Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/120316

Power Rayleigh Distribution for Fitting Total Deaths of Covid-19 in Egypt

N. M. Kilany[1,*], *Mohamed A. W. Mahmoud*[2] *and Lobna H. El-Refai*[1]

[1]Mathematics and Computer Science Department, Faculty of Science, Menoufia University, Menoufia, Egypt
[2]Mathematics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt

Received: 7 Sep. 2022, Revised: 21 Nov. 2022, Accepted: 23 Dec. 2022
Published online: 1 Sep. 2023

Abstract: The Rayleigh distribution incorporate the lifetime of an object or a service time. In this paper, a new model, called, Power Rayleigh distribution (PR) is submitted for specifying the confirmed total deaths of Corona virus (Covid-19) in Egypt. Statistical and reliability properties of the PR distribution such as survival function, failure rate function, mean residual life, order statistic and extreme value distribution are deduced and studied. Maximum likelihood method is used to evaluate the unknown parameters. Simulation Schemes are introduced. Finally, two sets of real-life data are construed and observed that the new model can provide a best fit to water runoff data and the total deaths of Covid-19 data than other well-known distributions.

Keywords: Rayleigh distribution; Hazard function; Maximum likelihood estimation; Goodness of fit; Covid-19

1 Introduction

Many research has provided variety models in order to investigate the lifetime of objects such as Rayleigh, Lomax, Weibull, Pareto distributions. Recently, these models have new extensions to have more reliability measures. Ahmad et al [1] discuss characterization of Weibull-Rayleigh Distribution. Weighted Lomax distribution is deduced by Kilany [2]. Kayid [3] estimate the parameters of one generalized mixture Pareto distribution by EM algorithm for complete and right-censored data.

The Rayleigh distribution is believed to be helpful lifetime distribution. It was first introduced by Rayleigh [4]. It has various applications, involving communication theory,reliability analysis,clinical studies, technology and applied statistics. If a random variable Y followed the Rayleigh distribution with one parameter then its PDF is obtained by:

$$f(y) = \frac{y}{\alpha^2} e^{\frac{-y^2}{2\alpha^2}}, \quad y > 0 \tag{1}$$

Many generalization and extensions of the Rayleigh distribution are introduced. Ogunsanya et al [5] introduced the Rayleigh-Cauchy Distribution to extend the two-parameter Cauchy distribution using T-RY families. Shrahili et al [6] proposed and studied the sine half-logistic inverse Rayleigh distribution as a new inverse Rayleigh distribution extension. To form the Reddit advertising and breast cancer data sets, Zhongjie et al [7] introduced a new extended shape of the generalized Rayleigh distribution . This distribution is named as a new generalized Rayleigh distribution and possesses heavy tailed properties. Falgore et al [8] apply the Inverse Lomax generator in order to have an extension of Rayleigh distribution named Inverse Lomax Rayleigh distribution . Properties of ILR were derived. Rivet et al. [9] discussed the Log Rayleigh distribution, Cordeiro et al. [10] derived Beta Generalized Rayleigh distribution, Weibull Rayleigh distribution is introduced by Merovci and Elbatal [11], Mahmoud and Ghazal [12] introduced exponentiated Rayleigh distribution. Several authors have considered extensions of Rayleigh distribution such as Inverse Rayleigh by Voda [13], Weighted Inverse Rayleigh distribution by Fatima and Ahmed [14] and also, Merovci [15] has constructed Transmuted Rayleigh distribution. Exponential Transformed Inverse Rayleigh distribution is discussed by Banerjee and Bhunia [16]. The quality of the procedures applied in statistical analysis relys heavily on the assumed probability distribution.The

* Corresponding author e-mail: neveenkilany@science.menofia.edu.eg.

Rayleigh distribution has been extended in this paper by using the power transformation $x = y^{\frac{1}{\beta}}$, see also Hassanein et al [17].The cumulative distribution function (CDF) of the X is presented by

$$F(x) = P(X \leq x) = P(Y \leq x^{\beta}) = F_y(x^{\beta})$$

Thus, the PDF of Power Rayleigh (PR) distribution is obtained as

$$f(x) = \frac{\beta}{\alpha^2} x^{2\beta-1} e^{\frac{-x^{2\beta}}{2\alpha^2}}, \qquad x > 0,\ \alpha,\ \beta > 0 \tag{2}$$

where α is a scale parameter and β is a shape parameter. The proper cumulative CDF is presented by

$$F(x) = 1 - e^{\frac{-x^{2\beta}}{2\alpha^2}} \tag{3}$$

This paper is arranged accordingly: in Section 2, statistical properties and reliability measures are discussed such as shapes of probability density function, shapes of failure rate function, the moments and some associated assess, the quantile function, skewness and kurtosis, mean residual life , Shannon entropy and stress-strength parameter. In Section 3, the maximum likelihood method and confidence interval are used to assess the two-parameter. In Section 4, the distribution of order statistic is introduced and limiting distribution of extreme value are derived. A simulation study is presented in Section 5. Finally, applications of the Power Rayleigh model to real-life data are illustrated in Section 6.

2 Statistical Concepts and Reliability Measures

In this section some statistical concepts and reliability assessment for the PR model are deduced and studied.

2.1 Behavior of Probability Density Function

At $x = 0$ and $x = \infty$, the attitude of PR distribution is provided

$$f(0) = \begin{cases} \infty, & \beta < \frac{1}{2} \\ \frac{1}{\alpha^2}, & \beta = \frac{1}{2} \\ 0, & \beta > \frac{1}{2} \end{cases},\ f(\infty) = 0.$$

The behavior of PDF of PR model is provided in the following theorem.

Theorem 1. *For all $\alpha > 0$ the PDF of Power Rayleigh model is*

(i) Decreasing if $\beta \leq \frac{1}{2}$.
(ii) Unimodal if $\beta > \frac{1}{2}$.

