The Power Rayleigh Distribution with an Application on Hydrological Data

Mohamed A.W. Mahmoud(1), N.M.Kilany(2)*, L.H.El-Refai(2)

(1)Department of Mathematics, Faculty of Science, Al-Azhar University, Egypt.

(2)Department of Mathematics, Faculty of Science, Menoufia University, Egypt.

February 12, 2020

Abstract

The Rayleigh distribution is used to model the lifetime of an object or a service time. In this paper, a new distribution with two parameters (Power Rayleigh distribution) is introduced. Statistical properties of the distribution such as density function, survival function, hazard function, moments, quantile function, residual life, order statistic and extreme value distribution are discussed. Maximum likelihood method is used to estimate the unknown parameters.Simulation Schemes are produced. Finally, an application of the model to real data set is presented to show the superiority of this new distribution by comparing the fitness with its special cases.

KeyWords: Rayleigh distribution, Hazard function, Mean residual life, Entropy, Maximum likelihood estimation.

1 Introduction

The Rayleigh distribution is considered to be a useful lifetime distribution. It was first introduced by *Rayleigh* [1]. It serves as an important model in communication theory, physical science, engineering and medical imaging science. In engineering, Rayleigh distribution measure the lifetime of an object where the lifetime depends on the object's age for example resistor, transformer and capacitors in aircrafts. It is a special case of the two parameter Weibull distribution with the shape parameter equal to 2. A random variable Y is said to have the Rayleigh distribution with parameter α if its probability density function (PDF) is given by:

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

Inference for model of Rayleigh distribution has been introduced by *Sinha* and *Howlader* [2], *Mishra* et al. [3] and *Abd Elfattah* et al. [4] Over the years the family of Rayleigh distribution is formed such as Generalized Rayleigh distribution is introduced by *Kundu* and *Raqab* [5], *Rivet* et al. [6] discussed the Log Rayleigh distribution, *Cordeiro* et al. [7] derived Beta Generalized Rayleigh distribution, Weibull Rayleigh distribution is introduced by *Merovci* and *Elbatal* [8], *Mahmoud* and *Ghazal* [9] introduced exponentiated Rayleigh distribution.

Several authors have considered extentions of Rayleigh distribution such as Inverse Rayleigh by *Voda* [10], Weighted Inverse Rayleigh distribution by *Fatima* and *Ahmed* [11] and Transmuted Rayleigh distribution by *Merovci* [12].

The quality of the procedures used in statistical analysis depends heavily on the assumed probability model or distribution. The Rayleigh distribution has been extended in this paper by using the power transformation $x = y^{\frac{1}{\beta}}$.The cumulative distribution function (CDF) of the X is given 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

$$\begin{aligned} f(x) &= \frac{d}{dx} F_y(x^{\beta}) \\ &= \frac{\beta}{\alpha^2} x^{2\beta-1} e^{\frac{-x^{2\beta}}{2\alpha^2}} \end{aligned} \tag{2}$$

Where α is a scale parameter and β is a shape parameter. The corresponding cumulative CDF is given by

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

This paper is outlined as follows: In Section 2, statistical properties and reliability measures are discussed such as shapes of probability density function, shapes of hazard function, the moments and some associated measures, the quantile function, skweness and kurtosis, mean residual life , shannon entropy and stress-strength parameter. In Section 3, the maximum likelihood method and confidence interval are used to estimate the two-parameter. In Section 4, the distribution of order statistic is derived and limiting distribution. A simulation study is produced to generate random samples follow PR distribution. Application of the distribution to data set is used and compared to the fit attained by some other distributions.

2 Statistical Properties and Reliability Measures

In this section some statistical properities and reliability measures for the PR distribution are derived and studied.

2.1 Shapes of Probability Density Function

The behavior of PDF of the PR distribution f(x) at $x = 0$ and $x = \infty$ is given by

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

The following theorem shows the shape of the PDF of PR distribution.

