The Pareto-Exponential Distribution: Theory and Real-life Applications

Md. Shohel Rana and Md. Mahabubur Rahman

Abstract— **In this article, Pareto-Exponential (PE) distribution, a sub model of the Pareto-X family which is introduced by Rana et al. [1], is explained. The maximum likelihood estimators of the model parameters are find out, and a simulation study is also done. Various distributional properties of the proposed model are illustrated. Some practical-life applications are appraised from different types of fields for examining the applicability of proposed model. The empirical implementation exhibits that the developed model allow greater suitability than any other models applied in this study.**

Keywords— **Family of distributions, T-X family, Pareto-X family, Pareto-Exponential distribution, Reliability function, Order statistics, Maximum likelihood estimation.**

I. Introduction

Development of a new probability distribution is a under way process to capture the multiplex character of the problems. Generalizations of probability model is considered to analysis different types of practical life facts and figures that are not covered by any specific model. so generalizations of probability model is a continuous process. A Swiss professor of economics, Vilfredo Pareto (1848-1923) introduced first the Pareto distribution which is applied in different exposure of socioeconomic conditions. Pareto [2] initiated the concept of the size distribution of income in his intimate economic texts. Burroughs and Tebbens [3] studied modeling earthquakes, forest fire areas, oil and gas field sizes.

To enhance the Pareto distribution, different generalizations of the models have been developed including the beta-Pareto distribution illustrated in [4]; Exponentiated Pareto distribution described in [5]; beta generalized Pareto distribution given in [6]; Gamma-Pareto distribution and its application explained in [7]; Kumaraswamy-Pareto distribution shown in [8]; Weibull-Pareto distribution and its applications expressed in [9]; Exponential-Pareto distribution presented in [10]; transmuted Pareto given in [11]; A new Weibull-Pareto distribution developed by the reference [12]; Kumaraswamy transmuted Pareto distribution presented in [13]; beta transmuted Pareto distribution expressed in [14] and cubic transmuted Pareto distribution proposed in [15] have been studied in the literature. All of these above distributions are the members of "Transformed-Transformer" (T-X) family developed in [16] as

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

If the random variable T follows Pareto distribution having probability density function (pdf) as $r(t) = \frac{\alpha x_m^\alpha}{t^{\alpha+1}}$ with $x_m \in \mathbb{R}^+$ and $\alpha \in \mathbb{R}^+$ are the scale and shape parameters respectively, and without loss of generality if $a = x_m$, and $W\left(G\left(x\right)\right) = x_m - log\left(1 - G\left(x\right)\right)$, then using (1), the cumulative density function (cdf) of Pareto-X family is derived by Rana et al. [1], and is 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 $G(x)$ is a base cdf of any random variable X. So it is possible to enhance the members of the family by using suitable standard base cdf $G(x)$ of random variable X in (2). If X follows exponential distribution, a new proposed PE distribution is developed, and studied in detail in the following sections.

The article is structured as follows: Pareto-Exponential is defined in section II, In section III, some distributional properties are illustrated. Distribution of several order statistics are presented in section IV. Maximum likelihood estimation of the model parameters besides a simulation study are presented in section V. Section VI shows the practicability of two different real-life applications. Finally, some concluding remarks are given.

ISSN: 2736-5484
DOI:10.24018/ejmath.2022.3.3.116
Published on May 24, 2022.
Md. Shohel Rana, Department of Statistics, Islamic University, Bangladesh. (e-mail: shohel.rana@iubat.edu).
Md. Mahabubur Rahman, Department of Statistics, Islamic University, Bangladesh. (corresponding e-mail: mmriu.stat@gmail.com).

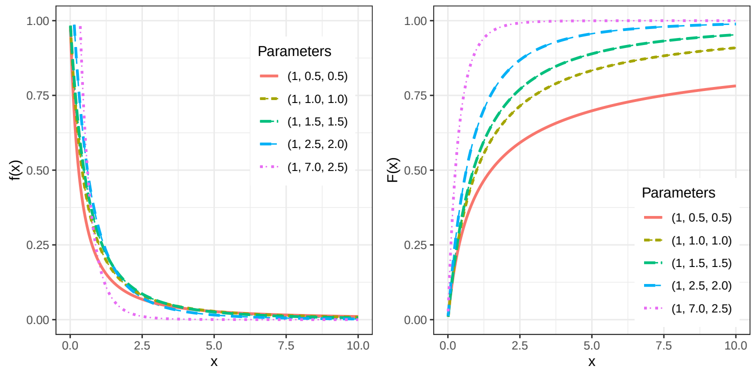

Fig. 1: The density function $f(x; x_m, \theta, \alpha)$ and distribution function $F(x; x_m, \theta, \alpha)$ of the proposed PE distribution.

