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

N. M. Kilany¹*, Mohamed A. W. Mahmoud ² and Lobna H. El-Refai¹

¹Mathematics and Computer Science Department, Faculty of Science, Menoufia University, Menoufia, Egypt
²Mathematics 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:

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


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 \( x = y^\beta \), 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
\[
f(x) = \frac{\beta \alpha^2 x^{2\beta - 1} e^{-\frac{x^2}{2\alpha^2}}}{2\alpha^2}, \quad x > 0, \alpha, \beta > 0 \tag{2}
\]
where \( \alpha \) is a scale parameter and \( \beta \) is a shape parameter. The proper cumulative CDF is presented by
\[
F(x) = 1 - e^{-\frac{x^2}{2\alpha^2}} \tag{3}
\]
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
\[
f(0) = \begin{cases} 
\infty, & \beta \leq \frac{1}{2} \\
\frac{1}{\alpha^2}, & \beta = \frac{1}{2}, \ f(\infty) = 0. \\
0, & \beta > \frac{1}{2}.
\end{cases}
\]
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
\[
f(x) = \frac{\Psi(x)}{\alpha} f(x),
\]
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 \( f(x) < 0 \). Hence, \( f(x) \) is decreasing. Also, if \( \beta < \frac{1}{2} \), then \( \Psi(x) < 0 \) and \( f(x) < 0 \) for all \( \alpha > 0 \). Hence \( f(x) \) is decreasing.

(ii) \( \forall \alpha > 0, f(x) = 0 \) if \( \int \Psi(x) = 0 \) which arises at the point
\[
x_0 = \left( \frac{2\alpha^2\beta - \alpha^2}{\beta} \right)^{\frac{1}{2\beta}}.
\]

Since, \( \hat{f}(x_0) = -\frac{2\beta^\beta}{\alpha^2} \left( \frac{2\beta - 1}{\beta} \right)^{\frac{1}{2\beta}} < 0 \), so \( f(x) \) has a mode at \( x_0 \).

Figure 1 shows the attitude of PDF of PR model for some selected choices of \( \alpha \) and \( \beta \).
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

\[
S(x) = 1 - F(x) = e^{-\frac{x^2}{2\alpha^2}}
\]

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

\[
h(x) = \frac{f(x)}{S(x)} = \frac{x^2}{\alpha^2} \beta^{2\beta - 1} \frac{\beta - 1}{\alpha^2} \beta
\]

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

\[
h'(x) = \frac{\beta}{\alpha^2} x^{2\beta - 1} (2\beta - 1)
\]

So,

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

(ii) if \( \beta > \frac{1}{2} \), then \( h'(x) > 0 \) and \( h(x) \) is increasing.

The different behavior of failure rate function for PR model are shown in Figure 2.

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

\[
\mu(x) = \frac{1}{S(x)} \int_x^\infty yf(y)dy - x
\]

\[
= e^{x^2/2} \int_x^\infty y \frac{\beta}{\alpha^2} y^{2\beta - 1} e^{-\frac{x^2}{2\alpha^2}} dy - x
\]

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

---

**Fig. 1:** Density function of the PR distribution.
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 (decreasing) 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.

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 fourth 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

$$\mu'_k = 2^{\frac{k}{\alpha}} \left( \frac{1}{\alpha} \right) \frac{\alpha}{\beta} \Gamma \left( 1 + \frac{k}{\beta} \right)$$

$$\mu''_k = 2^{\frac{k}{\alpha}} \left( \frac{1}{\alpha} \right) \frac{\alpha}{\beta} \Gamma \left( 1 + \frac{k}{\beta} \right)$$

$$\mu''''_k = 2^{\frac{k}{\alpha}} \left( \frac{1}{\alpha} \right) \frac{\alpha}{\beta} \Gamma \left( 1 + \frac{k}{\beta} \right)$$
The central moments of PR distribution are as follows:

\[
\begin{align*}
\mu_2 & = \frac{2^{3\beta} \left(\frac{1}{\alpha^2}\right)^{\beta} \left[-\beta \Gamma (1 + \frac{1}{2\beta})^2 + \Gamma (\frac{1}{\beta})\right]}{\beta} \\
\mu_3 & = \frac{2^{-2\beta} \left(\frac{1}{\alpha^2}\right)^{\beta} \left[\gamma (1 + \frac{1}{2\beta})^3 - 6 \beta \Gamma (\frac{1}{2\beta}) \Gamma (\frac{1}{\beta}) + 6 \beta^2 \Gamma (\frac{3}{2\beta})\right]}{\beta^3} \\
\mu_4 & = \frac{4^{-2\beta} \left(\frac{1}{\alpha^2}\right)^{\beta} \left[-3 \gamma (1 + \frac{1}{2\beta})^4 + 24 \beta \Gamma (\frac{1}{2\beta}) - 48 \beta^2 \Gamma (\frac{3}{2\beta}) + 32 \beta^3 \Gamma (\frac{5}{2\beta})\right]}{\beta^4}
\end{align*}
\]

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

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

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

\[
\rho = \frac{SD}{\mu} = \frac{2^{\beta} \left(\frac{1}{\alpha^2}\right)^{\beta} \left[-\beta \Gamma (1 + \frac{1}{2\beta})^2 + \Gamma (\frac{1}{\beta})\right]}{\Gamma (1 + \frac{1}{2\beta})}, \quad \beta > 0
\]

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:

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

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

\[
K = \frac{\mu_4}{\mu_2^2} = \frac{-3 \Gamma (\frac{1}{2\beta})^4 + 24 \beta \Gamma (\frac{1}{2\beta})^2 \Gamma (\frac{1}{\beta})^2 - 48 \beta^2 \Gamma (\frac{3}{2\beta}) \Gamma (\frac{1}{2\beta}) + 32 \beta^3 \Gamma (\frac{5}{2\beta})}{16 \beta^2 \left[-\beta \Gamma (1 + \frac{1}{2\beta})^2 + \Gamma (\frac{1}{\beta})^2\right]^2}, \quad \beta > 0
\]

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

\[
M_X(t) = \int_0^\infty e^{tx} f(x) dx = 1 + \sum_{r=1}^{\infty} \frac{t^r}{r!} \frac{2^\beta \left(\frac{1}{\alpha^2}\right)^{\beta} \Gamma (1 + \frac{r}{2\beta})}{\beta}
\]

The characteristic function of PR distribution is derived as follows:

\[
\phi_X(t) = \int_0^\infty e^{itx} f(x) dx = 1 + \sum_{r=1}^{\infty} \frac{(it)^r}{r!} \frac{2^\beta \left(\frac{1}{\alpha^2}\right)^{\beta} \Gamma (1 + \frac{r}{2\beta})}{\beta}
\]
2.5 Quantile Function

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

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

Quantile is a 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 \( u = \frac{1}{4} \) then the quantile function is called median, so the median of the PR distribution is given by
  \[ Q_1 = F^{-1}\left(\frac{1}{2}\right) = (-2\alpha^2 \ln\left(\frac{1}{2}\right))^{\frac{1}{\beta}} \]
- if \( u = \frac{1}{4} \) then the quantiles are called the first quartile, so the first quartile of the PR distribution is given by
  \[ Q_2 = F^{-1}\left(\frac{1}{4}\right) = (-2\alpha^2 \ln\left(\frac{1}{4}\right))^{\frac{1}{\beta}} \]
- if \( u = \frac{3}{4} \) then the quantiles are called the third quartile, so the third quartile of the PR distribution is given by
  \[ Q_3 = F^{-1}\left(\frac{3}{4}\right) = (-2\alpha^2 \ln\left(\frac{3}{4}\right))^{\frac{1}{\beta}} \]

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

![Quantile Function Plot](image)