Proof. Since

$$\acute{f}(x) = \frac{\Psi(x)}{x} f(x),$$

where,

$$\Psi(x) = \frac{-\beta}{\alpha^2} x^{2\beta} + 2\beta - 1.$$

(i) For $\beta = \frac{1}{2}$, then $\Psi(x) = \frac{-1}{2\alpha^2} x < 0$ and $\acute{f}(x) < 0$. Hence, $f(x)$ is decreasing. Also, if $\beta < \frac{1}{2}$, then $\Psi(x) < 0$ and $\acute{f}(x) < 0$ for all $\alpha > 0$. Hence $f(x)$ is decreasing.
(ii) $\forall \alpha > 0$, $\acute{f}(x) = 0\ iff\ \ \Psi(x) = 0$ which arises at the point

$$x_0 = \left(\frac{2\alpha^2\beta - \alpha^2}{\beta}\right)^{\frac{1}{2\beta}}$$

Since, $\acute{\acute{f}}(x_0) = \frac{-2\beta^{\frac{1}{\beta}+1}(2\beta-1)^{1-\frac{1}{\beta}}}{\alpha^{\frac{2}{3}}} < 0$, so $f(x)$ has a mode at x_0 .

Figure 1 shows the attitude of PDF of PR model for some selected choices of α and β.

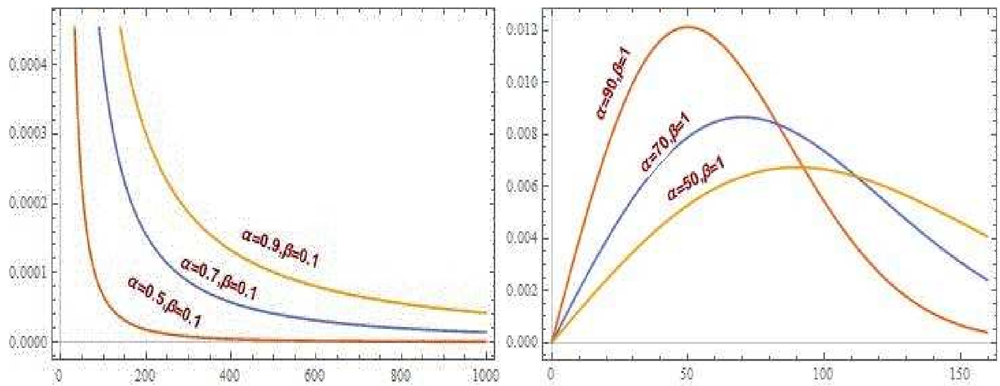

Fig. 1: Density function of the PR distribution.

2.2 Survival and Failure Rate Functions

The concerned event has not yet happened by time x. Thus, the survival function $S(x)$ denotes probability of surviving beyond time x which is defined by

$$S(x) = 1 - F(x) = e^{\frac{-x^{2\beta}}{2\alpha^2}} \tag{4}$$

The failure rate function is the conditional rate of failure at time x, assume that an individual remain alive until at least time x. The failure rate function of PR distribution is defined by

$$h(x) = \frac{f(x)}{S(x)} = \frac{x^{2\beta-1}}{\alpha^2}\beta \tag{5}$$

The shapes of hazard rate function of the PR distribution is discussed in the following theorem

Theorem 2. *For all $\alpha > 0$ the failure rate function of power Rayleigh distribution is*
(i) Decreasing if $\beta \leq \frac{1}{2}$.
(ii) Increasing if $\beta > \frac{1}{2}$.

Proof. The first derivative of $h(x)$ is obtained by
$\acute{h}(x) = \frac{\beta}{\alpha^2}x^{2\beta-2}(2\beta-1)$

So,

(i) if $\beta \leq \frac{1}{2}$, then $\acute{h}(x) \leq 0$ and this means that $h(x)$ is decreasing.
(ii) if $\beta > \frac{1}{2}$, then $\acute{h}(x) > 0$ and $h(x)$ is increasing.

The different behavior of failure rate function for PR model are shown in Figure2.

2.3 Mean Residual Life Function (MRL)

Suppose that $S(x)$ is the survival function of a continuous random variable X, the MRL function is specified as the expected value of the remaining lifetimes after a fixed time point x. The MRL of the PR distribution is given as follows

$$\mu(x) = \frac{1}{S(x)}\int_x^\infty y f(y)dy - x$$
$$= e^{\frac{x^{2\beta}}{2\alpha^2}}\int_x^\infty y \times \frac{\beta}{\alpha^2}y^{2\beta-1}e^{\frac{-y^{2\beta}}{2\alpha^2}}dy - x$$

Hence, $\mu(x) = 2\alpha^2 x^{1-2\beta}, \qquad x > 0.$

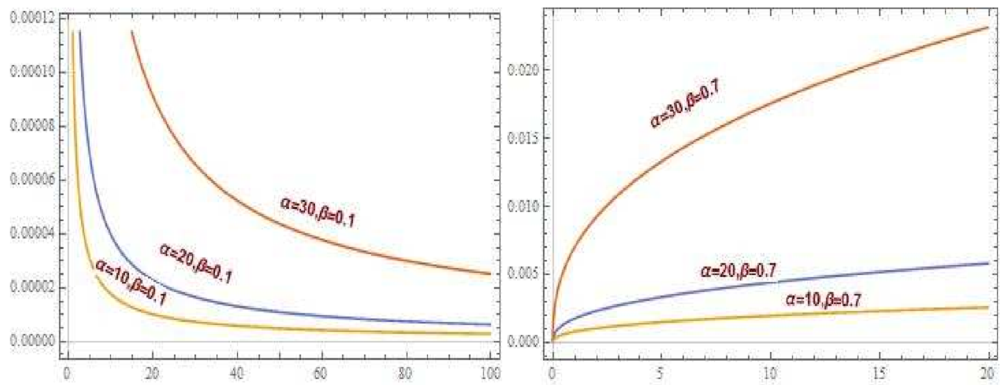

Fig. 2: Shapes of $h(x)$ of the PR model.

Theorem 3. *The MRL function $\mu(x)$ of the PR model is increasing if $\beta \leq \frac{1}{2}$and decreasing if $\beta > \frac{1}{2}$ for all $\alpha > 0$.*

Proof. Bryson and siddique [18] proved that when the $h(x)$ is increasing(deacreasing) then the corresponding mean residual life function will be increasing (decreasing). Thus, this provided the proof of the theorem.