II. Pareto-Exponential Distribution

Pareto-Exponential distribution, a member of Pareto-X family is defined here. Let X be a continuous random random variable having the exponential distribution function as

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

where $\theta \in \mathbb{R}^+$ is the scale parameters. Then, using (3) in (2), the cdf of the proposed PE distribution is obtained as

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

where $x_m, \theta \in \mathbb{R}^+$ are the scale parameters and $\alpha \in \mathbb{R}^+$ is the shape parameter of the model. The probability density function (pdf) of the proposed PE distribution is obtained by differentiating (4) with respect to x and further defined as follow.

Definition: The continuous random variable X is said to have a Pareto-Exponential distribution if its pdf is defined as

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

where $x_m, \theta \in \mathbb{R}^+$ are the scale parameters and $\alpha \in \mathbb{R}^+$ is the shape parameter of the model.

Special Case: The distribution function developed by Waseem et al. illustrated in [17] is a especial case of the proposed PE distribution function given in (4) for $\theta = \frac{1}{\lambda}$ and $x_m = 1$.

Some possible shapes for the pdf and cdf functions of the proposed PE distribution are presented in Fig. 1 that explained the proposed distribution has the capability to capture different behavior in datasets.

III. Distributional Properties

Several important distributional properties of the proposed PE distribution have been studied in the following subsections.

A. Moments

Let X be a random variable follows proposed PE distribution. Then, the r^{th} raw moment is given as

$$\mu_r' = \frac{\Gamma(r+1)\Gamma(\alpha - r)(\theta x_m)^r}{\Gamma(\alpha)}, \alpha > r. \tag{6}$$

Mean can be obtained by setting $r = 1$ in (6) as

$$\text{Mean} = \frac{\theta x_m \Gamma(\alpha - 1)}{\Gamma(\alpha)}; \alpha > 1, \tag{7}$$

and variance is obtained as

$$\text{Variance} = E(X^2) - E(X)^2 = \frac{\alpha\theta^2 x_m^2}{(\alpha-2)(\alpha-1)^2}; \alpha > 2, \tag{8}$$

where $E(X^r)$ for $i = 1, 2$ can be obtained from (6). One can obtain all other higher moments by using $r > 2$ in (6).

Using (7) and (8), Table I shows means and variances (in parenthesis) chart of the proposed PE distribution for different values of the model parameters respectively, and plotted in Fig. 2.

It is also investigated that when we increase the values of shape parameter α, remaining the constant value of scale parameter θ, it increases the value of means and variances of proposed PE distribution, and when we increase the values of scale parameter

TABLE I: Means and variances (in parenthesis) for the proposed PE distribution

		$\theta = 0.5$	$\theta = 1$	$\theta = 1.5$	$\theta = 2$	$\theta = 2.5$
$x_m = 1$	$\alpha = 2.01$	0.495 (49.260)	0.490 (24.270)	0.485 (15.946)	0.481 (11.788)	0.476 (9.297)
	$\alpha = 2.02$	0.990 (197.040)	0.980 (97.078)	0.971 (63.782)	0.962 (47.152)	0.952 (37.188)
	$\alpha = 2.03$	1.485 (443.339)	1.471 (218.426)	1.456 (143.510)	1.442 (106.093)	1.429 (83.673)
	$\alpha = 2.04$	1.980 (788.158)	1.961 (388.312)	1.942 (255.129)	1.923 (188.609)	1.905 (148.753)
	$\alpha = 2.05$	2.475 (1231.497)	2.451 (606.738)	2.427 (398.640)	2.404 (294.702)	2.381 (232.426)

θ, remaining the constant value of shape parameter α, it decreases the value of means and variances also.

We also know that the normalized r^{th} central or corrected moment of the random variable X can be written as

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

by (9), it is cleared that the 1^{st} and 2^{nd} standardized moments are 0 and 1. The third and fourth standardized moments are known as the measures of skewness (β_1) and kurtosis (β_2) respectively. So the measures of skewness and kurtosis of the proposed PE distribution are determined, and expressed as

$$\beta_1 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = \frac{2\theta^3 {x_m}^3 \left(\Gamma(\alpha-1)^3 - 3\Gamma(\alpha-2)\Gamma(\alpha)\Gamma(\alpha-1) + 12\theta x_m \Gamma(\alpha-4)\Gamma(\alpha)^2\right)}{\Gamma(\alpha)^3}, \tag{10}$$

and

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\theta^4 {x_m}^4 \Gamma(\alpha-1)^4 - 4\Gamma(\alpha-2)\Gamma(\alpha)\Gamma(\alpha-1)^2 - 8\Gamma(\alpha-4)\Gamma(\alpha)^3}{\Gamma(\alpha)^4} + \frac{32\theta x_m \Gamma(\alpha-4)\Gamma(\alpha)^2\Gamma(\alpha-1)}{\Gamma(\alpha)^4}, \tag{11}$$

respectively. Using (10) and (11), Table II shows the skewness and kurtosis (in parenthesis) chart of the proposed PE distribution for several values of the model parameters, and plotted in Fig. 3 that explain, it is positively skewed and leptokurtic.

