การแจกแจงการิมากำลัง-เรย์ลี: คุณสมบัติและการประยุกต์ใช้

อิศราภรณ์ เทียมศร[1], เอกภักดิ์ ตันประยูร[2] และ ศิรินภา อายุยืน[3,*]

Received: 14 November 2022; Revised: 15 January 2023; Accepted: 24 January 2023

บทคัดย่อ

งานวิจัยนี้ ได้นำเสนอตัวแบบช่วงชีวิตที่อยู่ในวงศ์ของตัวแบบวางนัยการิมากำลัง ซึ่งเรียกว่าการแจกแจงการิมากำลัง-เรย์ลี มีการศึกษาคุณสมบัติทางสถิติของการแจกแจงนี้ ได้แก่ ฟังก์ชันควอนไทล์ ฟังก์ชันอยู่รอด โมเมนต์ และสถิติอันดับ มีการประมาณค่าพารามิเตอร์ของการแจกแจกแจงการิมากำลัง-เรย์ลีด้วยวิธีภาวะน่าจะเป็นสูงสุด รวมทั้งมีการประยุกต์ใช้การแจกแจงการิมากำลัง-เรย์ลีกับข้อมูลจริง ซึ่งเป็นข้อมูลช่วงชีวิต จำนวน 2 ชุด ผลการศึกษาพบว่า การแจกแจงการิมากำลัง-เรย์ลี เป็นตัวแบบที่อธิบายลักษณะการแจกแจงความน่าจะเป็นของค่าสังเกตได้ใกล้เคียงกับข้อมูลจริง นั่นคือการแจกแจงนี้มีความยืดหยุ่นและครอบคลุมกับข้อมูลช่วงชีวิตมากกว่าการแจกแจงอื่น ๆ ที่เกี่ยวข้อง เช่น การแจกแจงเรย์ลี การแจกแจงการิมากำลัง-ลินด์เลย์ และการแจกแจงลินด์เลย์

คำสำคัญ: การแจกแจงการิมากำลัง-เรย์ลี, ข้อมูลช่วงชีวิต, การประมาณค่าภาวะน่าจะเป็นสูงสุด, ฟังก์ชันอยู่รอด

[*] *Corresponding Author Email : sirinapa_a@rmutt.ac.th*

[1,2,3] ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

Power Garima-Rayleigh distribution: properties and applications

Issaraporn Thiamsorn[1], Ekapak Tanprayoon[2] and Sirinapa Aryuyuen[3,*]

Received: 14 November 2022; Revised: 15 January 2023; Accepted: 24 January 2023

Abstract

This paper considers a new lifetime model of the class of power Garima generalized models, called the power Garima-Rayleigh (PG-R) distribution. Some statistical properties are provided, including quantile function, hazard function, moments, and order statistics. The maximum likelihood estimation is used to estimate the parameters of the PG-R distribution. In addition, we apply the PG-R distribution with two real-life data sets. The study found that the PG-R distribution is a model that describes the probability distribution of observation values close to the real data. The proposed distribution is more flexible and comprehensive with lifetime data than other related distributions, such as the Rayleigh, power Garima-Lindley, and Lindley distributions.

Keywords: Power Garima-Rayleigh distribution, lifetime data, maximum likelihood estimation, hazard function

[*] *Corresponding Author Email : sirinapa_a@rmutt.ac.th*

[1,2,3] Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi

1. Introduction

Many new statistical distributions have been derived using the commonly known distributions through different types of transformations, compounding, or mixing. Flexible distributions are obtained from extending the classical distributions by introducing one or more additional parameter(s) to the baseline distribution. Many generalized families of distributions have been proposed and studied for modelling data in many applied areas such as economics, engineering, biological studies, environmental sciences, medical sciences, and finance (Gupta et al., 1998). Some of the generated (G) families are: the beta-G (Eugene et al., 2002), gamma-G (Zografos & Balakrishnan, 2009; Ristic & Balakrishnan, 2012), transformed-transformer (T-X; Alzaatreh et al., 2013), Weibull-G (Bourguignon et al., 2014), exponentiated half-logistic-G (Cordeiro et al., 2014), half logistic-G family (Cordeiro et al., 2016; Soliman et al., 2017), exponentiated Weibull-G (Hassan & Elgarhy, 2016), generalized transmuted-G (Nofal et al., 2017), Gompertz-G (Alizadeh et al., 2017), beta Weibull-G (Yousof et al., 2017), new Weibull-G (Ahmad et al., 2018), and transmuted Gompertz-G (Reyad et al., 2018).