**Fig. 4:** Plots of quantile function of Power Rayleigh distribution for \( \alpha = 1 \).

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

\[
S_H = -\int_0^\infty f(x) \log f(x) dx \\
= -\int_0^\infty \frac{\beta}{\alpha x^2} e^{\frac{-\beta x^2}{2\alpha}} \log\left(\frac{\beta}{\alpha x^2} e^{\frac{-\beta x^2}{2\alpha}}\right) dx \\
= -\frac{1}{2} \beta \log\left(\frac{\alpha}{\beta}\right)^{\frac{1}{2}} \Gamma\left(2 - \frac{1}{2\beta}\right), \quad \beta > \frac{1}{2}.
\]
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 $x_1, x_2, \ldots, x_n$, then the sample likelihood function of this model is given by

$$L(\alpha, \beta, x) = \prod_{i=1}^{\infty} f(x_i; \alpha, \beta) = \left(\frac{\beta}{\alpha}\right)^n \prod_{i=1}^{\infty} x_i^{\beta-1} e^{-\frac{\beta}{2\alpha^2}}$$

Thus, the log-likelihood function is as follows,

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

(6)

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

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

(7)

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

(8)

From Eqs (7) and (8) we have

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

(9)

By substituting $\alpha^2$ in Eq (7), then we have

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

(10)

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

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

(11)

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

$$\beta_{k+1} = \beta_k - \frac{f(\beta_k)}{f'(\beta_k)}$$

where $k = 0, 1, 2, \ldots$
3.2 Interval Estimation

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

\[
\begin{align*}
I_{11} &= -E\left(\frac{\partial^2 \ln f(x)}{\partial \alpha^2}\right) = -\frac{2}{\alpha^2} - \frac{2\alpha}{\beta^2} F(1 + \frac{1}{\beta^2}) \\
I_{22} &= -E\left(\frac{\partial^2 \ln f(x)}{\partial \beta^2}\right) = 4 - \frac{2}{\beta^2} \\
I_{12} &= -E\left(\frac{\partial^2 \ln f(x)}{\partial \alpha \partial \beta}\right) = \frac{4\ln \alpha}{\beta}
\end{align*}
\]

Applying the large-sample theory of maximum likelihood estimators gives

\[
\sqrt{n}(\hat{\alpha} - \alpha, \hat{\beta} - \beta) \overset{d}{\longrightarrow} N_2(0, I^{-1}),
\]

where \(\overset{d}{\longrightarrow}\) indicates convergence in model and \(I^{-1}\) is the inverse of the matrix \(I\). The asymptotic variances and covariances of \(\hat{\alpha}\) and \(\hat{\beta}\) are provided by

\[
\begin{align*}
\text{var}(\hat{\alpha}) &= \frac{I_{11}}{n} \\
\text{var}(\hat{\beta}) &= \frac{I_{22}}{n} \\
\text{cov}(\hat{\alpha}, \hat{\beta}) &= -\frac{I_{12}}{n}
\end{align*}
\]

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

\[
\begin{align*}
\hat{\alpha} &\pm Z_{\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\alpha})}, \\
\hat{\beta} &\pm Z_{\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta})},
\end{align*}
\]

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

\[
f_Y(y) = \frac{n!}{(j-1)!(n-j)!} \times F^{j-1}(y) \times \{1 - F(y)\}^{n-j} \times f(y)
\]

\[
= \frac{(e^{\frac{2\beta}{\gamma^2}})^{j-1} \Gamma(j, 1 - e^{\frac{2\beta}{\gamma^2}})^{j-1} \beta^n}{\alpha^{j-1} \beta^{j-1} \Gamma(n-j, x)}, \quad y > 0
\]

The CDF of the order statistic is obviouse by

\[
F_Y(y) = \sum_{m=j}^{n} \frac{n!}{m!} \times F^m(y) \times \{1 - F(y)\}^{n-m}
\]

\[
= \frac{(e^{\frac{2\beta}{\gamma^2}})^{j-1} \Gamma(1 - e^{\frac{2\beta}{\gamma^2}}, x) \times \{1 - F(y)\}^{n-j}}{\beta^{j-1} \Gamma(n-j, x)}, \quad y > 0
\]

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 $[X_1, X_2, \ldots, X_n]$ from the Power Rayleigh distribution. Then we have