B. Moment Generating Function

All the moments of a distribution are possible to obtain by using Moment Generating Function (MGF). But, it does not way out always like the characteristic function for all random variables. The MGF for the proposed PE distribution is given in the following theorem.

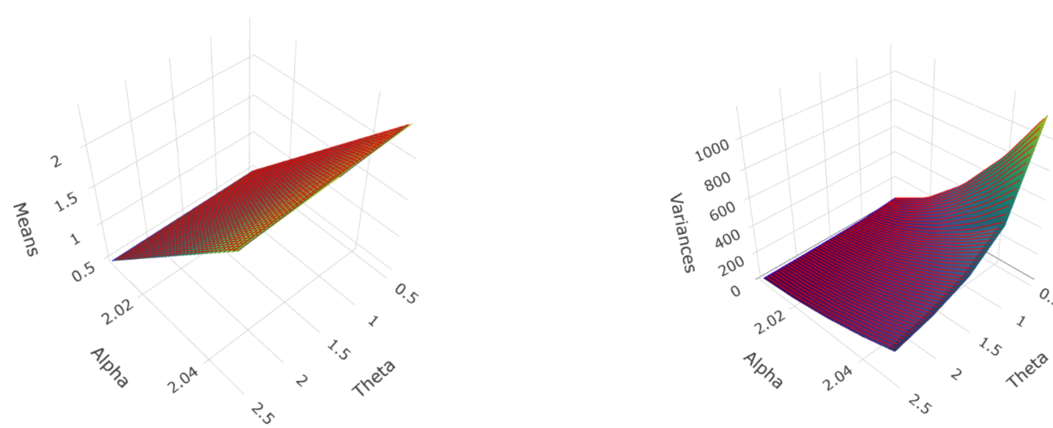

Fig. 2: The mean (left), and variance (right) plot of the proposed PE distribution.

TABLE II: Skewness and Kurtosis (in parenthesis) for the proposed PE distribution

		$\theta = 0.5$	$\theta = 1$	$\theta = 1.5$	$\theta = 2$	$\theta = 2.5$
$x_m = 1$	$\alpha = 2.01$	398.465 (786.149)	211.661 (412.175)	149.894 (288.254)	119.418 (226.888)	101.484 (190.580)
	$\alpha = 2.02$	6961.677 (27524.751)	3673.036 (14358.535)	2584.994 (9993.968)	2047.588 (7831.151)	1730.861 (6550.191)
	$\alpha = 2.03$	36232.765 (215010.019)	19078.158 (111993.966)	13401.569 (77840.374)	10596.943 (60912.507)	8943.240 (50883.689)
	$\alpha = 2.04$	116076.727 (918679.656)	61060.308 (478175.917)	42853.413 (332124.552)	33856.654 (259728.792)	28550.697 (216832.105)
	$\alpha = 2.05$	285680.432 (2826712.584)	150192.026 (1470692.372)	105351.843 (1021081.746)	83192.634 (798203.076)	70122.271 (666129.458)

Theorem: Let X follows the proposed PE distribution, then the moment generating function, $M_X(t)$ is

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

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

Proof: The moment generating function is defined 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 given in (5). Using the series representation of e^{tx} given in [18] , we have

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

Using $E(x^r)$ from (6) in (13), we have (12).

C. Reliability Analysis

The complement of cumulative distribution function is known as reliability function which is used by [19] or survival function used by [20] . Alternatively, it is defined by any other objects of interest will survive until a specified period of time. So for the proposed PE distribution, it is written as

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

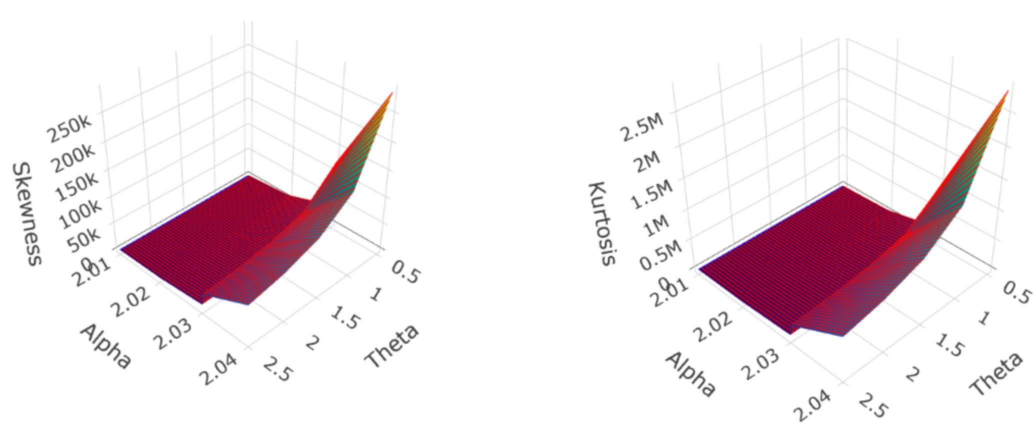

Fig. 3: The skewness (left) and kurtosis (right) plot of the proposed PE distribution.

