Pak.j.stat.oper.res. Vol.20 No.2 2024 pp 217-231 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.3845

Pakistan Journal of Statistics and Operation Research

A New Class of Probability Distributions With An Application to Engineering Data

Anwar Hassan[1], I. H. Dar[2*], M. A. Lone[3]

*Corresponding author

1. Department of Statistics, University of Kashmir,Srinagar, India, anwar.hassan5@gmail.com
2. Department of Statistics, University of Kashmir,Srinagar, India, ishfaqqh@gmail.com
3. Department of Statistics, University of Kashmir,Srinagar, India, murtazastat@gmail.com

Abstract

This manuscript introduces a novel class of probability distributions termed the New Exponentiated Transformation (NET), aimed at enhancing the flexibility of baseline distributions without adding complexity from extra parameters. The transformation is specialized on the exponentiated exponential distribution, resulting in the New Exponentiated Exponential (NEE) distribution. NEE offers increased flexibility in density function and features hazard rate functions with various shapes. The manuscript also highlights several mathematical properties of proposed distribution. To demonstrate the applicability of the proposed distribution, two engineering data-sets are analyzed, showing that NEE distribution provides a better fit than all other considered models.

Key Words:Exponentiated exponential distribution; hazard rate function; Renyi entropy; simulation study; maximum likelihood estimation.

Mathematical Subject Classification: 60E05, 62E15.

1. Introduction

In recent years, there has been a surge in the proposal of novel families of probability distributions, each seeking to augment the flexibility of existing distributions by introducing additional parameters. The overarching aim behind the introduction of these new distribution families is to enhance the ability to model complex data structures encountered across various domains. Among the researchers spearheading these endeavors, several notable contributions stand out. Marshall and Olkin (1997) presented a novel family of distribution functions and subsequently directed their focus towards enhancing the properties of the exponential distribution. Mahdavi and Kundu (2017) introduced the alpha power transformation (APT) family of distributions, aimed at extending the utility of the alpha power transformation methodology in statistical modeling. Building upon this work, Ijaz et al. (2021) proposed the novel alpha power transformed (NAPT) family of distribution functions, further expanding the repertoire of distributional options available to researchers. Cordeiro et al. (2013) ventured into a new class of distributions by incorporating two additional shape parameters, thereby offering greater flexibility in modeling a diverse range of data distributions. Hassan et al. (2021) proposed another family of probability distribution based on a trigonometric function. Similarly, Kumaraswamy (Kumaraswamy) introduced a novel methodology involving the addition of two shape parameters, further enriching the landscape of distributional possibilities. Additionally, Lone et al., Lone et al. (2022, 2024) introduced a new method and applied it to the Weibull distribution. Recenlty, Ahmad et al. (2024) proposed another method based on trigonometric functions. However, amidst this proliferation of parameter-rich distribution families, there remains a conspicuous lack of distributions in the statistical literature that provide flexibility without the burden of additional parameters. Recognizing this gap, researchers have endeavored to develop distribution families characterized by parsimony in parameters, thereby facilitating ease of interpretation and practical application. Mahmood and Chesneau (2019) introduced the new sine-G family of distributions, offering a streamlined alternative for modeling complex data structures. Maurya et al. (2017) proposed the DUS transformation, a novel transformation methodology aimed at simplifying the modeling process while maintaining flexibility. Expanding on this concept, Maurya et al. (2016) focused on the logarithmic transformation (LT), demonstrating its applicability in modeling exponential distributions. Subsequently, Maurya et al. (2018) introduced the new logarithmic transformation (NLT), featuring a single parameter with a decreasing failure rate, further enhancing its utility in practical applications.

In light of these developments, the need for distribution families that strike a balance between flexibility and complexity remains ever-present. In this manuscript, the aim is to address this need by introducing a new class of generating distributions characterized by enhanced flexibility without the inherent complexities associated with the addition of extra parameters. Through the presentation and analysis of the proposed distributions, the goal is to contribute to the ongoing evolution of distribution theory and its application in various fields of study.

