Hindawi
Computational and Mathematical Methods
Volume 2023, Article ID 9325679, 18 pages
https://doi.org/10.1155/2023/9325679

Research Article

A New Generated Family of Distributions: Statistical Properties and Applications with Real-Life Data

John Kwadey Okutu, Nana K. Frempong, Simon K. Appiah, and Atinuke O. Adebanji

Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Correspondence should be addressed to John Kwadey Okutu; jkokutu@st.knust.edu.gh

Received 28 March 2023; Revised 17 July 2023; Accepted 3 September 2023; Published 19 September 2023

Academic Editor: Pritpal Singh

Several standard distributions can be used to model lifetime data. Nevertheless, a number of these datasets from diverse fields such as engineering, finance, the environment, biological sciences, and others may not fit the standard distributions. As a result, there is a need to develop new distributions that incorporate a high degree of skewness and kurtosis while improving the degree of goodness-of-fit in empirical distributions. In this study, by applying the T-X method, we proposed a new flexible generated family, the Ramos-Louzada Generator (RL-G) with some relevant statistical properties such as quantile function, raw moments, incomplete moments, measures of inequality, entropy, mean and median deviations, and the reliability parameter. The RL-G family has the ability to model "right," "left," and "symmetric" data as well as different shapes of the hazard function. The maximum likelihood estimation (MLE) method has been used to estimate the parameters of the RL-G. The asymptotic performance of the MLE is assessed by simulation analysis. Finally, the flexibility of the RL-G family is demonstrated through the application of three real complete datasets from rainfall, breaking stress of carbon fibers, and survival times of hypertension patients, and it is evident that the RL-Weibull, which is a special case of the RL-G family, outperformed its submodels and other distributions.

1. Introduction

Choosing an appropriate statistical distribution for modeling and analyzing data is critical in order to draw more accurate conclusions. Many statistical distributions have been proposed to match different data forms over the years. Using conventional distributions for fitting these datasets may produce erroneous findings. As a result, there is a clear need for modifications to the standard distributions. The literature on probability distribution methods contains various extensions and generalizations of continuous, discrete, symmetric, and asymmetric distributions. Regarding the main methods of generating probability distributions and classes of probability distributions, Lee et al. [1] stated that the transformation technique, differential equation technique, and quantile method are three groups of methods developed prior to 1980, and those proposed after 1980 may be categorized as combination methods because these techniques attempt to develop new distributions through the combination of existing ones or by adding additional parameters to an existing distribution. Several studies have proposed using different generated classes to increase the number of parameters in distributions. The resulting distributions have found application in modeling data across various fields of study, such as environmental sciences, economics, and engineering. Some popular generators available in the literature include the exponentiated generated family by [2], the Marshall-Olkin-G by [3], the Kumaraswamy-G by [4], Beta-G by [5], Weibull-X by [6], Weibull-G by [7], the Lomax generator proposed by [8], the Topp-Leone generated family introduced by [9], the Lindley generator by [10], the Chen-G class by [11], the Burr III Topp-Leone-G by [12], the odd Burr-III family by [13], Marshall-Olkin Burr X family by [14], the Topp-Leone odd Lindley-G family by [15], and many others.

In [16], the Ramos-Louzada (RL) distribution, a one-parameter continuous distribution, was introduced for modeling lifetime data. The study demonstrated that the RL distribution performs better than some well-known lifetime distributions such as Lindley and exponential distributions. However, the RL distribution is limited to right-skewed lifetime data with an increasing failure rate. Therefore, it is essential to propose an extension or generalization of the RL to introduce flexibility in modeling different lifetime data with "symmetric" and "asymmetric" shapes and "monotonic" and "non-monotonic" failure rate functions. [17] produced the generalized Ramos-Louzada (GRL) distribution, which is the first extension of the RL distribution. In [18], the discrete RL distribution was developed and proposed. This study adopts the T-X method introduced by [19] to develop the Ramos-Louzada Generator (RL-G), which is capable of producing new distributions that are extensions or generalizations of the RL distribution. Therefore, for any continuous random variable, by applying the T-X method defined in (1), the cumulative distribution function (CDF) of random variable X can be expressed using the RL-G.

$$F(x,\omega)=\int_{c}^{Z(G(x,\boldsymbol{\epsilon}))} r(t;\theta)dt, x\in\mathbb{R}, \tag{1}$$

where $\omega=(\theta,\epsilon)$ is a parameter vector, $r(t;\theta)$ is the PDF generator of a random variable T, $Z(G(x,\epsilon)$ is an expression that depends on the CDF of the random variable X, and c is a real number.

The remaining part of the study is structured in the following manner: In Section 2, the CDF, PDF, and hazard rate function of the RL-G family of distributions are presented. Section 3 presents the mixture representation of the RL-G density functions. In Section 4, we have derived the statistical properties of the RL-G family. Parameter estimation for the proposed family of distributions is discussed in Section 5. Some special distributions of the RL-G family are discussed in Section 6. Section 7 presents Monte Carlo simulation analysis on the asymptotic performance of the MLE. Applications of the proposed distribution to three real datasets to demonstrate its flexibility and usefulness are captured in Section 8, while Section 9 presents the conclusion of the study.

2. The Ramos-Louzada Generated Family of Distributions

Given that equations (2) and (3) represent the CDF and PDF of the RL distribution,

$$R(t;\theta)=1-\frac{\left(\theta^2+t-\theta\right)}{\theta(\theta-1)}e^{-t/\theta}, \tag{2}$$

$$r(t;\theta)=\frac{\left(\theta^2+t-2\theta\right)}{\theta^2(\theta-1)}e^{-t/\theta}, \tag{3}$$

$$t>0,\theta\geq 2. \tag{4}$$

Let $G(x;\epsilon)$ represents the CDF of the baseline distribution and ϵ be a vector of parameters associated with the CDF. The proposed RL-G densities are obtained by using the T-X approach in (1) and letting $Z(G(x,\epsilon)=-\log(1-G(x,\epsilon)))$, $x>0$,thus

$$F_{\mathrm{RL-G}}(x,\theta,\epsilon)=\int_{0}^{-\log(1-G(x,\epsilon))} r(t)dt=\left[R(t)\right]_{0}^{-\log(1-G(x;\epsilon))}. \tag{5}$$

Substituting (2) and (3) into the above relation, we obtain the following:

$$\begin{aligned}F_{\mathrm{RL-G}}(x,\theta,\boldsymbol{\epsilon})&=\int_{0}^{-\log(1-G(x,\boldsymbol{\epsilon}))}\frac{\left(\theta^2+t-2\theta\right)}{\theta^2(\theta-1)}e^{-t/\theta}\,dt\\&=\left[1-\frac{\left(\theta^2+t-\theta\right)}{\theta(\theta-1)}e^{-t/\theta}\right]_{0}^{-\log(1-G(x;\boldsymbol{\epsilon}))}.\end{aligned} \tag{6}$$

Hence, the CDF of the proposed RL-G is expressed as

$$F_{\mathrm{RL-G}}(x,\theta,\epsilon)=1-\left(1-\frac{\log(1-G(x;\epsilon))}{\theta(\theta-1)}\right)(1-G(x;\epsilon))^{1/\theta}, x>0,\theta\geq 2. \tag{7}$$

The proposed RL-G family PDF is derived by finding the derivative of (7), thus

$$f_{\mathrm{RL-G}}(x;\theta,\epsilon)=\frac{g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta)-1}}{\theta^2(\theta-1)}\left(\theta^2-2\theta-\log(1-G(x;\epsilon))\right). \tag{8}$$

From which the survival and hazard functions are, respectively, obtained by

$$\begin{aligned}S_{\mathrm{RL-G}}(x;\theta,\epsilon)&=1-F_{\mathrm{RL-G}}(x,\theta,\epsilon)\\&=\left(1-\frac{\log(1-G(x;\epsilon))}{\theta(\theta-1)}\right)(1-G(x;\epsilon))^{1/\theta},\\h_{\mathrm{RL-G}}(x;\theta,\epsilon)&=\frac{f_{\mathrm{RL-G}}(x;\theta,\epsilon)}{S_{\mathrm{RL-G}}(x;\theta,\epsilon)}\\&=\frac{g(x,\epsilon)\left(\theta^2-2\theta-\log(1-G(x;\epsilon))\right.}{\theta(1-G(x;\epsilon))(\theta(\theta-1)-\log(1-G(x;\epsilon))}.\end{aligned} \tag{9}$$

Some basic motivations obtained when using RL-G densities are as follows:

(i) The properties of the baseline densities are enhanced

(ii) An extended form of the baseline model is generated with the introduction of extra parameter(s).

(iii) The kurtosis of the resulting distributions is more flexible compared to the baseline model

(iv) Special models with various forms of the hazard rate function are defined

Proposition 1. *The RL-G family is a valid PDF, which suffices that*

(i) $0 \le F_{RL-G}(x, \theta, \epsilon) \le 1$,

(ii) $\int_0^\infty f_{RL-G}(x;\theta,\epsilon) = 1$, *for all* $x > 0$

The proof of this proposition is shown in the appendix.

3. RL-G Family in Mixture Representation

This form of representation plays a very important role in deriving some statistical properties of the RL-G densities. Using the following generalized binomial series and the power series expansions on (8) where $|z| < 1$, $t > 0$ is a real noninteger

$$(1-z)^t = \sum_{a=0}^{\infty} \binom{t}{a} (-1)^a z^a,$$
$$\log(1-z) = -\sum_{b=0}^{\infty} \frac{z^{b+1}}{b+1}. \tag{10}$$

The pdf of the RL-G family, that is (8), now becomes

$$f_{\mathrm{RL-G}}(x;\theta,\epsilon) = \frac{(\theta^2-2\theta)g(x;\epsilon)}{\theta^2(\theta-1)} \sum_{a=0}^{\infty} \binom{\frac{1}{\theta}-1}{a} (-1)^a G(x;\epsilon)^a + \frac{g(x;\epsilon)}{\theta^2(\theta-1)} \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} \binom{\frac{1}{\theta}-1}{a} \frac{(-1)^a}{b+1} G(x;\epsilon)^{a+b+1}. \tag{11}$$

Multiplying the first term of the last equation by $a+1/a+1$, the second term by $(a+b+2)/(a+b+2)$ and rearranging, thus

$$f_{\mathrm{RL-G}}(x;\theta,\epsilon) = \sum_{a=0}^{\infty} \binom{\frac{1}{\theta}-1}{a} (-1)^a \frac{(\theta^2-2\theta)(a+1)g(x;\epsilon)G(x;\epsilon)^a}{\theta^2(\theta-1)(a+1)} + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} \binom{\frac{1}{\theta}-1}{a} (-1)^a \frac{(a+b+2)g(x;\epsilon)G(x;\epsilon)^{a+b+1}}{\theta^2(\theta-1)(b+1)(a+b+2)}, \tag{12}$$

$$f_{\mathrm{RL-G}}(x;\theta,\epsilon) = \sum_{a=0}^{\infty} \eta_a h_{a+1}(x) + \sum_{a,b=0}^{\infty} \kappa_{a,b} h_{a+b+2}(x), \tag{13}$$

where

$$\eta_a = \binom{\frac{1}{\theta}-1}{a} (-1)^a \frac{(\theta^2-2\theta)}{\theta^2(\theta-1)(a+1)},$$
$$\kappa_{a,b} = \binom{\frac{1}{\theta}-1}{a} \frac{(-1)^a}{\theta^2(\theta-1)(b+1)(a+b+2)},$$
$$h_{a+1}(x) = (a+1)g(x;\epsilon)G(x;\epsilon)^a,$$
$$h_{a+b+2}(x) = (a+b+2)g(x;\epsilon)G(x;\epsilon)^{a+b+1}. \tag{14}$$

Thus, (13) represents an infinite linear combination of exp-G densities of the baseline density. The linear representation form of the RL-G facilitates the derivation of other statistical properties of the RL-G density. Integrating (13) with respect to x produces the corresponding linear representation form of the CDF of the RL-G family.

$$F_{\mathrm{RL-G}}(x;\theta,\epsilon) = \sum_{a=0}^{\infty} \eta_a H_{a+1}(x) + \sum_{a,b=0}^{\infty} \kappa_{a,b} H_{a+b+2}(x), \tag{15}$$

where $H_{a+1}(x) = G(x;\epsilon)^{a+1}$, $H_{a+b+2}(x) = G(x;\epsilon)^{a+b+2}$, and $a+1, a+b+2$ are power parameters of the exponentiated-G distributions $H_{a+1}(x)$ and $H_{a+b+2}(x)$, respectively.