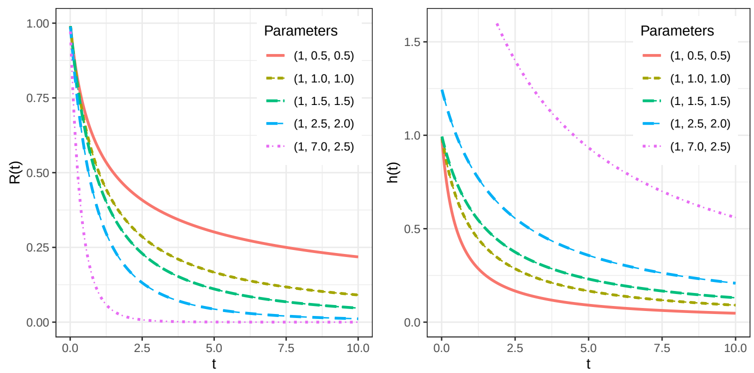

Fig. 4: The reliability function $R(t; x_m, \theta, \alpha)$ and hazard function $h(t; x_m, \theta, \alpha)$ of the proposed PE distribution.

where $x_m, \theta \in \mathbb{R}^+$ are the scale parameters and $\alpha \in \mathbb{R}^+$ is the shape parameter of the model. The hazard function is the ratio of the density function to the reliability function and is written as

$$h(t) = \frac{\alpha}{t + \theta x_m}; \; t \in [0, \infty). \tag{14}$$

Fig. 4 shows some possible shapes for the reliability and hazard functions of the proposed PE distribution with different values of the model parameters x_m, θ, and α respectively. The increasing then decreasing and decreasing hazard rates are observed from the figure.

D. Quantile Function and Median

The inverse of a probability distribution function of a random variable is known as quantile function, and is denoted by x_q. On the other hand, it is also known as percent point function. For proposed PE distribution, it is calculated by solving (4) for x and is gained as, see for example [15],

$$x_q = \theta \left(-\left((1-q)x_m^{-\alpha}\right)^{-1/\alpha} \right) \left(x_m \left((1-q)x_m^{-\alpha}\right)^{1/\alpha} - 1 \right), \tag{15}$$

and hence the median cab be written as

$$\text{Median} = x_{0.5} = -0.5^{-1/\alpha} \theta \left(x_m^{-\alpha}\right)^{-1/\alpha} \left(0.5^{1/\alpha} x_m \left(x_m^{-\alpha}\right)^{1/\alpha} - 1 \right). \tag{16}$$

The lower quartile, upper quartile can also be calculated by using $q = 0.25, 0.75$ in (15) respectively.

E. Generating Random Sample

When obtaining an unpredictable outcome is needed, random number generation is frequently used in various areas along with gambling, statistical sampling, and computer simulation also. For generating random sample from proposed PE distribution, the following expression can be used, see [15],

$$X = \theta \left(-\left((1-u)x_m^{-\alpha}\right)^{-1/\alpha} \right) \left(x_m \left((1-u)x_m^{-\alpha}\right)^{1/\alpha} - 1 \right), \tag{17}$$

where $u \sim U(0, 1)$. One can generate random sample from proposed PE distribution by using (17) for various known values of the model parameters.

IV. Order Statistics

Order statistics are the values of a random sample arranged in order of magnitude. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denotes the order statistic of a random sample $X_1, X_2, \cdots, X_n$ from a continuous population with distribution function $F_X(x)$, and density function $f_X(x)$. Then, the density function of $X_{r:n}$ is given by

$$f_{X_{r:n}}(x) = \frac{n!}{(r-1)!(n-r)!}[F(x)]^{r-1}[1-F(x)]^{n-r} f(x). \tag{18}$$