In this manuscript, a contribution is made to the ongoing discourse by proposing a novel class of probability distributions, referred to as the New Exponentiated Transformation (NET). The aim is to offer a distribution framework that provides enhanced flexibility in modeling while maintaining simplicity and ease of interpretation.
Motivated by the recognition of the limitations of existing distribution families, particularly in engineering field, this manuscript focuses on the application of one particular distribution from this family, namely, the new exponentiated exponential distribution (NEE) distribution. Engineering disciplines frequently encounter data sets with complex patterns arising from factors such as material properties, structural configurations, and environmental conditions. Standard distributions often struggle to capture these intricacies accurately, leading to inaccuracies in modeling and analysis. The proposed NEE distribution provides a promising solution to this challenge by offering a flexible and robust framework tailored to the needs of engineering field. By leveraging the inherent adaptability of the NEE distribution, engineers and researchers can effectively model a wide range of engineering data sets, encompassing parameters such as gauge lengths and failure stresses. Moreover, the simplicity of the NEE distribution facilitates its seamless integration into existing modeling methodologies, ensuring accessibility and usability in practical engineering applications.
In the subsequent sections of this manuscript, a detailed exposition of the NEE distribution is presented, including its theoretical underpinnings, key properties, and practical implications. Additionally, a comprehensive analysis of two engineering data sets using the proposed distribution is conducted, demonstrating its efficacy in capturing the complexities of real-world engineering phenomena. Through this endeavor, the aim is to advance the state-of-the-art in statistical modeling and contribute to the ongoing evolution of distribution theory.

2. New Exponentiated Transformation (NET)

Let $G(y)$ be the commulative distribution function (cdf) of any random variable Y. Then the cdf, $F(y)$ of the new exponentiated transformation is given by:

$$F(y) = 2^{G(y)} - 1 \quad ; y \in \mathbb{R},$$

The corresponding probability density function (pdf) is given by:

$$f(y) = log(2)g(y)2^{G(y)} \quad ; y \in \mathbb{R},$$

The survival function $S(y)$ for NET is given by:

$$S(y) = 1 - (2^{G(y)} - 1) = (2 - 2^{G(y)})$$

The hazard rate function $\lambda(y)$ is given by:

$$\lambda(y) = \frac{log(2)f(y)}{(2^{(1-G(y))} - 1)},$$

3. New Exponentiated Exponential Distribution

Let Y be a random variable following an exponentiated exponential distribution with a cumulative distribution function $G(y) = (1 - e^{-\theta y})^{\alpha}$; $y, \alpha, \theta > 0$ then, the cdf of the New Exponentiated Exponential distribution is defined as follows:

$$F(y) = 2^{(1-e^{-\theta y})^{\alpha}} - 1 \ ; \quad \alpha, y > 0$$

The corresponding pdf is given as follows:

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

The survival and hazard rate functions are, respectively, given by:

$$S(y) = 1 - \left(2^{(1-e^{-\theta y})^{\alpha}} - 1\right) \ ; \quad \alpha, y > 0$$
$$S(y) = 2 - 2^{(1-e^{-\theta y})^{\alpha}}$$

and

$$\lambda(y) = log(2^{\alpha})\theta e^{-\theta y} \frac{(1 - e^{-\theta y})^{(\alpha-1)}}{2^{(1-(1-e^{-\theta y})^{\alpha})}}; \ \alpha, y > 0$$

