The Pareto-Rayleigh Distribution: Theory with Some Real-life Applications

Md. Shohel Rana, Md. Matiur Rahman Molla, Md. Mahabubur Rahman

Abstract− **Pareto-Rayleigh (PR) distribution (a member of *Pareto-X* family proposed in [1]) is considered here to study. The estimated values of the model parameters are derived by the maximum likelihood estimation process. Different important properties of the introduced distribution are obtained. Four real-life applications are studied from various fields to evaluate the applicability of proposed PR model. An exclusive simulation has been conducted to observe the performance of the estimation technique. Finally, it is shown that the better fitness of newly developed model than others chosen models selected in this study.**

Keywords− **Transformed-Transformer (T-X) family, Pareto-X family, Pareto-Rayleigh distribution, Reliability analysis, Distribution of Order statistics, MLE (Maximum Likelihood Estimation).**

I. Introduction

In probability, distribution theory controlling several types of real-life facts is an important, and a very complex problem in our daily life. So, overcoming this crisis, many generalized family of distributions besides the specific distribution have already been expanded, and a constant attempt to expand many more as well. The size distribution of income named Pareto distribution is launched in [2]. Time to time several generalizations of the pareto distributions have been enrolled in literature like beta-Pareto developed by [3]; Gamma- Pareto, and its application done by [4]; Kumaraswamy-Pareto studied by [5]; Weibull-Pareto, and its applications illustrated in [6]; A new Weibull-Pareto introduced in[7]; Rayleigh-Pareto by [8]; cubic transmuted Pareto by [9]; and Composite Rayleigh-Pareto developed by [10] among others. *"Transformed-Transformer" or "T-X"* family of distribution developed by [11], is a generalization for above of all distributions defined as

$$F_{T-X}(x) = \int_a^{W(G(x))} r(t)\ \mathrm{d}t, \tag{1}$$

Now, for the random variable T follows Pareto distribution, using (1), Reference [1] proposed the cdf of *Pareto-X* family as

$$F_{P-X}(x) = 1 - \left[\frac{x_m}{x_m - \log\left(1 - G(x)\right)}\right]^{\alpha};\ x \in \mathbb{R}, \tag{2}$$

where, $x_m \in \mathbb{R}^+$ (scale) and $\alpha \in \mathbb{R}^+$ (shape) are parameters. $G(X)$ is a baseline distribution function of any X. So, by using any suitable standard baseline cdf $G(x)$ in (2), many new members of the family may possible to obtain. From this point of view, by considering the cumulative distribution function (cdf) $G(x)$ from the Rayleigh distribution, PR (a member of *Pareto-X* family) distribution is proposed and a detail illustration is given in section as below.

The paper is designed as: Pareto-Rayleigh (PR) is presented in *section II*, In *section III*, some properties of the proposed distribution are studied. *Section IV* presents the distributions of different order statistics. MLE of the model parameters, and the simulation study are described in *section V*. *Section VI* illustrates the applications of four real-life phenomena. Finally, summary and conclusions are drawn at the end.

II. Pareto-Rayleigh Distribution

Suppose, a random variable X follows Rayleigh distribution having cdf as

$$G(x) = 1 - e^{-\frac{x^2}{2\sigma^2}};\ x \in [0, \infty), \tag{3}$$

where, $\sigma \in \mathbb{R}^+$ is (scale) parameter. Using (3) in (2), the distribution function of the PR is as

$$F(x) = 1 - x_m^{\alpha}\left(\frac{x^2}{2\sigma^2} + x_m\right)^{-\alpha};\ x \in [0, \infty), \tag{4}$$

ISSN: 2736-5484
DOI: https://doi.org/10.24018/ejmath.2022.3.4.131
Published on July 28, 2022.
Md. Shohel Rana, Department of Statistics, Islamic University, Kushtia-7003, Bangladesh. (e-mail: shohel.rana@iubat.edu).
Md. Matiur Rahman Molla, Department of Statistics, Islamic University, Kushtia-7003, Bangladesh. (e-mail: matiuriustat@gmail.com).
Md. Mahabubur Rahman, Department of Statistics, Islamic University, Kushtia-7003, Bangladesh. (corresponding e-mail: mmriu.stat@gmail.com).

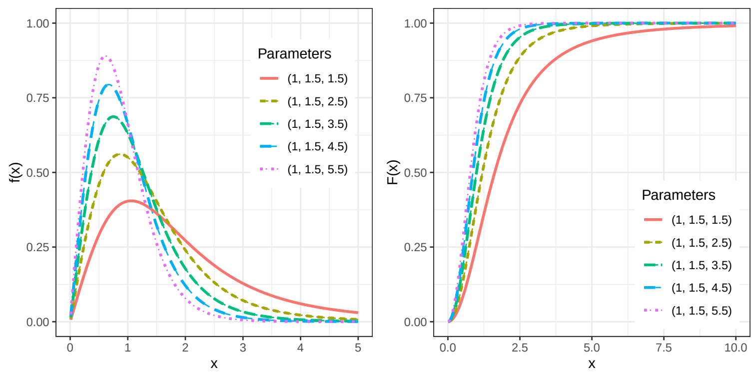

Fig. 1: The PDF $f(x)$ (Left) and CDF $F(x)$ (Right) for Several Values of the Parameters $x_m, \sigma, \alpha)$ of PR Model.

