



Exponentiated Rayleigh Poisson Distribution: Model, Properties and Applications

Ramesh Kumar Joshi^{1, *}, Govinda Prasad Dhungana²

¹Department of Statistics, Trichandra Multiple Campus Saraswoti Sadan, Tribhuvan University, Kathmandu, Nepal

²Govinda Prasad Dhungana, Birendra Multiple Campus, Tribhuvan University, Chitwan, Nepal

Email address:

ramesh.joshi@cit.edu.np (R. K. Joshi), govinda.dhungana@bimc.tu.edu.np (G. P. Dhungana)

*Corresponding author

To cite this article:

Ramesh Kumar Joshi, Govinda Prasad Dhungana. Exponentiated Rayleigh Poisson Distribution: Model, Properties and Applications. *American Journal of Theoretical and Applied Statistics*. Vol. 9, No. 6, 2020, pp. 272-282. doi: 10.11648/j.ajtas.20200906.13

Received: October 6, 2020; **Accepted:** October 24, 2020; **Published:** November 4, 2020

Abstract: In this research paper, a new class of life-time distribution is introduced by compounding A new generalization of Rayleigh distribution; properties and applications and The Exponentiated G Poisson model, the so-called Exponentiated Rayleigh Poisson distribution. Main aim of this research article is to enhance the flexibility of Exponentiated G. Poisson distribution by power transformation technique. The probability density function, the survival function and the hazard function of the new proposed model in graphical form are illustrated. We study the properties of this new distribution with special emphasis on its quantile function, mode, skewness, kurtosis and moments. We have discussed residual life function, the probability-weighted moments, order statistics, R'enyi and entropies. We also discussed parameter estimation considering the maximum likelihood estimation approach. We have calculated the value of log-likelihood, Akaike's information criteria, Bayesian information criteria, corrected Akaike's information criteria and Hennen-Quinn information criteria of Generalized Rayleigh distribution, Exponentiated Chen distribution, Exponentiated Exponential distribution, Exponentiated Inverted Weibull distribution, Compound Rayleigh distribution and newly proposed Exponentiated Rayleigh Poisson distribution and found that the newly proposed model has smaller values in comparison to other. We have studied the P-P plot, Q-Q plot Kolmogorov Smirnov test and TTT plot of the proposed distribution for model validation. We compared the empirical distribution CDF and estimated distributed function CDF of the proposed model with five other models. A real dataset is analyzed for illustrative purposes. The importance and flexibility of the new family is illustrated by applying different techniques and tools. A final conclusion concludes the paper.

Keywords: Exponentiated G Poisson Model, New Generalization of Rayleigh Distribution, Maximum Likely-hood Estimation (MLE), Probability Weighted Moments (PWM), Order Statistics

1. Introduction

In the last decades, many generalized distributions have been proposed based on different modification methods. These modification methods require the addition of one or more parameters to the base model which could provide better adaptability in the modeling of real-life data. Modern computing technology has made many of these techniques accessible even if analytical solutions are very complicated [8].

In many applied sciences such as medicine, engineering, and finance, amongst others, modeling and analyzing lifetime data are crucial. Several lifetime distributions have been used

to model such kinds of data. The quality of the analysis depends heavily on the assumed probability model or distributions. Because of this, considerable effort has been expended in the development of large classes of standard probability distributions along with relevant, statistical methodologies. However, there remain many important problems where real data does not follow any of the classical or standard probability models [20].

Over the years, new family of distributions have been proposed to generalize various distributions by compounding well-known distributions to provide greater flexibility in modeling data from practical viewpoints [12] introduced a general class of distribution generated from the logit of the

beta random variable. [6] proposed for generating families of continuous distributions proposed a random variable X the "transformer" is used to transform another random variable T, the "transformed ".[25] explored a new three-parameter distribution motivated mainly by lifetime issues. [18] proposed a generalization of Kumaraswamy distribution referred to as the exponentiated Kumaraswamy distribution. [24] introduced a new distribution generated by gamma random variables and [28] considers a new general class of distributions generated from the logit of the gamma random variable which is a special case of the gamma uniform distribution. [21] proposed a new family of continuous distributions called the exponentiated transmuted G family which extends the transmuted G family. [16] explored a two-parameter family of distributions on (0,1), which has many similarities to the beta distributions and a number of advantages in terms of tractability. Two general families of continuous distributions generated by a distribution F and two positive real parameters -- and -- which controls the skewness and tail weight of the distribution is introduced by [7]. [26] introduced a new family of continuous distributions called Marshall-Olkin Kumaraswamy distribution and [10] proposed a new class of continuous distributions with two extra shape parameters named the generalized odd log-logistic family of distributions.

To create the new probability model, different methods exist, among them, one most common method is to combine a valid probability distribution with a family of the probability distribution. In this paper, we proposed "Exponentiated Rayleigh Poisson distribution (ERP)", which is developed from the compounding of "Exponentiated G Poisson (EGP) Model" and "A new generalization of Rayleigh distribution: properties and applications". The main aim of this research article is to enhance the flexibility of Exponentiated G. Poisson distribution by power transformation technique. The induction of an extra parameter in the parent model usually provides greater flexibility and improves the goodness of fit.

The CDF and PDF functions, Reliability / Survival and Hazard Rate Functions (HRF) are explicitly presented in Section 2. Section 3 explores some important statistical properties such as quantile and median function, mode, skewness, Kurtosis, moments, moment-generating function, probability-weighted moments, residual life function, order statistics, and entropy. In section 4, the model parameters are estimated by the Maximum Likelihood Estimate (MLE) technique. In Section 5, real data analyses are explored to verify theoretical findings in different aspects. Some concluding remarks are given in section 6.

2. Exponentiated Rayleigh Poisson Distribution

The Exponentiated G Poisson family is obtained by compounding the Exponentiated G family and truncated Poisson distribution. The proposed family of distribution

extends several common distributions [15], having Cumulative distribution function (CDF) and Probability Density Function (PDF) are given as:

$$F(x) = \frac{1 - e^{-\beta[G(x,\lambda)]^\alpha}}{1 - e^{-\beta}}; x > 0, \alpha > 0, \beta > 0, \lambda > 0 \quad (1)$$

Where $\alpha > 0$ is the shape parameter and $\beta > 0, \lambda > 0$ are scale parameters.