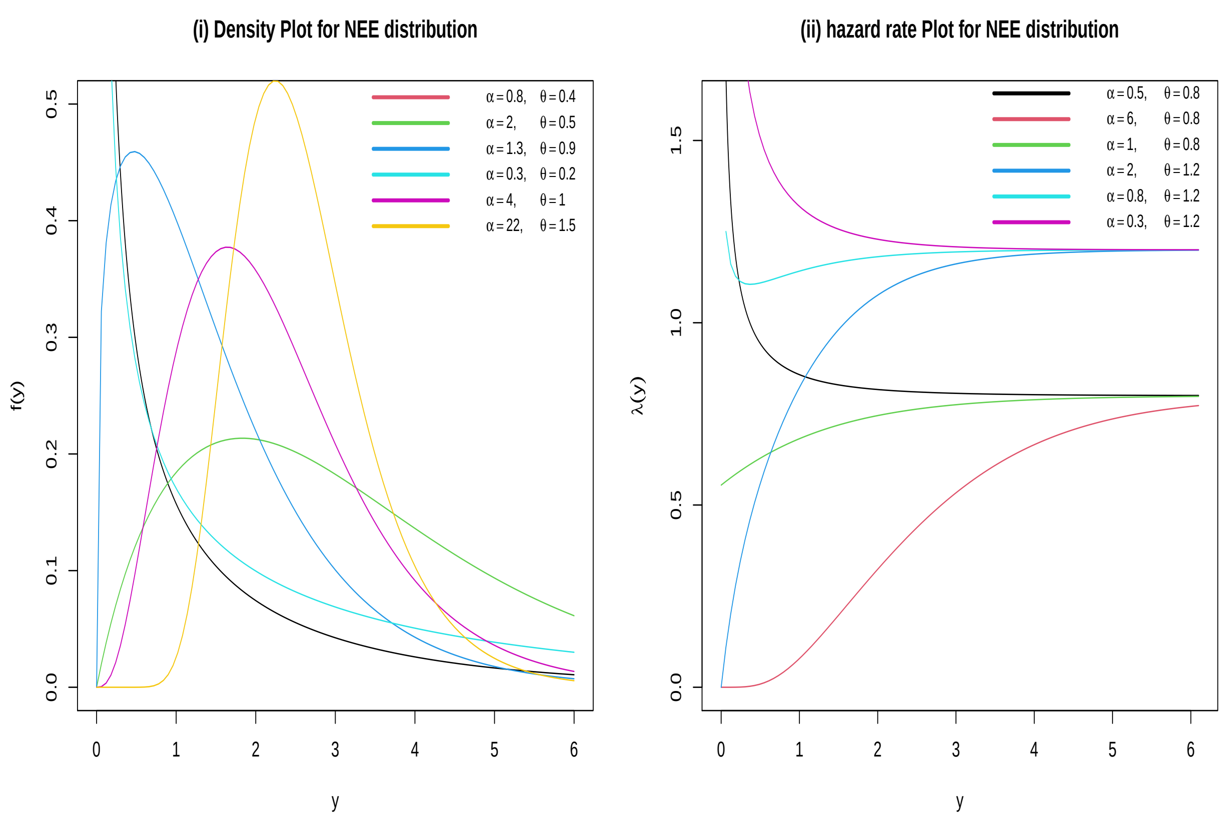

Figure 1: Plots of the NEE density and hazard rate function for different values of α and θ.

3.1. Behavior Of The Hazard Rate Function

To know the behaviour of the hazard rate function we go through Glaser (1980). He defined the term $\eta(t) = -\frac{f^{'}(t)}{f(t)}$ where $f(t)$ is density function and $f^{'}(t)$ is first derivative of $f(t)$ with respect to t and theorm is stated as.
Theorem 1:*1. If $\eta^{'}(t) > 0$ for all $t > 0$, then distribution has an increasing failure rate (IFR).*
2. If $\eta^{'}(t) < 0$ for all $t > 0$, then distribution has a decreasing failure rate (DFR)..
3. Suppose there exits $t^* > 0$ such that $\eta^{'}(t) < 0$, for all $t \in (0, t^*)$, $\eta^{'}(t^*) = 0$ and $\eta^{'}(t) > 0$ for all $t > t^*$ and $\epsilon = \lim_{t\to 0} f(t)$ exists.Then if
(i) $\epsilon = \infty$, distribution has bathtub failure rate.
(ii) $\epsilon = 0$, distribution has increasing failure rate (for more details see (5)).
Proof: Since , we have

$$\eta(t) = 2\left(\theta - \frac{\theta e^{-\theta t}(\alpha - 1) + log(2^{\alpha})\theta e^{-\theta t}(1 - e^{-\theta t})^{\alpha}}{(1 - e^{-\theta t})}\right)$$

and

$$\eta^{'}(t) = 2\frac{\theta^2 e^{-\theta t}}{(1 - e^{-\theta t})^2}\left((\alpha - 1) + log(2^{\alpha})(1 - e^{-\theta t})^{\alpha}(1 - \alpha e^{-\theta t})\right) \quad (2)$$