where, $x_m, \sigma \in \mathbb{R}^+$ are the (scale) and $\alpha \in \mathbb{R}^+$ is the(shape) parameters of the proposed model respectively. By differentiating (4) in terms of x, the pdf of PR distribution is find out and defined as

Definition: A random variable (continuous) X is named having a Pareto-Rayleigh distribution, if the pdf is given as

$$f(x) = \frac{\alpha x x_m^\alpha}{\sigma^2} \left(\frac{x^2}{2\sigma^2} + x_m \right)^{-\alpha-1} ; \; x \in [0, \infty), \tag{5}$$

where, $x_m \in \mathbb{R}^+, \sigma \in \mathbb{R}^+$ are(scale) and $\alpha \in \mathbb{R}^+$ is (shape) parameter of the proposed model.

Special Case:

(i) The cdf introduced by [12] is a special case of proposed PR model stated in (4) for $x = \sqrt{x}$, $\sigma = \frac{1}{\sqrt{2\lambda}}$ and $x_m = 1$.

(ii) The cdf introduced by [13] is taken into account as a special case of proposed PR model stated in (4) for $x_m = 1$.

Some different new shapes of the PR model are shown in *Fig. 1*. These shapes characteristics means that the proposed PR model can handle various characteristics in real-life phenomena.

III. Properties of the Proposed Model

The study about the properties is an important part of a probability model that allow us to know the real shape of a probability distribution. Some different properties of PR model are illustrated in below.

A. Moments of PR Model

Suppose X is a random variable having PR distribution. So the r^{th} raw moment is as

$$\mu_r' = \frac{2^{r/2} \sigma^r {x_m}^{r/2} \Gamma\left(\frac{r}{2}+1\right) \Gamma\left(\alpha - \frac{r}{2}\right)}{\Gamma(\alpha)}, \alpha > \frac{r}{2}. \tag{6}$$

Putting $r = 1$ in (6), we get

$$Mean = \frac{\sqrt{\frac{\pi}{2}} \sigma \sqrt{x_m} \Gamma\left(\alpha - \frac{1}{2}\right)}{\Gamma(\alpha)}, \alpha > \frac{1}{2}, \tag{7}$$

and

$$\begin{aligned} Variance &= E(X^2) - E(X)^2 \\ &= \frac{\sigma^2 x_m \left(4\Gamma(\alpha-1)\Gamma(\alpha) - \pi \Gamma\left(\alpha - \frac{1}{2}\right)^2\right)}{2\Gamma(\alpha)^2}, \alpha > 1, \end{aligned} \tag{8}$$

by setting $r > 2$ in (6), we can get all other greater moments along with measures of skewness and kurtosis of the PR model.

For various values of the parameters, we have been calculated means and variances (in bracket) in *Table I* and presented in *Fig. 2* by applying *(7)* and *(8)*. From here, we have been observed that for higher values of shape parameter α, the values of means and variances of PR model is being higher whereas for the higher values of scale parameter σ, the values of means and variances is also being higher.

For the random variable X, the normalized r^{th} central moment can be defined as

$$\frac{\mu_r}{\sigma_r} = \frac{E[x-\mu]^r}{\sigma^r}, \tag{9}$$

TABLE I: The Values of Means and Variances (in Bracket) for the PR Model

		$\sigma=0.5$	$\sigma=1$	$\sigma=1.5$	$\sigma=2$	$\sigma=2.5$
$x_m=1$	$\alpha=1.1$	0.981 (3.346)	0.886 (1.281)	0.813 (0.735)	0.755 (0.527)	0.707 (0.445)
	$\alpha=1.2$	1.962 (13.384)	1.772 (5.123)	1.626 (2.939)	1.510 (2.108)	1.414 (1.780)
	$\alpha=1.3$	2.943 (30.113)	2.658 (11.527)	2.439 (6.612)	2.264 (4.742)	2.121 (4.004)
	$\alpha=1.4$	3.924 (53.535)	3.544 (20.493)	3.252 (11.754)	3.019 (8.430)	2.828 (7.119)
	$\alpha=1.5$	4.905 (83.648)	4.430 (32.020)	4.065 (18.366)	3.774 (13.172)	3.536 (11.123)

from (9), we ca say that the value of 1^{st} and 2^{nd} standardized moments will be 0 and 1 where as 3^{rd} and 4^{th} standardized moments will be the measures of skewness, and kurtosis respectively. So for PR model measures of skewness, and kurtosis can be obtained, and defined as

$$\textit{Skewness}=\frac{\mu_3}{\mu_2^{\frac{3}{2}}}=\frac{\pi\sigma^3{x_m}^{3/2}\left(\sqrt{\pi}\Gamma\left(\alpha-\frac{1}{2}\right)^3-3\ 4^{2-\alpha}(\alpha-2)\Gamma(\alpha)\Gamma(2\alpha-3)\right)}{\sqrt{2}\Gamma(\alpha)^3}, \tag{10}$$

and

$$\textit{Kurtosis}=\frac{\mu_4}{\mu_2^2}=\frac{\sigma^4{x_m}^2\Big(-3\pi^2\Gamma\left(\alpha-\frac{1}{2}\right)^4+24\pi\Gamma(\alpha-1)\Gamma(\alpha)\Gamma\left(\alpha-\frac{1}{2}\right)^2\Big)}{4\Gamma(\alpha)^4}$$
$$-\frac{\sigma^4{x_m}^2\Big(24\pi\Gamma\left(\alpha-\frac{3}{2}\right)\Gamma(\alpha)^2\Gamma\left(\alpha-\frac{1}{2}\right)-32\Gamma(\alpha-2)\Gamma(\alpha)^3\Big)}{4\Gamma(\alpha)^4}, \tag{11}$$

respectively.