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

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

where,

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

**Proof.** For the PR distribution we obtain

(i) $\lim_{x \to -\infty} \frac{P(x)}{n(x)} = \left(\frac{a^2}{2}\right) \lim_{x \to -\infty} \frac{F(x)}{f(x)} = \left(\frac{a^2}{2}\right) \lim_{x \to -\infty} (1 - 2\beta)x^{-2\beta} = \left(\frac{a^2}{2}\right) (1 - 2\beta) = 0$

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,

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

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 experiment was Reiterated is $N = 1000$ with sample sizes: $30, 50, 70, 90$ for $(\alpha, \beta) = (0.5, 5)$ and $(0.7, 10)$. The next measures 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:

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

Table 1 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>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$n$</th>
<th>Bias($\alpha$)</th>
<th>MSE($\alpha$)</th>
<th>Bias($\beta$)</th>
<th>MSE($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>5</td>
<td>30</td>
<td>0.03288</td>
<td>0.02128</td>
<td>0.03052</td>
<td>0.00787</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.01486</td>
<td>0.01006</td>
<td>0.01459</td>
<td>0.00354</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70</td>
<td>0.00901</td>
<td>0.00720</td>
<td>0.01025</td>
<td>0.00242</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>0.00432</td>
<td>0.00532</td>
<td>0.00452</td>
<td>0.00169</td>
</tr>
<tr>
<td>0.7</td>
<td>10</td>
<td>30</td>
<td>0.00490</td>
<td>0.00551</td>
<td>0.61042</td>
<td>3.14893</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.00186</td>
<td>0.00297</td>
<td>0.29198</td>
<td>1.41796</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70</td>
<td>0.00028</td>
<td>0.00214</td>
<td>0.20501</td>
<td>0.96660</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>0.00011</td>
<td>0.00165</td>
<td>0.09048</td>
<td>0.67990</td>
</tr>
</tbody>
</table>

$c_2023$ NSP
Natural Sciences Publishing Corp.
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
  \[ f(x) = \frac{1}{\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0 \]

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

- Generalized Pareto (GP) distribution
  \[ f(x) = \begin{cases} \left(1 + k \left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{k}} - e^{-\left(1 + k \left(\frac{x-\mu}{\sigma}\right)\right)}, & \text{if } k \neq 0 \\ e^{-\left(\frac{x-\mu}{\sigma}\right)}, & \text{if } k = 0 \end{cases} \]
  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)
  \[ f(x) = \begin{cases} \left(1 + k \left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{k}} - e^{-\left(1 + k \left(\frac{x-\mu}{\sigma}\right)\right)}, & \text{if } k \neq 0, \ 1 + k \left(\frac{x-\mu}{\sigma}\right) > 0, \ \sigma > 0 \\ e^{-\left(\frac{x-\mu}{\sigma}\right)}, & \text{if } k = 0, \ -\infty < x < \infty, \ \sigma > 0 \end{cases} \]

- Generalized Logistic distribution (GL)
  \[ f(x) = \begin{cases} \sigma \left(1 + e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^{-\frac{1}{k}} - e^{-\left(\frac{x-\mu}{\sigma}\right)}, & \text{if } k \neq 0, \ 1 + k \left(\frac{x-\mu}{\sigma}\right) > 0, \ \sigma > 0 \\ \sigma \left(1 + e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^{-\frac{1}{k}}, & \text{if } k = 0, \ -\infty < x < \infty, \ \sigma > 0 \end{cases} \]