Figure 3 displays that the shapes of MRL function of the PR distribution.

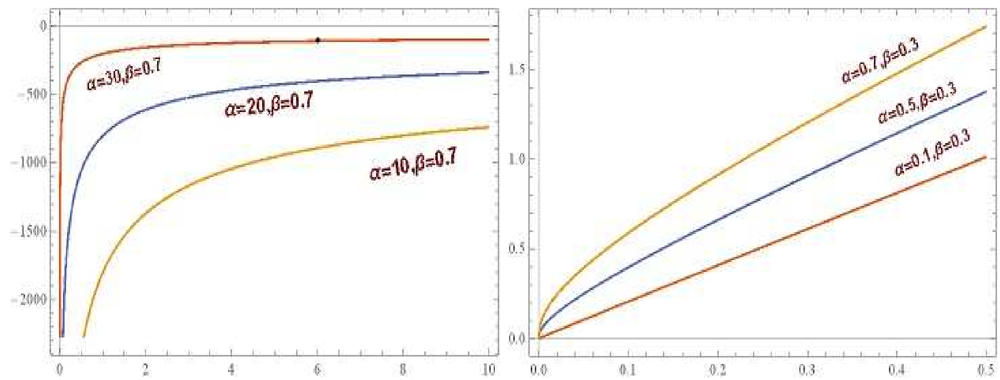

Fig. 3: Shapes of mean residual life of the Power Rayleigh model

2.4 Moments

A raw moment of order k is the average of all numbers in the set, with each number raised to the k^{th} power before you average it. The first and second raw moment provide some information about the location, variability and appearance of the distribution.The third and the forth raw moments provide some information on the shape of distribution. In this section we introduce the k^{th} moments. The k^{th} moment of Power Rayleigh distribution is defined by

$$\acute{\mu}_1 = 2^{\frac{1}{2\beta}} (\frac{1}{\alpha^2})^{\frac{-1}{2\beta}} \Gamma(1 + \frac{1}{2\beta})$$
$$\acute{\mu}_2 = 2^{\frac{1}{\beta}} (\frac{1}{\alpha^2})^{\frac{-1}{\beta}} \Gamma(1 + \frac{1}{\beta})$$
.
$$\acute{\mu}_n = 2^{\frac{n}{2\beta}} (\frac{1}{\alpha^2})^{\frac{-n}{2\beta}} \Gamma(1 + \frac{n}{2\beta})$$

The central moments of PR distribution are as follows

$\mu_2 = \frac{2^{\frac{1}{\beta}}(\frac{1}{\alpha^2})^{\frac{-1}{\beta}}[-\beta\,\Gamma(1+\frac{1}{2\beta})^2+\Gamma(\frac{1}{\beta})]}{\beta}$.

$\mu_3 = \frac{2^{-2+\frac{3}{2\beta}}(\frac{1}{\alpha^2})^{\frac{-3}{2\beta}}[\Gamma(\frac{1}{2\beta})^3-6\,\beta\,\Gamma(\frac{1}{2\beta})\,\Gamma(\frac{1}{\beta})+6\beta^2\Gamma(\frac{3}{2\beta})]}{\beta^3}$.

$\mu_4 = \frac{4^{-2+\frac{3}{2\beta}}(\frac{1}{\alpha^2})^{\frac{-2}{\beta}}[-3\Gamma(\frac{1}{2\beta})^4+24\beta\,\Gamma(\frac{1}{\beta})-48\beta^2\Gamma(\frac{3}{2\beta})+32\beta^3\Gamma(\frac{2}{\beta})]}{\beta^4}$

Hence, the standard deviation (SD) of the PR distribution is

$$SD = \sqrt{\mu_2} = \sqrt{\frac{2^{\frac{1}{\beta}}(\frac{1}{\alpha^2})^{\frac{-1}{\beta}}[-\beta\,\Gamma(1+\frac{1}{2\beta})^2+\Gamma(\frac{1}{\beta})]}{\beta}}, \qquad \beta > 0$$

The coefficient of variation performs the proportion of the standard deviation to the mean, and it is a helpful relation for matching the degree of variation from one data series to another, even if the means are seriously diverse from one another. Thus, the coefficient of variation of the PR distribution is given by

$$\rho = \frac{SD}{\mu} = \frac{2^{\frac{-1}{2\beta}}(\frac{1}{\alpha^2})^{\frac{1}{2\beta}}\sqrt{\frac{2^{\frac{1}{\beta}}(\frac{1}{\alpha^2})^{\frac{-1}{\beta}}[-\beta\,\Gamma(1+\frac{1}{2\beta})^2+\Gamma(\frac{1}{\beta})]}{\beta}}}{\Gamma(1+\frac{1}{2\beta})}, \qquad \beta > 0$$

The skewness and kurtosis statistics are based on the sample size. Smaller sample sizes can accord results that are very deceptive. Skewness measures the relative size of the two tails. Skewness of the PR distribution can be obtained as follows :

$$S = \sqrt{\frac{\mu_3^2}{\mu_2^3}} = \frac{1}{4}\sqrt{\frac{[\Gamma(\frac{1}{\beta})^3-6\beta\Gamma(\frac{1}{2\beta})\Gamma(\frac{1}{\beta})\,-\,6\beta^2\Gamma(\frac{3}{2\beta})]^2}{\beta^3(-\beta\,\Gamma(1+\frac{1}{2\beta})^2+\Gamma(\frac{1}{\beta}))^3}}, \qquad \beta > 0$$

Kurtosis is the measurement of the joint weight of a distribution's tails relative to the center of the model. The kurtosis of the PR model is presented by

$$K = \frac{\mu_4}{\mu_2^2} = \frac{-3\,\Gamma(\frac{1}{2\beta})^4+24\beta\,\Gamma(\frac{1}{2\beta})^2\,\Gamma(\frac{1}{\beta})\,-\,48\,\beta^2\Gamma(\frac{1}{2\beta})\,\Gamma(\frac{3}{2\beta})\,+\,32\Gamma(\frac{2}{\beta})}{16\beta^2\,[-\beta\,\Gamma(1+\frac{1}{2\beta})^2-\Gamma(\frac{1}{\beta})]^2}, \qquad \beta > 0$$