4. Some Relevant Statistical Properties of the RL-G Family

In this section, we have derived some relevant statistical properties of the RL-G family. These include the quantile function, the raw (noncentral) moments, measures of inequality, the entropy measure, mean and median deviations, and the reliability parameter.

4.1. The Quantile Function. By definition, the quantile function $Q(p)$ of the RL-G family is $F_{\mathrm{RL-G}}(Q(p);\theta,\epsilon) = p$, where $0 \le p \le 1$. Now from (7), we have $1 - (1 - (\log(1 - G(Q(p);\epsilon))/\theta(\theta-1)))(1 - G(Q(p);\epsilon))^{1/\theta} = p$, if we let $1 - G(Q(p);\epsilon) = y$, solving for y gives; $y = \{W[-(\theta-1)(1-p)e^{-(\theta-1)}]/-(\theta-1)(1-p)\}^{-\theta}$, where the negative branch of the Lambert function, is denoted by $W(.)$; see [20].

But $1 - G(Q(p);\epsilon) = y$; and hence, the RL-G family quantile function $Q(p)$ is represented as

$$Q(p) = G^{-1}\left[1 - \left\{\frac{W\left[-(\theta-1)(1-p)e^{-(\theta-1)}\right]}{-(\theta-1)(1-p)}\right\}^{-\theta}\right], \tag{16}$$

where the baseline distribution $G(x;\epsilon)$ has its inverse denoted as $G^{-1}(.)$.

4.2. Moments. In statistical analysis, the kurtosis, mean, skewness, and variance are measures that can be computed using the noncentral moments of a distribution.

If $X \sim RL - G$ random variable, then the rth moment is defined as follows:

$$\mu'_r = \int_0^{\infty} x^r f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx. \tag{17}$$

Substituting (13) into the above definition and simplifying, the rth noncentral moments can be expressed as

$$\begin{aligned}\mu'_r &= \sum_{a=0}^{\infty} \eta_a{}^* \int_0^{\infty} x^r g(x;\epsilon)G(x;\epsilon)^a dx \\ &+ \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \int_0^{\infty} x^r g(x;\epsilon)G(x;\epsilon)^{a+b+1} dx,\end{aligned} \tag{18}$$

which can be simplified as

$$\mu'_r = \sum_{a=0}^{\infty} \eta_a{}^* \ell_{r,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \ell_{r,a+b+1}, \tag{19}$$

where $\eta_a{}^* = \eta_a(a+1)$ and $\kappa_{a,b}{}^* = \kappa_{a,b}(a+b+2)$:

$$\begin{aligned}\ell_{r,a} &= \int_0^{\infty} x^r g(x;\epsilon)G(x;\epsilon)^a dx, \\ \ell_{r,a+b+1} &= \int_0^{\infty} x^r g(x;\epsilon)G(x;\epsilon)^{a+b+1} dx.\end{aligned} \tag{20}$$

Alternatively, (18) can be expressed in terms of the baseline quantile function, supposed $G(x;\epsilon) = z$ in (18), then $G^{-1}(x;\epsilon) = Q(z) = x$, , $dz = g(x;\epsilon)dx$.

As $x \longrightarrow 0, z \longrightarrow 0$ and as $x \longrightarrow \infty, z \longrightarrow 1$.

From (18), we have the rth noncentral moments expressed as

$$\mu'_r = \sum_{a=0}^{\infty} \eta_a{}^* \int_0^1 Q^r(z) z^a dz + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \int_0^1 Q^r(z) z^{a+b+1} dz. \tag{21}$$

4.3. Incomplete Moment. This statistical property plays an essential role in the computation of the mean and medium deviations, inequality and entropy measures, and residual life of a random variable.

The RL-G family rth incomplete moment is defined as $m_r(t) = \int_0^t x^r f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx$.

Setting (13) into the definition, we obtain the following:

$$m_r(t) = \sum_{a=0}^{\infty} \eta_a{}^* \zeta_{r,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \zeta_{r,a+b+1}, \tag{22}$$

where $\zeta_{r,a} = \int_0^t x^r g(x;\epsilon)G(x;\epsilon)^a dx, \zeta_{r,a+b+1} = \int_0^t x^r g(x;\epsilon)G(x;\epsilon)^{a+b+1} dx$, $\eta_a{}^*$, and $\kappa_{a,b}{}^*$ are defined in (18).

Alternatively, the rth incomplete moment is expressed in terms of the baseline quantile function. Supposed $G(x;\epsilon) = z$ in (22), then $G^{-1}(x;\epsilon) = Q(z) = x$, $dz = g(x;\epsilon)dx$. As $x \longrightarrow 0, z \longrightarrow 0$ and as $x \longrightarrow t, z \longrightarrow G(t;\epsilon)$. From (22), we have the rth incomplete moments expressed as

$$m_r(t) = \sum_{a=0}^{\infty} \eta_a{}^* \delta_{r,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \delta_{r,a+b+1}, \tag{23}$$

where $\delta_{r,a} = \int_0^{G(t;\epsilon)} Q^r(z) z^a dz, \delta_{r,a+b+1} = \int_0^{G(t;\epsilon)} Q^r(z) z^{a+b+1} dz$, $\eta_a{}^*$, and $\kappa_{a,b}{}^*$ as defined before.

4.4. Measures of Inequality. The Bonferroni and Lorenz curves are two of the most commonly used measures of inequality that are applied in various fields such as insurance, demography, reliability engineering, and economics.

By definition, the Lorenz curve $L_F(x)$ of the RL-G family is defined by $L_F(x) = 1/E(X)\int_0^t x f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx$, where $E(X)$ is the mean and $\int_0^t x f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx$ is the first incomplete moment of the RL-G family obtained by setting $r = 1$ into the incomplete moment's expression, that is;

$$\int_0^t x f_{\mathrm{RL-G}}(x;\theta,\boldsymbol{\epsilon})dx = \sum_{a=0}^{\infty} \eta_a{}^* \zeta_{1,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \zeta_{1,a+b+1}. \tag{24}$$

Substituting into the definition for $L_F(x)$ produces;

$$L_F(x) = \frac{1}{E(X)} \left(\sum_{a=0}^{\infty} \eta_a{}^* \zeta_{1,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \zeta_{1,a+b+1} \right), \tag{25}$$

where $\zeta_{1,a} = \int_0^t x g(x;\epsilon)G(x;\epsilon)^a dx$ and $\zeta_{1,a+b+1} = \int_0^t x g(x;\epsilon) G(x;\epsilon)^{a+b+1} dx$. By definition, the Bonferroni curve $B_F(x)$ of the RL-G family is;

$$\begin{aligned}B_F(x) &= \frac{L_F(x)}{F_{\mathrm{RL-G}}(x;\theta,\epsilon)} \\ &= \frac{1}{E(X)F_{\mathrm{RL-G}}(x;\theta,\epsilon)} \int_0^t x f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx B_F(x) \\ &= \frac{1}{E(X)F_{\mathrm{RL-G}}(x;\theta,\epsilon)} \left(\sum_{a=0}^{\infty} \eta_a{}^* \zeta_{1,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \zeta_{1,a+b+1} \right).\end{aligned} \tag{26}$$

4.5. Mean and Median Deviations. The mean deviation denoted by π_1 for RL-G random variable is defined by

$$\begin{aligned}\pi_1 &= \int_0^{\infty} |x-\mu| f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx \\ &= 2\mu F_{\mathrm{RL-G}}(\mu;\theta,\epsilon) - 2\int_0^{\mu} x f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx,\end{aligned} \tag{27}$$

where $\int_0^{\mu} x f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx$ is the first incomplete moment.

Hence, the mean deviation μ is

$$\pi_1 = 2\mu F_{\mathrm{RL-G}}(\mu;\theta,\epsilon) - 2\left(\sum_{a=0}^{\infty} \eta_a{}^* \zeta_{1,a} + \sum_{a,b=0}^{\infty} \kappa_{a,b}{}^* \zeta_{1,a+b+1} \right), \tag{28}$$

where $\zeta_{1,a} = \int_0^{\mu} xg(x;\epsilon)G(x;\epsilon)^a dx, \zeta_{1,a+b+1} = \int_0^{\mu} xg(x;\epsilon)G(x;\epsilon)^{a+b+1} dx$, ${\eta_a}^*$, and ${\kappa_{a,b}}^*$ as defined before.

The median deviation about the median M denoted by π_2 for RL-G random variable is defined by $\pi_2 = \int_0^{\infty} |x - M| f_{\text{RL-G}}(x;\theta,\boldsymbol{\epsilon})dx = \mu - 2\int_0^{M} x f_{\text{RL-G}}(x;\theta,\epsilon)dx$.

Hence, the medium deviation about the median is expressed as

$$\pi_2 = \mu - 2\left(\sum_{a=0}^{\infty} {\eta_a}^* \zeta_{1,a} + \sum_{a,b=0}^{\infty} {\kappa_{a,b}}^* \zeta_{1,a+b+1}\right), \tag{29}$$

where $\zeta_{1,a} = \int_0^{M} xg(x;\epsilon)G(x;\epsilon)^a dx$ and $\zeta_{1,a+b+1} = \int_0^{M} xg(x;\epsilon)G(x;\epsilon)^{a+b+1} dx$, ${\eta_a}^*$, and ${\kappa_{a,b}}^*$ as defined before.

4.6. Entropy Measure. The randomness in the RL-G random variable is measured by using the following measures of entropy: Renyi [21], Shannon [22], Havrda and Charvat [23], and Tsallis [24].

The RL-G family has the Renyi entropy denoted by $I_R(\gamma)$ and is defined by

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \int_0^{\infty} f_{RL-G}^{\gamma}(x;\theta,\epsilon)dx, \tag{30}$$

where $\gamma > 0, \gamma \neq 1$.