For different values of the parameters, *Table II* provides the calculated values of skewness, and kurtosis (in bracket) of PR model by using *(10)* and *(11)*, and plotted in *Fig. 3* that illustrate, PR model is positively skewed and leptokurtic.

B. Moment Generating Function (MGF) of PR Model

The following theorem represents the MGF of PR model.

Theorem: Suppose X be the random variable of PR model, then the mgf, $M_X(t)$ is

$$M_X(t)=\sum_{r=0}^{\infty}\frac{t^r}{r!}\frac{2^{r/2}\sigma^r{x_m}^{r/2}\Gamma\left(\frac{r}{2}+1\right)\Gamma\left(\alpha-\frac{r}{2}\right)}{\Gamma(\alpha)},\alpha>\frac{r}{2}, \tag{12}$$

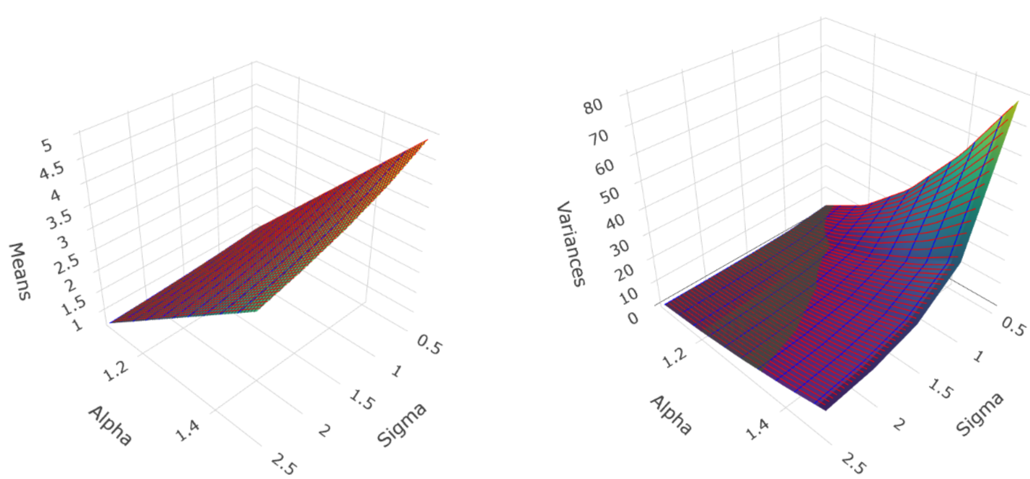

Fig. 2: The Plot of Mean (Left) and Variance (Right) for PR Model.

TABLE II: The Values of Skewness and Kurtosis (in Bracket) for the PR Model

		$\sigma = 0.5$	$\sigma = 1$	$\sigma = 1.5$	$\sigma = 2$	$\sigma = 2.5$
	$\alpha = 1.51$	11.458	5.154	3.086	2.075	1.485
	$(\alpha = 1.10)$	(25.553)	(22.717)	(20.357)	(18.362)	(16.651)
	$\alpha = 1.52$	91.668	41.232	24.686	16.597	11.881
	$(\alpha = 1.11)$	(408.852)	(363.480)	(325.717)	(293.787)	(266.421)
$x_m = 1$	$\alpha = 1.53$	309.379	139.158	83.314	56.015	40.098
	$(\alpha = 1.12)$	(2069.815)	(1840.117)	(1648.944)	(1487.296)	(1348.758)
	$\alpha = 1.54$	733.343	329.856	197.484	132.776	95.047
	$(\alpha = 1.13)$	(6541.636)	(5815.678)	(5211.479)	(4700.590)	(4262.743)
	$\alpha = 1.55$	1432.311	644.251	385.711	259.328	185.639
	$(\alpha = 1.14)$	(15970.792)	(14198.432)	(12723.336)	(11476.050)	(10407.087)

where, $t \in \mathbb{R}$.

Proof: The mgf can be written as

$$M_x(t) \quad = \quad E\left(e^{\text{tx}}\right) = \int_0^\infty e^{\text{tx}} f(x)\, dx,$$

where, $f(x)$ is stated in (5). From the expansion of e^{tx} illustrated in [14], we can get

$$M_x(t) \quad = \quad \int_0^\infty \sum_{r=0}^{\infty} \frac{t^r}{r!} x^r f(x) dx = \sum_{r=0}^{\infty} \frac{t^r}{r!} E(x^r). \tag{13}$$

From (6), setting $E(X^r)$ in (13), we can prove (12).

C. The Characteristic Function (CF) of PR Model

The following theorem defines the CF of PR distribution.