Let X be a random variable from our model, therefore the moment generating function of PR distribution is defined by

$$\begin{aligned} M_X(t) &= \int_0^\infty e^{tx} f(x)dx \\ &= 1+\sum_{r=1}^{\infty}\frac{t^r}{r!}2^{\frac{r}{2\beta}}(\frac{1}{\alpha^2})^{\frac{-r}{2\beta}}\Gamma(1+\frac{r}{2\beta}) \end{aligned}$$

The characteristic function of PR distribution is derived as follows:

$$\begin{aligned} \phi_X(t) &= \int_0^\infty e^{itx} f(x)dx \\ &= 1+\sum_{r=1}^{\infty}\frac{(it)^r}{r!}2^{\frac{r}{2\beta}}(\frac{1}{\alpha^2})^{\frac{-r}{2\beta}}\Gamma(1+\frac{r}{2\beta}) \end{aligned}$$

2.5 Quantile Function

For the Power Rayleigh model, the quantile function is obtained by

$F^{-1}(u) = (-2\alpha^2 \ln(1-u))^{\frac{1}{2\beta}}$.

Quantile is useful measure because it is less susceptible to long tailed distribution and it is clear that the quantile function is more helpful descriptive statistics than means and other moments. Some quantiles have special names (see [19]).

– if $u = \frac{1}{2}$ then the quantile function is called median, so the median of the PR distribution is given by

$Q_1 = F^{-1}(\frac{1}{2}) = (-2\alpha^2 \ln(\frac{1}{2}))^{\frac{1}{2\beta}}$

– if $u = \frac{1}{4}$ then the quantiles are called the first quartile, so the first quartile of the PR distribution is given by

$Q_2 = F^{-1}(\frac{1}{4}) = (-2\alpha^2 \ln(\frac{3}{4}))^{\frac{1}{2\beta}}$

– if $u = \frac{3}{4}$ then the quantiles are called the third quartile, so the third quartile of the PR distribution is given by

$Q_3 = F^{-1}(\frac{3}{4}) = (-2\alpha^2 \ln(\frac{1}{4}))^{\frac{1}{2\beta}}$

Figure 4 shows the quantiles at $\alpha = 1$ for the PR distribution.

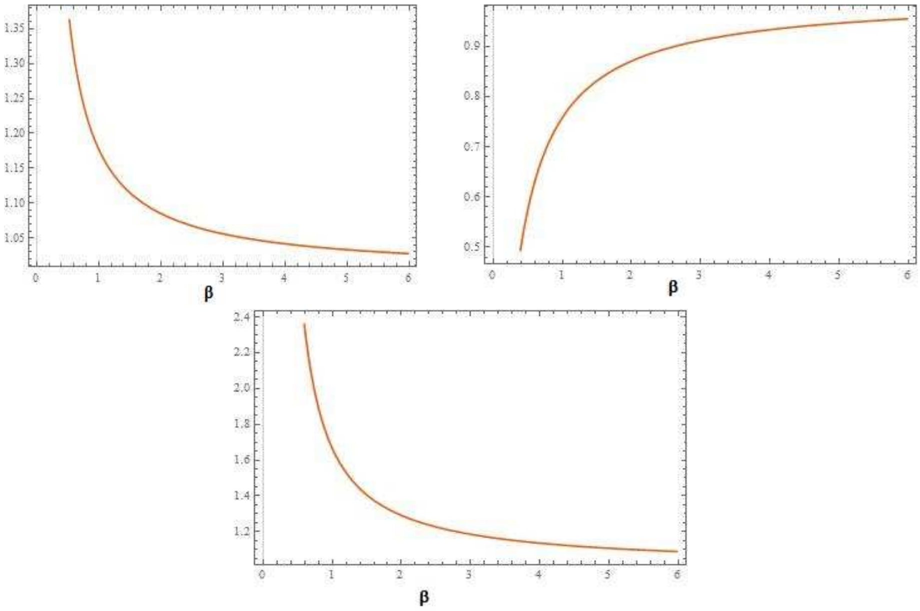

Fig. 4: Plots of quantile function of Power Rayleigh distribution for α=1.

2.6 Shannon Entropy

An entropy is interpreted as the degree of randomness in the system and many fields employ Shannon entropy such as biology, physics and chemistry as a Motivating force for protein unfolding . The Shannon entropy of random variable X is specified by

For the PR model, Shannon entropy follows as

$$S_H = -\int_0^\infty f(x)\log f(x)dx$$
$$= -\int_0^\infty \frac{\beta}{\alpha^2} x^{2\beta-1} e^{\frac{-x^{2\beta}}{2\alpha^2}} \log(\frac{\beta}{\alpha^2} x^{2\beta-1} e^{\frac{-x^{2\beta}}{2\alpha^2}})dx$$
$$= -\frac{1}{2}\beta \log(\frac{1}{\alpha^2})^{\frac{1}{2\beta}} \Gamma(2-\frac{1}{2\beta}), \qquad \beta > \frac{1}{4}.$$

3 Methods of Estimations

In this section, we study the evaluation of the parameters for the PR distribution by applying the maximum likelihood method. Moreover, interval estimation is discussed based on Fisher information matrix.

3.1 Maximum likelihood estimation

Suppose a random sample of size n from PR distribution as $x_1, x_2, ..., x_n$, then the sample likelihood function of this model is given by

$$L(\alpha,\beta,x) = \prod_{i=1}^{\infty} f(x_{i,\alpha,\beta})$$

$$= (\frac{\beta}{\alpha^2})^n \prod_{i=1}^{\infty} x_i^{2\beta-1} e^{-\sum_{i=1}^{n} \frac{x_i^{2\beta}}{2\alpha^2}}$$

Thus, the log-likelihood function is as follows,

$$l(\alpha,\beta,x) = n\ln(\beta) - n\ln(\alpha^2) + 2\beta\sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\ln x_i - \frac{\sum_{i=1}^{n} x_i^{2\beta}}{2\alpha^2} \tag{6}$$