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

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

Proof. The first derivative of f(x) is given by ∎

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

where,

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

(i) If $\beta = \frac{1}{2}$, then $g(x) = \frac{-1}{2\alpha^2} x < 0$ and $\acute{f}(x) < 0$. Hence, $f(x)$ is decreasing. Also, if $\beta < \frac{1}{2}$,then $g(x) < 0$ and $\acute{f}(x) < 0$ for all $\alpha > 0$. Hence $f(x)$ is decreasing.

(ii) $\forall \alpha > 0$, $\acute{f}(x) = 0 \; iff \;\; g(x) = 0$ which occurs at the point

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

Since,

$$\acute{f}(x_0) = \frac{-2\beta^{\frac{1}{\beta}+1}(2\beta-1)^{1-\frac{1}{\beta}}}{\alpha^{\frac{2}{3}}} < 0$$

So $f(x)$ has a local maximum at x_0 . Figure1 shows the behavior of PDF of PR distribution for some selected choices of α and β.

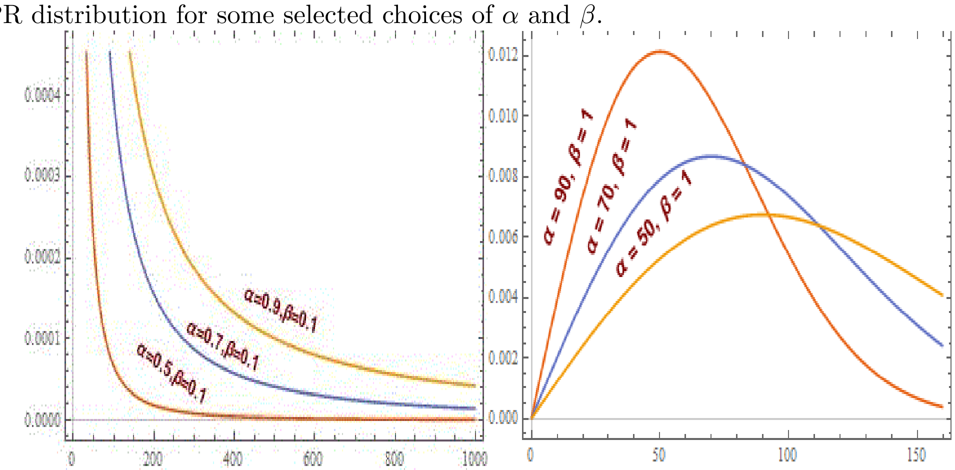


*Fig*1:Density function of the Power Rayleigh distribution.

2.2 Survival and Hazard Rate Functions

The event of interest has not yet occured 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 hazard rate function is the conditional rate of failure at time x, given that an individual has survived until at least time x. The hazard rate function of PR distribution is given by

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

The behavior of $h(x)$ of the PR distribution is discussed in the following theorem

Theorem 2 *For all $\alpha > 0$ the hazard 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 given 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.
Figure2 shows the shapes of hazard rate function of PR distribution.

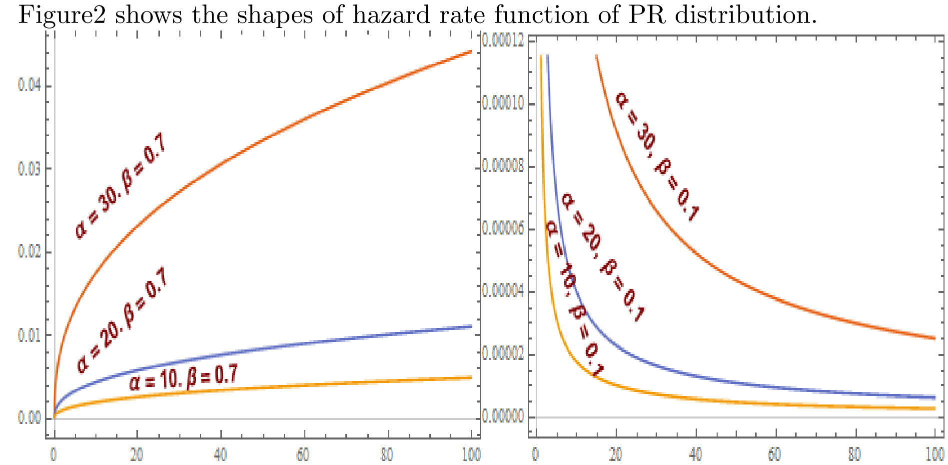


*Fig*2: Plots of hazard rate function of the Power Rayleigh distribution

■

2.3 Mean Residual Life Function

Let X be a continuous random variable with survival function $S(x)$, the mean residual life function is defined as the expected value of the remaining lifetimes after a fixed time point x. The mean residual life of the PR distribution is given as follows