Theorem: Suppose X be the random variable of PR model, then the cf, $\phi_X(t)$ is

$$\phi_X(t) = \sum_{r=0}^{\infty} \frac{(\text{i}t)^r}{r!} \frac{2^{r/2} \sigma^r {x_m}^{r/2} \Gamma\left(\frac{r}{2}+1\right) \Gamma\left(\alpha - \frac{r}{2}\right)}{\Gamma(\alpha)}, \alpha > \frac{r}{2},$$

where, $t \in \mathbb{R}$ *and* $i = \sqrt{-1}$ *be the imaginary unit.*

Proof: It is very easy to prove.

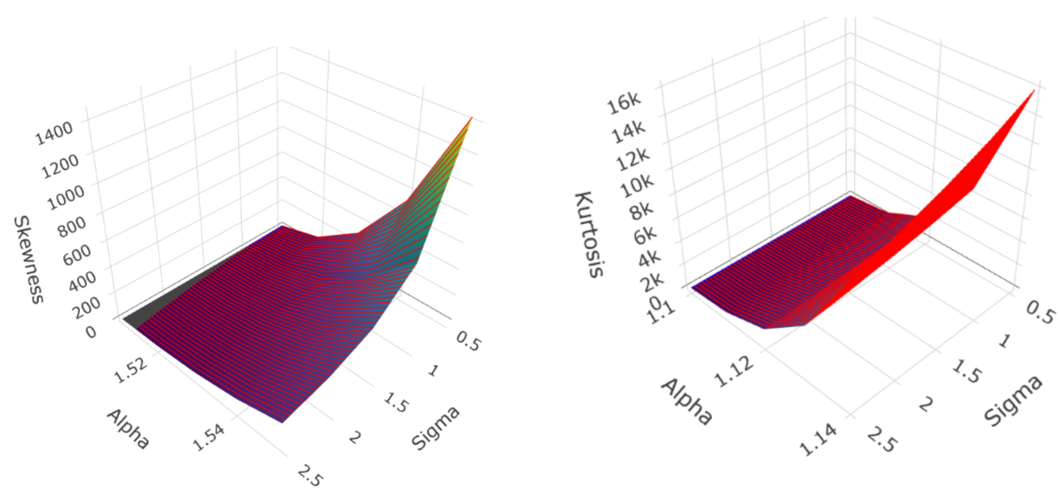

Fig. 3: The plot of Skewness (Left) and Kurtosis (Right) for PR Model.

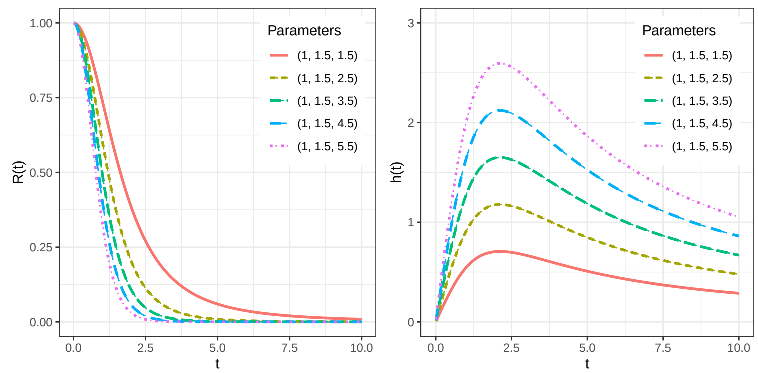

Fig. 4: The Reliability $R(t)$ (Left) and Hazard $h(t)$ (Right) Function for Several Values of (x_m, σ, α) for PR Model.

D. Analysis of Reliability Function of PR Model

Reliability is meant by any objects like patients or other devices will alive beyond a certain time frame. The reliability given by [15] whereas survival named by [16]. It is simply called the supplement of cdf. So it is given for PR model as

$$R(t) = 1 - F(t) = x_m^\alpha \left(\frac{t^2}{2\sigma^2} + x_m\right)^{-\alpha}; \; t \in [0, \infty),$$

where, $x_m, \lambda \in \mathbb{R}^+$ are (scale) and $\alpha, k \in \mathbb{R}^+$ are (shape) parameters of PR model. The ratio of density to reliability function is known as hazard function and for PR model it is defined by

$$h(t) = \frac{\alpha t}{\sigma^2 \left(\frac{t^2}{2\sigma^2} + x_m\right)}; \; t \in [0, \infty).$$

The different hazard rates along with increasing and decreasing are noticed from the *Fig. 4*.

E. Functional form of Quantile and Median of PR Model

This function is represented by x_q and explained by a function associated to a probability distribution of a random variable that is opposite of its distribution function. The quantile function for PR distribution is derived by solving (4) for x, and is gotten as, for example (see [9]),

$$x_q = \sqrt{2}\sqrt{\sigma^2 \left((1-q)x_m{}^{-\alpha}\right)^{-1/\alpha} - \sigma^2 x_m}, \tag{14}$$

and

$$\textit{Median} = x_{0.5} = \sqrt{2(0.5)^{-1/\alpha}\sigma^2 \left(x_m{}^{-\alpha}\right)^{-1/\alpha} - 2.\sigma^2 x_m}.$$

By setting $q = 0.25$ and $q = 0.75$ in (14), we can calculate lower quartile and upper quartile respectively.