<table>
<thead>
<tr>
<th>Distributions</th>
<th>-log L</th>
<th>AIC</th>
<th>BIC</th>
<th>AICC</th>
<th>HQIC</th>
<th>CAIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumbel (μ, σ)</td>
<td>169.564</td>
<td>343.127</td>
<td>346.18</td>
<td>343.514</td>
<td>344.168</td>
<td>343.514</td>
</tr>
<tr>
<td>Pareto (α, β)</td>
<td>145.389</td>
<td>294.778</td>
<td>297.831</td>
<td>295.165</td>
<td>295.819</td>
<td>295.165</td>
</tr>
<tr>
<td>GP (μ, σ, k)</td>
<td>148.974</td>
<td>303.948</td>
<td>308.527</td>
<td>304.748</td>
<td>305.51</td>
<td>304.748</td>
</tr>
<tr>
<td>GEV (μ, σ, k)</td>
<td>150.814</td>
<td>307.628</td>
<td>312.207</td>
<td>308.428</td>
<td>309.189</td>
<td>308.428</td>
</tr>
<tr>
<td>GL (μ, σ, k)</td>
<td>151.498</td>
<td>308.996</td>
<td>313.575</td>
<td>309.796</td>
<td>310.557</td>
<td>309.796</td>
</tr>
<tr>
<td>PR (α, β)</td>
<td>140.349</td>
<td>284.698</td>
<td>287.75</td>
<td>285.085</td>
<td>285.739</td>
<td>283.085</td>
</tr>
</tbody>
</table>

Table 2: The Model Selection Criteria for the runoff data.
Table 3: Estimates of the parameters for runoff data

<table>
<thead>
<tr>
<th>Distributions</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumbel ($\mu, \sigma$)</td>
<td>12.194, 25.186</td>
</tr>
<tr>
<td>Pareto ($\alpha, \beta$)</td>
<td>0.3862, 0.6</td>
</tr>
<tr>
<td>GP ($\mu, \sigma, k$)</td>
<td>$-3.1511, 16.411, 0.52449$</td>
</tr>
<tr>
<td>GEV ($\mu, \sigma, k$)</td>
<td>$6.4599, 12.399, 0.59636$</td>
</tr>
<tr>
<td>GL ($\mu, \sigma, k$)</td>
<td>$11.726, 11.298, 0.61583$</td>
</tr>
<tr>
<td>PR ($\alpha, \beta$)</td>
<td>$1.6997, 0.2957$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Distributions</th>
<th>K-S</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumbel ($\mu, \sigma$)</td>
<td>0.47722</td>
<td>0.0000003</td>
</tr>
<tr>
<td>Pareto ($\alpha, \beta$)</td>
<td>0.15072</td>
<td>0.38418</td>
</tr>
<tr>
<td>GP ($\mu, \sigma, k$)</td>
<td>0.19416</td>
<td>0.13453</td>
</tr>
<tr>
<td>GEV ($\mu, \sigma, k$)</td>
<td>0.17514</td>
<td>0.22048</td>
</tr>
<tr>
<td>GL ($\mu, \sigma, k$)</td>
<td>0.18030</td>
<td>0.19381</td>
</tr>
<tr>
<td>PR ($\alpha, \beta$)</td>
<td>0.13869</td>
<td>0.52756</td>
</tr>
</tbody>
</table>

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) to compare PR distribution with other models. The model with minimum AIC, BIC, AICC, CAIC and HQIC value is selected as the best model to fit the data. From Table 2, we determine that the power Rayleigh distribution is the best model 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-Smirnov (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-Smirnov statistics for the Power Rayleigh distribution than other distributions, thus, for the runoff data, the PR model is the best 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 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, 4, 6, 6, 7, 8, 10, 14, 19, 21, 24, 30, 36, 40, 41, 46, 52, 56, 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
  \[ f(x) = \frac{x^2}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x > 0, \sigma > 0 \]

- Pareto distribution
  \[ f(x) = \frac{a^\alpha}{(x+b)^{\alpha+1}}, \quad x > 0, \alpha, \beta > 0 \]

- Log-logistic distribution
  \[ f(x) = \frac{(\beta/\alpha)^{\beta-1}(x/\alpha)^{\beta-1}}{(1+(x/\alpha)^{\beta})^2}, \quad x > 0, \alpha, \beta > 0 \]

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

© 2023 NSP
Natural Sciences Publishing Co.
- Gamma distribution
  \[ f(x) = \frac{\beta^\alpha e^{\beta - x}}{\Gamma(\alpha)}, \quad x > 0, \alpha, \beta > 0 \]

- Exponential distribution
  \[ f(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \lambda > 0 \]