$$\mu(x) = E(X - x \mid X > x) = \frac{1}{S(x)} \int_x^\infty y f(y) dy - x$$

$$= e^{\frac{x^{2\beta}}{2\alpha^2}} \int_x^\infty y \times \frac{\beta}{\alpha^2} y^{2\beta-1} e^{\frac{-y^{2\beta}}{2\alpha^2}} dy - x$$

$$= 2\alpha^2 x^{1-2\beta} \qquad x > 0.$$

Calabria and *Pulcini* 1987 [13] showed that the behavior of mean residual life function.The following lemma is useful to determine the shape of mean residual life function $\mu(x)$.

Lemma 3 *(Bryson and siddique* [14]*) Let X be a non-negative continuous random variable with hazard rate function $h(x)$ and mean residual life function $\mu(x)$.If $h(x)$ is increasing (decreasing), then $\mu(x)$ is increasing(decreasing).*

Theorem 4 *The mean residual life function $\mu(x)$ of the PR distribution is increasing if $\beta \le \frac{1}{2}$and decreasing if $\beta > \frac{1}{2}$ for all $\alpha > 0$.*

Figure3 shows that the shapes of mean residual life of the PR distribution.

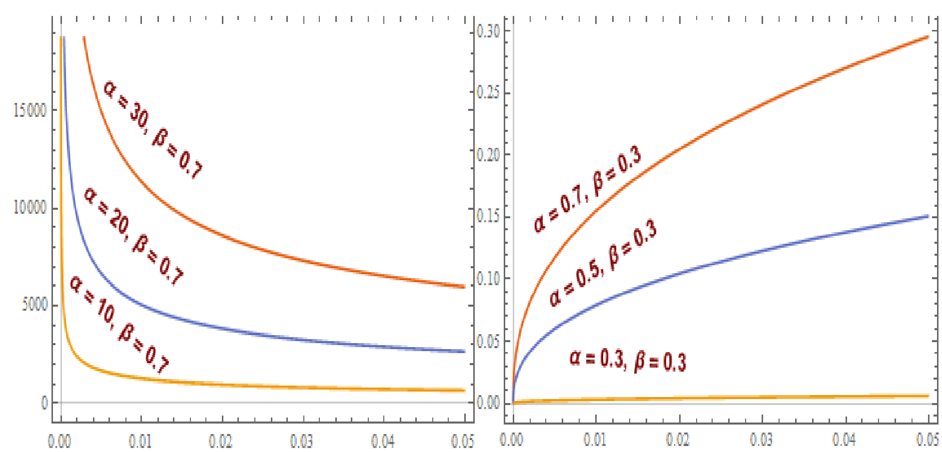


*Fig*3: Plots of mean residual life of the Power Rayleigh distribution

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 raw moment and the second raw moment provides 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 about the mean are given by

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

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

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

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

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

The coefficient of variation represents the ratio of the standard deviation to the mean, and it is a useful statistic for comparing the degree of variation from one data series to another, even if the means are drastically different from one another. Thus, the coefficient of variation of the PR distribution is given by

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

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

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

Kurtosis is the measure of the a mount probability in two tails. The kurtosis of the PR distribution is given by

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

Let X be a random variable from Power Rayleigh distribution then the moment generating function of power Rayleigh distribution is defined by

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

Similarly the characteristic function of PR distribution can be 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

The quantile function of the power rayleigh distribution is given 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 may be more useful descriptive statistics than means and other moments-related statistics. Some quantiles have special names (see *Ghitany* 2013[15]) :