F. Random Number Generation for PR Model

Random number is usually used for getting an uncertain outcome in many areas along with computer simulation. In order to generate random number for PR model, we have been used the expression, as below for example (see, [9]).

$$X = \sqrt{2}\sqrt{\sigma^2 \left((1-u)x_m{}^{-\alpha}\right)^{-1/\alpha} - \sigma^2 x_m}, \tag{15}$$

where, u follows uniform distribution. By using (15) we can easily generate random number from PR model for different familiar values of parameters.

IV. Distribution of Order Statistics of PR Model

This distribution of different order statistics are normally applied in wide areas besides study of floods and other extreme meteorological circumstances. Suppose $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ represents the order statistic for the random sample $X_1, X_2, \cdots, X_n$ having *cdf* $F_X(x)$ and *pdf* $f_X(x)$ then the *cdf* of $X_{r:n}$ is defined as

$$f_{X_{r:n}}(x) = \frac{n!}{(r-1)!(n-r)!}[F(x)]^{r-1}[1 - F(x)]^{n-r} f(x). \tag{16}$$

The cdf of rth order statistic in terms of PR model is defined by applying (16), as

$$\begin{aligned} f_{X_{r:n}}(x) &= \frac{n!}{(r-1)!(n-r)!}\left[1 - x_m^{\alpha}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha}\right]^{r-1} \\ &\times \left[x_m^{\alpha}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha}\right]^{n-r}\frac{\alpha x x_m^{\alpha}}{\sigma^2}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha-1}, \end{aligned} \tag{17}$$

where, $r=1,2,\cdots,n$. Setting $r=1$ in (17), the cdf of the first order statistic $X_{1:n}$ for PR model is

$$f_{X_{1:n}}(x) = n\left[x_m^{\alpha}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha}\right]^{n-1}\frac{\alpha x x_m^{\alpha}}{\sigma^2}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha-1},$$

and for setting $r=n$ in (17), the cdf of the last order statistic $X_{n:n}$ for PR model is

$$f_{X_{n:n}}(x) = n\left[1 - x_m^{\alpha}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha}\right]^{n-1}\frac{\alpha x x_m^{\alpha}}{\sigma^2}\left(\frac{x^2}{2\sigma^2}+x_m\right)^{-\alpha-1}.$$

For the PR model, the kth order moment of $X_{r:n}$ is derived as below.

$$E(X_{r:n}^k) = \int_0^{\infty} x_r^k \cdot f_{X_{r:n}}(x)\cdot \mathrm{d}x,$$

where, $f_{X_{r:n}}(x)$ is available in (17).

V. Parameter Estimation and Simulation study for PR Model

In statistics, MLE is defined by estimating the parameters of a probability model by the maximization of likelihood function. In the following two subsections, we have estimated parameters of the PR distribution by using this method, and hence, a simulation study is also done.

A. Estimating Parameter and Inference of PR Model

Consider a random sample, $x_1, x_2, \cdots, x_n$ of size n from PR model. So the likelihood and corresponding log-likelihood function is written by

$$\mathcal{L}(x) = \left[\frac{1}{\sigma^2}.x_m^{\alpha}.\alpha\right]^n\left(\prod_{i=1}^n x_i\right)\prod_{i=1}^n\left(\frac{x_i^2}{2\sigma^2}+x_m\right)^{-(\alpha+1)},$$

and

$$\begin{aligned} \ell(x) &= -2n\log(\sigma) + \alpha n\log(x_m) + n\log(\alpha) + \sum_{i=1}^n \log(x_i) \\ &\quad -(\alpha+1)\sum_{i=1}^n \log\left(\frac{x_i^2}{2\sigma^2}+x_m\right). \end{aligned} \tag{18}$$

respectively.

The MLE of x_m is $x_{(1)}$, known as first-order statistic . Then σ and α are determined through the maximization of (18). The first derivatives of (18) with respect to σ and α as

$$\frac{\delta\ell}{\delta\alpha} = n\left(\frac{1}{\alpha}+\log(x_m)\right) - \sum_{i=1}^n \log\left(\frac{x_i^2}{2\sigma^2}+x_m\right),$$

and

$$\frac{\delta\ell}{\delta\sigma} = (\alpha+1)\sum_{i=1}^n \frac{x_i^2}{\sigma^3\left(\frac{x_i^2}{2\sigma^2}+x_m\right)} - \frac{2n}{\sigma},$$

respectively.

So, applying $\frac{\partial l}{\partial\alpha}=0$, and $\frac{\partial l}{\partial\sigma}=0$, and finding the solution of system of nonlinear equations, we can obtain the MLE as $\hat{\Theta}=(\hat{\alpha},\hat{\sigma})'$ of $\Theta=(\alpha,\sigma)'$. Hence, as $n\to\infty$, the asymptotic model of the $MLEs$ $(\hat{\alpha},\hat{\sigma})$ are obtained as, for example (see, [17]),

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\sigma}\end{pmatrix} \sim N\left[\begin{pmatrix}\alpha\\ \sigma\end{pmatrix}, \begin{pmatrix}\hat{V}_{11} & \hat{V}_{12}\\ \hat{V}_{21} & \hat{V}_{22}\end{pmatrix}\right],$$

where, $\hat{V}_{ij} = V_{ij}|_{\Theta=\hat{\Theta}}$. V (asymptotic variance-covariance matrix), estimated values of $\hat{\alpha}$ and $\hat{\sigma}$ are got through inverting Hessian matrix; (see Appendix). $100(1-\alpha)\%$ both sided approximate confidence limits for α and σ are written by

$$\hat{\alpha} \pm Z_{\alpha/2}\sqrt{\hat{V}_{11}} \quad \text{and} \quad \hat{\sigma} \pm Z_{\alpha/2}\sqrt{\hat{V}_{22}}$$

respectively, where, Z_α indicates the αth percentile of the standard normal model.