Substituting the density of the RL-G into the above definition and applying the generalized binomial expansion, the following expression is obtained:

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \int_0^{\infty} \frac{(g(x,\boldsymbol{\epsilon}))^{\gamma}}{\theta^{2\gamma}(\theta-1)^{\gamma}} \sum_{a=0}^{\infty} \binom{\frac{\gamma}{\theta} - \gamma}{a} (-1)^i (G(x;\boldsymbol{\epsilon}))^a \times \sum_{b=0}^{\infty} \binom{\gamma}{b} (\theta^2 - 2\theta)^{\gamma-b} (-\log(1 - G(x;\epsilon))^{\gamma}. \tag{31}$$

Applying the following log power series expansion in the last expression

$$(-\log(1 - G(x;\epsilon))^{\gamma} = \gamma \sum_{d=0}^{\infty} \binom{d-\gamma}{d} \sum_{c=0}^{d} \frac{(-1)^{c+d} P_{c,d}}{\gamma - c} \cdot \binom{d}{c} (G(x;\epsilon))^{\gamma+c}, \tag{32}$$

where the constants $P_{k,l}$ are obtained recursively by using the following relation:

$$P_{c,d} = \frac{1}{d} \sum_{z=0}^{d} \frac{(-1)^z [z(c+1) - l]}{z+1} P_{c,d-z}, \quad P_{c,0} = 1. \tag{33}$$

And after simplifying, we obtain the Renyi entropy as;

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \left[\frac{1}{\theta^{2\gamma}(\theta-1)^{\gamma}} \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \sum_{d=0}^{\infty} \sum_{c=0}^{d} \binom{\frac{\gamma}{\theta} - \gamma}{a} \binom{\gamma}{b} \binom{d-\gamma}{d} \binom{d}{c} \frac{(-1)^{a+c+d} (\theta^2 - 2\theta)^{\gamma-b} \gamma P_{c,d}}{\gamma - c} \times \int_0^{\infty} (g(x,\epsilon))^{\gamma} (G(x;\epsilon))^{a+\gamma+c} dx \right]. \tag{34}$$

The RL-G family Shannon entropy is defined by;

$$I_X = E(-\log f_{\text{RL-G}}(x;\theta,\epsilon)). \tag{35}$$

By setting (13) into the above definition, I_X is obtained as;

$$I_X = E\left[-\log\left(\sum_{a=0}^{\infty} \eta_a h_{a+1}(x) + \sum_{a,b=0}^{\infty} \kappa_{a,b} h_{a+b+2}(x)\right)\right], \tag{36}$$

where η_a, $\kappa_{a,b}$, $h_{a+1}(x)$, and $h_{a+b+2}(x)$ are defined in (13).

The Havrda and Charvat entropy for the RL-G family is represented by;

$$I(\gamma) = \frac{1}{2^{1-\gamma} - 1} \left[\int_0^{\infty} f_{\text{RL-G}}^{\gamma}(x;\theta,\epsilon)dx - 1\right], \tag{37}$$

where $\gamma > 0$ and $\gamma \neq 1$, and the expression in the integral is similar to the one used in Renyi entropy. Thus, the Havrda and Charvat entropy for the RL-G family can be expressed:

$$I(\gamma) = \frac{1}{2^{1-\gamma} - 1} \left(\left[\frac{1}{\theta^{2\gamma}(\theta-1)^{\gamma}} \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \sum_{d=0}^{\infty} \sum_{c=0}^{d} \binom{\frac{\gamma}{\theta} - \gamma}{a} \binom{\gamma}{b} \binom{d-\gamma}{d} \binom{d}{c} \frac{(-1)^{a+c+d} (\theta^2 - 2\theta)^{\gamma-b} \gamma P_{c,d}}{\gamma - c} \times \int_0^{\infty} (g(x,\epsilon))^{\gamma} (G(x;\epsilon))^{a+\gamma+c} dx \right] - 1 \right). \tag{38}$$

The Tsallis's generalized entropy for RL-G random variable is obtained by using the following formula:

$$I_T(\gamma) = \frac{1}{1-\gamma}\left[1 - \int_0^\infty f_{\mathrm{RL-G}}^{\gamma}(x;\theta,\epsilon)dx\right], \tag{39}$$

where $\gamma > 0$ and $\gamma \neq 1$, from which we obtain

$$I_T(\gamma) = \frac{1}{1-\gamma}\left(1 - \left[\frac{1}{\theta^{2\gamma}(\theta-1)^{\gamma}}\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\sum_{d=0}^{\infty}\sum_{c=0}^{d}\binom{\frac{\gamma}{\theta}-\gamma}{a}\binom{\gamma}{b}\binom{d-\gamma}{d}\binom{d}{c}\frac{(-1)^{a+c+d}\left(\theta^2-2\theta\right)^{\gamma-b}\gamma P_{c,d}}{\gamma-c} \times \int_0^\infty (g(x,\epsilon))^{\gamma}(G(x;\epsilon))^{a+\gamma+c}dx\right]\right). \tag{40}$$

4.7. Reliability Parameter. Let $X_1 \sim \mathrm{RL-G}(x,\theta_1,\boldsymbol{\epsilon})$ and $X_2 \sim \mathrm{RL-G}(x,\theta_1,\boldsymbol{\epsilon})$; X_2 and X_1 are strength and stress random variables. The stress-strength reliability parameter of RL-G family of distribution is defined by

$$\begin{aligned} R &= P(X_2 < X_1) = \int_0^\infty f_{\mathrm{RL-G}}(x;\theta_1,\epsilon)F_{\mathrm{RL-G}}(x;\theta_2,\epsilon)dx \\ &= 1 - \int_0^\infty f_{\mathrm{RL-G}}(x;\theta_1,\epsilon)S_{\mathrm{RL-G}}(x;\theta_2,\epsilon)dx \\ &= 1 - \int_0^\infty \left[\frac{g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta_1)+(1/\theta_2)-1}}{\theta_1^2(\theta_1-1)}\left(\theta_1^2 - 2\theta_1 - \log\left(1-G(x;\epsilon)\right)\right)\right. \\ &\quad \left.\times\left(1 - \frac{\log\left(1-G(x;\epsilon)\right)}{\theta_2(\theta_2-1)}\right)dx\right]. \end{aligned} \tag{41}$$

The simplified result from the last expression is

$$\begin{aligned} R &= 1 - \frac{\left(\theta_1^2-2\theta_1\right)}{\theta_1^2(\theta_1-1)}\int_0^\infty g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta_1)+(1/\theta_2)-1}dx \\ &\quad + \frac{\left(\theta_1^2-2\theta_1\right)}{\theta_1^2(\theta_1-1)\theta_2(\theta_2-1)}\int_0^\infty g(x,\epsilon) \\ &\quad \cdot (1-G(x;\epsilon))^{(1/\theta_1)+(1/\theta_2)-1}\log\left(1-G(x;\epsilon)\right)dx \\ &\quad + \frac{1}{\theta_1^2(\theta_1-1)}\int_0^\infty g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta_1)+(1/\theta_2)-1}\log\left(1-G(x;\epsilon)\right)dx \\ &\quad - \frac{1}{\theta_1^2(\theta_1-1)\theta_2(\theta_2-1)}\int_0^\infty g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta_1)+(1/\theta_2)-1} \\ &\quad \cdot \left[\log\left(1-G(x;\epsilon)\right)\right]^2 dx. \end{aligned} \tag{42}$$

Evaluating each of the integrals and using the generalized binomial series expansion, the result of [25] for a power series raised to a positive integer n;

$$\left[\log\left(1-G(x;\epsilon)\right)\right]^2 = \left[-\sum_{b=0}^{\infty}\frac{[G(x;\epsilon)]^{b+1}}{b+1}\right]^2 = \sum_{c=0}^{\infty}b_{2,c}[G(x;\epsilon)]^{c+2}, \tag{43}$$

where $b_{2,0} = a_0^2$ and for any integer

$$\begin{aligned} c \geq 1, b_{2,c} &= \frac{1}{ca_0}\sum_{b=1}^{c}[3b-c]a_c b_{2,c-b}, \\ a_c &= \frac{1}{b+1}. \end{aligned} \tag{44}$$

The reliability parameter after simplification is expressed as

$$\begin{aligned} R &= 1 - \frac{\left(\theta_1^2-2\theta_1\right)}{\theta_1^2(\theta_1-1)}\left[\frac{\theta_1\theta_2}{\theta_1+\theta_2} - \frac{1}{\theta_2(\theta_2-1)}\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\frac{(-1)^{a+1}}{(b+1)(a+b+2)}\right. \\ &\quad \left.\cdot\binom{\frac{1}{\theta_1}+\frac{1}{\theta_2}-1}{a}\right] + \frac{1}{\theta_1^2(\theta_1-1)} \\ &\quad \cdot\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\frac{(-1)^{a+1}}{(b+1)(a+b+2)}\binom{\frac{1}{\theta_1}+\frac{1}{\theta_2}-1}{a}\right. \\ &\quad \left. - \frac{1}{\theta_2(\theta_2-1)}\sum_{a=0}^{\infty}\sum_{c=0}^{\infty}(-1)^a\binom{\frac{1}{\theta_1}+\frac{1}{\theta_2}-1}{a}\frac{b_{2,c}}{a+c+3}\right]. \end{aligned} \tag{45}$$

5. Maximum Likelihood Estimation of the RL-G Family

Suppose the RL-G family has a random sample of size n given by $x_1, x_2, \cdots, x_n$, then the log-likelihood function for the parameter vector is given by

$$\begin{aligned} \boldsymbol{\ell}(\Xi) &= \sum_{i=1}^{n}\log g(x_i;\epsilon) + \left(\frac{1}{\theta}-1\right)\sum_{i=1}^{n}\log(1-G(x_i,\epsilon)) \\ &\quad - 2n\log\theta - n\log(\theta-1) \\ &\quad + \sum_{i=1}^{n}\log\left(\theta^2 - 2\theta - \log\left(1-G(x_i,\epsilon)\right)\right). \end{aligned} \tag{46}$$

Taking derivatives with respect to θ and $\boldsymbol{\epsilon}$

$$\frac{\partial \ell(\Xi)}{\partial \theta} = -\frac{1}{\theta^2}\sum_{i=1}^{n}\log(1-G(x_i,\boldsymbol{\epsilon})) - \frac{2n}{\theta} - \frac{n}{\theta-1} + (2\theta-2)\sum_{i=1}^{n}\frac{1}{\theta^2-2\theta-\log(1-G(x_i,\boldsymbol{\epsilon}))},$$

$$\frac{\partial \ell(\Xi)}{\partial \boldsymbol{\epsilon}} = \sum_{i=1}^{n}\frac{1}{g(x_i;\boldsymbol{\epsilon})}\frac{\partial g(x_i;\boldsymbol{\epsilon})}{\partial \boldsymbol{\epsilon}} - \left(\frac{1}{\theta}-1\right)\sum_{i=1}^{n}\frac{1}{(1-G(x_i,\boldsymbol{\epsilon}))}\frac{\partial G(x_i;\boldsymbol{\epsilon})}{\partial \boldsymbol{\epsilon}} + \sum_{i=1}^{n}\frac{1}{(\theta^2-2\theta-\log(1-G(x_i,\boldsymbol{\epsilon}))(1-G(x_i,\boldsymbol{\epsilon}))}\frac{\partial G(x_i;\boldsymbol{\epsilon})}{\partial \boldsymbol{\epsilon}}. \tag{47}$$

By using numerical techniques, the above equations are set to zero and simultaneously solved to obtain the maximum likelihood estimates.