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

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

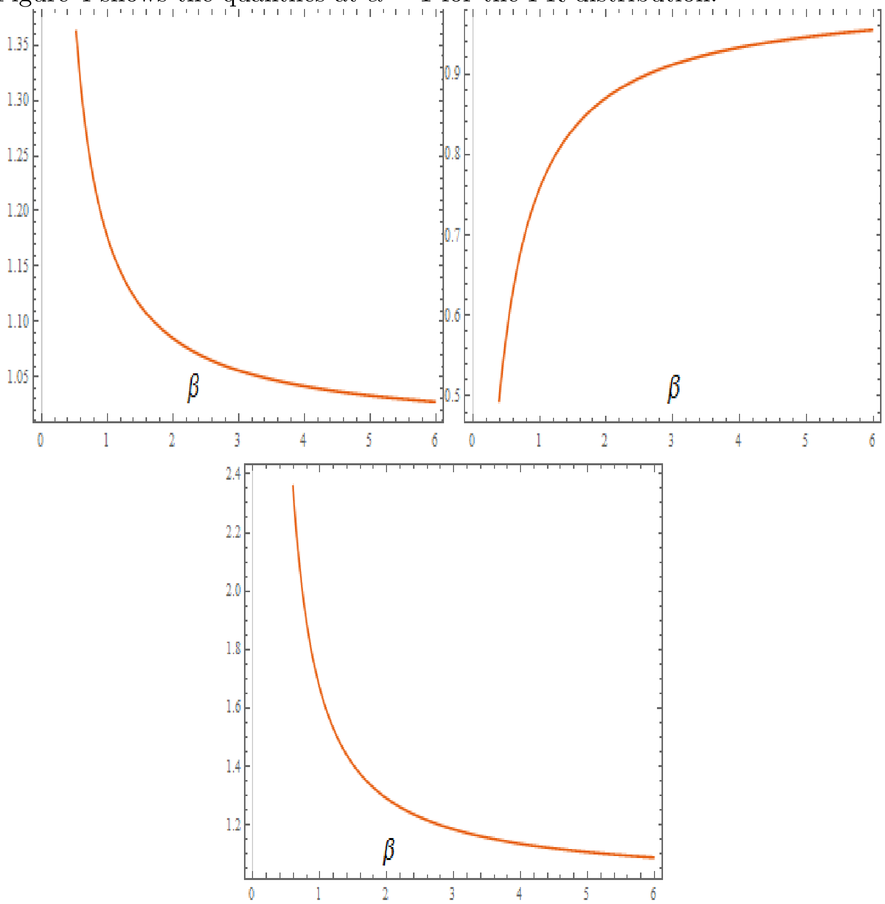
- if $u = \frac{1}{4}$ then the quantiles are called the first quartile, so the first quartile of the PR distribution is given by

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

- 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.



$Fig4$: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 it can be used in many fields such as chemistry, physics and biology as a driving force for protein unfloding or catalysis enzymes.The shannon entropy of random variable X is defined by

$S_H = -\int_0^\infty f(x)\log f(x)dx$

Shannnon entropy for the PR distribution is given by

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

3 Methods of Estimations

In ihis section, we consider the maximum likelihood method to estimate the involved parameters of PR distribution. Moreover, the interval estimation is discussed based on Fisher information matrix.

3.1 Maximum likelihood estimation

Let $x_1, x_2, ..., x_n$ be arandom sample of size n from PR distribution, then the sample likelihood function of this model can be given by

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

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

with respective sample log-likelihood function

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

The maximum likelihood estimators $\hat{\beta}$ and $\hat{\alpha}$ are obtained by solving 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

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

By substituding 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

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

3.2 Interval Estimation

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

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

$I_{22} = -E(\frac{\partial^2 \ln f(x)}{\partial \beta^2}) = -4 - \frac{1}{\beta^2}$

$I_{12} = -E(\frac{\partial^2 \ln f(x)}{\partial \alpha \partial \beta}) = \frac{4 \ln x}{\alpha}$