B. The Study of Simulation for PR Model

Different sample of sizes like, 50, 100, 200, 500 and 1000 along with PR model have been applied to conduct a Monte Carlo simulation. For generating random sample, MLE are obtained using starting values of $\sigma = 1.5$ and $\alpha = 1.5$. 10000 times repeated this process and hence average estimated values besides Mean-Squared Errors ($MSE's$) are evaluated under these estimates. The obtained results are shown in *Table III*, and the shapes of $MSE's$ are ploted in *Fig. 5*. It is examined that the estimated values stay very near about around the real values (initial values) of the parameters that proves the estimating process is efficient enough. It is also examined that the estimated $MSE's$ are decreasing consistently in terms of increasing the sample size.

TABLE III: Average Estimated Vales and $MSE's$ for PR Models

Sample Size	*Estimate*		MSE	
	σ	α	σ	α
50	1.500	1.520	0.019	0.077
100	1.500	1.420	0.010	0.031
200	1.500	1.460	0.005	0.017
500	1.500	1.610	0.002	0.009
1000	1.500	1.470	0.001	0.003

VI. Applications of some Real-life Data for PR Model

Four different real-life datasets have been used for the applications of PR model as follows.

Sources of Datasets: Breast cancer dataset represents the survival times of 121 patients used by [18] and has recently been studied by [19]. Carbon Fibers dataset is available in [20]. Failure Times dataset are reported in the book "Weibull Models" by [21] and is also used by [22] and [23] . Wheaton river dataset is used by [24] and [3]. The descriptive statistics for four datasets are shown in Table IV.

TABLE IV: Descriptive Statistics for the Selected Datasets.

Dataset	*Min.*	Q_1	*Median*	*Mean*	Q_3	*Max.*	*Skewness*	*Kurtosis*	*Variance*
Breast Cancer	0.300	17.500	40.000	46.330	60.000	154.000	1.030	0.346	1244.464
Carbon Fibers	0.390	1.840	2.700	2.621	3.220	5.560	0.362	0.043	1.027
Failure Times	0.040	1.870	2.390	2.563	3.376	4.663	0.085	-0.689	1.239
Wheaton River	0.100	2.125	9.500	12.200	20.125	64.000	1.442	2.725	151.264

We have been considered Pareto-Exponential distribution proposed by [12], and Pareto distribution by [2] for testing the performance of the proposed PR model.

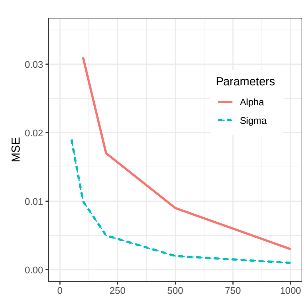

Fig. 5: $MSE's$ of Model Parameters vs Different Numbers of Sample Sizes for PR Model.

TABLE V: MLE of Parameters and Respective SE for the Selected Models.

Datasets	*Distribution*	*Parameter*	*Estimate*	*logLike*	*SE*
Breast Cancer	*Pareto-Rayleigh*	x_m	0.300	-511.969	–
		σ	1.500		0.093
		α	0.150		0.014
	Pareto-Exponential	x_m	0.300	-672.202	–
		λ	1.500		0.126
		α	0.233		0.023
	Pareto	x_m	0.300	-726.834	–
		α	0.214		0.019
Carbon Fibers	*Pareto-Rayleigh*	x_m	0.390	-46.125	–
		σ	1.500		0.093
		α	0.653		0.076
	Pareto-Exponential	x_m	0.390	-221.007	–
		λ	1.500		0.132
		α	0.607		0.066
	Pareto	x_m	0.390	-247.564	–
		α	0.549		0.055
Failure times	*Pareto-Rayleigh*	x_m	0.040	-82.248	–
		σ	1.500		0.107
		α	0.292		0.034
	Pareto-Exponential	x_m	0.040	-231.773	–
		λ	1.500		0.149
		α	0.272		0.030
	Pareto	x_m	0.040	-270.459	–
		α	0.249		0.027
Wheaton River	*Pareto-Rayleigh*	x_m	0.100	-168.068	–
		σ	1.500		0.118
		α	0.218		0.026
	Pareto-Exponential	x_m	0.100	-272.748	–
		λ	1.500		0.163
		α	0.266		0.032
	Pareto	x_m	0.100	-303.012	–
		α	0.244		0.028

TABLE VI: Selection Criteria Estimated for the Selected Models.