The density function of the rth order statistic in case of proposed PE distribution is obtained by using (18), as,

$$f_{X_{r:n}}(x) \quad = \quad \frac{n!}{(r-1)!(n-r)!} \left[1 - x_m^{\alpha} \left(\frac{x}{\theta} + x_m \right)^{-\alpha} \right]^{r-1} \left[x_m^{\alpha} \left(\frac{x}{\theta} + x_m \right)^{-\alpha} \right]^{n-r} \frac{\alpha x_m^{\alpha} \left(\frac{x}{\theta} + x_m \right)^{-\alpha-1}}{\theta}, \tag{19}$$

where $r = 1, 2, \cdots, n$. Using $r = 1$ in (19), we can obtain the density function of the lowest order statistic $X_{1:n}$, as

$$f_{X_{1:n}}(x) = n\frac{\alpha x_m^{\alpha}\left(\frac{x}{\theta}+x_m\right)^{-\alpha-1}}{\theta}\left[x_m^{\alpha}\left(\frac{x}{\theta}+x_m\right)^{-\alpha}\right]^{n-1}, \tag{20}$$

and for using $r = n$ in (19), the density function of the highest order statistic $X_{n:n}$, is

$$f_{X_{n:n}}(x) = n\frac{\alpha x_m^{\alpha}\left(\frac{x}{\theta}+x_m\right)^{-\alpha-1}}{\theta}\left[1-x_m^{\alpha}\left(\frac{x}{\theta}+x_m\right)^{-\alpha}\right]^{n-1}. \tag{21}$$

The kth order moment of $X_{r:n}$ in case of the proposed PE distribution is obtained by using the following formula:

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

where $f_{X_{r:n}}(x)$ is presented in (19).

V. Estimation of Parameters and Simulation study

Maximum Likelihood Estimation (MLE) method is used to estimate the parameters of a probability distribution by maximizing a likelihood function of a probability model. The technique of estimating the model parameters along with the simulation study for the proposed PE distribution is illustrated as follows.

A. Estimation and Inference

Let a random sample, $x_1, x_2, \cdots, x_n$ of size n is drawn from the proposed PE distribution. Now the likelihood function is given by

$$\mathcal{L}(x) = \left(\frac{\alpha x_m^{\alpha}}{\theta}\right)^n \prod_{i=1}^{n}\left(x_m+\frac{x}{\theta}\right)^{-(\alpha+1)},$$

with corresponding log-likelihood (Log-like.) function is

$$\ell(x) = \alpha n\log(x_m) + n\log(\alpha) - \mathrm{nlog}\theta - (\alpha+1)\sum_{i=1}^{n}\log\left(\frac{x_i}{\theta}+x_m\right). \tag{22}$$

The maximum likelihood estimator of x_m is the first-order statistic $x_{(1)}$. Then α and θ are obtained by maximizing (22). The derivatives of (22) with respect to the unknown parameters are given as

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

Now, setting, $\frac{\partial l}{\partial\alpha} = 0$ and $\frac{\partial l}{\partial\theta} = 0$, and solving the resulting nonlinear system of equations gives the maximum likelihood estimate $\hat{\Theta} = \left(\hat{\alpha}, \hat{\theta}\right)^{'}$ of $\Theta = (\alpha, \theta)^{'}$. Hence as $n \to \infty$, the asymptotic distribution of the $MLEs$ $\left(\hat{\alpha}, \hat{\theta}\right)$ are given by: see for example [21],

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\theta}\end{pmatrix} \sim N_2\left[\begin{pmatrix}\alpha\\ \theta\end{pmatrix}, \begin{pmatrix}\hat{V}_{11} & \hat{V}_{12}\\ \hat{V}_{21} & \hat{V}_{22}\end{pmatrix}\right], \tag{23}$$

where $\hat{V}_{ij} = V_{ij}|_{\theta=\hat{\theta}}$. The asymptotic variance-covariance matrix V, of the estimates $\hat{\alpha}$ and $\hat{\theta}$ are obtained by inverting Hessian matrix; see Appendix I. An approximate $100(1-\alpha)\%$ two sided confidence intervals for α and θ are respectively given by:

$$\hat{\alpha} \pm Z_{\alpha/2}\sqrt{\hat{V}_{11}}, \quad \hat{\theta} \pm Z_{\alpha/2}\sqrt{\hat{V}_{22}}$$

where Z_{α} is the α^{th} percentile of the standard normal distribution.

B. Simulation Study

We have conducted a Monte Carlo simulation study for proposed PE distribution. Several sample of sizes like, 50, 100, 200, 500 and 1000 have been used to perform . We have drawn random samples for the initial values of $\theta = 0.5$ and $\alpha = 0.5$, and hence using this parameters maximum likelihood estimators are calculated. The process is repeated 10000 times. The mean and mean squared errors (MSE's) are evaluated for the estimates. The obtained outcomes are presented in Table III, and the shape of MSE's of the model parameters are shown in Fig. 5. It is investigated that the resulting estimates stay very close to the real values of the parameters. So, it proves that the estimation process is accurate enough. Further, it is also examined that the estimated MSE's are decreasing consistently with respect to increasing the number of sample size. Finally, we have also been observed very clearly about the accuracy of the estimation methods.