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 $\xrightarrow{d}$ denotes convergence in distribution and I^{-1} is the inverse of the matrix I. The asymototic variances and covariances of α and β are given 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 given 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 distribution of order statistic is used to known sometimes about how the order of the data behaved. For a sample of independent observation $x_1, x_2, ..., x_n$ from the Power Rayleigh distribution , the ordered sample values $x_{(1:n)} \leq x_{(2:n)} \leq .... \leq x_{(n:n)}$ are called the order statistic. Let $Y = X_{j:n}$ then the probability density function is given 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 cumulative distribution function of the order statistic is given by

$$F_y(y) = \sum_{m=j}^{n} \binom{n}{m} \times F^m(y) \times \{1 - F(y)\}^{n-m}$$

$$= \frac{(e^{-\frac{y^{2\beta}}{2\alpha^2}})^{n-j} \times (1-e^{-\frac{y^{2\beta}}{2\alpha^2}})^{j} \times n \times 2F1[1, j-n, 1+j, 1-e^{-\frac{y^{2\beta}}{2\alpha^2}}]}{j!(n-j)!}, \qquad y > 0$$

where $2F1$ is the hypergeometric function (see [*16*]).

4.2 Limiting Distribution of Extreme Values

Let $M_n = X_{(n:n)} = \max[X_1, X_2 ...., X_n], m_n = X_{(1:n)} = \min[X_1, X_2, ..., X_n]$ from the Power Rayleigh distribution. The limiting distributions of M_n and m_n can be introduced in the following theorem.

Theorem 5 *Let M_n and m_n be the maximum and the minimum of a random sample from the Power Rayleigh distribution, then*

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

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

where

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

Proof. For the PR distribution we have

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

$$= \frac{\alpha^2}{\beta}\left(\frac{1-2\beta}{\infty}\right) = 0$$

Therefore, by Theorm 8.3.3 of *Arnold* et al. [17], the maximal domain of attraction of the PR distribution is the standard Gumbel distribution providing part (i).

(ii) Using L'Hospital rule , we have

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

Therefore by Theorem 8.3.6 of *Arnold* et al. [17], the minimal domain of attraction of the PR distribution is the standard Weibull distribution providing part (ii). ■

5 Simulation Study

The equation $F(x) - u = 0$, where u is an observation of the uniform distribution $(0, 1)$ and $F(X)$ is the cumulative distribution function of the PR distribution, is used to implement the simulation study by creating random samples follow PR distribution. The simulation expriment was repeated 1000 times each with sample sizes: 30, 50, 70, 90 for $(\alpha, \beta) = (0.5, 5)$ and $(0.7, 10)$. The following measurs are calculated:

(i) Average bias of $\overset{\wedge}{\alpha}$ and $\overset{\wedge}{\beta}$ of the parameters α and β are respectively:

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

(ii) The Mean square error (MSE) of $\overset{\wedge}{\alpha}$ and $\overset{\wedge}{\beta}$ of the parameters α and β are respectively:

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

Table1 shows the average bias and the MSE of the estimates. The values of the bias are small, possitive and the values of the MSEs decreases while the sample size increases.

$Table1:$ Bias and MSE for 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

This section is devoted to illustrate the proposed distribution, PR distribution, by fitting it to real set data. A 34 storm events was observed from a watersheds . The data set as reported in Kang et al. (2013) [18] describes the water runoff (mm) of a study storms, which represents to one of the hydrological characteristics observed from a mall watershed in Korea (west of city of Suwon), the observations 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.

A number of probability distribution models viz., Gumbel, Pareto, Generalized Extreme Value (GEV) and Generalized Logistic (GL) distribution are in use in the hydrological data analysis. Hydrological studies are useful in designing, planning, and managing water resources. The selection of the most suitable probability distribution and associated parameter estimation procedure are the fundamental step in hydrology analysis.

The PR distribution was fitted to runoff data using MLE, Kolmogorov–Smirnov test statistics and compared via goodness of fit criteria among the following well-known distributions which fitted to these data,

- Gumbel distribution with density function

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

- Pareto distribution with density function

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

- Generalized Pareto (GP) distribution with density function

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

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) with density function

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

- Generalized Logistic distribution (GL) with density function

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

Table 2: The Measures AIC, BIC, AICC, HQIC and CAIC 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 [19])to compare PR distribution with other models . The model with minimum AIC, BIC, AICC, CAIC and HQIC value is chosen as the best model to fit the data. From Table 2, we conclude that the power Rayleigh distribution is the best comparable to the other models.