The estimated parameters $\overset{\wedge}{\beta}$ and $\overset{\wedge}{\alpha}$ are specified by having a solution for the following equations

$$\frac{n}{\beta} + 2\sum_{i=1}^{n}\ln x_i - \frac{\sum_{i=1}^{n} x_i^{2\beta}\ln x_i}{\alpha^2} = 0 \tag{7}$$

$$\frac{-2n}{\alpha} + \frac{\sum_{i=1}^{n} x_i^{2\beta}}{\alpha^3} = 0 \tag{8}$$

From Eqs (7) and (8) we have

$$\overset{\wedge}{\alpha} = \sqrt{\frac{\sum_{i=1}^{n} x_i^{2\beta}}{2n}} \tag{9}$$

By substituting α^2 in Eq (7), then we have

$$2\beta\frac{\sum_{i=1}^{n} x_i^{2\beta}\ln x_i}{\sum_{i=1}^{n} x_i^{2\beta}} - \frac{2\beta}{n}\sum_{i=1}^{n}\ln x_i = 1 \tag{10}$$

Taking log-function for Eq (10), we obtain

$$f(\beta) = \ln n[\sum_{i=1}^{n}(x_i^{2\beta}\ln x_i - x_i^{2\beta} - \ln x_i)] = 0 \tag{11}$$

The estimation of β can obtained by solving Eq (11) in one-variable Newton-Raphson optimization algorithm as follows

$$\beta_{k+1} = \beta_k - \frac{f(\beta_k)}{\acute{f}(\beta_k)}, \qquad \text{where } k = 0,1,2,.... \tag{12}$$

3.2 Interval Estimation

For finding the interval estimation of (α,β), the elements of Fisher information matrix $I=[I_{ij}], i,j=1,2$ are given by

$$I_{11}=-E\left(\frac{\partial^2 \ln f(x)}{\partial\alpha^2}\right)=\frac{2}{\alpha^2}-\frac{2^{\frac{1}{2}}\left(\frac{1}{\alpha^2}\right)^{\frac{-1}{\beta}}\Gamma(1+\frac{1}{\beta})}{\alpha^4}$$
$$I_{22}=-E\left(\frac{\partial^2 \ln f(x)}{\partial\beta^2}\right)=-4-\frac{1}{\beta^2}$$
$$I_{12}=-E\left(\frac{\partial^2 \ln f(x)}{\partial\alpha\partial\beta}\right)=\frac{4\ln x}{\alpha}$$

Applying the large-sample theory of maximum likelihood estimators gives

$$\sqrt{n}\binom{\hat{\alpha}-\alpha}{\hat{\beta}-\beta}\xrightarrow{d} N_2(0,I^{-1}),$$

where $\xrightarrow{d}$ indicates convergence in model and I^{-1} is the inverse of the matrix I. The asymototic variances and covariances of α and β are provided by

$$var(\hat{\alpha})=\frac{I_{22}}{n\triangle}\quad var(\hat{\beta})=\frac{I_{11}}{n\triangle}\quad cov(\hat{\beta})=\frac{-I_{12}}{n\triangle}$$

where $\triangle=I_{11}I_{22}-I_{12}^2$ is the determinant of matrix I. The corresponding asymptotic $100(1-\alpha)\%$ confidence interval of $\hat{\alpha}$ and $\hat{\beta}$, respectively, are provided by

$$\hat{\alpha}\pm Z_{\frac{\alpha}{2}}\sqrt{var(\hat{\alpha})},\qquad \hat{\beta}\pm Z_{\frac{\alpha}{2}}\sqrt{var(\hat{\beta})},$$

where$Z_{\frac{\alpha}{2}}$ is the upper $\frac{\alpha}{2}$quantile of the standard normal distribution.

4 Order Statistic and Extreme Values

4.1 Distribution of Order Statistic

The order statistic of this model is used to known sometimes about how the order of the data behaved. For a sample observation $x_1,x_2,.....,x_n$ from the Power Rayleigh distribution which are independent, the sample values $x_{(1:n)}\leq x_{(2:n)}\leq ....\leq x_{(n:n)}$ which are ordered, named as the order statistic. Let $Y=X_{j:n}$ then the probability density function is defined by

$$f_y(y)=\frac{n!}{(j-1)!(n-j)!}\times F^{j-1}(y)\times\{1-F(y)\}^{n-j}\times f(y)$$
$$=\frac{(e^{-\frac{y^{2\beta}}{2\alpha^2}})^{n-j+1}\times(1-e^{-\frac{y^{2\beta}}{2\alpha^2}})^{j-1}y^{2\beta}\beta n}{\alpha^2(j-1)!(n-j)!},\qquad y>0$$

The CDF of the order statistic is obvioused by

$$F_y(y)=\sum_{m=j}^{n}\binom{n}{m}\times F^m(y)\times\{1-F(y)\}^{n-m}$$
$$=\frac{(e^{-\frac{y^{2\beta}}{2\alpha^2}})^{n-j}\times(1-e^{-\frac{y^{2\beta}}{2\alpha^2}})^{j}\times n\times{}_2F_1[1,j-n,1+j,1-e^{-\frac{y^{2\beta}}{2\alpha^2}}]}{j!(n-j)!},\qquad y>0$$

where ${}_2F_1$ is the hypergeometric function (see [20]).

4.2 Limiting Distribution of Extreme Values

Suppose that M_n and m_n is the maximum and minimum of the sample created from the Power Rayleigh model. The following theorem presented the limiting distributions of the extreme values.

Theorem 4. *Let M_n and m_n be the maximum and the minimum of a random sample$[X_1, X_2, ..., X_n]$ from the Power Rayleigh distribution. Then we have*

$$(i)\ \lim_{n\to\infty} P\left(\frac{M_n - a_n}{b_n} \leq x\right) = \exp(-e^{-x}), \qquad -\infty < x < \infty$$

$$(ii)\ \lim_{n\to\infty} P\left(\frac{m_n - c_n}{d_n} \leq x\right) = 1 - \exp(-x^{2\beta}), \qquad x > 0$$

where,

$$a_n = F^{-1}(1 - \frac{1}{n}), b_n = \frac{1}{nf(a_n)}, c_n = 0, d_n = \frac{1}{F^{-1}(n)}$$

Proof. For the PR distribution we obtain

$$(i)\ \lim_{x\to\infty} \frac{d}{dx}\left\{\frac{1}{h(x)}\right\} = \frac{\alpha^2}{\beta} \lim_{x\to\infty} \frac{d}{dx}\{x^{1-2\beta}\} = \frac{\alpha^2}{\beta} \lim_{x\to\infty} (1 - 2\beta)x^{-2\beta}$$
$$= \frac{\alpha^2}{\beta}\left(\frac{1-2\beta}{\infty}\right) = 0$$

From Arnold et al. [21], Theorem 8.3.3,, the maximal domain of attraction of the PR distribution is the standard Gumbel distribution and thus part (i) is proved.