TABLE III: Average estimate of model parameters and MSE's for proposed PE

Sample Size	Estimate		MSE	
	θ	α	θ	α
50	0.580	0.640	0.597	0.018
100	0.520	0.530	0.031	0.006
200	0.480	0.510	0.026	0.003
500	0.620	0.540	0.009	0.002
1000	0.560	0.540	0.003	0.001

VI. Real-life Applications

In various aspects of our daily life, the different distributions are frequently used for real-life datasets. In terms of proposed PE distribution, applications of two different real-life datasets have been studied as below.

Sources of Datasets: The Airplane dataset is used by [22], as failure times of the air conditioning system of an airplane as: 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52 and 95. Also, the Leukemia dataset about leukemia is considered from [23] as: 1, 3, 3, 6, 7, 7,10, 12, 14, 15, 18, 19, 22, 26, 29, 34, 40, 1, 1, 2, 2, 3, 4, 5, 8, 8, 9, 11, 12, 14, 16, 18, 21, 31 and 44. The summary statistics for the above two datasets are presented in Table IV.

TABLE IV: Summary statistics for the selected datasets

Dataset	Min.	Q_1	Median	Mean	Q_3	Max	Skewness	Kurtosis	Variance
Airplane	1.000	12.500	22.000	59.600	83.000	261.000	1.609	1.641	5167.421
Leukemia	1.000	4.500	11.000	13.600	18.500	44.000	0.985	0.136	130.423

The Transmuted-Pareto distribution developed by the reference [11] and Pareto distribution by the reference [2] are used for examining the fitness of the proposed PE distribution.

Fig. 6, the Probability-Probability (P-P) and Quantile-Quantile (Q-Q) plot represents that the proposed PE distribution strictly allow the real data. We have also been used the TTT plot to justify the pattern of hazard rate function. See [24], [25], and [26] for further study. TTT plot and estimated hazard curve of proposed PE distribution for Aireplane dataset are shown first at left and for Leukemia dataset are shown then at right of Fig. 7 respectively. We have been investigated that for both the datasets, a series of convex and concave TTT plot with upward trend whereas estimated hazard curve shows the slight decreasing failure rate, and then constantly decreasing failure rate.

Table V shows the determined values of the model parameters, and their corresponding standard errors besides log-likelihood for the other comparison models. Also some various 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 Statistic (C-vM) of the proposed PE model besides other chosen models for the considered datasets are presented in Table VI. The estimated cdf of the proposed PE distribution along with other comparison models for the Airplane, and Leukemia datasets are plotted over empirical distribution function, and are shown at left and right side of Fig. 8 respectively. The overall result analysis presented in different tables, and the estimated plots of the proposed PE

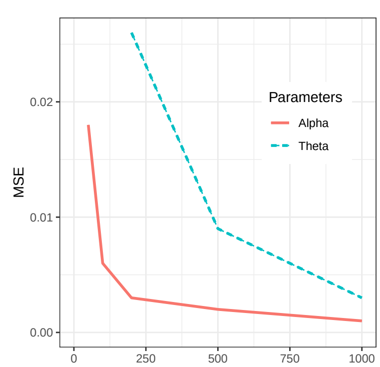

Fig. 5: The MSE's of model parameters versus different combination of sample observations for proposed PE distribution.

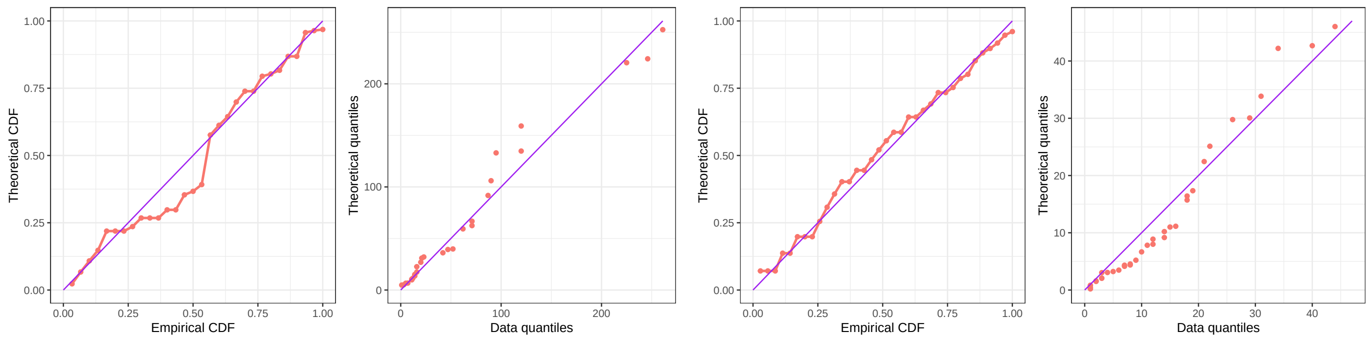

Fig. 6: P-P and Q-Q plots of the proposed PE distribution for the Aireplane (left) and Leukemia (right) datasets.