Recently, Aryuyuen et al. (2021) proposed a power Garima-G (PG-G) family of distributions, which is obtained by using the concept of T-X family by Alzaatreh et al. (2013) when a generator T distributed as the power Garima (PG) distribution. The cumulative density function (cdf) and the probability density function (pdf) of a random variable X distributed as the PG-G family are represented, respectively.

$$\mathrm{F}(x;\alpha,\beta,\lambda)=1-\left[1+\frac{\beta}{2+\beta}\left(\frac{\mathrm{G}^{\alpha}(x;\xi)}{1-\mathrm{G}^{\alpha}(x;\xi)}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{\mathrm{G}^{\alpha}(x;\xi)}{1-\mathrm{G}^{\alpha}(x;\xi)}\right)^{\lambda}\right\},\ x>0 \tag{1}$$

and

$$\mathrm{f}(x;\alpha,\beta,\lambda)=\frac{\alpha\beta\lambda\,\mathrm{g}(x;\xi)\,\mathrm{G}^{\alpha-1}(x;\xi)}{(2+\beta)\left[1-\mathrm{G}^{\alpha}(x;\xi)\right]^{2}}\left[(1+\beta)+\beta\left(\frac{\mathrm{G}^{\alpha}(x;\xi)}{1-\mathrm{G}^{\alpha}(x;\xi)}\right)^{\lambda}\right]$$
$$\times\left(\frac{\mathrm{G}^{\alpha}(\mathrm{x};\xi)}{1-\mathrm{G}^{\alpha}(\mathrm{x};\xi)}\right)^{\lambda-1}\exp\left\{-\beta\left(\frac{\mathrm{G}^{\alpha}(\mathrm{x};\xi)}{1-\mathrm{G}^{\alpha}(\mathrm{x};\xi)}\right)^{\lambda}\right\}, \tag{2}$$

where $\alpha,\beta,\lambda>0,$ and $\mathrm{G}(x;\xi)$ is the cdf of any existing (baseline) distribution with a parameter vector ξ.

In practice, the PG-G family of distributions: (a) is the primary method of introducing an additional parameter(s) to generate an extended version of the baseline distribution; (b) improve the characteristics of the traditional distributions; (c) make the kurtosis more flexible compared to the baseline distribution; (d) generated distributions that have the pdf with various shapes, i.e., symmetric, right-skewed, left-skewed, and reversed-J shaped; (e) define special models with all types of hazard rate function; (f) define special models having a closed-form for cdf, survival function as well as hazard rate function; (g) provide consistently better fits than other generated distributions having the same or higher number of parameters. Thus, we are interested in constructing a new distribution to analyse lifetime data using the PG-G family of distributions using some baseline distribution. This study uses a Rayleigh distribution as the baseline distribution (Aryuyuen et al., 2021).

Rayleigh (1880) derived the Rayleigh distribution; it was introduced in connection with a problem in the field of acoustics. The cdf and pdf of the Rayleigh distribution are given by

$$\mathrm{G}(x;\sigma)=1-e^{-x^{2}/(2\sigma^{2})}\ \text{and}\ \mathrm{g}(x;\sigma)=\frac{1}{\sigma^{2}}xe^{-x^{2}/(2\sigma^{2})}\ \text{for}\ x>0\ \text{and}\ \sigma>0\,. \tag{3}$$

In nature, physical phenomena in many fields of science (for example, noise theory, lethality, radar return, and others) have amplitude distributions that the Rayleigh density function can characterize. Since then, extensive work has occurred related to this distribution in different areas of science and technology. It has some relations with well-known distributions like the Weibull, chi-square, or extreme value distributions. The hazard function of the Rayleigh distribution is an increasing function of time (Rayleigh, 1880; Dey et al., 2014; Al-Babtain, 2020). Statistical inference of one parameter Rayleigh distribution, such as parameter estimation, predictions, and testing, have been extensively studied by several authors (See Johnson et al., 1994; Abd-Elfattah et al., 2006; Dey & Das, 2007; Dey, 2009; Al-Babtain, 2020).

This paper proposes a new lifetime model of the class of power Garima generalized models called the power Garima-Rayleigh distribution. Some statistical properties of the proposed distribution are provided. The algorithm for generating a random number from the proposed distribution is shown. We estimate the model parameters using the maximum likelihood (ML) method. Moreover, applications are illustrated. Finally, we offer some concluding remarks.

2. Methods

The methods of this research are:

2.1 The cdf and pdf of the proposed distribution and its mathematical properties are studied.

2.2 The appropriate parameter estimation method for the proposed distribution using the ML method is provided.

2.3 Application of the proposed distribution to real data sets has been studied by comparing it to some existing distributions using the criteria of Akaike information criterion (AIC), corrected Akaike information criterion (CAIC), and Kolmogorov-Smirnov (D*) statistics.

3. Results and Discussion

This section proposes a new distribution, namely the power Garima-Rayleigh distribution. We provided some statistical properties, parameter estimation, and applications of the proposed distribution. The remaining part of this article is presented as follows:

3.1 The power Garima-Rayleigh distribution

Let X be a random variable with the cdf of the PG-G family of distributions as in (1), when $\mathrm{G}(x;\sigma)$ is the cdf of the Rayleigh distribution as (3). Then the cdf of X is

$$\mathrm{F}(x;\boldsymbol{\theta})=1-\left[1+\frac{\beta}{2+\beta}\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right\},\ x>0 \tag{4}$$

where $\boldsymbol{\theta}=(\alpha,\beta,\lambda,\sigma)^T$ and $\alpha,\beta,\lambda,\sigma>0.$ Its corresponding pdf is

$$\mathrm{f}(x;\boldsymbol{\theta})=\frac{\alpha\beta\lambda x e^{-x^2/(2\sigma^2)}}{(2+\beta)\sigma^2}\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha-1}}{\left(1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}\right)^2}\left[(1+\beta)+\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]$$
$$\times\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda-1}\exp\left\{-\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right\}. \tag{5}$$

We have a random variable X with the cdf as (4) and pdf as (5) distributed as the power Garima-Rayleigh (PG-R) distribution with parameters α, β, λ and σ, denoted by $X \sim \mathrm{PG\text{-}R}\,(\alpha,\beta,\lambda,\sigma)$. Some pdf plots of PG-R distribution with different values of α, β, λ and σ are shown in Figure1.