Datasets	*Distribution*	*-2logLike*	*AIC*	*AICc*	*BIC*	*KS*	*C-vM*
Breast Cancer	*Pareto-Rayleigh*	1023.938	1027.938	1028.039	1033.529	0.369	0.064
	Pareto-Exponential	1344.404	1348.404	1348.506	1353.996	0.416	0.064
	Pareto	1453.668	1455.669	1455.702	1458.464	0.428	0.068
Carbon Fibers	*Pareto-Rayleigh*	92.252	96.249	96.373	101.460	0.296	0.682
	Pareto-Exponential	442.014	446.015	446.138	451.225	0.420	0.752
	Pareto	495.128	497.128	497.169	499.733	0.395	0.851
Failure Times	*Pareto-Rayleigh*	164.496	168.497	168.643	173.382	0.678	0.021
	Pareto-Exponential	463.546	467.545	467.691	472.430	0.761	0.033
	Pareto	540.918	542.917	542.966	545.360	0.765	0.327
Wheaton River	*Pareto-Rayleigh*	336.136	340.136	340.310	344.689	0.198	0.058
	Pareto-Exponential	545.496	549.495	549.669	554.048	0.307	0.236
	Pareto	606.024	608.024	608.081	610.301	0.319	0.255

Table V shows the evaluated values of the model parameters, log-likelihood as well as the corresponding standard errors for the chosen models. Different model selection criteria like -2Log-Likelihood, Akaike's Information Criterion *(AIC)*, Corrected Akaike's Information Criterion *(AICc)*, Bayesian Information Criterion *(BIC)*, Kolmogorov-Smirnov Statistic *(KS)*, and Cramer-von Mises Staistic *(C-vM)* are presented in *Table VI*. The estimated *cdf* of the proposed PR distribution along with selected models for four different datasets like Breast Cancer, Carbon Fibers, Failure Times, and Wheaton River datasets are plotted over empirical cdf and as shown starting from the left to right sides of *Fig. 6* subsequently . The overall results obtained in *Table (VI), (V)*, and the estimated plots presented in *Fig. 6* demonstrate the better fitness of the conformation in favor of proposed PR distribution than other chosen models used in this study.

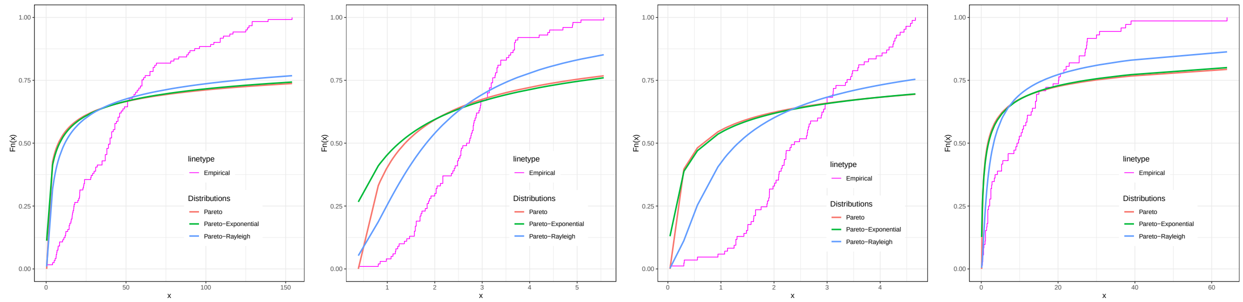

Fig. 6: Estimated CDF Plotted Over Empirical CDF for the Selected Datasets.

VII. Summary and Conclusion

In this article, Pareto-Rayleigh (PR) distribution, a special case of *Pareto-X* family of distributions has been proposed and studied in detail. Some of the important properties of the PR distribution are illustrated. The random number generating and reliability function besides the distribution of different order statistics are explained. The MLE technique has been applied to evaluate the model parameters of the PR distribution, and a large sample test of the model parameters is also done. At the end, four different practical life datasets have been studied for testing the practicability of PR model. Overall, it is proved that the proposed Pareto-Raleigh (PR) model acts better than other chosen models used under the study.

Appendix: Hessian matrix

The Hessian Matrix of PR Model is

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix},$$

where, the V (Variance-Covariance) Matrix is Calculated by

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}^{-1},$$

and the Elements of Hessian Matrix can be Calculated as

$$H_{11} = -\frac{\delta^2 \ell}{\delta \alpha^2} = \frac{n}{\alpha^2}, \quad H_{12} = -\frac{\delta^2 \ell}{\delta \alpha \ \delta \sigma} = -\sum_{i=1}^{n} \frac{x_i^2}{\sigma^3 \left(\frac{x_i^2}{2\sigma^2} + x_m \right)},$$

and

$$H_{22} = -\frac{\delta^2 \ell}{\delta \sigma^2} = (\alpha + 1) \sum_{i=1}^{n} \left(\frac{3x_i^2}{\sigma^4 \left(\frac{x_i^2}{2\sigma^2} + x_m \right)} - \frac{x_i^4}{\sigma^6 \left(\frac{x_i^2}{2\sigma^2} + x_m \right)^2} \right) - \frac{2n}{\sigma^2}.$$

Acknowledgment

Md. Shohel Rana (1'st author), is very grateful to the supervisor, Dr. Md. Mahabubur Rahman, professor (Associate) in Statistics for his continuous support and important suggestions to develop the quality of a standard research paper. I'd also like to thank my honorable supervisor sir to give me a scope to enroll in PhD program in Statistics at Islamic University, Kushtia-7003, Bangladesh.

Conflict of interest

Without any conflict of interest said by authors.