The corresponding PDF is

$$f(x) = \frac{\alpha\beta g(x)[G(x,\lambda)]^{\alpha-1} e^{-\beta[G(x,\lambda)]^\alpha}}{1 - e^{-\beta}}; x > 0, \alpha > 0, \beta > 0, \lambda > 0 \quad (2)$$

Among the probability distributions, Rayleigh is one of the most widely used distributions. Rayleigh's distribution, introduced by Rayleigh in 1880, claimed to be a special case of Weibull's distribution. It plays a key role in modeling and analyzing life-time data such as project loading, survival and reliability analysis, communication theory, etc [4]. With regard to this importance and the desire to make this distribution more flexible, many researchers have developed extensive extensions to Rayleigh distribution having CDF and PDF are

$$G(x) = 1 - e^{-(\lambda x)^2}; x > 0, \lambda > 0 \quad (3)$$

$$g(x) = 2\lambda^2 x e^{-(\lambda x)^2}; x > 0, \lambda > 0 \quad (4)$$

Here we compounded the CDF of Rayleigh distribution in Exponentiated G Poisson family, it becomes, Exponentiated Rayleigh Poisson distribution having CDF

$$F(x) = \frac{1 - e^{-\beta[1 - e^{-(\lambda x)^2}]^\alpha}}{1 - e^{-\beta}}; x > 0, \alpha > 0, \beta > 0, \lambda > 0 \quad (5)$$

and PDF of Exponentiated Rayleigh Poisson distribution is

$$f(x) = \frac{2\alpha\beta\lambda^2 x e^{-(\lambda x)^2} [1 - e^{-(\lambda x)^2}]^{\alpha-1} e^{-\beta[1 - e^{-(\lambda x)^2}]^\alpha}}{1 - e^{-\beta}}; x > 0, \alpha > 0, \beta > 0, \lambda > 0 \quad (6)$$

2.1. Survival Function

The survival function is the probability of the non-failure occurring before time t. Therefore, the survival function of the proposed model is,

$$R(x) = 1 - F(x) = 1 - \frac{1 - e^{-\beta[1 - e^{-(\lambda x)^2}]^\alpha}}{1 - e^{-\beta}}; x > 0, \alpha > 0, \beta > 0, \lambda > 0 \quad (7)$$

In figure 1 (left panel) various graphs of the density when $\lambda=10$, and for different values of α and β are displayed. This indicates that the PDF. of the ERP distribution is well suited for non-normal and skewed. The greater the value of α , the more positively skewed the distribution. For $\alpha < 1$ and $\beta < 1$ the density is reversed J shaped. Figure 1 (right panel) shows

the survival function of the ERP distribution has decreasing failure properties.

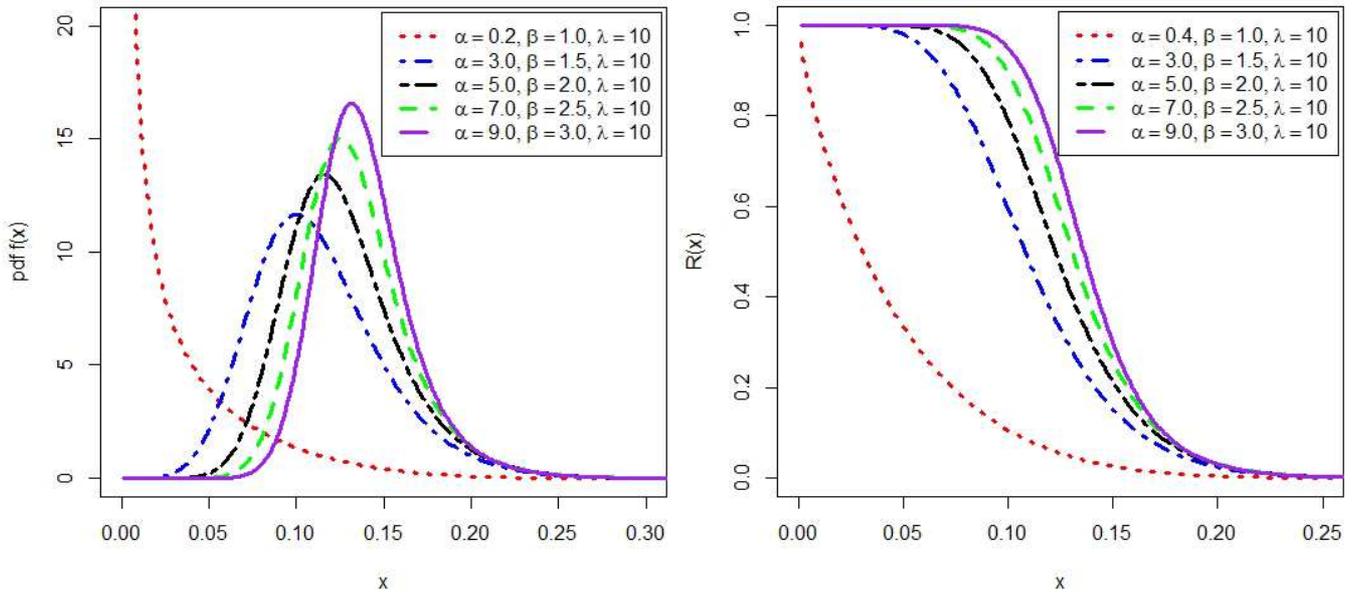


Figure 1. Probability density function (left panel) and survival function (right panel) of ERP.

2.2. Hazard Function

The hazard function is the instant rate of failure at a given time t. Therefore, the hazard function of the proposed model is

$$h(x) = \frac{f(x)}{R(x)} = \frac{2\alpha\beta\lambda^2 x e^{-(\lambda x)^2} [1 - e^{-(\lambda x)^2}]^{\alpha-1} e^{-\beta[1 - e^{-(\lambda x)^2}]}^\alpha}{e^{-\beta[1 - e^{-(\lambda x)^2}]^\alpha} - e^{-\beta}}; x > 0, \alpha > 0, \beta > 0, \lambda > 0 \tag{8}$$

Figure 2 shows that the hazard function of the ERP distribution has upside down (concave and convex) properties.

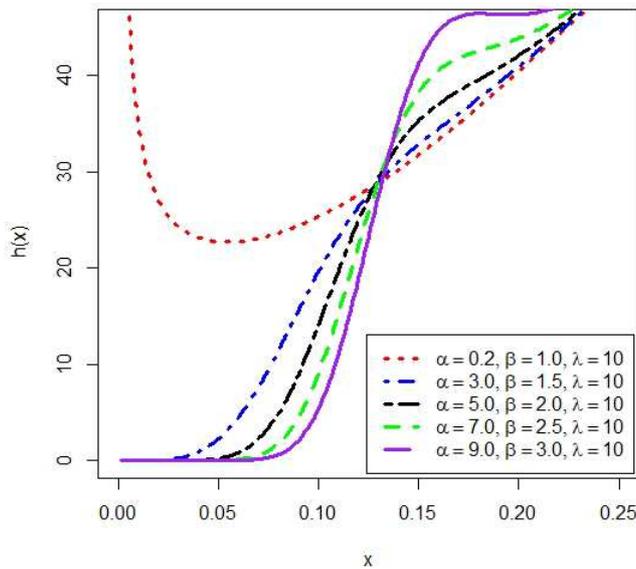


Figure 2. Hazard function of ERP distribution with different values of alpha, beta, and lambda.