3.2 Some statistical properties

We derive statistical properties of the PG-R distribution, including survival and hazard functions, moments, quantile function, skewness and kurtosis, algorithm for generating random variables, and order statistics.

3.2.1 Survival and hazard functions

The survival function $\mathrm{S}(x)$, is the probability that a subject survives longer than time x, $\mathrm{S}(x)=1-\mathrm{F}(x).$ The survival function of the PG-R distribution is

$$\mathrm{S}(x;\boldsymbol{\theta})=\left[1+\frac{\beta}{2+\beta}\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right\},\ x>0\,. \tag{6}$$

From the expression of the hazard function $\mathrm{h}(x)$, $\mathrm{h}(x)=\mathrm{f}(x)/\mathrm{S}(x)$, we have the hazard function of the PG-R distribution, that is

$$\mathrm{h}(x;\boldsymbol{\theta})=\frac{\alpha\beta\lambda x e^{-x^2/(2\sigma^2)}}{(2+\beta)\sigma^2}\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha-1}}{\left(1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}\right)^2}\left[(1+\beta)+\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]$$
$$\times\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda-1}\left[1+\frac{\beta}{2+\beta}\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}. \tag{7}$$

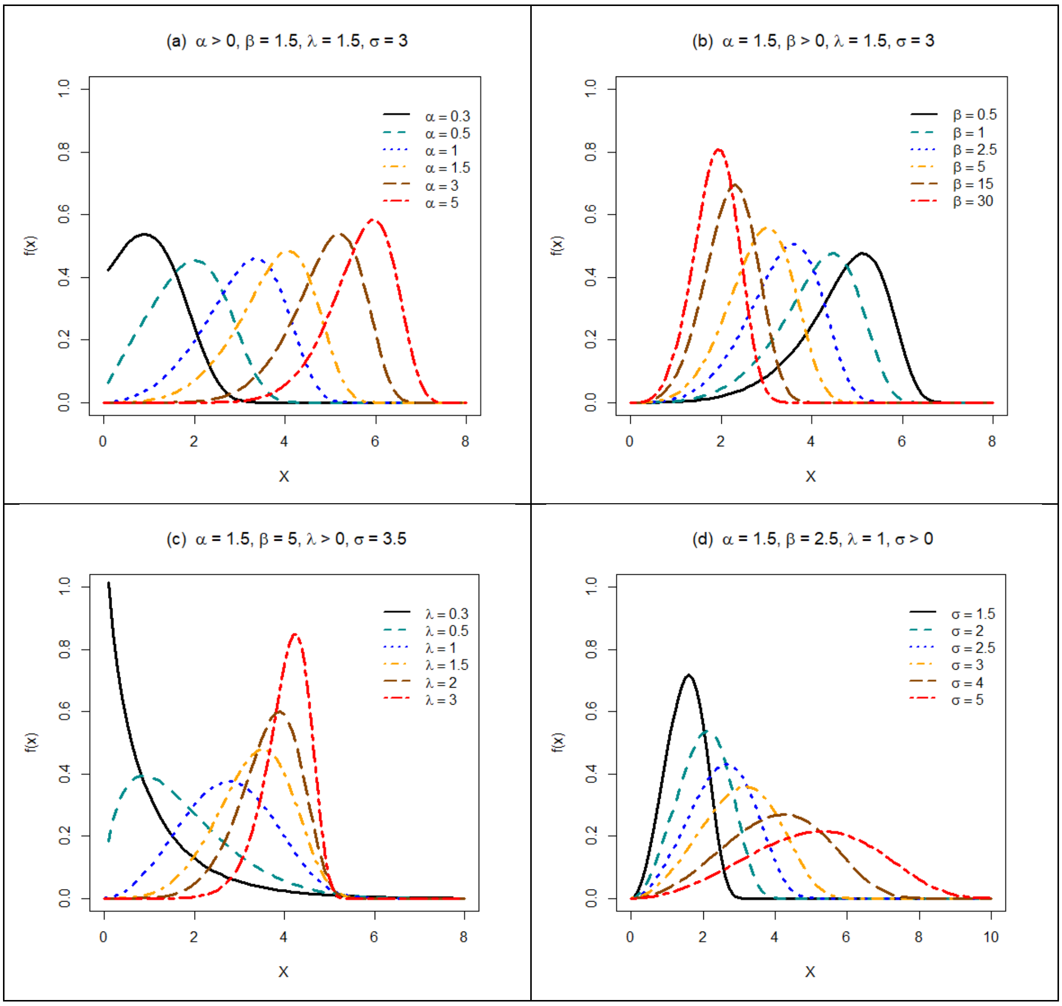

Figure 1 Some pdf plots of the PG-R distribution with different values of α, β, λ and σ.