The following three cases arises.
(1) For $\alpha \geq 1$, then from equation (2), we have $\eta^{'}(t) > 0$ for all $t > 0$, hence distribution has an increasing failure rate (IFR).
(2) For $\alpha \leq 0.5$, then from equation (2), we have $\eta^{'}(t) < 0$ for all $t > 0$, hence distribution has a decreasing failure rate (DFR).
(3) For $0.5 < \alpha < 1$, then there exists a t^* such that $\eta^{'}(t) < 0$, for $t \in (0, t^*)$, $\eta^{'}(t^*) = 0$ and $\eta^{'}(t) > 0$ for all $t > t^*$, where t^* depends on the value of α and θ but it is not possible to find the exact functional form of t^* in terms of α and λ
Now from equation (1), we have

$$\lim_{t\to 0} f(t) = \lim_{t\to 0} log(2^{\alpha})\theta e^{-\theta t}(1 - e^{-\theta t})^{(\alpha-1)}2^{(1-e^{-\theta t})^{\alpha}} = \infty$$

Hence the distribution has bathtub hazard rate.
The behavior of hazard rate function at extremes for different parameter values.

$$\lim_{y\to 0}\lambda(y) = \begin{cases} \infty & for\ 0 < \alpha < 0.5, \\ k & for\ 0.5 < \alpha < 1 \\ log(2^{\alpha\theta}) & for\ \alpha = 1, \\ 0 & for\ \alpha > 1, \end{cases} \qquad \lim_{y\to\infty}\lambda(y) = \begin{cases} \theta & for \quad \alpha > 0, \end{cases}$$

where k is finite.

3.2. Quantile Function

The quantile function of NEE is given by:

$$Y = -\frac{1}{\theta}log\left[1 - \left(\frac{log(1+U)}{log2}\right)^{\frac{1}{\alpha}}\right]$$

where U $\sim (0, 1)$ distribution. The q^{th} quantile of NEE distribution is given by:

$$y_q = -\frac{1}{\theta}log\left[1 - \left(\frac{log(1+q)}{log2}\right)^{\frac{1}{\alpha}}\right]$$

The median is obtained as:

$$y_{0.5} = -\frac{1}{\theta}log\left[1-\left(\frac{log(1.5)}{log2}\right)^{\frac{1}{\alpha}}\right]$$

3.3. Moments

The r^{th} moment of the NEE distribution is given by:

$$E(Y^r) = \int_0^\infty y^r f(y)dy$$

by using

$$(\alpha)^x = \sum_{k=0}^{\infty}\frac{(log\alpha)^k}{k!}x^k, \tag{3}$$

and

$$(1-x)^n = \sum_{k=0}^{n}\binom{n}{k}(-1)^k x^k\ ; \qquad |x|<1 \tag{4}$$

and solving, we get the final expression for r^{th} moment as:

$$E(Y^r) = \frac{1}{\theta^r}log(2^\alpha)\sum_{j=0}^{\infty}\sum_{k=0}^{\alpha(j+1)-1}(-1)^k\frac{(log2)^j}{j!}\binom{\alpha(j+1)-1}{k}\frac{\Gamma(r+1)}{(k+1)^{r+1}}$$

3.4. Moment Generating Function

The moment generating function of NEE distribution is defined as:

$$M_Y(t) = \int_0^\infty e^{ty}f(y)dy$$

again using (3) and (4), the final expression of moment generating function is given as follows:

$$M_Y(t) = log(2^\alpha)\sum_{r=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\alpha(j+1)-1}\frac{(-1)^k(log2)^j}{j!(k+1)^{r+1}}\left(\frac{t}{\theta}\right)^r\binom{\alpha(j+1)-1}{k}\ ;t<\theta$$

3.5. Mean Residual Life And Mean Waiting Time

Suppose that Y is a continuous random variable with survivial function $S(y)$ then, the mean residual life is the expected additional lifetime given that a component has survived until time t. The mean residual life function, say $\mu(t)$, is defined as:

$$\mu(t) = \frac{1}{S(t)}\left(E(t) - \int_0^t yf(y)dy\right) - t$$

The mean residual life of NEE distribution is given as:

$$\mu(t)=\frac{1}{(2-2^{(1-e^{-\theta t})^{\alpha}})}\frac{1}{\theta}log(2^{\alpha})\sum_{j=0}^{\infty}\sum_{k=0}^{\alpha(j+1)-1}\frac{(-1)^{k}(log2)^{j}}{j!(k+1)^{2}}\binom{\alpha(j+1)-1}{k}\left[1-\gamma(\theta t(k+1),2)\right]-t$$

where $\gamma(a,b)=\int_{0}^{a}y^{b-1}e^{-y}dy$ is the lower incomplete gamma function.