3. Statistical Properties

In this section, major properties of ERP distribution have been derived.

3.1. Useful Expansions

Distribution is derived from the generalized binomial series. For, |z| < 1, n > 0 we have

$$(1 - z)^n = \sum_{i=0}^{\infty} (-1)^i \binom{n}{i} Z^i \text{ for } , n > 0 \tag{9}$$

The power series expansion of corresponding to an exponential function is

$$e^{-az} = \sum_{j=0}^{\infty} \frac{(-1)^j (az)^j}{j!} \tag{10}$$

Using the binomial theorem (9) and (10) in equation (6), the PDF of the proposed model is

$$f(x) = \sum_{i=0}^{\infty} \xi_i x e^{-(1+i)(\lambda x)^2} \tag{11}$$

Where, $\xi_i = \sum_{j=0}^{\infty} \frac{(-1)^{i+j} 2\alpha\lambda^2\beta^{j+1}}{(1-e^{-\beta})^j j!} \binom{\alpha(j+1)-1}{i}$

$$Q(u) = \left[-\frac{1}{\lambda^2} \ln \left[1 - \left\{ -\frac{1}{\beta} \ln [1 - u(1 - e^{-\beta})] \right\}^{\frac{1}{\alpha}} \right] \right]^{\frac{1}{2}} \tag{12}$$

3.2. Quantile Function

The quantile functions are used in theoretical aspects of probability theory. It is an alternative of PDF and CDF, used to obtain statistical measures like median, skewness, and kurtosis. It can also be used to generate random variables. The quantile function is given by $Q(u) = F^{-1}(u)$; Therefore, the corresponding quantile function for the purposed model is given as;

Where u has uniform U (0, 1) distribution. We obtained the median by substituting u=0.5 in equation (12)

3.3. Mode

The maximum repetitive value of the given PDF is the mode. To calculate the mode the necessary and sufficient condition is; $\frac{df(x)}{dx} = 0$ and $\frac{d^2f(x)}{dx^2} < 0$ After applying necessary condition, we get,

$$\frac{1}{x} - 2\lambda^2 x \left[1 - e^{-(\lambda x)^2} \left\{ \frac{1}{1 - e^{-(\lambda x)^2}} \left\{ (\alpha - 1) + \alpha\beta \left(1 - e^{-(\lambda x)^2} \right)^\alpha \right\} \right\} \right] = 0 \tag{13}$$

Equation (13) is a nonlinear equation, and its solution cannot be found analytically. It can be found numerically by using Newton-Raphson method.

Where Q (.) is the quantile function defined in equation (13).

3.4. Skewness and Kurtosis

For statistical analysis, skewness and kurtosis are used to describe the characteristics of the distribution. Bowley's skewness [5] takes the form

To calculate the statistical measure of the proposed distribution, 100 random samples are generated from the equation (12) at the initial value $\alpha = 5.0$, $\beta = 2.0$ and $\lambda = 10$.

$$S_k = \frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)}$$

Then, we measured the mean, median, mode, standard deviation, skewness, and kurtosis from these random samples, which illustrates the characteristics of the intended model. The model value is increasing at the initial value of parameters, $\alpha = 3.5$, $\beta = 6.0$ and $\lambda = 3.5$ and then gradually decreasing. The standard deviation of the proposed model is decreasing when the value of α and β are increasing and the value of λ is decreasing. Likewise, it is observed that different values of skewness and kurtosis are not symmetrical, non-normal (non-mesokurtic) in characteristics. Therefore, the characteristic of the proposed model is unimodal skewed and non-normal (Figure 1, Table 1).

Moors' kurtosis [5] is based on Octiles and could be written as

$$K_u = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)}$$

Table 1. The mean, median, mode, skewness, Kurtosis of Proposed distribution with different values of parameters.

Parameters			Mean	Median	Mode	SD	Skewness	Kurtosis
α	β	λ						
1.0	1.0	6.0	4.93190	5.18646	3.47842	1.14209	-1.50479	5.13026
1.5	2.0	5.5	5.63586	6.06342	4.53186	1.17217	-2.29426	8.16931
2.0	3.0	5.0	5.52849	5.96922	5.93371	1.37618	-1.41377	4.98233
2.5	4.0	4.5	4.34110	4.40433	6.83482	1.65981	-0.18166	1.98246
3.0	5.0	4.0	2.61113	2.26396	5.67804	2.16712	0.70480	2.44391
3.5	6.0	3.5	3.31396	0.70057	7.64840	0.95582	2.13028	7.44750
4.0	7.0	3.0	0.37266	0.13161	0.85208	0.40571	4.04105	20.8869
4.5	8.0	2.5	0.07164	0.01353	0.13079	0.06885	5.28873	32.4093
5.0	9.0	2.0	0.00651	0.00062	0.00884	0.00319	5.91808	38.6905

3.5. Moments

In mathematics, the moments are quantitative measurements in the form of a function which describe the characteristic of the proposed distribution. Since $X \sim ERP(\alpha, \beta, \lambda)$ the r^{th} raw moment is defined as (used value of $f(x)$ is used from equation (11))

$$\mu_r' = E(X^r) = \int_0^{\infty} x^r f(x) dx = \sum_{i=0}^{\infty} \xi_i \int_0^{\infty} x^{r+1} e^{-(1+i)(\lambda x)^2} dx \tag{14}$$

After the integration of (14), the r^{th} raw moment of the proposed model is;

$$\mu'_r = \sum_{i=0}^{\infty} \frac{1}{2\lambda^{r+2}(1+i)^{\binom{r}{2}+1}} \xi_i \Gamma\left(\frac{r}{2}+1\right)$$

Particularly, when $r = 1$, the mean of the proposed model is

$$\mu'_1 = \sum_{i=0}^{\infty} \frac{\xi_i \sqrt{\pi}}{4\lambda^3(1+i)^{3/2}}$$

Similarly, when $r = 2$, the second-order raw moment of the proposed model is

$$\mu'_2 = \sum_{i=0}^{\infty} \frac{\xi_i}{2\lambda^4(1+i)^2}$$

By using the relation, $Var(X) = \mu'_2 - (\mu'_1)^2$ we can calculate the variance of the proposed model. The mean and standard deviation for the proposed model with different values of the parameters are presented in (Table 1).