3.2.2 Moments

The r^{th} moments of the PG-G family of distributions (Aryuyuen et al., 2021) is

$$\mathrm{E}(X^r)=\frac{\alpha\beta\lambda}{2+\beta}\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_r\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right],\ \text{for}\ \phi_r=\int_{-\infty}^{\infty}x^r b_{i,j}^{\alpha}(x;\xi)\,dx,\ \text{where}$$

$$b_{i,j}^{\alpha}(x;\xi)=\frac{g(x;\xi)}{\left[1-G^{\alpha}(x;\xi)\right]G(x;\xi)}\left(\frac{G^{\alpha}(x;\xi)}{1-G^{\alpha}(x;\xi)}\right)^{j}\ \text{for}\ i,j=0,1,2,\ldots\ \text{and}$$

$\delta_k(i,j) = \frac{\Gamma[\lambda(i+k)+1]}{\Gamma(j+1)\Gamma[\lambda(i+k)-j+1]}\left(\frac{\beta^i(-1)^{i+j}}{i!}\right)^j$ for $k=1,2$ and $\Gamma(.)$ is a gamma function.

From the cdf and pdf in (3), the r^{th} moments of the PG-R distribution is

$$\mathrm{E}(X^r) = \frac{\alpha\beta\lambda}{(2+\beta)} \sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty} \phi_r \left[(1+\beta)\delta_1(i,j) + \beta\delta_2(i,j)\right], \text{ where} \tag{8}$$

$$\phi_r = \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\left\{\frac{x^{r+1}e^{-x^2/(2\sigma^2)}}{(1-e^{-x^2/(2\sigma^2)})\left[1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}\right]}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^j\right\}dx .$$

From the moment of the PG-R distribution, we have the mean and variance are respectively

$$\mathrm{E}(X) = \frac{\alpha\beta\lambda}{2+\beta}\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_1\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right], \text{ and}$$

$$V(X) = \frac{\alpha\beta\lambda}{(2+\beta)}\left\{\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_2\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right] - \frac{\alpha\beta\lambda}{2+\beta}\left(\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_1\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right]\right)^2\right\}.$$

3.2.3 Quantile function and related measurements

Let X be a random variable distributed as the PG-R distribution with cdf as (4). Theoretically, $\mathrm{F}(x;\alpha,\beta,\lambda,\sigma) = U$, where U is a uniform random variable on the interval (0,1). Then the quantile function (qf) of the PG-R distribution is obtained by inverting equation (4) as follows:

$$\mathrm{Q}(u;\alpha,\beta,\lambda,\sigma) = \sigma\left\{-2\log\left[1-\left(1+\left[-\frac{1}{\beta}\left(\mathrm{W}_{-1}\left[-(1-u)(2+\beta)\exp(-2-\beta)\right]+\beta+2\right)\right]^{-1/\lambda}\right)^{-1/\alpha}\right]\right\}^{1/2} \tag{9}$$

where $\mathrm{W}_{-1}[\cdot]$ denotes the negative branch of the `LambertW` function (see Corless et al., 1996; Veberic, 2012), i.e., $\mathrm{W}_{-1}[z]\exp\{\mathrm{W}_{-1}[z]\} = z$, where z is a complex number.

Median: In particular, the median can be derived from (9) be setting $u = 0.5$. Then, the median is given by

$$\text{Median} = \sigma\left\{-2\log\left[1-\left(1+\left[-\frac{1}{\beta}\left(\mathrm{W}_{-1}\left[-\left(1+\frac{\beta}{2}\right)\exp(-2-\beta)\right]+\beta+2\right)\right]^{-1/\lambda}\right)^{-1/\alpha}\right]\right\}^{1/2} .$$

Algorithm for generating variate: We can generate a random variable of X_i from the PG-R distribution using the following algorithm.

Step 1: Generate U_i from the uniform distribution on the interval (0, 1) for $i = 1,2,...,n.$

Step 2: Generate X_i by using the qf of the PG-R distribution in (9), $X_i = \mathrm{Q}(u_i;\alpha,\beta,\lambda,\sigma)$.

For $\mathrm{W}_{-1}[\cdot]$ in equation (9), it computes with `lambertWm1` function in the **LamW** package (Adler, 2015) in contribution package in R (R Core Team, 2022).

Since the moments of the PG-R distribution is not a close form. Thus, the expression of mean, standard deviation, skewness, and kurtosis is in the close form, too. However, these values are shown in terms of the simulation study. Some simulations are performed for illustration purposes of how the skewness of the PG-R distribution with different values of each parameter that is appeared in Figure1. Random samples with size 1,000 are generated 10,000 times, that is, $X_i = \mathrm{Q}(u_i;\alpha,\beta,\lambda,\sigma)$. The results, including sample mean, variance, skewness value (SV), and kurtosis value (KV), are shown in Table 1. Skewness can be used as a measure of the symmetry of distribution for SV = 0. It is defined as a symmetrical distribution (left-skewed for SV < 0 and right-skewed for SV > 0). Meanwhile, the KV is often compared to the kurtosis of the normal distribution that is KV = 3. If the KV

> 3, the dataset has heavier tails than a normal distribution. Moreover, the dataset has lighter tails than a normal distribution for KV < 3.