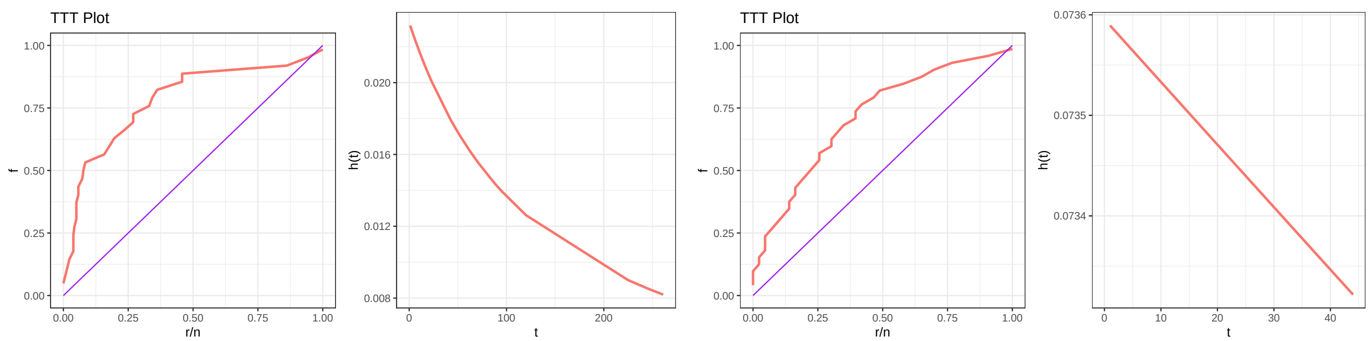

Fig. 7: Estimated TTT and Hazard Curves of the proposed PE distribution for the Aireplane (left) and Leukemia (right) datasets.

TABLE V: MLE of parameters and respective SE for the selected models for both the datasets

Datasets	Distribution	Parameter	Estimate	SE	Log-like.
Airplane	Pareto-Exponential	x_m	1.000	–	-151.837
		θ	141.218	168.235	
		α	3.297	3.074	
	Transmuted-Pareto	x_m	1.000	–	-161.474
		θ	-0.888	0.111	
		α	0.419	0.059	
	Pareto	x_m	1.000	–	-167.084
		α	0.298	0.054	
Leukemia	Pareto-Exponential	x_m	1.000	–	-126.359
		θ	11779.329	104388.254	
		α	866.907	7675.908	
	Transmuted-Pareto	x_m	1.000	–	-134.937
		θ	-0.696	0.171	
		α	0.606	0.087	
	Pareto	x_m	1.000	–	-138.235
		α	0.460	0.078	

distribution besides the plots of other selected models for the selected datasets shown in various figures demonstrate the better fit of the performances in favor of proposed PE distribution than other chosen comparison models used in this study.

VII. Concluding Remarks

In this study, a two-parameters sub model of the Pareto-X family named as the Pareto-Exponential distribution has been illustrated in detail. The distributional properties of the proposed Pareto-Exponential distribution including moments, moment generating function, characteristics function, quantile function, random number generator and reliability function along with the distribution of several order statistics are studied here. The maximum likelihood estimation method has been used to estimate the model parameters of the distribution, and a simulation study has been done also. Finally, two different real-life datasets have been used to analysis for illustrative purposes. Overall, we have investigated that the proposed Pareto-Exponential (PE) distribution performs better fit than other selected comparison models used in this study.

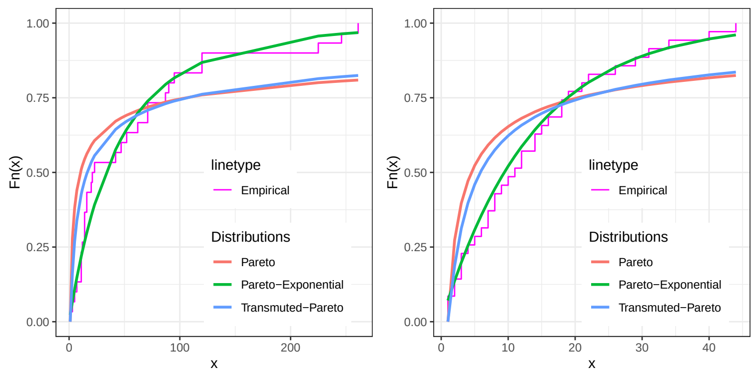

Fig. 8: Estimated cdf (left) for Airplane and estimated cdf (right) for Leukemia datasets respectively over empirical cdf.