The lower incomplete moments, say $\varphi_s(t)$, is given by;

$$\varphi_s(t) = \int_0^t x^s f(x) dx \tag{15}$$

Using the relation (11) in equation (15), and applying lower incomplete gamma function in $\gamma(s, t) = \int_0^t x^{s-1} e^{-x} dx$ in equation (15), and integrating equation (15), we get the value of lower incomplete moment $\varphi_s(t)$ as

$$\varphi_s(t) = \sum_{i=0}^{\infty} \frac{1}{2[(1+i)\lambda^2]^{\binom{s}{2}+1}} \xi_i \gamma\left\{\left(\frac{s}{2}+1\right), (1+i)t^2\lambda^2\right\}$$

Similarly, the conditional moment is defined as

$$\tau_s(t) = \int_t^{\infty} x^s f(x) dx \tag{16}$$

Using the relation (11) in equation (18), and applying upper incomplete gamma function, $\Gamma(s, t) = \int_0^t x^{s-1} e^{-x} dx$ in equation (16), and integrating equation (16), we get the value of conditional moment as

$$\tau_s(t) = \sum_{i=0}^{\infty} \frac{1}{2[(1+i)\lambda^2]^{\binom{s}{2}+1}} \xi_i \Gamma\left\{\left(\frac{s}{2}+1\right), (1+i)t^2\lambda^2\right\}$$

Likewise, Moment Generating Function (MGF) is

$$M_X(t) = E[e^{tX}] = \sum_{r=0}^{\infty} \frac{t^r}{r!} E(X^r) \tag{17}$$

Using the result of the equation (14) in equation (17), the MGF is

$$M_X(t) = \sum_{i=0}^{\infty} \sum_{r=0}^{\infty} \frac{1}{2r!\lambda^{r+2}(1+i)^{\binom{r}{2}+1}} \xi_i t^r \Gamma\left(\frac{r}{2}+1\right)$$

3.6. Residual Life Function

The n^{th} moment of the residual life of X is given by

$$m_n(t) = \frac{1}{R(t)} \int_t^{\infty} (x-t)^n f(x) dx \tag{18}$$

Applying the binomial expansion of $(x-t)^n = \sum_{d=0}^n (-1)^d \binom{n}{d} x^{n-d} t^d$ into the equation (18), then,

$$m_n(t) = \frac{1}{R(t)} \sum_{i=0}^{\infty} \sum_{d=0}^n \xi_i (-t)^d \binom{n}{d} \int_t^{\infty} x^{n-d} e^{-(1+i)(\lambda x)^2} dx \tag{19}$$

Applying the upper incomplete gamma function in equation (19), n^{th} moment of the residual life of X is

$$m_n(t) = \frac{1}{2R(t)} \sum_{i=0}^{\infty} \sum_{d=0}^n \xi_i (-t)^d \binom{n}{d} \frac{\Gamma\left[\left\{\frac{(n-d+1)}{2}\right\}, (1+i)(t\lambda)^2\right]}{[(1+i)\lambda^2]^{(n-d+1)/2}}$$

The n th moment of the revised residual life function of X is found as

$$M_n(t) = \frac{1}{F(t)} \int_0^t (t-x)^n f(x) dx \tag{20}$$

Similarly, applying the binomial expansion of $(x-t)^n$ into the above formula, we get

$$M_n(t) = \frac{1}{2F(t)} \sum_{i=0}^{\infty} \sum_{d=0}^n \xi_i (-1)^{n+d} t^d \binom{n}{d} \frac{\Gamma\left[\left\{\frac{(n-d+1)}{2}\right\}, (1+i)(t\lambda)^2\right]}{[(1+i)\lambda^2]^{(n-d+1)/2}}$$

3.7. The Probability Weighted Moments (PWM)

The Probability Weighted Moments have been obtained from the following relation

$$\tau_{r,s} = E(X^r F(x)^s) = \int_{-\infty}^{\infty} x^r f(x) F(x)^s dx \tag{21}$$

Now, we have applied the expansion of

$$[F(x)]^s = \sum_{k=0}^{\infty} \eta_k e^{-k(\lambda x)^2} \tag{22}$$

Where, $\eta_k = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{i+j+k} (\beta i)^j}{(1-e^{-\beta})^s j!} \binom{\alpha j}{k} \binom{s}{j}$

Now, using the relation (11) and (22) in equation (21), it becomes,

$$\tau_{r,s} = \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \xi_i \eta_k \int_0^{\infty} x^{r+1} e^{-(1+i+k)(\lambda x)^2} dx \tag{23}$$

After integrating the equation (22), the PWM of the proposed model is

$$\tau_{r,s} = \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{2 \left[(1+i+k)\lambda^2 \right]^{\left(\frac{r}{2}+1\right)}} \xi_i \eta_k \Gamma\left(\frac{r}{2}+1\right)$$

3.8. Order Statistics

Let $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ the ordered statistics of a random sample of size n from the following ERP distribution with parameters $(\alpha, \beta$ and $\lambda)$. The PDF of m^{th} order statistic [David (1981)], has defined as

$$f(X_{(m)}(x_{(m)})) = \frac{f(x_{(m)})}{B(m, n-m+1)} \sum_{v=0}^{n-m} (-1)^v \binom{n-m}{v} F(x_{(m)})^{v+m-1} \tag{24}$$

Where, $B(.,.)$ is the beta function. By substituting (11) and (23) in equation (24) where S is replaced by $v+m-1$. Then, equation (24) becomes

$$f(X_{(m)}(x_{(m)})) = \frac{1}{B(m, n-m+1)} \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{n-m} \eta^* x_{(m)} e^{-(i+k+1)\lambda^2 x_{(m)}^2} \tag{25}$$

Where, $\eta^* = (-1)^v \binom{n-m}{v} \xi_i \eta_k$ then, the moment of order statistics is

$$E(X_{(m)}^r) = \int_0^{\infty} x_{(m)}^r f(X_{(m)}(x_{(m)})) dx_{(m)} \tag{26}$$

Using the value of (25) in equation (26) and integrating equation (26) we get the value of the r^{th} moment of order statistics as

$$E(X_{(m)}^r) = \frac{1}{B(m, n-m+1)} \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \sum_{v=0}^{n-m} \eta^* \frac{\Gamma(r/2+1)}{\left[(1+i+k)\lambda^2 \right]^{\left(r/2+1\right)}}$$