References

[1] Rana MS, Shahbaz MQ, Rahman MM. Pareto-Weibull Distribution with Properties and Applications: A Member of Pareto-X Family. *Pakistan Journal of Statistics and Operation Research*. 2022; 18(1): 121-132.
[2] Steiner P. XXV. Vilfredo Pareto et l'économie politique walrasienne. *Nouvelle histoire de la pensée économique. La Découverte*. 1993; 2: 498-521.
[3] Akinsete A, Famoye F, Lee C. The beta-Pareto distribution. *Statistics*. 2008; 42(6): 547-563.
[4] Alzaatreh A, Famoye F, Lee C. Gamma-Pareto distribution and its applications. *Journal of Modern Applied Statistical Methods*. 2012; 11(1): 7.
[5] Bourguignon M, Silva RB, Zea LM, Cordeiro GM. The Kumaraswamy Pareto distribution. *J. Stat. Theory Appl.* 2013; 12(2): 129-144.
[6] Alzaatreh A, Famoye F, Lee C. Weibull-Pareto distribution and its applications. *Communications in Statistics-Theory and Methods*. 2013; 42(9): 1673-1691.
[7] Tahir MH, Cordeiro GM, Alzaatreh A, Mansoor M, Zubair M. A new Weibull–Pareto distribution: properties and applications. *Communications in Statistics-Simulation and Computation*. 2016; 45(10): 3548-3567.
[8] Al-Kadim, KA, Mohammed, BDA. Rayleigh pareto distribution. *Journal of University of Babylon for Pure and Applied Sciences*. 2018; 26(1): 84-95.

[9] Rahman MM, Al-Zahrani B, Shahbaz MQ. Cubic transmuted Pareto distribution. *Annals of Data Science*. 2020; 7(1): 91-108.
[10] Benatmane C, Zeghdoudi H, Shanker R, Lazri N. Composite Rayleigh-Pareto distribution: Application to real fire insurance losses data set. *Journal of Statistics and Management Systems*. 2021; 24(3): 545-557.
[11] Alzaatreh A, Lee C, Famoye F. A new method for generating families of continuous distributions. *Metron*. 2013; 71(1): 63-79.
[12] Waseem N, Bashir S. On generating T-X family of Pareto-exponential distribution with properties and applications. *International Journal of Applied Mathematics & Statistics*. 2019; 58(3): 90-106.
[13] G Prasad1, R Subba Rao2*, RRL Kantam3. Limited Failure Censored Life Test Sampling Plan: Pareto-Rayleigh Model. *International Journal of Engineering and Technology*, IJET. 2018; 136–140.
[14] Zwillinger D, Jeffrey A. Table of integrals, series, and products. *Elsevier*. 2007.
[15] Ebeling CE. An introduction to reliability and maintainability engineering. *Tata McGraw-Hill Education*. 2004.
[16] Carpenter M. Survival analysis: a self-learning text. *Taylor & Francis*. 1997.
[17] Aryal GR, Tsokos CP. Transmuted Weibull distribution: A generalization of the Weibull probability distribution. *European Journal of pure and applied mathematics*. 2011; 4(2): 89-102.
[18] Al-Kadim KA Mahdi AA. Exponentiated transmuted exponential distribution. *Journal of University of Babylon for Pure and Applied Sciences*. 2018; 26(2): 78-90.
[19] Mohammed AS, Ugwuowo FI. The trunsmuted exponential-exponential distribution with application to breast cancer data. *Pak. J. Statist*. 2021; 37(1): 1-18
[20] Ijaz M, Asim SM. Lomax exponential distribution with an application to real-life data. *PloS one*. 2019; 14(12): e0225827.
[21] Murthy and Xie, Jiang. Weibull model. 2004; *John Wiley and Sons*
[22] Vijay R, Sundaresan R, Maiya MP, Murthy SS, Fu Y, Klein HP, Groll M. Characterisation of Mg–x wt. FeTi (x= 5–30) and Mg–40 wt. FeTiMn hydrogen absorbing materials prepared by mechanical alloying. *Journal of alloys and compounds*. 2004; 384(1-2): 283-295
[23] Kim CS, Djamaludin I, Murthy DNP. Warranty and discrete preventive maintenance. *Reliability Engineering and System Safety*. 2004; 84(3): 301-309
[24] Choulakian V, Stephens MA. Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics*. 2001; 43(4): 478-484.

Md. Shohel Rana, is a researcher/professor(Assistant), at Islamic University, kushtia-7003, Bangladesh, with over 8 research publications, and studied Statistics discipline. His research interest/focus is in the areas of Stochastic Modeling, Distribution Theory, Statistical Modeling, Biostatistics, R Programming, Computational Statistics, and R Statistical Package.

Md. Matiur Rahman Molla, is a researcher/professor(Assistant), at Islamic University, kushtia-7003, Bangladesh, with over 4 research publications, and studied Statistics discipline. His research interest/focus is in the areas of Stochastic Modeling, Time Series Analysis, Distribution Theory, Statistical Modeling, Python, R Programming, Computational Statistics, and R Statistical Package.

Md. Mahabubur Rahman (PhD), is an Associate Professor of Statistics, Islamic University, Kushtia-7003, Bangladesh. He received the PhD degree in Statistics from the King Abdulaziz University, KSA. His research interests are in the areas of mathematical statistics as well as distribution theory. He has published several research articles extensively in internationally refereed journals. He is the referee of several mathematical journals.