TABLE VI: Selection criteria estimated for the selected models for both the datasets

Datasets	Distribution	-2Log-like.	AIC	AICc	BIC	KS	C-vM
Airplane	Pareto-Exponential	303.675	307.675	308.119	310.477	0.141	0.001
	Transmuted-Pareto	322.948	326.948	327.392	329.750	0.261	0.007
	Pareto	334.168	336.168	336.311	337.569	0.344	0.015
Leukemia	Pareto-Exponential	252.718	256.718	257.093	259.828	0.060	0.001
	Transmuted-Pareto	269.874	273.874	274.249	276.9852	0.202	0.018
	Pareto	276.47	278.470	278.591	280.026	0.249	0.038

Appendix I: Hessian Matrix

The Hessian matrix for the proposed Pareto-Exponential distribution is given as

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

where the variance-covariance matrix V is obtained by

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

with the elements of Hessian matrix are obtained as

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

and

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

Acknowledgment

As a first author, Md. Shohel Rana, would like to express the gratefulness to the counselor/supervisor, Associate Prof. Dr. Md. Mahabubur Rahman, for keeping worthy rolls to enrich the quality output of research article by giving anonymous suggestions and for his continuous supports. I'd also like to take this scope to thank my respectable supervisor for funding my PhD program through the Islamic University, Kushtia-7003, Bangladesh.

Conflict of Interest

Having no conflict of interest declared 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] Burroughs SM, Tebbens SF. Upper-truncated power law distributions. *Fractals*. 2001; 9(02): 209-222.
[4] Akinsete A, Famoye F, Lee C. The beta-Pareto distribution. *Statistics*. 2008; 42(6): 547-563.
[5] Shawky AI, Abu-Zinadah HH. Exponentiated Pareto distribution: different method of estimations. *International Journal of Contemporary Mathematical Sciences*. 2009; 4(14): 677-693.
[6] Mahmoudi E. The beta generalized Pareto distribution with application to lifetime data. *Math. and Computers in Simulation*. 2011; 81(11): 2414-2430.
[7] Alzaatreh A, Famoye F, Lee C. Gamma-Pareto distribution and its applications. *Journal of Modern Applied Statistical Methods*. 2012; 11(1): 7.
[8] Bourguignon M, Silva RB, Zea LM, Cordeiro GM. The Kumaraswamy Pareto distribution. *J. Stat. Theory Appl.* 2013; 12(2): 129-144.
[9] Alzaatreh A, Famoye F, Lee C. Weibull-Pareto distribution and its applications. *Communications in Statistics-Theory and Methods*. 2013; 42(9): 1673-1691.
[10] Al-Kadim KA, Boshi MA. Exponential Pareto distribution. *Mathematical Theory and Modeling*. 2013; 3(5): 135-146.
[11] Merovci F, Puka L. Transmuted Pareto distribution. *In ProbStat Forum*. 2014; 7(1): 1-11.
[12] 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.
[13] Chhetri SB, Akinsete AA, Aryal G, Long H. The Kumaraswamy transmuted Pareto distribution. *Journal of Statistical Distributions and Applications*. 2017; 4(1): 1-24.
[14] Chhetri S, Long H, Aryal G. The beta transmuted Pareto distribution: theory and applications. *Journal of Statistics Applications & Probability*. 2017; 6(2): 243-258.
[15] Rahman MM, Al-Zahrani B, Shahbaz MQ. Cubic transmuted Pareto distribution. *Annals of Data Science*. 2020; 7(1): 91-108.
[16] Alzaatreh A, Lee C, Famoye F. A new method for generating families of continuous distributions. *Metron*. 2013; 71(1): 63-79.
[17] 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.
[18] Zwillinger D, Jeffrey A. Table of integrals, series, and products. *Elsevier*. 2007.
[19] Ebeling CE. An introduction to reliability and maintainability engineering. *Tata McGraw-Hill Education*. 2004.
[20] Carpenter M. Survival analysis: a self-learning text. *Taylor & Francis*. 1997.
[21] 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.
[22] Gupta RD, Kundu D. Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*. 2001; 43(1): 117-130.
[23] Abinash MJ, Vasudevan V. A study on wrapper-based feature selection algorithm for leukemia dataset. In Intelligent Engineering Informatics. *Springer, Singapore*. 2018.
Abinash MJ, Vasudevan V. A study on wrapper-based feature selection algorithm for leukemia dataset. *Intelligent Engineering Informatics*. Springer, Singapore, 2018; 695: 311-321.
[24] Aarset MV. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*. 1987; 36(1): 106-108.
[25] Barlow RE, Campo R. Total time on test processes and applications to failure data analysis. *Reliability and fault tree analysis*. 1975.
[26] Cho M, Schenker N, Taylor JM, Zhuang D. Survival analysis with long-term survivors and partially observed covariates. *Canadian Journal of Statistics*. 2001; 29(3): 421-436.

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. Mahabubur Rahman (PhD), is an Associate Professor of Statistics, Islamic University, Kushtia, 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.