3.9. R`enyi and q-entropies

The entropy of a random variable X is a measure of the variation of uncertainty and has been used in many fields

such as physics, engineering, and economics among others. The R`enyi entropy is defined as

$$I_{\delta}(X) = \frac{1}{1-\delta} \log \int_{-\infty}^{\infty} [f(x)]^{\delta} dx, \quad \delta > 0 \text{ \& } \delta \neq 1 \tag{27}$$

We have

$$[f(x)]^{\delta} = \left(\frac{2\alpha\beta\lambda^2}{1-e^{-\beta}} \right)^{\delta} x^{\delta} e^{-\delta(\lambda x)^2} \left[1-e^{-(\lambda x)^2} \right]^{\delta(\alpha-1)} e^{-\beta\delta \left[1-e^{-(\lambda x)^2} \right]^{\alpha}} \tag{28}$$

Apply the equation (9) and (10) in equation (28), we get

$$[f(x)]^{\delta} = \sum_{i=0}^{\infty} \xi_i^* x^{\delta} e^{-(\delta+i)(\lambda x)^2}$$

Where, $\xi_i^* = \sum_{j=0}^{\infty} \frac{(-1)^{i+j} \beta^{j+\delta}}{j!} \left(\frac{2\alpha\lambda^2}{1-e^{-\beta}} \right)^{\delta} \binom{\alpha j + \delta(\alpha-1)}{i}$

Used this expression in equation (27) then R`enyi entropy is

$$I_{\delta}(X) = \frac{1}{1-\delta} \log \left[\sum_{i=0}^n \frac{\xi_i^*}{2\lambda^{\delta+1}} (\delta+i)^{\frac{\delta+1}{2}} \Gamma\left(\frac{\delta+1}{2}\right) \right]$$

Similarly, q-entropy is defined as

$$H_q(x) = \frac{1}{1-q} \log \left[1 - \int_{-\infty}^{\infty} \{f(x)\}^q \right], q > 0 \text{ and } q \neq 1 \tag{29}$$

The q-entropy is obtained by substituting the result (28) into (29), where δ is replaced by q, we get

$$H_q(x) = \frac{1}{1-q} \log \left[1 - \left[\sum_{i=0}^n \frac{\xi_i^*}{2\lambda^{\delta+1}} (q+i)^{\frac{q+1}{2}} \Gamma\left(\frac{q+1}{2}\right) \right] \right]$$

4. Maximum Likelihood Estimation

The maximum likelihood estimates (MLEs) of the unknown parameters of the distribution based on $x = (x_1, \dots, x_n)$ observed sample value with the set of parameters $\ell(\alpha, \beta, \lambda | x)$ the log-likelihood function of the parameter $\ell(\alpha, \beta, \lambda)$ is given by

$$\ell = n \ln(2\alpha\beta\lambda^2) - n \ln(1-e^{-\beta}) + \sum_{i=1}^n \ln(x_i) - \lambda^2 \sum_{i=1}^n x_i^2 + (\alpha-1) \left[\sum_{i=1}^n \ln\left(1-e^{-(\lambda x_i)^2}\right) \right] - \beta \sum_{i=1}^n \left(1-e^{-(\lambda x_i)^2}\right)^{\alpha} \tag{30}$$

Maximum likelihood estimators of the parameters have been obtained by differentiating with respect to parameters and equating to zero, Let,

$$\omega_i = \frac{1}{(e^{(\lambda x_i)^2} - 1)} ; \text{ and } \phi_i = 1 - e^{-(\lambda x_i)^2}$$

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \ln(\phi_i) - \beta \sum_{i=1}^n \phi_i \ln(\phi_i) = 0 \tag{31}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} - \frac{n}{e^\beta - 1} - \sum_{i=1}^n \phi_i^\alpha = 0 \tag{32}$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{2n}{\lambda} - 2\lambda \sum_{i=1}^n x_i^2 + 2\lambda(\alpha - 1) \sum_{i=1}^n \omega_i x_i^2 - 2\alpha\beta\lambda \sum_{i=1}^n \omega_i \phi_i^\alpha x_i^2 = 0 \tag{33}$$

To estimate the unknown parameters α, β and λ by solving non-linear equations (31), (32) and (33), it is clear that these equations cannot be solved analytically. Therefore, we estimate the value of unknown parameters by applying the Newton-Raphson's iterative technique in the log-likelihood function of equation (30) directly, using *optim()* function in R software [9] and [23].

Let us denote the parameter vector by $\delta = (\alpha, \beta, \lambda)$ and the corresponding MLE of δ as $\hat{\delta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})$, then asymptotic normality result is $(\hat{\delta} - \delta) \rightarrow N_3 \left(0, (I(\delta))^{-1} \right)$, where $I(\delta)$, is the Fisher's Information Matrix. It is useless that the MLE has an asymptotic variance $(I(\delta))^{-1}$. Therefore, the common procedure is used for Observed Information Matrix $O(\hat{\delta})$ (as an estimation of the Fisher's Information Matrix $I(\delta)$) where the element of $O(\hat{\delta})$ are expresses in the following form

$$O(\hat{\delta}) = - \left(\begin{array}{ccc} \frac{\partial^2 \ell}{\partial \alpha^2} & \frac{\partial^2 \ell}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 \ell}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell}{\partial \beta^2} & \frac{\partial^2 \ell}{\partial \beta \partial \lambda} \\ \frac{\partial^2 \ell}{\partial \alpha \partial \lambda} & \frac{\partial^2 \ell}{\partial \beta \partial \lambda} & \frac{\partial^2 \ell}{\partial \lambda^2} \end{array} \right)_{(\hat{\alpha}, \hat{\beta}, \hat{\lambda})} = -H(\delta) \Big|_{\delta = \hat{\delta}} \tag{34}$$

Where, H is the Hessian matrix, $\delta = (\alpha, \beta, \lambda)$ and $\hat{\delta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})$. The variance-covariance matrix is obtained by the Newton Raphson algorithm to maximize the likelihood function of the Observed Information Matrix in (34).

$$(-H(\delta)_{|\delta=\hat{\delta}})^{-1} = \begin{pmatrix} \text{var}(\hat{\alpha}) & \text{cov}(\hat{\alpha}, \hat{\beta}) & \text{cov}(\hat{\alpha}, \hat{\lambda}) \\ \text{cov}(\hat{\alpha}, \hat{\beta}) & \text{var}(\hat{\beta}) & \text{cov}(\hat{\beta}, \hat{\lambda}) \\ \text{cov}(\hat{\alpha}, \hat{\lambda}) & \text{cov}(\hat{\beta}, \hat{\lambda}) & \text{var}(\hat{\lambda}) \end{pmatrix} \tag{35}$$

Finally, we construct an approximate 100 (1- γ) % confidence interval for α, β and λ is

$$\hat{\alpha} \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\alpha})} ; \hat{\beta} \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\beta})} \text{ and } \hat{\lambda} \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\lambda})} \tag{36}$$

where, $Z_{\gamma/2}$ is the upper percentile of standard normal variate.