(ii) By applying L'Hospital rule ,

$$\lim_{\varepsilon\to o} \frac{F(F^{-1}(0)+\varepsilon x)}{F(F^{-1}(0)+\varepsilon)} = \lim_{\varepsilon\to 0} \frac{F(\varepsilon x)}{F(\varepsilon)} = \lim_{\varepsilon\to 0} \frac{x^{2\beta} e^{-\frac{(\varepsilon x)^{2\beta}}{2\alpha^2}}}{e^{-\frac{x^{2\beta}}{2\alpha^2}}} = x^{2\beta}$$

From Arnold et al. [21], Theorem 8.3.6 , the minimal domain of attraction of the PR distribution is the standard Weibull distribution and this proved part (ii).

5 Simulation Study

The equation $F(x) - u = 0$ is used for implementing the simulation study by creating random samples from PR model. The simulation expriment was Reiterated is N= 1000 with sample sizes: 30, 50, 70, 90 for $(\alpha, \beta) = (0.5, 5)$ and $(0.7, 10)$. The next measurs are calculated:

(i) Average bias of the estimated parameters can be calculated from:

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha} - \alpha) \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{N}(\hat{\beta} - \beta)$$

(ii) The Mean square error (MSE) of the estimated parameter can be calculated from:

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha} - \alpha)^2 \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{N}(\hat{\beta} - \beta)^2$$

Table1 shows the average bias and the MSE of the estimates. It is clear that the positive and smallest values of the bias for PR model and where the sample size increases, the values of the MSE decreases.

*Table*1 : Bias and MSE for PR parameters α, β.

α	β	n	**Bias**(α)	**MSE**(α)	**Bias**(β)	**MSE**(β)
1	0.5	30	0.03288	0.02128	0.03052	0.00787
		50	0.01486	0.01006	0.01459	0.00354
		70	0.00901	0.00720	0.01025	0.00242
		90	0.00432	0.00532	0.00452	0.00169
0.7	10	30	0.00490	0.00551	0.61042	3.14893
		50	0.00186	0.00297	0.29198	1.41796
		70	0.00028	0.00214	0.20501	0.96660
		90	0.00011	0.00165	0.09048	0.67990

6 Application

The aim of this section is the two real data set which are fitted to the proposed distribution.

6.1 Hydrologic Data

A 34 storm events was realized from a watersheds. The data as described in Kang et al. [22] describes the water runoff (mm) of a study storms, which represents to one of the hydrological characteristics realized from a mall watershed in Korea (west of city of Suwon), the results being available since 1996. The data are as follows: 0.9, 0.6, 16.8, 59.3, 2, 78.2, 30.7, 146.8, 1.8, 3.4, 1.1, 0.8, 2.5, 6.1, 17, 5.1, 216.2, 8.1, 1.6, 2, 2, 0.8, 0.8, 2.9, 7.3, 13.3, 181.7, 20.5, 24.1, 33.5, 89.1, 7.2, 6, 75.9.

Many probability models such as Gumbel, Pareto, Generalized Extreme Value (GEV) and Generalized Logistic (GL) distribution are in use in the hydrological data analysis. Hydrological studies are lucrative in controlling water resources, planning and projection . The chosen of the more convenient probability distribution and related estimation procedure for parameter are the essential stage in hydrology analysis.

The PR distribution was fitted to runoff data using MLE, Kolmogorov-Smirnov test statistics and contrasted through goodness of fit measures for the next distributions which appropriate to these data,

– Gumbel distribution

$$f(x) = \frac{e^{-e^{-\frac{x-\mu}{\sigma}} - \frac{x-\mu}{\sigma}}}{\sigma}, \qquad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0$$

– Pareto distribution

$$f(x) = \frac{\alpha\beta^{\alpha}}{(x+\beta)^{\alpha+1}}, \qquad x > 0,\ \alpha,\ \beta > 0$$

– Generalized Pareto (GP) distribution

$$f(x) = \begin{cases} \frac{(1+k(\frac{x-\mu}{\sigma}))^{-\frac{1}{k}-1}}{\sigma}, & k \neq 0 \\ \frac{e^{-\left(\frac{x-\mu}{\sigma}\right)}}{\sigma}, & k = 0 \end{cases}$$

for $x \geq \mu$ when $k \geq 0$, and $\mu \leq x \leq \mu - \frac{\sigma}{k}$ when $k < 0$, where $-\infty < \mu, k < \infty, \sigma > 0$

– Generalized Extreme Value distribution (GEV)

$$f(x) = \begin{cases} \frac{(1+k(\frac{x-\mu}{\sigma}))^{-\frac{1}{k}-1} e^{-\left(1+k\left(\frac{x-\mu}{\sigma}\right)\right)^{\frac{-1}{k}}}}{\sigma}, & k \neq 0,\ 1+k\left(\frac{x-\mu}{\sigma}\right) > 0, \quad \sigma > 0 \\ \frac{e^{-\left(\frac{x-\mu}{\sigma} - e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)}}{\sigma}, & k = 0,\ -\infty < x < \infty, \quad \sigma > 0 \end{cases}$$

– Generalized Logistic distribution (GL)

$$f(x) = \begin{cases} \frac{(1+k(\frac{x-\mu}{\sigma}))^{-\frac{1}{k}-1}}{\sigma\left(\left(1+(1+k\left(\frac{x-\mu}{\sigma}\right)\right)\right)^{\frac{-1}{k}}\right)^{2}}, & k \neq 0,\ 1+k\left(\frac{x-\mu}{\sigma}\right) > 0, \quad \sigma > 0 \\ \frac{e^{-\left(\frac{x-\mu}{\sigma}\right)}}{\sigma\left(1+e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^{2}}, & k = 0,\ -\infty < x < \infty, \quad \sigma > 0 \end{cases}$$

Table 2: The Model Selection Criteria for the runoff data.