6. Special Distributions of the RL-G Family

In this section, two special members of the RL-G family, the Ramos-Louzada Weibull (RLW) distribution and the Ramos-Louzada Kumaraswamy (RLKum) distribution, are derived, and the flexibility of these distributions is illustrated by displaying plots of their hazard rate and density functions at some parameter values. Simulation analysis and applications to real datasets of the RLW distribution are studied in the latter section.

(1) Assuming the distribution for the baseline is the Weibull, whose CDF and PDF are, respectively, given by $G(x,\alpha,\beta) = 1-e^{-\alpha x^{\beta}}$ and $g(x,\alpha,\beta) = \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}}$, $x>0$, $\alpha>0$, $\beta>0$, α, and β are scale and shape parameters, respectively. The Ramos-Louzada Weibull (RLW) distribution is obtained by substituting $G(x,\alpha,\beta)$ and $g(x,\alpha,\beta)$ into equations (7) and (8). Thus, the CDF and PDF of the new RLW distribution are, respectively, obtained below:

$$F_{\mathrm{RLW}}(x,\theta,\alpha,\beta) = 1-\left(1+\frac{\alpha x^{\beta}}{\theta(\theta-1)}\right)e^{-\left(\alpha x^{\beta}/\theta\right)},$$

$$f_{\mathrm{RLW}}(x,\theta,\alpha,\beta) = \frac{\alpha\beta x^{\beta-1}\left(\theta^2-2\theta+\alpha x^{\beta}\right)e^{-\left(\alpha x^{\beta}/\theta\right)}}{\theta^2(\theta-1)}. \tag{48}$$

The hazard rate function is expressed as

$$h_{\mathrm{RLW}}(x,\theta,\alpha,\beta) = \frac{\alpha x^{\beta}\left(\theta^2-2\theta+\alpha x^{\beta}\right)}{\theta\left(\theta(\theta-1)+\alpha x^{\beta}\right)}. \tag{49}$$

Figures 1 and 2, respectively, display the plots of the PDF and hazard rate function of the RLW distribution with various selections of parameter values. From Figure 1, the RLW distribution can take several forms, such as "left-skewed," almost "symmetric," "reversed J-shapes," and "right-skewed," and plots of the hazard rate function in Figure 2 illustrate various forms, such as "increasing," "decreasing," "J-shape," and "reversed J-shape."

Submodels of the RLW distribution are as follows:

(i) When $\alpha=\beta=1$, we have the RL distribution given in (2)

(ii) When $\alpha=1$, the Generalized RL distribution proposed by [17] is obtained. The GRL density function is expressed as

$$F_{\mathrm{GRL}}(x,\theta,1,\beta) = 1-\left(1+\frac{x^{\beta}}{\theta(\theta-1)}\right)e^{-\left(x^{\beta}/\theta\right)} \tag{50}$$

(iii) When $\beta=1$, we obtain the Ramos-Louzada Exponential (RLE) distribution. Its density is defined by

$$F_{\mathrm{RLE}}(x,\theta,\beta,\alpha) = 1-\left(1+\frac{\alpha x}{\theta(\theta-1)}\right)e^{-(\alpha x/\theta)} \tag{51}$$

(iv) When $\beta=2$, we obtain the Ramos-Louzada Raleigh (RLR) density defined by

$$F_{\mathrm{RLR}}(x,\theta,\alpha,2) = 1-\left(1+\frac{\alpha x^2}{\theta(\theta-1)}\right)e^{-\left(\alpha x^2/\theta\right)} \tag{52}$$

(2) Supposed that the baseline is the Kumaraswamy distribution whose density is defined as $G(x) = 1-(1-x^{\beta})^{\alpha}$, and $g(x) = \alpha\beta x^{\beta-1}(1-x^{\beta})^{\alpha-1}$, $0<x<1$, $\alpha>0\beta>0$, β, and α are, respectively, shape and scale parameters, and equations (53), (54), and (55), respectively, express the CDF, PDF, and failure rate function of the RLKum distribution:

$$F_{\mathrm{RLKum}}(x,\theta,\beta,\alpha) = 1-\left(1-\frac{\alpha\log\left(1-x^{\beta}\right)}{\theta(\theta-1)}\right)\left(1-x^{\beta}\right)^{\alpha/\theta}, \tag{53}$$

$$f_{\mathrm{RLKum}}(x,\theta,\beta,\alpha) = \frac{\alpha\beta x^{\beta-1}}{\theta^2(\theta-1)}\left(1-x^{\beta}\right)^{(\alpha/\theta)-1}\cdot\left(\theta^2-2\theta-\alpha\log\left(1-x^{\beta}\right)\right), \tag{54}$$

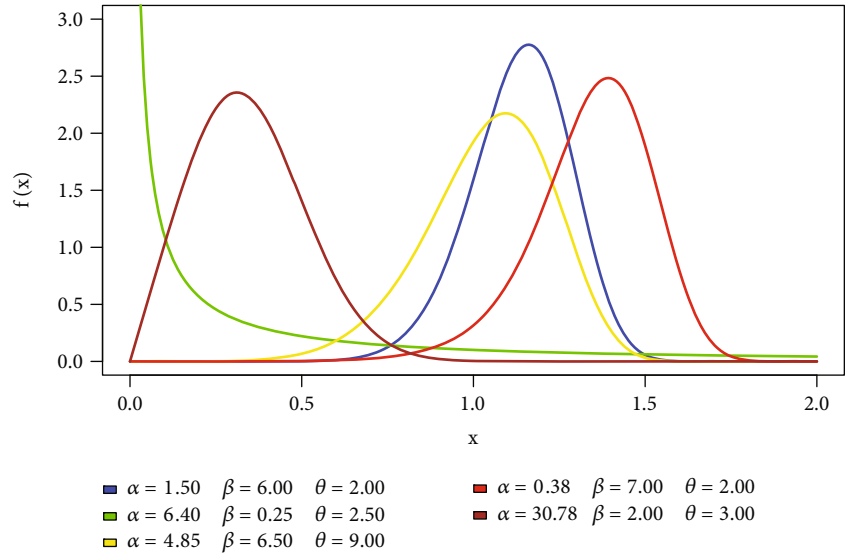

Figure 1: Plots of the PDF of the RLW distribution for different parameter values.

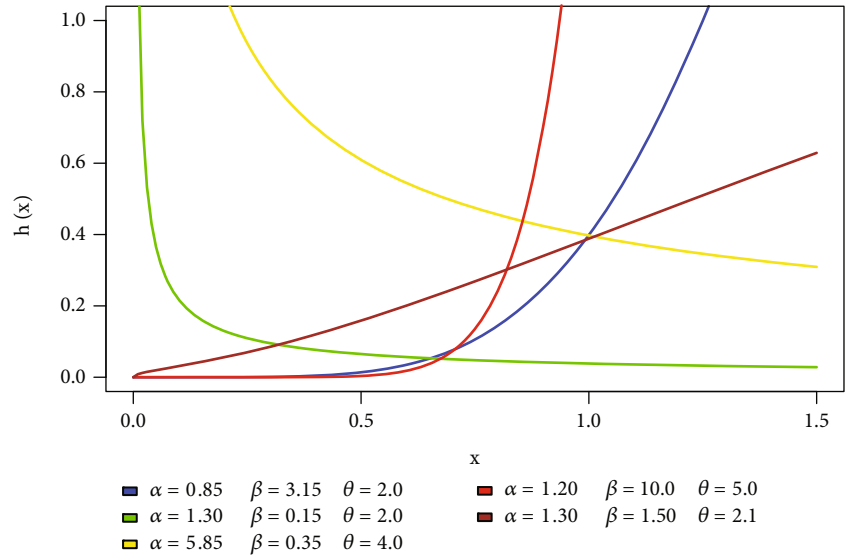

Figure 2: Plots of the hazard rate function of the RLW distribution for different parameter values.

$$h_{\mathrm{RLKum}}(x,\theta,\beta,\alpha) = \frac{\alpha\beta x^{\beta-1}\left(\theta^2 - 2\theta - \alpha\ \log\left(1-x^{\beta}\right)\right)}{\theta\left(1-x^{\beta}\right)\left(\theta(\theta-1) - \alpha\ \log\left(1-x^{\beta}\right)\right)}. \tag{55}$$

Figures 3 and 4, respectively, display plots of the RLKum PDF and hazard rate function at various selections of parameter values. From Figure 3, the RLKum distribution can take various forms, such as a "reversed J-shape," a "left-skewed" distribution, or a "J-shape." The hazard rate function plots in Figure 4 illustrate various shapes such as "decreasing," "increasing," "J-shape," "bathtub," and "inverted bathtub." Thus, the RLKum distribution is capable of modeling data with "non-monotonic" and "monotonic" hazard rate functions.

7. Monte Carlo Simulation

In this section, simulation analysis with sample sizes, $n = 50, 150, 200, 500, 800, 1000$, was performed to evaluate the properties of the ML estimators for the RLW distribution parameters by examining the average estimates (AV), the average bias (AB), and the root mean square (RMSE) for the estimated parameters. The analysis was repeated for $N = 1500$ times, with initial parameter values: (I) $\alpha = 2.6$, $\beta = 1.5$, and $\theta = 2.1$; and (II) $\alpha = 1.6$, $\beta = 0.5$, and $\theta = 2.1$. The random number generation is produced by solving the CDF of the RLW with the uniroot function in R software, and the estimations are obtained with the optim function in the same software. The AB, RMSE, and AV were estimated using the following expressions:

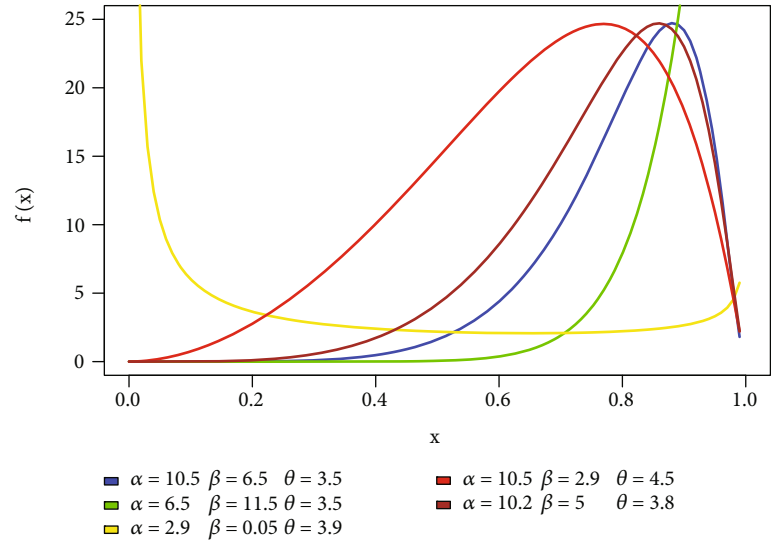

Figure 3: PDF plots of the RLKum distribution for different parameter values.

