution had the best fit for lower quantiles, while for the 75th quantile, the RELL distribution showed a superior fit. For future studies, we intend to propose a regression model based on this class of distributions

Figure 5: Histogram with estimated densities for the REW and RELL distributions for the dataset on the breaking stress of carbon fibers (in GBa).

capable of modeling quantiles of data with different characteristics.

ACKNOWLEDGMENTS

This research work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico ($CNPq$) and FACEPE from Brazil. The authors would like to thank the Associate Editor and anonymous referees for their constructive comments on an earlier version of the manuscript, which resulted in this improved version.

REFERENCES

- A. El-Gohary, Ahmad Alshamrani, A. N. A.-O. (2013). The generalized gompertz distribution. Applied Mathematical Modelling, 37(1):13–24.
- Abdul-Moniem, I. and Abdel-Hameed, H. (2012). On exponentiated lomax distribution. International Journal of Mathematical Education, 33:1–7.
- Akaike, H. (1973). Information theory and an extention of the maximum likelihood principle. In 2nd International Symposium on Information Theory, 1973, pages 267–281. Akademiai Kiado.
- Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. Journal of the American Statistical Association, 49(268):765–769.
- Cordeiro, G. M., Lemonte, A. J., and Ortega, E. M. (2013). An extended fatigue life distribution. Statistics, 47(3):626–653.
- Cordeiro, G. M., Ortega, E. M. M., and Silva, G. O. (2011). Modelos de regressão de sobrevivência estendidos. In Escola de Modelos de Regressão. UFC/DEMA/ABE.
- Cox, D. and Hinkley, D. (1974). Theoretical Statistics (1st ed.). Chapman and Hall UK.
- Cramér, H. (1928). On the composition of elementary errors. Scandinavian Actuarial Journal, 1928(1):13–74.
- Gallardo, D. I. and Santos-Neto, M. (2021). Quantile regression for positive data using a general class of distributions. arXiv preprint arXiv:2109.09281.
- Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: An alternative to gamma and weibull distributions. Biometrical Journal, 43(1):117–130.
- Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge didistribuzione. Giorn Dell'inst Ital Degli Att, 4:89–91.
- Kundu, D. and Raqab, M. Z. (2005). Generalized rayleigh distribution: different methods of estimations. Computational Statistics and Data Analysis, 49(1):187–200.
- Lehmann, E. L. (1953). The Power of Rank Tests. The Annals of Mathematical Statistics, 24(1):23 – 43.
- Lehmann, E. L. and Casella, G. (1998). Theory of Point Estimation. Springer-Verlag, New York, NY, USA, second edition.
- Marinho, P. R. D., Silva, R. B., Bourguignon, M., Cordeiro, G. M., and Nadarajah, S. (2019). Adequacymodel: An r package for probability distributions and general purpose optimization. PLOS ONE, 14(8):1–30.
- Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated weibull family for analyzing bathtub failure-rate data. IEEE Transactions on Reliability, 42(2):299–302.
- Nadarajah, S. (2006). The exponentiated gumbel distribution with climate application. Environmetrics, 17(1):13–23.
- Nadarajah, S. and Kotz, S. (2003). The exponentiated fréchet distribution. Interstat Electronic Journal, 14:1–7.
- Nichols, M. D. and Padgett, W. J. (2006). A bootstrap control chart for weibull percentiles. Quality and Reliability Engineering International, 22(2):141–151.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992). Numeric recipes in c: the art of scientific computing. Camb. Univ. Press Camb.
- Rosaiah, K., Kantam, R., and Kumar, S. (2006). Reliability test plans for exponentiated log-logistic distribution.
- Sakamoto, Y., Ishiguro, M., and Kitagawa, G. (1983). Akaike Information Criterion Statistics. Mathematics and its applications. D. Reidel Publishing Company.
- Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, pages 461–464.
- Team, R. D. C. (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Von Mises, R. (1928). Statistik und wahrheit. Julius Springer, 20.
- Weibull, W. (1939). The Phenomenon of Rupture in Solids. Handlingar / Ingeniörsvetenskapsakademien. Generalstabena Litografiska Anstalts Forlag.