We can consider the effects of the shape parameters on the pdf-plotted of PG-R distribution in Figure 1, and the values of skewness and kurtosis values in Table 1, which the results are follows. For fixed other parameters but α changed, the distribution shapes are right-skewed and left-skewed when $\alpha < 1$ and $\alpha \geq 1$, respectively (see Figure 1 (a)). For fixed other parameters but β changed, the distribution shapes are left-skewed, and is close to symmetric when β increase (see Figure 1 (b)). Figure 1 (c) shows that the various shape when different values of λ (for fixed other parameters); the distribution is decreasing function when $\lambda \leq 0.5$, but it is unimodal distribution when $\lambda \geq 0.5$; for $0.5 \leq \lambda < 1$ the distribution is right-skewed; for $\lambda \geq 1$ the distribution is left-skewed. Figure 1 (d) shows that the values for σ increase and the distribution is flatten but is still left-skewed.

Table 1 Results of the average values of the sample mean, variance, skewness, and kurtosis of PG-R distribution.

Figure 1	α	β	λ	σ	sample mean	variance	SV	KV
(a) $\alpha\uparrow$	*5*	1.5	1.5	3	5.6669	0.5127	-0.6228	3.3300
	3	1.5	1.5	3	4.8972	0.5995	-0.6045	3.2596
	1.5	1.5	1.5	3	3.7586	0.7113	-0.5175	3.0176
	1	1.5	1.5	3	3.0525	0.7416	-0.4039	2.7756
	0.5	1.5	1.5	3	1.8474	0.6360	-0.0525	2.3674
	0.3	1.5	1.5	3	1.0642	0.4123	0.3676	2.4358
(b) $\beta\uparrow$	1.5	*30*	1.5	3	1.8720	0.2296	-0.1768	2.7719
	1.5	*15*	1.5	3	2.2005	0.3124	-0.2073	2.7692
	1.5	*5*	1.5	3	2.8518	0.4901	-0.3089	2.8009
	1.5	*2.5*	1.5	3	3.3514	0.6198	-0.4168	2.8867
	1.5	*1*	1.5	3	4.0974	0.7743	-0.6058	3.1648
	1.5	*0.5*	1.5	3	4.6923	0.8586	-0.7637	3.5045
(c) $\lambda\uparrow$	1.5	5	*3*	3.5	4.0460	0.2569	-0.7412	3.7268
	1.5	5	*2*	3.5	3.6649	0.4671	-0.5139	3.1472
	1.5	5	*1.5*	3.5	3.3270	0.6666	-0.3089	2.7987
	1.5	5	*1*	3.5	2.7644	0.9777	0.0555	2.5559
	1.5	5	*0.5*	3.5	1.7010	1.2950	0.9253	3.6533
	1.5	5	*0.3*	3.5	1.0197	1.2068	1.9518	7.7386
(d) $\sigma\uparrow$	1.5	2.5	1	*5*	5.0469	2.9459	-0.0967	2.4988
	1.5	2.5	1	*4*	4.0365	1.8857	-0.0961	2.4986
	1.5	2.5	1	*3*	3.0278	1.0605	-0.0959	2.4981
	1.5	2.5	1	*2.5*	2.5227	0.7364	-0.0965	2.4979
	1.5	2.5	1	*2*	2.0185	0.4715	-0.0961	2.4971
	1.5	2.5	1	*1.5*	1.5141	0.2650	-0.0969	2.4976

3.2.4 Order statistics

Let $X_{(1)},\ldots,X_{(n)}$ denote the order statistics of a random sample X_i for $i=1,2,...,n$ from the PG-R distribution with the pdf in equation (5) and cdf in equation (4); then the pdf of $X_{(j)}$ for $j=1,2,...,n$ is

$$f_{X_{(j)}}(x)=\frac{\Gamma(n+1)\alpha\beta\lambda x(1-e^{-x^2/(2\sigma^2)})^{\alpha-1}e^{-x^2/(2\sigma^2)}}{\Gamma(j)\Gamma(n-j+1)\sigma^2(2+\beta)\left[1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}\right]^2}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda-1}$$

$$\times\left[(1+\beta)+\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right\}$$

$$\times\left(1-\left[1+\frac{\beta}{2+\beta}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right\}\right)^{j-1}$$

$$\times\left(\left[1+\frac{\beta}{2+\beta}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right\}\right)^{n-j}.$$

3.3 Parameter estimation

The ML estimates of the unknown parameters for the PG-R distribution are determined based on complete samples. Let $X_1,\ldots,X_n$ be a random sample of size n from the PG-R distribution with set of parameters $\boldsymbol{\theta}=(\alpha,\beta,\lambda,\sigma)^T$. The log-likelihood function of $\boldsymbol{\theta}$ can be expressed as

$$\ell(x_i;\boldsymbol{\theta}) = n\log(\alpha)+n\log(\beta)+n\log\lambda-n\log(2+\beta)-2n\log(\sigma)+\sum_{i=1}^{n}\log x_i-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2$$
$$+(\alpha\lambda-1)\sum_{i=1}^{n}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]-(\lambda+1)\sum_{i=1}^{n}\log\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)$$
$$+\sum_{i=1}^{n}\log\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]-\beta\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}.$$