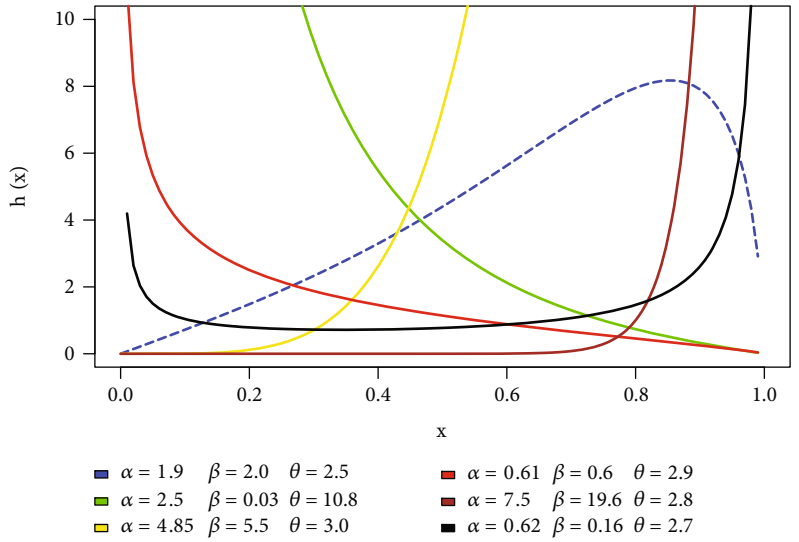

Figure 4: Plots of the RLKum hazard rate function for arbitrary parameter values.

RMSE $= \sqrt{(1/N)\sum_{i=1}^{N}(\widehat{\omega}_i - \omega)^2}$, AB $= (1/N)\sum_{i=1}^{N}(\widehat{\omega}_i - \omega)$, and $Av = (1/N)\sum_{i=1}^{N}\widehat{\omega}_i$, where $\omega = \alpha, \beta, \theta$.

Table 1 displays the simulated results of AB, RMSE, and AV for the parameter values of the RLW distribution. It can be observed that, in all cases, the AB and RMSE decrease to zero with increasing sample size. Furthermore, the AV of the estimators is quite close to the actual values. Hence, the maximum likelihood estimation and their asymptotic results perform well in estimating the RLW parameters. Similarly, alternative parameter choices can yield similar results.

8. Application

Application to three datasets of the RLW distribution is demonstrated in this section. The goodness-of-fit via Cramer-von Mises distance values (CVM), the Anderson-Darling statistic (AD), the Kolmogorov-Smirnov statistics (KS), and model selection criteria such as Bayesian information criteria (BIC), consistent Akaike information criteria (CAIC), and Akaike information criteria (AIC) of the RLW distribution, its nested models, and some other competing distributions were compared. In the first two applications, the RLW distribution was compared with its submodels, Nakagami (NAK) by [26]), inverse Weibull (INW) by [27], Nadarajah and Haghighi (NH) by [28], and modified extended Chen (MEC) by [29]. In the third application, the following nonnested models were used: Marshall-Olkin exponential (MOEx) by [3], generalized exponential (GE) by [30], and generalized inverse Weibull (GIW) by [31].

The CDF of the nonnested models are given below:

(i) Generalized Inverse Weibull: $F(x) = e^{(-\theta(\alpha/x)^{\beta})}$, $x > 0, \alpha > 0, \beta > 0$

Table 1: Simulated results of AB, RMSE, and AV for RLW distribution.

Parameter	n	$\alpha = 2.6, \beta = 1.5, \theta = 2.1$			$\alpha = 1.6, \beta = 0.5, \theta = 2.1$		
		AB	RMSE	Av	AB	RMSE	Av
α	50	3.7878	7.4370	55.7878	-2.3974	2.6659	29.6026
	150	1.8761	3.2108	18.7104	-0.8261	2.2767	15.1739
	200	0.6427	2.1650	13.6427	-0.4792	0.8368	7.5208
	500	0.4377	1.0995	8.6610	-0.0795	0.7929	3.1205
	800	0.1487	0.9620	2.7123	-0.0899	0.3987	1.9100
	1000	0.0775	0.0916	2.6776	-0.0853	0.2577	1.5146
β	50	7.4123	2.0823	37.4123	2.450	0.6921	12.4500
	150	3.9078	1.2336	11.5419	1.1450	0.4379	6.1449
	200	1.4904	0.8109	8.9904	0.5004	0.2696	3.0004
	500	0.5422	0.4669	2.0312	0.1513	0.1348	1.1513
	800	0.2271	0.2854	1.6795	0.0807	0.0915	0.7057
	1000	0.0839	0.0381	1.6839	0.0609	0.0787	0.5609
θ	50	27.1127	13.5530	69.1127	23.7908	12.8938	65.7908
	150	10.0151	7.2365	20.4318	13.5720	11.1659	34.5720
	200	6.1930	5.6416	16.6930	5.9125	5.0713	16.4125
	500	2.5672	3.5417	7.8732	2.1320	4.2120	6.3320
	800	0.9231	1.5608	3.5199	0.9522	2.3619	3.5773
	1000	0.0697	0.1241	2.1697	0.6582	1.6670	2.7583

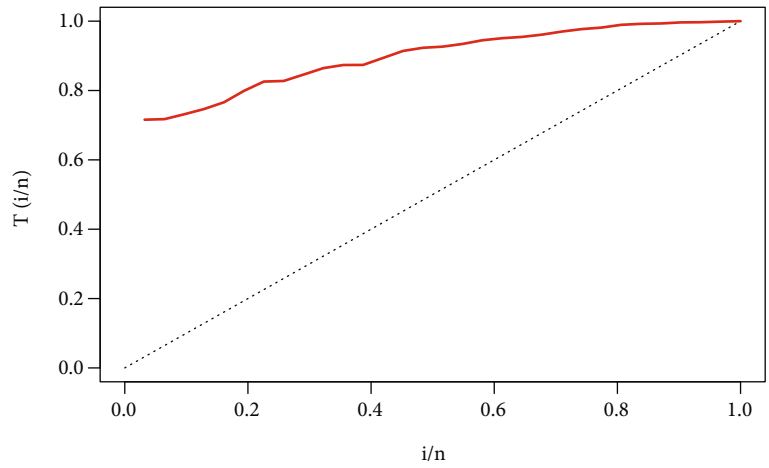

Figure 5: Plots of the TTT transform for the rainfall data.

(ii) Inverse Weibull: $F(x) = e^{-(-1/\theta(1/x-\alpha)^{\beta})}$, ,$x > 0, \alpha > 0, \beta > 0, \theta > 0$

(iii) Nadarajah and Haghighi distribution: $F(x) = 1 - e^{(1-(1+\alpha x)^{\beta})}$

(iv) Nakagami distribution: $F(x) = \gamma(\alpha, (\alpha/\beta)x^2)/\Gamma(\alpha)$, $x > 0, \alpha > 0, \beta > 0$

(v) Modified extended Chen distribution: $F(x) = [\theta(e^{x^{-\beta}} - 1) + 1]^{-\alpha}, x > 0, \alpha, \beta > 0\, \theta > 0$

(vi) Generalized exponential: $F(x) = (1 - e^{-\alpha x})^{\beta}, x > 0, \alpha > 0, \beta > 0$

(vii) Marshall-Olkin exponential: $F(x) = 1 - (\alpha e^{-\beta x}/(1 - (1 - \alpha)e^{-\beta x})), x > 0, \alpha > 0, \beta > 0$

8.1. Dataset 1: Rainfall Data. The information displays the highest annual average monthly rainfall (in inches) that was seen in Ghana's Ashanti region between 1989 and 2019. The dataset can be found in [32]. The dataset contains the following:

Table 2: ML estimates, p values, and standard errors of parameters for the rainfall data.

Distribution	Parameter	Estimate	Standard error	z-value	p value
RLW	$\widehat{\alpha}$	0.021088	0.047615	0.44289	0.6578
	$\widehat{\beta}$	6.55145	0.92802	7.0596	$<2.2\times10^{-12}$
	$\widehat{\theta}$	1.0×10^{5}	3.2035×10^{-8}	3.12×10^{12}	$<2.2\times10^{-16}$
GRL	$\widehat{\alpha}$	4.89283	0.073309	66.7426	$<2.2\times10^{-16}$
	$\widehat{\theta}$	1.0×10^{5}	3.0244×10^{-7}	3.31×10^{11}	$<2.2\times10^{-16}$
RL	$\widehat{\theta}$	8.3350	1.8249	4.5674	<0.0001
NAK	$\widehat{\alpha}$	100.3082	6.2502	16.0488	$<2.2\times10^{-16}$
	$\widehat{\beta}$	8.93087	2.0694	4.3157	<0.0001
MEC	$\widehat{\alpha}$	156.9674	132.3598	1.1859	0.2357
	$\widehat{\beta}$	4.6075	0.4665	9.8764	$<2.2\times10^{-16}$
	$\widehat{\theta}$	155.7753	135.3867	1.1506	0.2499
INW	$\widehat{\alpha}$	3.8818	1.8358	2.1145	0.0345
	$\widehat{\beta}$	3.1846	1.2936	2.4619	0.0138
	$\widehat{\theta}$	0.0064	0.0209	0.3067	0.7591
NH	$\widehat{\alpha}$	3.5585×10^{-3}	4.155×10^{-4}	8.5642	$<2.2\times10^{-16}$
	$\widehat{\beta}$	22.700	6.595×10^{-8}	3.442×10^{8}	$<2.2\times10^{-16}$

Table 3: Information criteria and goodness-of-fit statistics.

Model	ℓ	CVM	AD	KS	AIC	CAIC	BIC
RLW	-60.7947	0.0500	0.4276	0.1063	127.5894	128.5125	131.1914
GRL	-62.9598	0.0598	0.5801	0.1329	129.1957	129.6401	132.0637
RL	-101.6749	2.1434	10.0240	0.5220	205.3499	205.4928	206.7839
NAK	-60.8008	0.0558	0.4369	0.0902	128.4296	128.9616	131.4060
MEC	‘-64.7592	0.1100	0.8852	0.1699	135.5185	136.4416	139.8205
INW	-64.8004	0.1446	0.9720	0.1370	135.6009	136,5240	139.9029
NH	-92.2319	2.3334	10.6640	0.5308	188.4639	188.9083	191.3319

Table 4: LRT statistics for the rainfall data.