Distributions	**-log L**	**AIC**	**BIC**	**AICC**	**HQIC**	**CAIC**
Gumbel (μ, σ)	169.564	343.127	346.18	343.514	344.168	343.514
Pareto (α, β)	145.389	294.778	297.831	295.165	295.819	295.165
GP (μ, σ, k)	148.974	303.948	308.527	304.748	305.51	304.748
GEV (μ, σ, k)	150.814	307.628	312.207	308.428	309.189	308.428
GL (μ, σ, k)	151.498	308.996	313.575	309.796	310.557	309.796
PR (α, β)	140.349	284.698	287.75	285.085	285.739	285.085

Table 3: Estimates of the parameters for runoff data

Distributions	Estimates		
Gumbel (μ,σ)	12.194	25.186	–
Pareto (α,β)	0.3862	0.6	–
GP (μ,σ,k)	−3.1511	16.411	0.52449
GEV (μ,σ,k)	6.4599	12.399	0.59636
GL (μ,σ,k)	11.726	11.298	0.61583
PR (α,β)	1.6997	0.2957	–

Table 4: K-S goodness of fit test and P-Value for runoff data

Distributions	K-S	P-Value
Gumbel (μ,σ)	0.47722	0.0000003
Pareto (α,β)	0.15072	0.38418
GP (μ,σ,k)	0.19416	0.13453
GEV (μ,σ,k)	0.17514	0.22048
GL (μ,σ,k)	0.18030	0.19381
PR (α,β)	0.13869	0.52756

We use the criteria AIC (Akaike information criterion), BIC (Bayesian information criterion), AICC (Corrected Akaike information criterion), CAIC (consistent Akaike information criteria) and HQIC(Hannan-Quinn information criterion) (see Chen 1995 [23])to compare PR distribution with other models . The model with minimum AIC, BIC, AICC, CAIC and HQIC value is selected as the bestead model to fit the data. From Table 2, we determine that the power Rayleigh distribution is the bestead compared to the other models.

The maximum likelihood method is applied for evaluating the parameters of all the matched distributions and the evaluation of parameter are obtained in Tables 3. Further, Kolmogrov -Simrnov (K-S) goodness of fit test statistics utilized to check the fitting model of data set. The K-S statistics are detected for each distribution and written in Table 4. It obvious that the greatest P-value and the smallest Kolmogrov- Simrnov statistics for the Power Rayleight distribution than other distributions, thus, for the runoff data, the PR model is the bestead model.

6.2 COVID-19 Data

The rapid spread of Corona Virus (COVID-19) from China to other countries, makes its extremely important to study the number of all cases or deaths. The following data from the National Vital Statistics in Egypt [24] represents the overall deaths of Covid-19 since its begin in Egypt until May, 27th 2020. The data are as follows:1, 2, 2, 2, 2, 2, 4, 6, 6, 7, 8, 10, 14, 19, 21, 21, 24, 30, 36, 40, 41, 46, 52, 52, 66, 71, 78, 85, 103, 118, 135, 146, 159, 164, 178, 183, 196, 205, 224, 239, 250, 264, 264, 276, 294, 307, 337, 359, 380, 392, 406, 415, 429, 436, 452, 469, 482, 503, 514, 525, 533, 544, 556, 571, 592, 612, 630, 645, 659, 680, 696, 707, 735, 764, 783, 797.

The PR distribution yields a better fit for Covid-19 deaths data compared while the following well-known distributions

–Rayleigh distribution with density function

$$f(x)=\frac{x}{\alpha^2}e^{\frac{-x^2}{2\alpha^2}},\qquad x>0,\ \alpha>0$$

–Pareto distribution

$$f(x)=\frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}},\qquad x>0,\ \alpha,\ \beta>0$$

–Log-logistic distribution

$$f(x)=\frac{(\beta/\alpha)\ (x/\alpha)^{\beta-1}}{(1+(x/\alpha)^\beta)^2},\qquad x>0,\ \alpha,\ \beta>0$$

–Lognormal distribution

$$f(\ln x)=\frac{1}{\sigma\sqrt{2\pi}}\exp[\frac{-(\ln x-\mu)^2}{2\sigma^2}],\quad x>0,\ \mu,\ \sigma>0$$

–Gamma distribution
$f(x) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, \qquad x > 0, \alpha, \beta > 0$

–Exponential distribution
$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0, \lambda > 0$

–Lomax distribution
$f(x) = \frac{\alpha}{\beta}[1 + \frac{x}{\beta}]^{-(\alpha+1)}, \quad x \geq 0, \alpha, \beta > 0$

Table 5: Model Selection Criteria for the COVID-19 deaths data.

Distributions	**-log L**	**AIC**	**BIC**	**AICC**	**HQIC**
Rayleigh (σ)	578.04	1158.08	1160.42	1158.13	1159.02
Pareto (α, β)	558.658	1121.32	1126.0	1121.48	1123.19
Log-logistic (α, β)	520.223	1044.45	1049.13	1044.61	1046.32
Lognormal (μ, σ)	518.376	1040.75	1045.44	1040.91	1042.63
Gamma (α, β)	517.502	1039.0	1043.69	1039.17	1040.88
Exponential (λ)	510.172	1022.34	1024.69	1022.4	1023.28
Lomax (α, β)	510.272	1024.54	1029.23	1024.71	1026.42
PR (α, β)	506.967	1017.93	1022.62	1018.1	1019.81

Table 6: Estimates of the parameters for COVID-19 deaths data.

Distributions	**Estimates**	
Rayleigh (σ)	221.37	–
Pareto (α, β)	0.21249	0.999
Log-logistic (α, β)	0.85112	107.79
Lognormal (μ, σ)	4.7062	1.8349
Gamma (α, β)	1.234	224.83
Exponential (λ)	0.0036	–
Lomax (α, β)	89.194	24072.0
PR (α, β)	6.25624	0.395029

Table 7: K-S goodness of fit test and P-Value for COVID-19 deaths data.