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

<table>
<thead>
<tr>
<th>Distributions</th>
<th>-log L</th>
<th>AIC</th>
<th>BIC</th>
<th>AICC</th>
<th>HQIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayleigh ((\sigma))</td>
<td>578.01</td>
<td>1158.02</td>
<td>1160.42</td>
<td>1158.13</td>
<td>1159.02</td>
</tr>
<tr>
<td>Pareto ((\alpha, \beta))</td>
<td>558.65</td>
<td>1121.32</td>
<td>1126.00</td>
<td>1121.48</td>
<td>1123.19</td>
</tr>
<tr>
<td>Log-logistic ((\alpha, \beta))</td>
<td>520.22</td>
<td>1044.45</td>
<td>1049.13</td>
<td>1044.61</td>
<td>1046.32</td>
</tr>
<tr>
<td>Lognormal ((\mu, \sigma))</td>
<td>518.37</td>
<td>1040.75</td>
<td>1045.44</td>
<td>1040.91</td>
<td>1042.63</td>
</tr>
<tr>
<td>Gamma ((\alpha, \beta))</td>
<td>517.50</td>
<td>1039.00</td>
<td>1043.09</td>
<td>1039.17</td>
<td>1040.88</td>
</tr>
<tr>
<td>Exponential ((\lambda))</td>
<td>510.17</td>
<td>1022.34</td>
<td>1024.69</td>
<td>1022.42</td>
<td>1023.28</td>
</tr>
<tr>
<td>Lomax ((\alpha, \beta))</td>
<td>510.27</td>
<td>1024.54</td>
<td>1029.25</td>
<td>1024.71</td>
<td>1026.42</td>
</tr>
<tr>
<td>PR ((\alpha, \beta))</td>
<td>506.97</td>
<td>1017.93</td>
<td>1022.02</td>
<td>1018.10</td>
<td>1019.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distributions</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayleigh ((\sigma))</td>
<td>221.37</td>
</tr>
<tr>
<td>Pareto ((\alpha, \beta))</td>
<td>0.21249</td>
</tr>
<tr>
<td>Log-logistic ((\alpha, \beta))</td>
<td>0.85112</td>
</tr>
<tr>
<td>Lognormal ((\mu, \sigma))</td>
<td>4.7062</td>
</tr>
<tr>
<td>Gamma ((\alpha, \beta))</td>
<td>1.234</td>
</tr>
<tr>
<td>Exponential ((\lambda))</td>
<td>0.0036</td>
</tr>
<tr>
<td>Lomax ((\alpha, \beta))</td>
<td>89.194</td>
</tr>
<tr>
<td>PR ((\alpha, \beta))</td>
<td>6.25624</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distributions</th>
<th>K-S</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayleigh ((\sigma))</td>
<td>0.29257</td>
<td>3.76838 \times 10^{-6}</td>
</tr>
<tr>
<td>Pareto ((\alpha, \beta))</td>
<td>0.29935</td>
<td>2.03129 \times 10^{-6}</td>
</tr>
<tr>
<td>Log-logistic ((\alpha, \beta))</td>
<td>0.166381</td>
<td>0.0281364</td>
</tr>
<tr>
<td>Lognormal ((\mu, \sigma))</td>
<td>0.162764</td>
<td>0.0338222</td>
</tr>
<tr>
<td>Gamma ((\alpha, \beta))</td>
<td>0.18272</td>
<td>0.0115675</td>
</tr>
<tr>
<td>Exponential ((\lambda))</td>
<td>0.140966</td>
<td>0.0937465</td>
</tr>
<tr>
<td>Lomax ((\alpha, \beta))</td>
<td>0.136609</td>
<td>0.11293</td>
</tr>
<tr>
<td>PR ((\alpha, \beta))</td>
<td>0.116209</td>
<td>0.249452</td>
</tr>
</tbody>
</table>

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- Smirnov statistics for the Power Rayleigh distribution than other distributions. Thus, the Kolmogrov- Smirnov (K-S) goodness of fit statistics indicates the proposed distribution its the bestade 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