Model	Hypotheses	LRT	Critical value
GRL	$H_0 : \beta = 1$ vrs $H_1 : H_0$ is false	4.3302	3.841
RL	$H_0 : \alpha = \beta = 1$ vrs $H_1 : H_0$ is false	81.7604	5.991

12.469, 7.079, 11.929, 11.370, 12.906, 8, 7.394, 7.063, 12.213, 9.654, 8.327, 7.228, 10.689, 10.413, 10.039, 8.984, 10.508, 7.614, 12.165, 11.201, 8.988, 8.594, 10.961, 8.350, 9.882, 11.720, 10.272, 9.311, 8.854, 9.819, and 11.863. A graphical representation of the dataset using the hazard function is displayed in Figure 5. The total test on time (TTT) plot indicates that the curve has an increasing hazard rate.

Table 2 shows the ML estimates, standard errors, and p values of the parameters of the fitted distributions for the rainfall data. Two parameters of the RLW are statistically significant at the 5% significance level, except for the INW. The GRL, RL, NAK, and NH have all their estimated parameters statistically significant at the 5% significance level.

From Table 3, a better fit is provided by the RLW to the rainfall data compared to its submodels and the nonnested models because it has the maximum value of log-likelihood (ℓ) and the smallest CVM, AD, AIC, CAIC, and BIC. A close competitive model to the RLW is the NAK.

From the likelihood ratio test (LRT) results in Table 4, it is obvious that significant differences exist between RLW

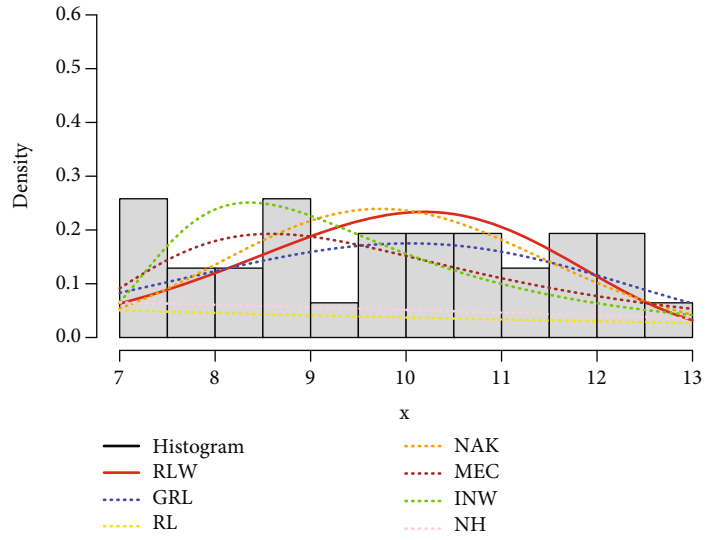

Figure 6: Fitted PDFs vs. histogram of the maximum annual rainfall data.

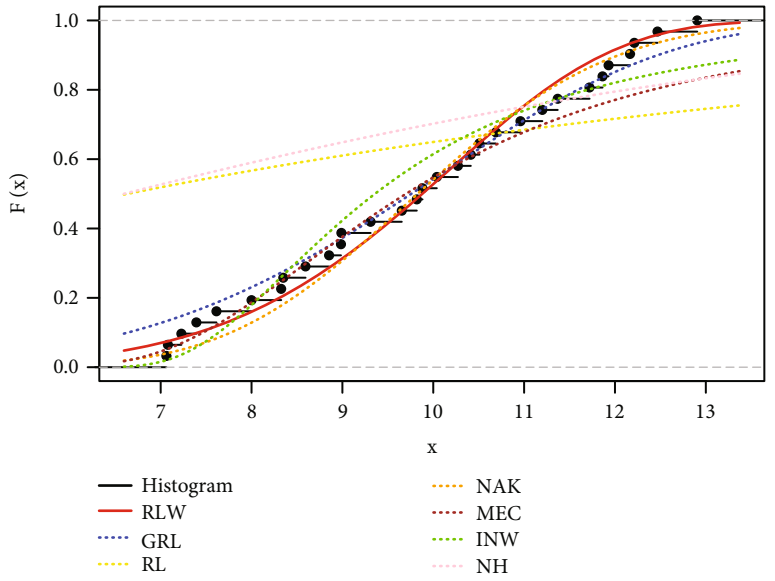

Figure 7: Empirical vs. fitted CDFs for the maximum annual rainfall data.

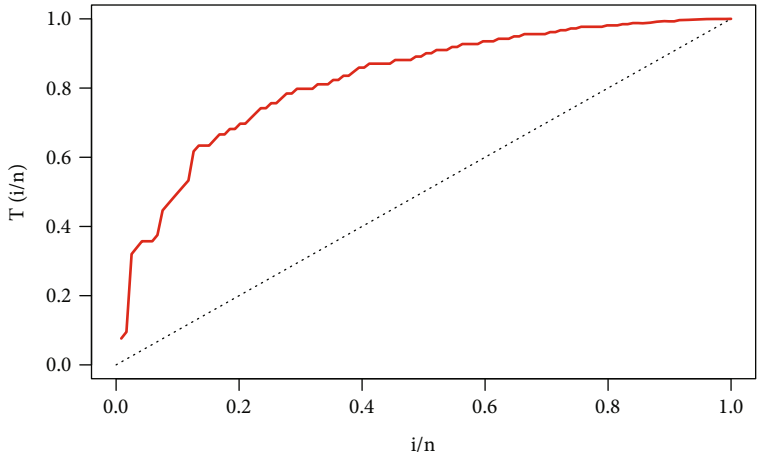

Figure 8: Plots of the TTT transform for the data.

Table 5: ML estimates, p values, and standard errors of parameters for the data.

Distribution	Parameter	Estimate	Standard error	z-value	p value
RLW	$\widehat{\alpha}$	2.73×10^{-2}	3.4093×10^{-2}	0.8006	0.4064
	$\widehat{\beta}$	3.7015	0.2894	12.7903	$<2 \times 10^{-16}$
	$\widehat{\theta}$	100000	1.9966×10^{-8}	5.009×10^{12}	$<2 \times 10^{-16}$
GRL	$\widehat{\alpha}$	1.7531	2.1340×10^{-2}	82.149	$<2 \times 10^{-16}$
	$\widehat{\theta}$	1.098×10^{3}	5.0791×10^{-6}	1.9882×10^{8}	$<2 \times 10^{-16}$
RL	$\widehat{\theta}$	51.3680	4.8074	10.685	$<2 \times 10^{-16}$
NAK	$\widehat{\alpha}$	1840.4489	123.0302	14.9593	$<2 \times 10^{-16}$
	$\widehat{\beta}$	1.4171	0.2182	6.4948	8.31510^{-11}
MEC	$\widehat{\alpha}$	3.9728	0.7949	4.9978	<0.001
	$\widehat{\beta}$	1.7714	0.0837	21.1678	$<2 \times 10^{-16}$
	$\widehat{\theta}$	197.938	71.1715	2.7811	0.005417
INW	$\widehat{\alpha}$	-4.6592	1.8744	-2.4857	0.01293
	$\widehat{\beta}$	1.5767	0.1538	10.2547	$<2 \times 10^{-16}$
	$\widehat{\theta}$	0.00254	0.0016	1.5530	0.12043
NH	$\widehat{\alpha}$	2.7387×10^{-3}	1.7333×10^{-4}	1.5801	$<2 \times 10^{-16}$
	$\widehat{\beta}$	5.6071	9.0352×10^{-8}	6.2058×10^{7}	$<2 \times 10^{-16}$

Table 6: Information criteria and goodness-of-fit statistics.

Model	ℓ	CVM	AD	KS	AIC	CAIC	BIC
RLW	-502.8805	0.0896	1.2456	0.073658	1011.761	1011.9697	1020.099
GRL	-538.491	3.3364	16.124	0.27391	1080.759	1080.8624	1086.317
RL	-590.1535	5.2102	0.36118	0.36118	1182.307	1182.3412	1185.086
NAK	-541.4105	7.7603	40.551	0.38797	1086.821	1086.9244	1092.379
MEC	-564.8485	2.1847	12.104	0.2873	1135.697	1135.9057	1144.035
INW	-575.887	2.5781	14.001	0.31293	1157.774	1157.9827	1166.112
NH	-562.205	5.4959	25.963	0.3626	1128.41	1128.5134	1133.968

Table 7: LRT statistics for the hypertension data.

Model	Hypotheses	LRT	Critical value
GRL	$H_0 : \beta = 1$ vrs $H_1 : H_0$ is false	71.221	3.841
RL	$H_0 : \alpha = \beta = 1$ vrs $H_1 : H_0$ is false	174.546	5.991

and its submodels based on the LRT test since their LRT statistics values are, respectively, greater than the critical values at the 5% level of significance.

The graphs of the fitted PDFs versus the histogram of the data are displayed in Figure 6, and the fitted CDFs versus the empirical data are displayed in Figure 7. It is noted that the plots of the densities of the RLW depict the empirical density and CDF of the maximum annual rainfall data more closely than the other models.

8.2. Dataset 2: Hypertension Data. This dataset shows the survival periods in years before the development of hypertension for 119 patients randomly selected from the Bolgatanga Regional Hospital in Ghana's Upper East region. The dataset is in [33], and it has the following items:

71, 5, 39, 62, 52,71, 38, 56, 35, 69,34,71,66,70, 52, 37, 35, 71, 73, 19, 74, 74, 75, 51, 76,49, 19, 76, 78, 76, 76, 49, 47, 48, 48, 46, 46, 46, 41, 40, 43, 45, 47, 47, 44, 45, 46, 42, 43, 42, 20, 28, 26, 60, 27, 24, 29, 60, 25, 60, 69, 36, 69, 69, 68, 68, 67, 67,

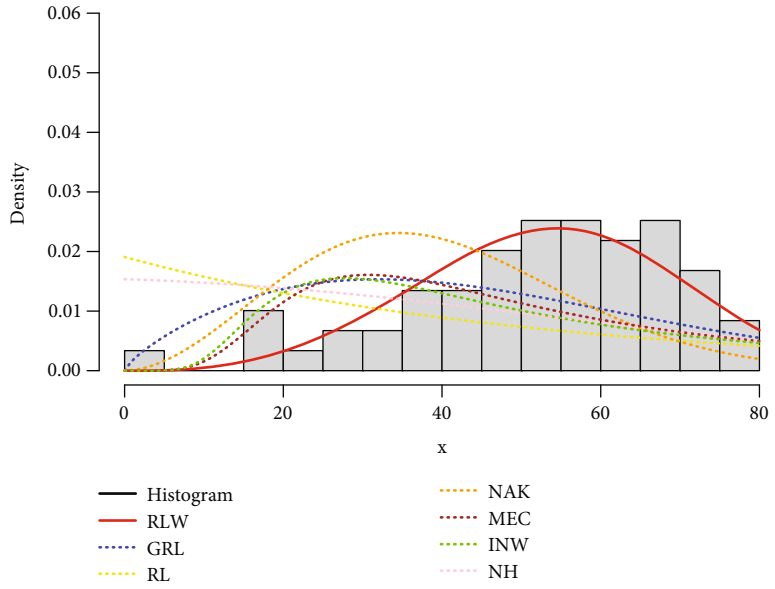

Figure 9: Fitted PDFs vs. histogram of the hypertension data.