The maximum likelihood method is used for estimating the parameters of all the compared distributions and the parameter estimates are given in Tables 3. Further, Kolmogrov -Simrnov (K-S) goodness of fit test statistics used to test the fitting model of data set. The K-S statistics are determined for each distribution and listed in Table 4. It can be observed that the PR distribution has the smallest statistics and the largest p -Value. Accordingly, we can conclude that the PR distribution represents the best fit among the compared distributions for the runoff data.

7 Conclusion

A new distribution with two parameters called the Power Rayleigh distribution is proposed by power transformation .This model is more flixeble than the Rayleigh distribution in the area of reliability studies. In terms of the probability density, hazard rate and mean residual life functions are studied. An application to real data to show that the superiority of this new distribution by comparing it to other distributions.

References

[1] 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.

[2] S. K. Sinha, H. A. Howlader, Credible and HPD Intervals of The Parameter and Reliability of Rayleigh Distribution. IEEE Transactions on Reliability., 32, 217-220, 1993.

[3] S. Lalitha, A. Mishra, Modified Maximum Likelihood Estimation for Rayleigh Distribution, Communications in Statistics-Theory and Methods, 25, 389-401, 1996

[4] A. M. Abd Elfattah, A. S. Hassan, D. M. Ziedan, Efficiency of Maximum Likelihood Estimators under Different Censored Sampling Schemes for Rayleigh distribution, Interstat, Elctronic Journal., 1, 1-16, 2006.

[5] D. Kundu, M. Z. Raqab, Generalized Rayleigh Distribution: Different Methods of Estimation, Computational Statistics and Data Analysis., 1, 187-200, 2005.

[6] B. Rivet, L. Girin, 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.

[7] G. M. Cordeiro, C. T. Cristino, E. M. Hashimoto, E. M. M. Ortega, The Beta Generalized Rayleigh Distribution with Applications to Lifetime Data. Statistical Papers, 54, 133-161, 2011.

[8] F. Merovci, I. Elbatal, Weibull Rayleigh Distribution:Theory and Applications. Applied Mathematics & Information Sciences, 9, 2127-2137, 2015.

[9] M. A. W. Mahmoud, 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.

[10] V. G. Voda, On The Inverse Rayleigh Distributed Random Variable. Reports of Statistical Application Research., 19, 13-21, 1972.

[11] K. Fatima, S. P. Ahmed, Weighted Inverse Rayleigh Distribution. International Journal of Statistic and Systems, 12, 119-137, 2017.

[12] F. Merovci, Transmuted Rayleigh Distribution. Austrian Journal of Statistics, 42, 21-31, 2013.

[13] R. Calabria, G. Pulcini, On the Asymptotic Behaviour of Mean Residual Life Function. Reliability Engineering, 19, 165-170, 1987.

[14] M. C. Bryson, M. M. Siddique, Some Criteria for Aging. Journal of the American Statistical Association., 64, 1472–1483, 1969.

[15] M. E. Ghitany, D. K. AL-Mutairi, N. Balakrishnan, L. J. AL-ENezi, Power Lindely Distribution and Associated Inference. Computational Statistics and Data Analysis, 64, 20-33, 2013.

[16] N. N. Lebedev, Special Functions and their Application. Prentice-Hall, 1965.

[17] B. C. Arnold, N. Balakrishnan, H. N. Nagaraja, A Frist Course in Order Statistics. Wiley, New York, 1992.

[18] M. S. Kang, J. H. Goo, I. Song , J. A. Chun , Y .G. Her, S. W. Hwang, S. W. Park, Estimating design floods based on the critical storm duration for small watersheds, Journal of Hydro-environment Research, 7, 209–218, 2013.

[19] G. Chen, N. Balakrishnan, A general Purpose Approximate Goodness-of-Fit Test. Journal of Quality Technology., 27, 154–161, 1995.