5. Data analysis

5.1. Real DATA

In this section, we represent the analysis of one real data set to verify our proposed model. We consider a collection of real data provided in [2], which was the result of a test on the endurance of deep groove ball bearings. They were originally discussed by [19]. The data are; 17.88, 28.92, 33.0, 41.52, 42.12, 45.60, 48.40, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.4.

5.2. Parameter Estimation

We have computed the value of parameters by maximizing the log-likelihood function of our proposed model along with five alternative models named Generalized Rayleigh (GR) [17], Exponentiated Chen (EC) [11], Exponentiated Exponential (EE) [22], Exponentiated Inverted Weibull (EIW) [13] and Compound Rayleigh (CR) [27], by direct using *optim()* function in R software [9] and [23]. The PDF of compared models are shown in the Appendix. The values of MLE of each parameter with standard error are presented in table 2.

Table 2. The MLE value of each parameter with SE of different models.

Models	MLE (SE)			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$
ERP	1.01577 (0.21181)	0.50728 (0.72696)	0.01245 (0.00128)	-
GR	1.01073 (0.21075)	-	0.01238 (0.00128)	-
EC	1.69426 (0.86005)	0.34706 (0.03430)	0.01415 (0.01446)	-
EE	1.02866 (0.26864)	-	0.01414 (0.00362)	-
EIW	-	0.69531 (0.08791)	-	11.51702 (3.6530)
CR	0.12839 (0.02866)	1.68679 (1.05129)	-	-

5.3. Model Comparisons

We have found out the values of log-likelihood along with different criteria of goodness of fit (i) Akaike's information

criterion (AIC), (ii) Bayesian information criterion (BIC), (iii) Corrected Akaike's information criterion (CAIC), and (iv) Hannan-Quinn Information Criterion (HQIC): AIC=

$$-2\ell(\hat{\theta}) + 2k; \text{ BIC} = -2\ell(\hat{\theta}) + k \log(n); \text{ CAIC} = \text{AIC} + \frac{2k(k+1)}{n-k-1};$$

$$\text{and HQIC} = -2\ell(\hat{\theta}) + 2k \ln(\ell n(n))$$

Where k is the number of parameters in the model and n is the total sample under consideration.

According to AIC, BIC, CIAC, and HQIC, the lowest model value is the best model among the models of comparison. The intended model has a smaller value in comparison to other models mentioned in the table. Therefore, our model is better as compared to others.

Table 3. The value of log-likelihood, AIC, BIC, CAIC, and HQIC of different models.

Probability Models	AIC	BIC	CAIC	HQIC	$\ell(\hat{\theta})$
ERP	221.5938	225.0003	222.857	222.4506	-107.7924
GR	230.9916	233.2625	231.5916	231.5627	-113.4922
EC	234.5079	237.9144	235.7711	235.3646	-114.2539
EE	246.3234	248.5944	246.9234	246.8946	-121.1587
EIW	263.6078	265.8788	264.2078	264.1790	-129.7975
CR	303.4312	305.7022	304.0312	304.0024	-149.7066

5.4. Model Validation

The probability-probability (P-P) plots and quantile-quantile (Q-Q) plots can be inspected to verify this result. P-P and Q-Q plots show that the theoretical distribution versus empirical distribution. A P-P plot depicts the points:

$$\left(F(x_{(i)}), F(x_{(i)}; \hat{\delta}) \right); i = 1, 2, \dots, n$$

Where $\hat{\delta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ and $x_{(i)}$ is the order statistic of the

proposed model, $F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X \leq x)$ is the empirical distribution function, and $I(\cdot)$ is the indicator function.

Similarly, the Q-Q plot depicts the points:

$$\left(x_{(i)}, F^{-1}\left(\frac{i}{n+1}; \hat{\delta}\right) \right); i = 1, 2, \dots, n$$

In our proposed model it is observed that there is a good fit of theoretical distribution versus empirical distribution in both plots (Figure 3).

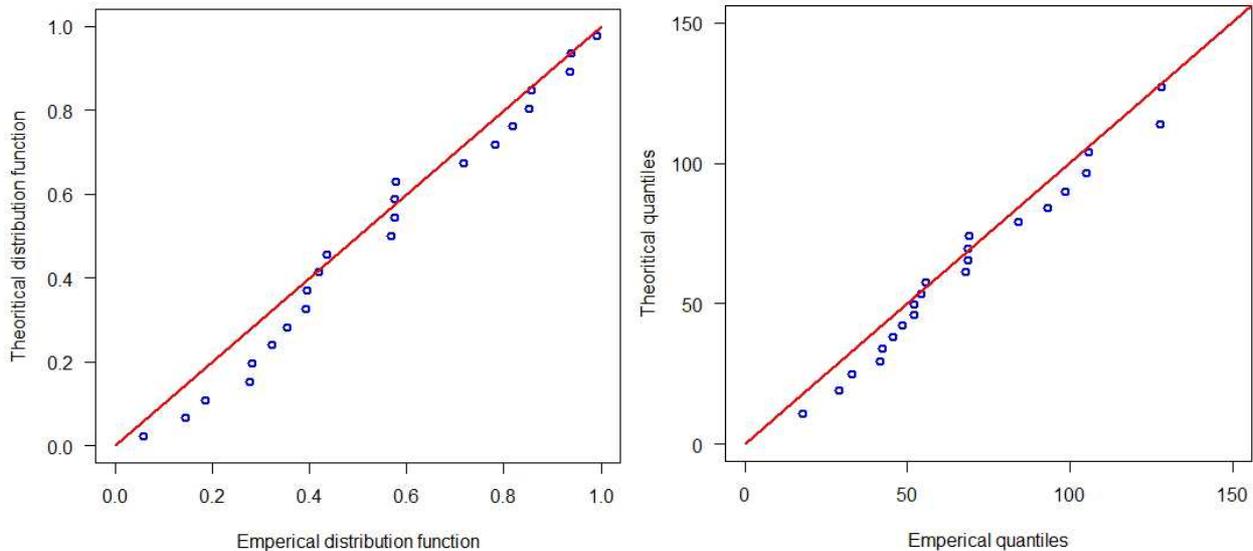


Figure 3. P-P plot (left panel) and Q-Q plot (right panel) of the proposed model.