Distributions	**K-S**	**P-Value**
Rayleigh (σ)	0.29257	3.76838×10^{-6}
Pareto (α, β)	0.29935	2.03129×10^{-6}
Log-logistic (α, β)	0.166381	0.0281564
Lognormal (μ, σ)	0.162764	0.0338222
Gamma (α, β)	0.18272	0.011675
Exponential (λ)	0.140966	0.0937465
Lomax (α, β)	0.136609	0.11293
PR (α, β)	0.116209	0.249452

From Table 5, the information criteria AIC, BIC, CAIC and HQIC show that the proposed model exhibits the least loss of information behaviour for Covid-19 data. In Table 7 clarifies the greatest P-value and the smallest Kolmogrov- Simrnov statistics for the Power Rayleight distribution than other distributions. Thus, the Kolmogrov- Simrnov (K-S) goodness of fit statistics indicates the proposed distribution its the bestead one for fitting the total number of Covid-19 deaths in Egypt comparable to other models.

7 Conclusion

The two parameters distribution called the Power Rayleigh distribution is proposed by power transformation.This model is more flexible than the Rayleigh distribution in the area of reliability studies. In terms of the probability density, failure rate and mean residual life functions are studied. Two numerical examples of realistic data are utilized to display the importance of this new model by matching it to other distributions.

References

[1] A. Ahmad, S.P. Ahmad and A. Ahmed, Characterization and Estimation of Weibull-Rayleigh Distribution with Applications to Life Time Data, Applied Mathematics and Information Sciences Letters, 5(2), 71-79 (2017).
[2] N.M. Kilany, Weighted Lomax distribution. SpringerPlus .5(1): 1862 (2016).
[3] M. Kayid, One Generalized Mixture Pareto Distribution and Estimation of the Parameters by the EM Algorithm for Complete and Right-Censored Data.IEEE Access, 9, 149372-149382 (2021)
[4] J. Rayleigh, On the Resultant of A large Number of Vibrations of The Same Pitch and of Arbitrary Phase. Philos. Mag., 10, 73-78 (1980).
[5] A. S. Ogunsanya, E. E. Akarawak and M. I. Matthew, On some properties of Rayleigh-Cauchy distribution, Journal of Statistics and Management Systems, 1213-1231 (2021).
[6] M. Shrahili, I. Elbatal and M. Elgarhy, Sine Half-Logistic Inverse Rayleigh Distribution: Properties, Estimation, and Applications in Biomedical Data, Journal of Mathematics, vol. 2021, ID 4220479, 10 pages, https://doi.org/10.1155/2021/4220479 (2021).
[7] S. Zhongjie, A. Alrumayh, Z. Ahmad, R. Abu-Shanab, M. Al - Mutairi, R. and Aldallal, A new generalized rayleigh distribution with analysis to big data of an online community, Alexandria Engineering Journal, 61,12, 11523-11535 (2022).
[8] J. Y. Falgore, M. N. Isah and H. A. Abdulsalam, Inverse Lomax-Rayleigh distribution with application, Heliyon, 7, 11, e08383, https://doi.org/10.1016/j.heliyon.2021.e08383 (2021).
[9] B. Rivet, L. Girin and C. Jutten, Log-Rayleigh Distribution: A Simple and Efficient Statistical Representation of Log-Spectral Coefficients. IEEE Transactions on Audio, Speech, and Language Processing, Institute of Electrical and Electronics Engineers, 15, 796-802 (2007).
[10] G. M. Cordeiro, C. T. Cristino, E. M. Hashimoto and E. M. Ortega, M. The Beta Generalized Rayleigh Distribution with Applications to Lifetime Data. Statistical Papers, 54, 133-161 (2011).
[11] F. Merovci, and I. Elbatal, Weibull Rayleigh Distribution:Theory and Applications. Applied Mathematics & Information Sciences, 9, 2127-2137 (2015).
[12] M. A. W. Mahmoud and M. G. M. Ghazal, Estimations from The Exponentiated Rayleigh Distribution Based on Generalized Type-II Hybrid Censored Data. Journal of the Egyption Mathematical Society , 25, 71-78 (2016).
[13] V. G. Voda, On The Inverse Rayleigh Distributed Random Variable. Reports of Statistical Application Research., 19, 13-21 (1972).
[14] K. Fatima and S. P. Ahmed, Weighted Inverse Rayleigh Distribution. International Journal of Statistic and Systems, 12, 119-137 (2017).
[15] F. Merovci, Transmuted Rayleigh Distribution. Austrian Journal of Statistics, 42, 21-31 (2013).
[16] P. Banerjee and S. Bhunia, Exponential Transformed Inverse Rayleigh Distribution: Statistical Properties and Different Methods of Estimation. Austrian Journal of Statistics, 51, 60-75 (2022).
[17] A. El-Houssainy Rady, W. A. Hassanein and T. A. Elhaddad, The power Lomax distribution with an application to bladder cancer data. Springerplus. 5(1): 1838 (2016).
[18] M. C. Bryson, and M. M. Siddique, Some Criteria for Aging. Journal of the American Statistical Association., 64, 1472–1483 (1969).
[19] M. E. Ghitany, D. K. AL-Mutairi, N. Balakrishnan and L. J. AL-ENezi, Power Lindely Distribution and Associated Inference. Computational Statistics and Data Analysis, 64, 20-33 (2013).
[20] N. N. Lebedev, Special Functions and their Application. Prentice-Hall (1965).
[21] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, A Frist Course in Order Statistics. Wiley, New York (1992).
[22] M. S. Kang, J. H. Goo, I. Song , J. A. Chun , Y .G. Her, S. W. Hwang and S. W. Park, Estimating design floods based on the critical storm duration for small watersheds, Journal of Hydro-environment Research, 7, 209–218 (2013).
[23] G. Chen and N. Balakrishnan, A general Purpose Approximate Goodness-of-Fit Test. Journal of Quality Technology., 27, 154–161, 1995 (1995).
[24] https://ourworldindata.org/coronavirus/country/egypt.country= EGY.