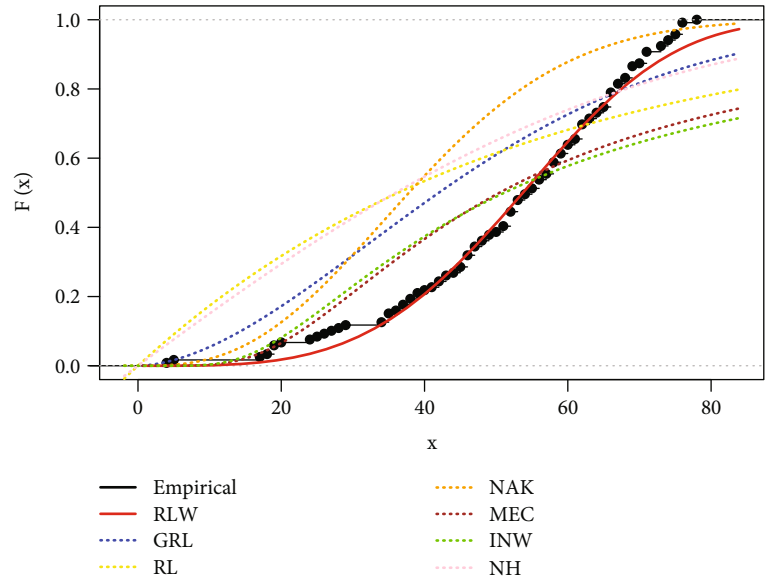

Figure 10: Empirical vs. fitted CDFs of the hypertension data.

67, 52, 35, 66, 55, 66, 61, 61, 64, 64, 65, 65, 63, 63, 62, 39, 62, 62, 62, 59, 59, 59, 58, 58, 58, 18, 57, 57, 56, 56, 37, 53, 53, 53, 53, 54, 54, 66, 17, 50, 75, 51, 38, 52, 66, 4, 52, 55, 19, 58, and 73. Figure 8 shows a graphic depiction of the dataset using the hazard function. The RLW distribution can therefore be used to represent the curve because the TTT plot shows that the hazard rate is growing.

The parameter estimates for the fitted models for the hypertension data are shown in Table 5, along with their standard errors and p values. In addition, the estimated parameters for the GRL, RL, and NH models are all statistically significant at the 5% level of significance. Two other estimated parameters of the RLW are also significant at this level. Based on Table 6, the RLW distribution offers a better match to the hypertension data compared to its nested models and the other distributions since it has the least CVM, AD, KS, AIC, and CAIC as well as the highest log-likelihood value.

It is clear from Table 7's likelihood ratio test (LRT) results that there are significant differences between RLW and its submodels based on the LRT test, as each of their LRT statistical values exceeds the critical values at the 5% level of significance.

Figures 9 and 10, respectively, are graphs showing the fitted PDFs against the data's histogram and the fitted CDFs against the empirical data. It should be observed that the RLW plots of densities more accurately represent the empirical density and CDF of the hypertension data than the other models.

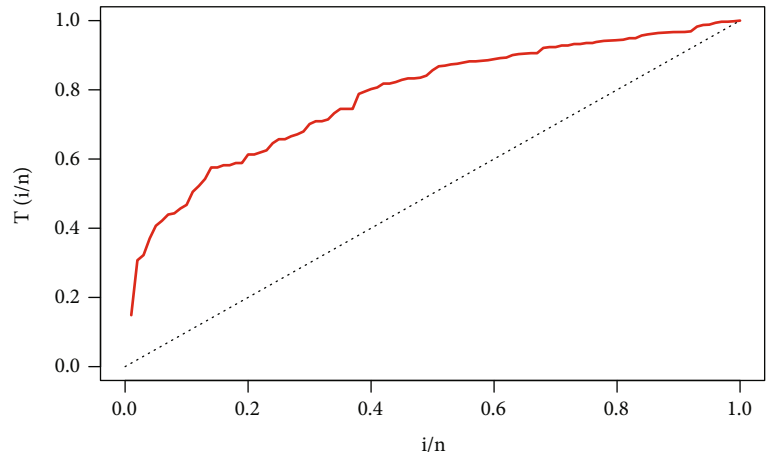

Figure 11: TTT transform plot of the carbon fiber data.

Table 8: ML estimates, standard errors, and p values of parameters for the carbon fiber data.

Model	Parameter	Estimate	Standard error	z-value	p value
RLW	$\widehat{\alpha}$	2.8205	0.7956	3.5451	<0.001
	$\widehat{\beta}$	2.7922	0.2141	13.0441	<0.001
	$\widehat{\theta}$	56.4575	0.0390	1448.397	<0.001
GRL	$\widehat{\widehat{\beta}}$	2.7880	0.2162	12.8956	<0.001
	$\widehat{\theta}$	19.2252	5.8102	3.3089	<0.001
RLRa	$\widehat{\alpha}$	151.5764	14.9180	10.1610	<0.001
	$\widehat{\theta}$	1194.8486	1.8909	631.901	<0.001
GE	$\widehat{\alpha}$	1.0132	0.0875	11.5824	<0.001
	$\widehat{\beta}$	7.7883	1.4962	5.2054	<0.001
INW	$\widehat{\alpha}$	-1.6811	0.6092	-2.7595	<0.01
	$\widehat{\beta}$	3.8282	0.7188	5.3257	<0.001
	$\widehat{\theta}$	0.0066	0.0103	0.6404	0.5219
GIW	$\widehat{\alpha}$	0.5217	0.0433	12.0440	<0.001
	$\widehat{\beta}$	1.7690	0.1119	15.8060	<0.001
	$\widehat{\theta}$	9.764	0.0013	7464.547	<0.001

Table 9: Information criteria and goodness-of-fit statistics.

Model	ℓ	CVM	AD	KS	AIC	CAIC	BIC
RLW	-141.5308	0.06334	0.4179	0.0605	289.0615	289.3089	296.8770
GRL	-141.5417	0.06340	0.4195	0.0607	289.1934	289.3159	296.2937
RLRa	-149.5009	0.6340	3.5460	0.1383	303.0018	303.1242	308.2121
GE	–146.1823	0.2292	1.2246	0.1077	296.3646	296.4870	301.5749
INW	-154.1823	0.3876	2.4332	0.1286	314.3646	314.6120	322.1801
GIW	-173.1440	0.8875	5.3496	0.1777	352.2879	352.5353	360.1034

Table 10: Likelihood ratio test statistics for the carbon fiber data.

Model	Hypotheses	LRT	Critical value
GRL	$H_0 : \beta = 1$ vrs $H_1 : H_0$ is false	0.0218	3.841
RLRa	$H_0 : \beta = 2$ vrs $H_1 : H_0$ is false	7.9701	3.841

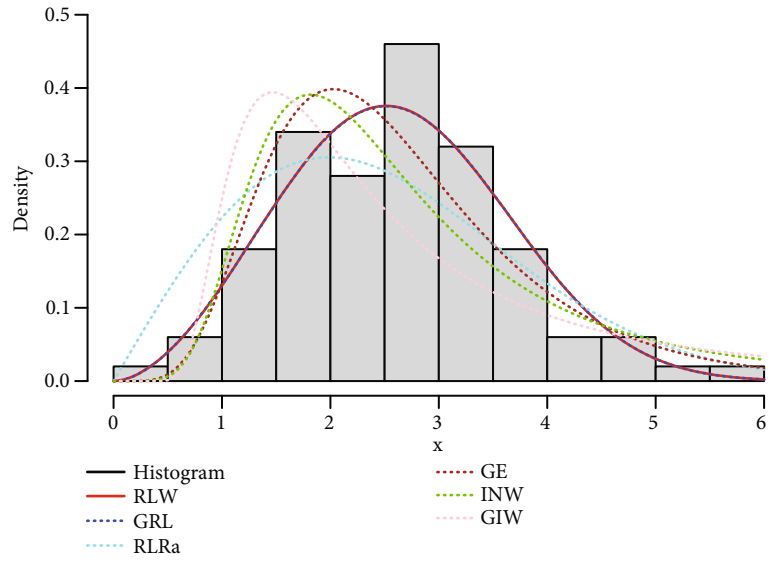

Figure 12: Fitted PDFs vs. histogram of the carbon fiber data.

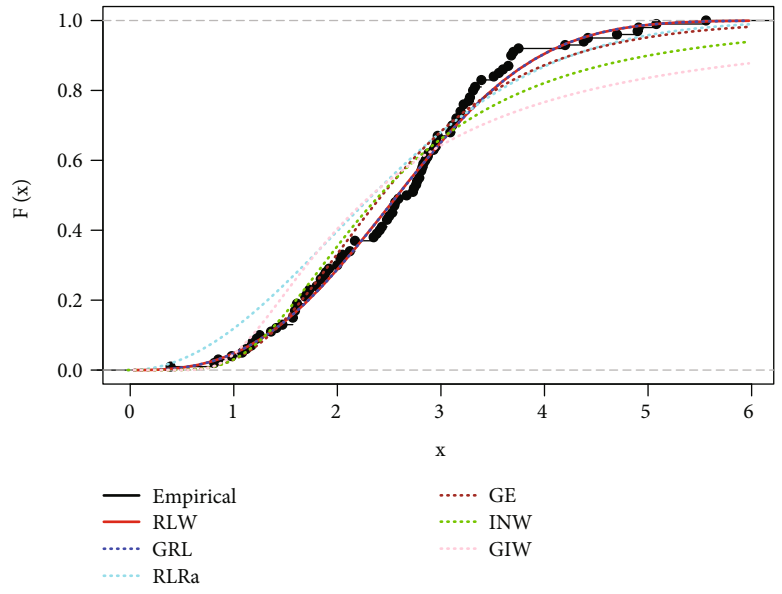

Figure 13: Empirical vs. fitted CDFs of the carbon fiber data.

8.3. Dataset 3: Carbon Fiber Data. The third application of the RLW with other competing models and its submodels is demonstrated in this section with the breaking stress of carbon fibers. It consists of the breaking stress of 50 mm-long carbon fibers. The dataset is in [34], and it includes the following:

3.70, 2.74, 2.73, 2.5, 3.6, 3.11, 3.27, 2.87, 1.47, 3.11, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56, 3.15, 2.35, 2.55, 2.59, 2.38, 2.81, 2.77, 2.17, 2.83, 1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59, 3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2.00, 1.22, 1.12, 1.71, 2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.80, 1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, and 3.65

A graphical representation of the dataset using the hazard function is displayed in Figure 11. The plot of the TTT indicates that the curve has an increasing hazard rate and hence can be modeled using the RLW distribution and the other competing models.

Table 8 exhibits the maximum likelihood estimates, standard errors, and p values of the parameters of the fitted models for the carbon fiber data. The estimated parameters of the RLW and the other distributions are significant at the 5% level of significance, except for one parameter of the INW. From Table 9, the RLW distribution provides a better fit to the carbon fiber data compared to its nested models and the other distributions because it has the smallest CVM, AD, KS, AIC, and CAIC and the greatest value of the log-likelihood.

The LRT results of Table 10 indicate that there is no significant difference between the RLW and the GRL distribution since the LRT statistics values are less than the critical values at the 5% level of significance. On the other hand, there is a significant difference between the RLW and the RLRa distribution.