Further, validation of the one-sample Kolmogorov-Smirnov test is applied. The K-S test (D=0.14543) is not statistically significant, suggesting that it fits well with the theoretical distribution versus empirical distribution (p-value=0.7155). We have plotted the empirical distribution function and fitted distribution function from Figure 4 (left panel), it is clear that the proposed model provides a satisfactory fit to the given data.

Total Time Test (TTT) plot is an important graphical method to check whether or not our data set can be applied to a

particular model. Due to [1], the empirical version of the

$$\text{TTT plot is given by plotting: } T\left(\frac{r}{n}\right) = \frac{\sum_{i=1}^r y_{i:n} + (n-r)y_{r:n}}{\sum_{i=1}^r y_{i:n}};$$

where, $(r=1, 2, \dots, n)$ $y_{i:n} (i=1, 2, \dots, n)$ and are the order statistics of the sample. Hence, the TTT plot of the data set is concave which indicates that increasing the hazard rate shape of the proposed distribution (Figure 4, right panel).

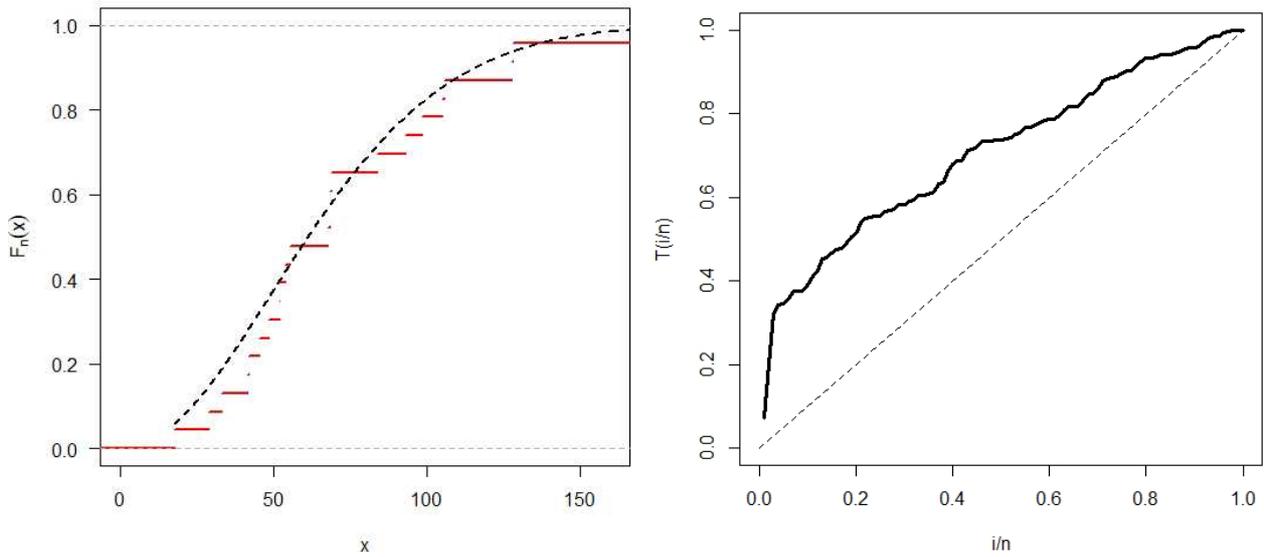


Figure 4. The graph of the empirical and fitted distribution of function (left panel), TTT plot.

Likewise, we compared the empirical distribution CDF with the estimated distribution function CDF of the proposed model along with five other models CR, GR, EE, EC, and EIW. Again, we compared the estimated PDF of the intended model with the estimated PDF of other well-known

distributions from the same data set. The built model is well adapted in both cases than other distributions. The developed model, therefore, represents an alternative model for real data modeling in different fields, with enormous advantages (Figure 5).

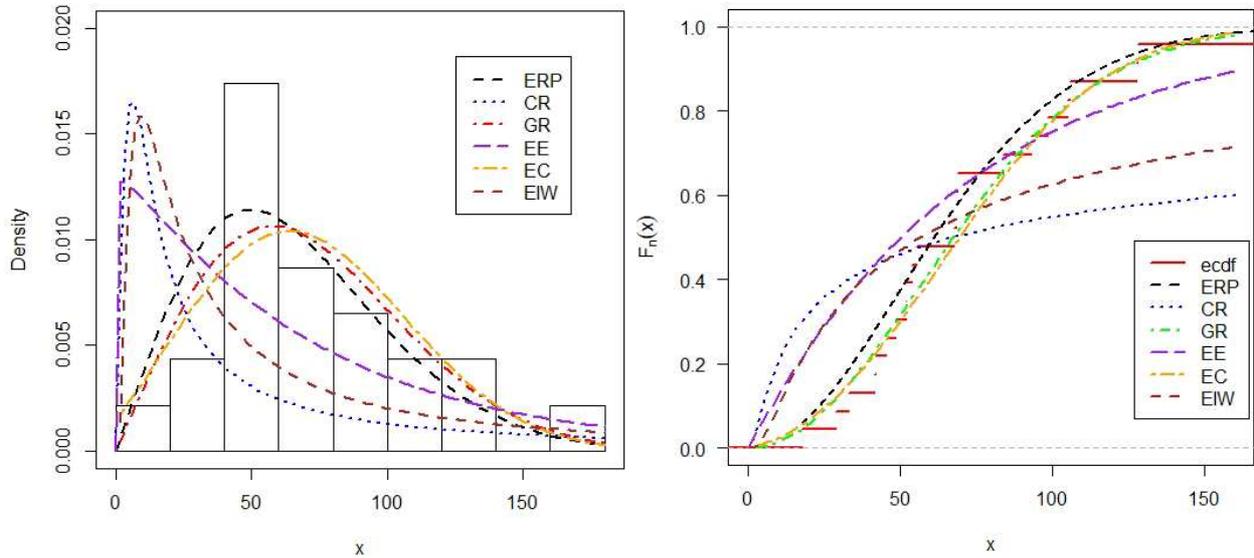


Figure 5. Estimated fitted CDF with EDF (left panel) and estimated fitted PDF (right panel).

6. Conclusion

In our study, we have introduced a new distribution named Exponentiated Rayleigh Poisson distribution in G family of distributions. Some mathematical properties of the new distribution including probability weighted moments, order statistics and their moments, residual life function, quantile function, skewness, kurtosis, and survival function entropy are derived well. The hazard function shows the upside curve (concave) shape. In model comparison, AIC, BIC, CIAC, and HQIC have been calculated. The proposed model has smaller

values in comparison to other models, indicates the better model. During the graphical comparison of empirical distribution CDF with the estimated distribution function (EDF) and PDF of the proposed model with other models, the proposed model is adapted well in both cases. Although this generalization technique can be used to generalize many other distributions for illustration purposes, we have chosen the Exponentiated Exponential Poisson G family model and Rayleigh distribution as base distributions. The importance and flexibility of the new family are illustrated by different examples mentioned above. We hope that this study will

serve as a reference and help to advance future research in the subject area.