The elements of the score function $\mathrm{U}(\boldsymbol{\theta})=(\mathrm{U}_{\alpha},\mathrm{U}_{\beta},\mathrm{U}_{\lambda},\mathrm{U}_{\sigma})$ are given by

$$\mathrm{U}_{\alpha}=\frac{n}{\alpha}+\lambda\sum_{i=1}^{n}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]$$
$$+(\lambda+1)\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]\right)-\beta\lambda\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]}{\left[1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right]^{\lambda+1}}\right)$$
$$+\beta\lambda\sum_{i=1}^{n}\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]}{\left[1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right]^{\lambda+1}}\right),$$

$$\mathrm{U}_{\beta}=\frac{n}{\beta}-\frac{n}{2+\beta}$$
$$+\sum_{i=1}^{n}\left[(1+\beta)+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\left[1+\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]-\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda},$$

$$\mathrm{U}_{\lambda}=\frac{n}{\lambda}+\alpha\sum_{i=1}^{n}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]-\sum_{i=1}^{n}\log\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)$$
$$+\beta\sum_{i=1}^{n}\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\log\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)$$
$$-\beta\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\log\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right),$$

$$U_\sigma = -\frac{2n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n} x_i^2 - (\alpha\lambda - 1)\sum_{i=1}^{n}\left(\frac{x_i^2 e^{-x_i^2/(2\sigma^2)}}{1 - e^{-x_i^2/(2\sigma^2)}}\right)$$

$$+\frac{\alpha(\lambda+1)}{\sigma^3}\sum_{i=1}^{n}\left(\frac{x_i^2\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha-1} e^{-x_i^2/(2\sigma^2)}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right) + \frac{\alpha\beta\lambda}{\sigma^3}\sum_{i=1}^{n}\frac{x_i^2 e^{-x_i^2/(2\sigma^2)}\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda-1}}{\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)^{\lambda+1}}$$

$$-\frac{\alpha\beta\lambda}{\sigma^3}\sum_{i=1}^{n} x_i^2\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda-1} e^{-x_i^2/(2\sigma^2)}}{\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)^{\lambda+1}}.$$

Then the ML estimates of the parameters α, β, λ and σ are obtained by setting the last two equations to be zero and solving them. Clearly, it is difficult to solve them, therefore computer packages. In this study, the `nlm` function in the **stats** package and contribution package in R (R Core Team, 2022) is used to find the ML estimates.

3.4 Application study

In this section, we demonstrate the flexibility and the potentiality of the PG-R distribution through two real data sets, which data set are:

Data I: The data set is gauge lengths of 10 mm (Kundu & Raqab, 2009). This data set consists of 63 observations: 1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

Data II: The second data is the strength data set originally reported by Badar and Priest in 1982 (see Dey, 2014). It describes the strength measured in GPA for single-carbon fibers and impregnated 1000-carbon fiber tows. This data set consists of 62 observations: 0.5620, 0.5640, 0.7290, 0.8020, 0.9500, 1.0530, 1.1110, 1.1940, 1.2080, 1.2160, 1.2470, 1.2560, 1.2710, 1.2777, 1.3050, 1.3130, 1.3900, 1.4290, 1.4740, 1.4900, 1.5030, 1.5200, 1.5220, 1.5240, 1.5510, 1.6090, 1.6320, 1.6320, 1.6760, 1.6840, 1.6850, 1.7280, 1.7400, 1.7640, 1.7850, 1.8040, 1.8160, 1.8240, 1.8360, 1.8790.

We use these real data sets to illustrate the importance and flexibility of the PG-R distribution and compare it with some lifetime distributions which related the PG-R distribution and the PG-G family of distributions with four-parameter, such as the Rayleigh (R), power Garima-Lindley (PG-L; Aryuyuen et al., 2021), and Lindley (L) distributions. All the model parameters are estimated by the ML method for the proposed model. These applications will be used to determine the estimated parameters of each distribution. The value of ML estimates of each distribution is obtained using the `nlm` function in the stats package and contribution package in R (R Core Team, 2022). The results are presented in Tables 2-3. In each application, we compare the fit of the PG-G distribution with the fit of the corresponding baseline distribution (R distribution), the PG-G family with a four-parameter distribution (PG-L distribution), and the L distribution. The criteria of the Akaike information criterion (AIC), the corrected Akaike information criterion (CAIC), and Kolmogorov-Smirnov (D*) statistics are used. The model gives the smallest values of AIC, CAIC, and D*. Therefore, it is the best model for fitting data.

Table 2 ML estimates of the model parameters and goodness of measures for their estimates for Data I.

Distributions	ML estimates	AIC	CAIC	D* (p-value)
PG-R	$\hat{\alpha}=40.3156,\ \hat{\beta}=2.3421,$ $\hat{\lambda}=0.3076,\ \hat{\sigma}=1.3118$	121.19	121.88	0.0770 (0.8492)
R	$\hat{\sigma}=2.2067$	189.04	189.11	0.3607 (<0.0001)
PG-L	$\hat{\alpha}=6.5264,\ \hat{\beta}=216.0021,$ $\hat{\lambda}=0.7011,\ \hat{\tau}=0.2743$	129.65	130.34	0.0810 (0.8031)
L	$\hat{\tau}=0.5392$	244.72	244.78	0.4308 (<0.0001)

Table 3 ML estimates of the model parameters and goodness of measures for their estimates for Data II.