The histogram of the data against the PDFs of the fitted models and the fitted CDFs vs. the empirical carbon fiber data are, respectively, exhibited in Figures 12 and 13. It is observed from the plots that the RLW densities depict the empirical density and CDF of the carbon fiber data more closely than the other models.

9. Conclusion

In reality, the lack of developed family of distributions based on the RL distribution in the scientific literature served as motivation for our study. In this paper, a new family of distributions, the RL-G family, with its statistical properties such as the quantile function, the raw moments, the incomplete moments, measures of inequality, entropy, mean and median deviations, and the reliability parameter, is studied. The parameters of the proposed generator were estimated using the ML estimation method. The RLW and the RLKum are two special members of the RL-G family. The outcome of the simulation analysis indicates that the ML estimation method and its asymptotic properties performed quite well. Applications of the RLW from the RL-G family were carried out on three complete real datasets, and it is evident that the RLW outperformed its submodels and other distributions. As a result, the newly suggested family of distributions has a broader range of applications in a variety of areas. Despite the RL-G model's numerous advantages, such as flexibility, generalization of the RL distribution, and the ability to provide superior fits to the dataset in comparison to other compared models available in the literature, it cannot be employed for assessing discrete datasets, and expressions of its estimators are difficult to reduce to a simple, closed-form.

Appendix

A. Proof of Proposition

A.1. Proof

(i) From the CDF of the RL-G family in (7), as $x \longrightarrow 0$, $G(x;\epsilon) \longrightarrow 0$ and $F_{\mathrm{RL-G}}(x,\theta,\epsilon) \longrightarrow 1-(1-(\log(1-0)/\theta(\theta-1)))(1-0)^{1/\theta} \longrightarrow 0$, and as $x \longrightarrow \infty$, $G(x;\epsilon) \longrightarrow 1$ and $F_{\mathrm{RL-G}}(x,\theta,\epsilon) \longrightarrow 1-(1-(\log(1-1)/\theta(\theta-1)))(1-1)^{1/\theta}$ thus $F_{\mathrm{RL-G}}(x,\theta,\boldsymbol{\epsilon}) \longrightarrow 1$.

Hence, $0 \le F_{\mathrm{RL-G}}(x,\theta,\epsilon) \le 1$, for all $x > 0$

(ii) Demonstrating that the integration over the support $(0,\infty)$ is 1, that is

$$\int_0^\infty f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx = \int_0^\infty \frac{g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta)-1}}{\theta^2(\theta-1)} \cdot \left(\theta^2 - 2\theta - \log(1-G(x;\epsilon))\right)dx. \tag{A.1}$$

Let $y = -\log(1-G(x;\boldsymbol{\epsilon})) \Rightarrow (1-G(x;\boldsymbol{\epsilon})) = e^{-y}$, as $x \longrightarrow 0, y \longrightarrow 0$ and

$$\text{as } x \longrightarrow \infty, y \longrightarrow \infty, \frac{dy}{dx} = \frac{g(x,\boldsymbol{\epsilon})}{1-G(x;\boldsymbol{\epsilon})}, dx = \frac{(1-G(x;\boldsymbol{\epsilon}))dy}{g(x,\boldsymbol{\epsilon})}. \tag{A.2}$$

Thus

$$\begin{aligned}\int_0^\infty f_{\mathrm{RL-G}}(x;\theta,\epsilon)dx &= \int_0^\infty \frac{g(x,\epsilon)(1-G(x;\epsilon))^{(1/\theta)-1}}{\theta^2(\theta-1)} \cdot \left(\theta^2 - 2\theta - \mathrm{y}\right)\frac{(1-G(x;\epsilon))dy}{g(x,\epsilon)} \\ &= \int_0^\infty \frac{\left(\theta^2 - 2\theta - \mathrm{y}\right)e^{-(y/\theta)}}{\theta^2(\theta-1)}dy = 1.\end{aligned} \tag{A.3}$$

The function in the integral is the PDF of the RL distribution. Thus, from the above discussions, the RL-G family is a legitimate PDF for the continuous random variable X.

Data Availability

The (rainfall, hypertension, and carbon fiber) data used for this study is duly cited and is available in the text.

Conflicts of Interest

The authors declare that they have no conflicting interests.

Authors' Contributions

All authors have significantly contributed to the research and preparation of the manuscript.

References

[1] C. Lee, F. Famoye, and A. Y. Alzaatreh, "Methods for generating families of univariate continuous distributions in the recent decades," *WIREs Computational Statistics*, vol. 5, no. 3, pp. 219–238, 2013.

[2] G. S. Mudholkar and D. K. Srivastava, "Exponentiated Weibull family for analyzing bathtub failure-rate data," *IEEE Transitions on Reliability*, vol. 42, no. 2, pp. 299–302, 1993.

[3] A. W. Marshall and I. Olkin, "A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families," *Biometrika*, vol. 84, no. 3, pp. 641–652, 1997.

[4] G. M. Cordeiro and M. de Castro, "A new family of generalized distributions," *Journal of Statistical Computation and Simulation*, vol. 81, no. 7, pp. 883–898, 2011.

[5] N. Eugene, C. Lee, and F. Famoye, "Beta-normal distribution and its applications," *Communucation in Statistics-Theory and Methods*, vol. 31, no. 4, pp. 497–512, 2002.

[6] A. Alzaatreh and I. Ghosh, "On the Weibull-X family of distributions," *Journal of Statistical Theory and Applications*, vol. 14, no. 2, pp. 169–183, 2015.

[7] M. Bourgignon, R. B. Silva, and G. M. Cordeiro, "The Weibull-G family of probability distributions," *Journal of Data Sciences*, vol. 12, no. 1, pp. 53–68, 2014.

[8] G. M. Cordeiro, M. M. Otega Edwin, V. Popovic Bozidor, and R. R. Pescim, "The Lomax generator of distributions: properties, minification process and regression model," *Applied Mathematics and Computations*, vol. 247, pp. 465–486, 2014.

[9] W. Bodhisuwan and Y. Sangsanit, "The Topp-Leonn generator of distributions: properties and inferences," *Songklanakarin Journal of Science and Technology*, vol. 38, no. 5, pp. 537–548, 2016.

[10] S. Cakmakyapan and G. Ozel, "Lindley family of distributions: properties and applications," *Hacettepe Journal of Mathematics and Statistics*, vol. 46, no. 6, pp. 1113–1137, 2017.

[11] L. Anzagra, S. Sarpong, and S. Nasiru, "Chen-G class of distributions," *Cogent Mathematics and Statistics*, vol. 7, no. 1, 2020.

[12] F. Chipepa, B. Oluyede, and P. O. Peter, "The Burr III-Topp-Leone-G family of distributions with applications," *Heliyon*, vol. 7, no. 4, article e06534, 2021.

[13] F. Jamal, M. A. Nasir, M. H. Tahir, and N. H. Montazeri, "The odd Burr-III family of distributions," *Journal of Statistics Applications and Probability*, vol. 6, no. 1, pp. 105–122, 2017.

[14] F. Jamal, M. H. Tahir, M. Alizadeh, and M. A. Nasir, "On Marshall-Olkin Burr X family of distribution," *Tbilisi Mathematical Journal*, vol. 10, no. 4, pp. 175–199, 2017.

[15] H. Reyad, M. Alizadeh, F. Jamal, and S. Othman, "The Topp Leone odd Lindley-G family of distributions: properties and applications," *Journal of Statistics and Management Systems*, vol. 21, no. 7, pp. 1273–1297, 2018.

[16] P. L. Ramos and F. Louzada, "A distribution for instantaneous failures," *Stat*, vol. 2, pp. 247–256, 2019.

[17] H. Al-Mofleh, A. Z. Afify, and N. A. Ibrahim, "A new extended two-parameter distributions: properties, estimation methods, and applications in medicine and geology," *Mathematics*, vol. 8, no. 1578, 2020.

[18] A. S. Eldeeb, M. Ahsan-ul-Haq, and M. Eliwa, "A discrete Ramos-Louzada distribution for asymmetric and over-dispersed data with leptokurtic-shaped: properties and various estimation techniques with inference," *AIMS Mathematics*, vol. 7, no. 2, pp. 1726–1741, 2022.

[19] A. Alzaatreh, C. Lee, and F. Famoye, "A new method for generating families of continous distributions," *Metron*, vol. 71, pp. 63–79, 2013.

[20] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and G. E. Knuth, "On the Lambert W function," *Advances in Computation Mathematics*, vol. 5, pp. 329–359, 1996.

[21] A. Rényi, "On measures of entropy and information," *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, contributions to the theory of statistics*, , pp. 547–561, University of California Press, 1961.

[22] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[23] J. Havrda and F. Charvat, "Quantification method of classification processes: concept of structural a-entropy," *Kybernetika*, vol. 3, no. 1, p. 30, 1967.

[24] C. Tsallis, "Possible generalization of Boltzmann Gibbs statistics," *Journal of Statistical Physics*, vol. 52, no. 1-2, pp. 479–487, 1988.

[25] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, New York, 2007.

[26] M. Nakagami, "The m-distributions-a general formula of intensity distribution of rapid fading," *Proceedings of a symposium held at the University of California*, , pp. 7–36, Science-Direct, California, 1960.

[27] F. G. Akgul, B. Seno'glu, and T. Arslan, "An alternative distribution to Weibull for modeling the wind speed data: inverse Weibull distribution," *Energy Conversion and Management*, vol. 114, pp. 234–240, 2016.

[28] S. Nadarajah and F. Haghighi, "An extension of the exponential distribution," *A Journal of Theoretical and Applied Statistics*, vol. 45, no. 6, pp. 543–558, 2011.

[29] Y. A. Anafo, I. Brew, and S. Nasiru, "Modified extended Chen distribution: properties and applications," *Applied Mathematics and Information Sciences*, vol. 16, no. 5, pp. 711–728, 2022.

[30] R. D. Gupta and D. Kundu, "Generalised exponential distribution: different method of estimations," *Journal of Statistical Computation and Simulation*, vol. 69, no. 4, pp. 315–337, 2001.

[31] F. R. De Gusmao, E. M. Ortega, and G. M. Cordeiro, "The generalized inverse Weibull distribution," *Statistical Papers*, vol. 52, no. 3, pp. 591–619, 2011.

[32] J. M. Camarillo-Naranjo, J. I. Alvarez-Francoso, N. Limones-Rodriguez, and M. Pita-Lopez, "The global climate monitor system: from climate data-handling to knowledge dissemination," *International Journal of Digital Earth*, vol. 12, no. 4, pp. 394–414, 2019.

[33] E. Zamanah, S. Nasiru, A. Luguterah, and A. Luguterah, "Harmonic mixture Weibull-G family of distributions: properties, regression and applications to medical data," *Computational and Mathematical Methods*, vol. 2022, Article ID 2836545, 24 pages, 2022.

[34] M. Ijaz, *The Development of New Life-Time Distributions and Investigation of Their Statistical Properties, [Ph.D. Thesis]*, University of Peshwar, 2020, http://prr.hec.gov.pk/jspui/handle/123456789/16633.