Appendix

Generalized Rayleigh (GR);

$$f(x; \alpha, \lambda) = 2\alpha\lambda^2 x e^{-(\lambda x)^2} \left(1 - e^{-(\lambda x)^2}\right)^{\alpha-1}$$

Exponentiated Chen (EC);

$$f(x; \alpha, \beta, \lambda) = \alpha\beta\lambda x^{\beta-1} e^{x^\beta} e^{\lambda(1-e^{x^\beta})} \left(1 - e^{\lambda(1-e^{x^\beta})}\right)^{\alpha-1}$$

Exponentiated Exponential (EE);

$$f(x; \alpha, \lambda) = 2\alpha\lambda e^{-(\lambda x)} \left(1 - e^{-(\lambda x)}\right)^{\alpha-1}$$

Exponentiated Inverted Weibull (EIW);

$$f(x; \beta, \theta) = \beta\theta e^{-(\beta+1)x} \left(1 - e^{-x^\beta}\right)^\theta$$

Compound Rayleigh (CR);

$$f(x; \alpha, \beta) = 2\alpha\beta^\alpha x \left(\beta + x^2\right)^{-(\alpha+1)}$$

Conflicts of Interest

The authors declare that they have no competing interests.

Acknowledgements

The authors are grateful to the University Grant Commission, Sanathimi, Nepal for the partial support of the completion of this paper. The authors are also thankful to the editor and anonymous reviewer for their constructive comments and valuable suggestions which certainly improves the presentation and quality of the paper.

References

- [1] Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36 (1), 106-108.
- [2] Abouammoh, A. M., & Alshingiti, A. M. (2009). Reliability estimation of generalized inverted exponential distribution. *Journal of Statistical Computation and Simulation*, 79 (11), 1301-1315.
- [3] Abouelmagd, T. H. M., Hamed, M. S., & Abd El Hadi, N. E. (2017). The Poisson-G family of distributions with applications. *Pakistan Journal of Statistics and Operation Research*, 13 (2), 313-326.
- [4] Al-Noor, N. H., & Assi, N. K. (2020). Rayleigh-Rayleigh Distribution: Properties and Applications. In *Journal of Physics: Conference Series* (Vol. 1591, No. 1, p. 012038). IOP Publishing.
- [5] Al-saiary, Z. A., Bakoban, R. A., & Al-zahrani, A. A. (2020). Characterizations of the Beta Kumaraswamy Exponential Distribution. *Mathematics*, 8 (1), 23.
- [6] Alzaatreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions, *Metron*, 71, 63-79.
- [7] Amini, M., MirMostafae, S., and Ahmadi, J. (2013). Log-gamma-generated families of distributions, *Statistics*, 48 (4), 913-932.
- [8] Aryal, G. R., & Yousof, H. M. (2017). The exponentiated generalized-G Poisson family of distributions. *Stochastics and Quality Control*, 32 (1), 7-23.
- [9] Braun, W. J., & Murdoch, D. J. (2016). *A first course in statistical programming with R*. Cambridge University Press.
- [10] Cordeiro, G. M., Alizadeh, M., Ozel, G., Hosseini, B., Ortega, E. M. M., Altun, E. (2016). The generalized odd log-logistic family of distributions: Properties, regression models and applications, *Journal of Statistical Computation and Simulation*, 87 (5), 908-932.
- [11] Dey, S., Kumar, D., Ramos, P. L., & Louzada, F. (2017). Exponentiated Chen distribution: Properties and estimation. *Communications in Statistics-Simulation and Computation*, 46 (10), 8118-8139.
- [12] Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications, *Communications in Statistics-Theory and Methods*, 31, 497-512.
- [13] Flaih, A., Elsalloukh, H., Mendi, E., & Milanova, M. (2012). The exponentiated inverted Weibull distribution. *Appl. Math. Inf. Sci*, 6 (2), 167-171.
- [14] Gauss, M. C., Alizadeh, M., Ozel, G., Hosseini, B. Ortega, E. M. M., and Altunc, E. (2017).
- [15] Gomes, A. E., Da-Silva, C. Q., & Cordeiro, G. M. (2015). The exponentiated G Poisson model. *Communications in Statistics-Theory and Methods*, 44 (20), 4217-4240.
- [16] Jones, M. C. (2009). Kumaraswamy's distribution: a beta-type distribution with some tractability advantages, *Statistical Methodology*, 6, 70-81.
- [17] Kundu, D., & Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. *Computational statistics & data analysis*, 49 (1), 187-200.
- [18] Lemonte, A. J., Barreto-Souza, W., and Cordeiro, G. M. (2013). The exponentiated Kumaraswamy distribution and its log-transform, *Brazilian Journal of Probability and Statistics*, 27, 31-53.
- [19] Lieblein, J., & Zelen, M. (1956). Statistical investigation of the fatigue life of deep-groove ball bearings. *Journal of research of the national bureau of standards*, 57 (5), 273-316.
- [20] Merovci, F. (2014). Transmuted generalized Rayleigh distribution. *Journal of Statistics Applications & Probability*, 3 (1), 9.
- [21] Merovci, F., Alizadeh, M., Yousof, H. M., and Hamedani, G. G. (2017). The exponentiated transmuted-G family of distributions: Theory and applications, *Communications in Statistics-Theory and Methods*, 46 (21), 10800-10822.
- [22] Nadarajah, S. (2011). The exponentiated exponential distribution: a survey.
- [23] R Core Team R (2019). *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- [24] Ristic, M. M. and Balakrishnan, N. (2012). The gamma exponentiated exponential distribution, *Journal of Statistical Computation and Simulation*, 82, 1191-1206.
- [25] Ristic, M. M. and Nadarajah, S. (2014). A new lifetime distribution, *Journal of Statistical Computation and Simulation*, 84 (1), 135-150.
- [26] Roshini, G. and Thobias, S. (2017). Marshall-Olkin Kumaraswamy Distribution, *International Mathematical Forum*, 12 (2), 47-69.
- [27] Tahani, A. (2011). Estimation of the unknown parameters for the compound Rayleigh distribution based on progressive first-failure-censored sampling. *Open Journal of Statistics*, 2011.
- [28] Torabi, H. and Montazeri, N. H. (2012). The gamma-uniform distribution and its applications, *Kybernetika*, 48, 16-30.