Distribution	ML estimates	AIC	CAIC	D* (p-value)
PG-R	$\hat{\alpha}=2.5351,\ \hat{\beta}=114.5274,$ $\hat{\lambda}=0.7918,\ \hat{\sigma}=4.1840$	95.81	86.51	0.0490 (0.9984)
R	$\hat{\sigma}=$ 1.2464	121.60	121.67	0.2551 (<0.0001)
PG-L	$\hat{\alpha}=1.4544,\ \hat{\beta}=21.2383,$ $\hat{\lambda}=2.1026,\ \hat{\tau}=0.4522$	95.98	96.68	0.0559 (0.9902)
L	$\hat{\tau}=$ 0.9015	176.00	176.06	0.5454 (<0.0001)

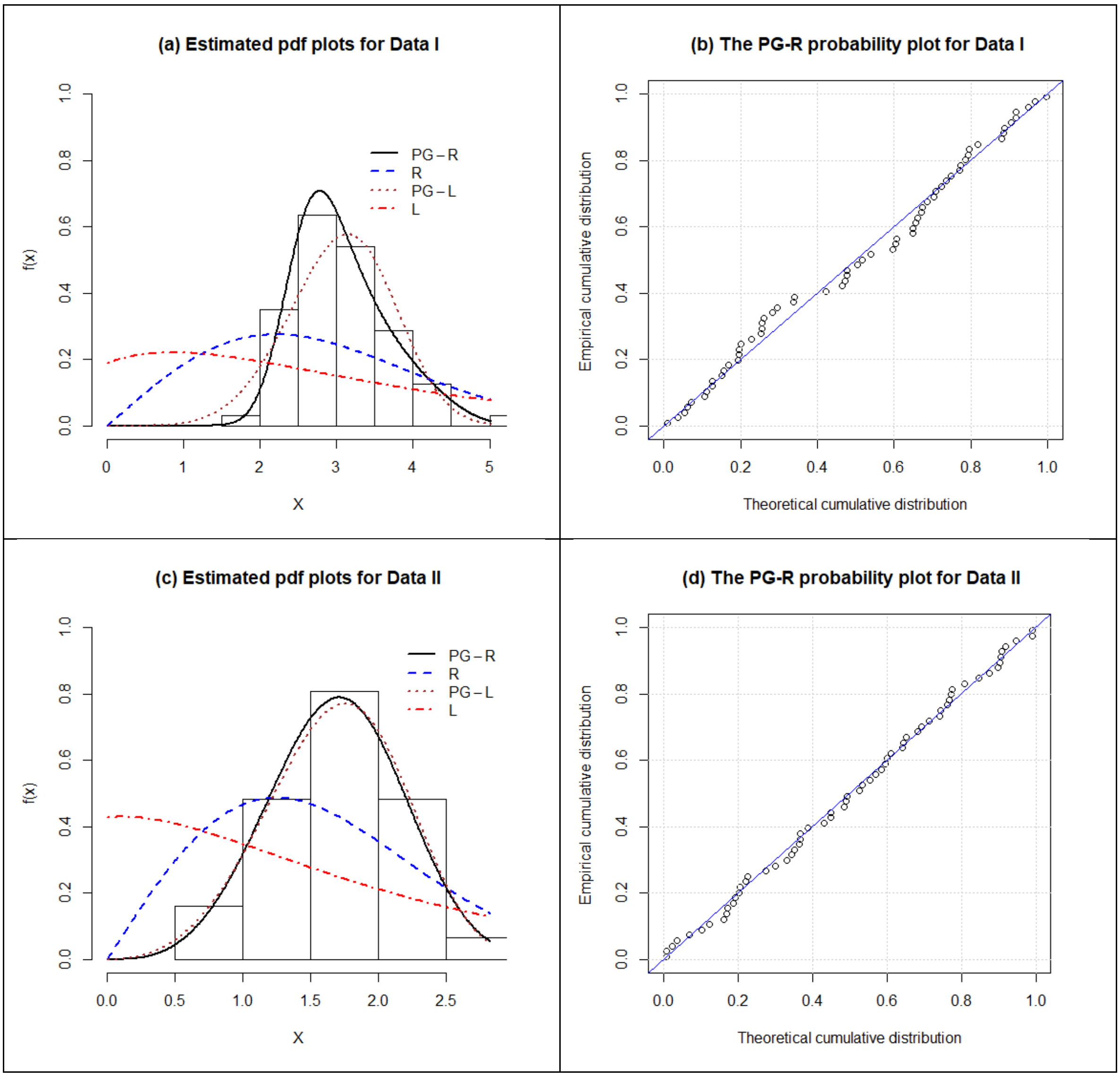

Figure 2 Plots of the estimated pdf for each distribution and the PG-R probability plot for the data sets.

Figure 2 shows the plots of the estimated pdf of the fitted models for the data sets. The numerical values of the AIC, CAIC, and D* statistics are listed in Tables 2-3. The PG-R model gives the lowest values for the AIC, CAIC, and D* statistics for the data set among the fitted model. The PG-R distribution could fit the data better than the PG-L, R, and L distributions. A density plot compares the fitted densities of the models with the empirical histogram of the observed data (Figure 2).

4. Conclusion

This paper proposes a new four-parameter distribution, namely the PG-R distribution. The PG-R distribution is motivated by the wide use of the Rayleigh distribution in practice and the fact that generalization provides more flexibility to analyze positive real-life data. We derive explicit expressions for the quantile function and order statistics. We also provide a foundation for some mathematical properties of this distribution, including the derivation of the hazard rate function, moments, quantile function, generating numbers, and order statistics. The model parameters are estimated by maximum likelihood. Applying the PG-R distribution to a real data set indicates that the new distribution outperforms several distributions, including the PG-L, Rayleigh, and Lindley distributions.

Acknowledgement

The authors gratefully acknowledge participating in the Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Thailand. We also thank those who could not be mentioned here for their kindness and encouragement. Moreover, the authors thank the anonymous reviewers for their comments and suggestions.

References

Abd Elfattah, A. M., Hassan, A. S., & Ziedan, D. M. (2006). Efficiency of maximum likelihood estimators under different censored sampling schemes for Rayleigh distribution. **Interstat**, 1, 1-16.

Adler A. (2015). **lamW: Lambert-W Function**. CRAN.R-project. Retrieved November 1, 2022 from https://CRAN.R-project.org/package=lamW.

Ahmad, Z., Elgarhy, M., & Hamedani, G. G. (2018). A new Weibull-X family of distributions: properties, characterizations and applications. **Journal of Statistical Distributions and Applications**, 5(1), 1-18.

Al-Babtain, A. A. (2020). A new extended Rayleigh distribution. **Journal of King Saud University-Science**, 32(5), 2576-2581.

Alizadeh, M., Cordeiro, G. M., Pinho, L. G. B., & Ghosh, I. (2017). The Gompertz-G family of distributions. **Journal of Statistical Theory and Practice**, 11(1), 179-207.

Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. **Metron**, 71(1), 63-79.

Aryuyuen, S., Bodhisuwan, W., & Ngamkham, T. (2021). Power Garima-generated family of distributions: properties and application. **Lobachevskii Journal of Mathematics**, 42(2), 287-299.

Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. **Journal of data science**, 12(1), 53-68.

Cordeiro, G. M., Alizadeh, M., & Diniz Marinho, P. R. (2016). The type I half-logistic family of distributions. **Journal of Statistical Computation and Simulation**, 86(4), 707-728.

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

Corless, R. M., Gonnet, G. H., Hare, D. E., Jeffrey, D. J., & Knuth, D. E. (1996). On the LambertW function. **Advances in Computational mathematics**, 5(1), 329-359.

Dey, S. (2009). Comparison of Bayes estimators of the parameter and reliability function for Rayleigh distribution under different loss functions. **Malaysian Journal of Mathematical Sciences**, 3, 247-264.

Dey, S., & Das, M. K. (2007). A note on prediction interval for a Rayleigh distribution: Bayesian approach. **American Journal of Mathematical and Management Sciences**, 27(1-2), 43-48.

Dey, S., Dey, T., & Kundu, D. (2014). Two-parameter Rayleigh distribution: different methods of estimation. **American Journal of Mathematical and Management Sciences**, 33(1), 55-74.

Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. **Communications in Statistics-Theory and methods**, 31(4), 497-512.

Gupta, R.C., Gupta P.L., Gupta R.D. (1998). Modeling failure time data by Lehman alternatives. **Communications in Statistics-Theory and methods**, 27(4), 887-904.

Hassan, A. S., & Elgarhy, M. (2016). A new family of exponentiated Weibull - generated distributions. **International Journal of Mathematics and Its Applications**, 4(1-D), 135-148.

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1994). **Continuous Univariate Distribution**. (2nd ed.). John Wiley & Sons.

Kundu, D., & Raqab, M. Z. (2009). Estimation of R= P (Y< X) for three-parameter Weibull distribution. **Statistics & Probability Letters**, 79(17), 1839-1846.

Nofal, Z. M., Afify, A. Z., Yousof, H. M., & Cordeiro, G. M. (2017). The generalized transmuted-G family of distributions. **Communications in Statistics-Theory and Methods**, 46(8), 4119-4136.

R Core Team (2022). **R: A language and environment for statistical computing**. R-project. Retrieved November 1, 2022 from https://www.R-project.org/.

Rayleigh, L. (1880). XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, 10(60), 73-78.

Reyad, H., Jamal, F., Othman, S., & Hamedani, G. G. (2018). The transmuted Gompertz-G family of distributions: properties and applications. **Tbilisi Mathematical Journal**, 11(3), 47-67.

Ristić, M. M., & Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution. **Journal of statistical computation and simulation**, 82(8), 1191-1206.

Soliman, A. H., Elgarhy, M. A. E., & Shakil, M. (2017). Type II half logistic family of distributions with applications. **Pakistan Journal of Statistics and Operation Research**, 245-264.

Veberič, D. (2012). Lambert W function for applications in physics. **Computer Physics Communications**, 183(12), 2622-2628.

Yousof, H. M., Rasekhi, M., Afify, A. Z., Ghosh, I., Alizadeh, M., & Hamedani, G. G. (2017). The beta Weibull-G family of distributions: theory, characterizations and applications and applications. **Pakistan Journal of Statistics**, 33(2), 95-116.

Zografos, K., & Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. **Statistical methodology**, 6(4), 344-362.