The mean waiting time represents the waiting time elapsed since the failure of an object on condition that this failure had occurred in the interval [0, t]. The mean waiting time of Y, say $\bar{\mu}(t)$, is defined by:

$$\bar{\mu}(t)=t-\frac{1}{F(t)}\int_{0}^{t}yf(y)dy$$

$$\bar{\mu}(t)=t-\frac{1}{(2^{(1-e^{-\theta t})^{\alpha}}-1)}\frac{1}{\theta}log(2^{\alpha})\sum_{j=0}^{\infty}\sum_{k=0}^{\alpha(j+1)-1}\frac{(-1)^{k}(log2)^{j}}{j!(k+1)^{2}}\binom{\alpha(j+1)-1}{k}\gamma(\theta t(k+1),2)$$

3.6. Rényi Entropy

The entropy of a random variable measures the variation of the uncertainty. A large value of entropy indicates the greater uncertainty in the data. The Rényi entropy, say $RE_{Y}(\beta)$ is defined as:

$$RE_{Y}(\beta)=\frac{1}{1-\beta}log\left(\int_{-\infty}^{\infty}f(y)^{\beta}dy\right);\quad \beta>0,\quad \beta\neq 1.$$

The Rényi entropy of NEE distribution is given by:

$$RE_{Y}(\beta)=\frac{1}{1-\beta}log\left((log(2^{\alpha}))^{\beta}\sum_{j=0}^{\infty}\sum_{k=0}^{\alpha j+\beta(\alpha+1)}\frac{(-1)^{k}(log2^{\beta})^{j}}{j!(k+1)}\binom{\alpha j+\beta(\alpha+1)}{k}\right)$$

3.7. Order Statistics

Let $Y_1, Y_2, ..., Y_n$ be a random sample of size n, and let $Y_{i:n}$ denote the i^{th} order statistic, then, the pdf of $Y_{i:n}$, say $f_{i:n}(y)$ is given by:

$$f_{i:n}(y)=\frac{n!}{(i-1)!(n-i)!}F(y)^{i-1}f(y)(1-G(y))^{n-i}.$$

The pdf $f_{i:n}(y)$ of i^{th} order statistic of NEE distribution is given as:

$$f_{i:n}(y)=\frac{log(2^{\alpha})}{B(i,n-i+1)}\theta e^{-\theta y}(1-e^{-\theta y})^{\alpha-1}2^{(1-e^{-\theta y})^{\alpha}}(2^{(1-e^{-\theta y})^{\alpha}}-1)^{i-1}(2-2^{(1-e^{-\theta y})^{\alpha}})^{n-1}$$

Where $B(a,b)$ is the beta function.

3.8. Stress Strength Parameter

Let Y_1 and Y_2 be independent strength and stress random variables respectively, where $Y_1\sim NEE(\alpha_1,\theta)$ and $Y_2\sim NEE(\alpha_2,\theta)$, then the stress strength parameter $\mathbb{P}(Y_1>Y_2)$, say R , is defined as:

$$R=\int_{-\infty}^{\infty}f_1(y)F_2(y)dy$$

The stress strength parameter R, is defined as

$$R = \int_0^\infty \theta e^{-\theta y}(1-e^{-\theta y})^{(\alpha_1-1)}2^{(1-e^{-\theta y})^{\alpha_1}}(2^{(1-e^{-\theta y})^{\alpha_2}}-1)dy,$$

after solving the above expression, the stress strength parameter R is given by:

$$R = log(2^{\alpha_1})\sum_{j=0}^{\infty}\frac{(log2)^j}{j!}\left[\sum_{k=0}^{(j(\alpha_1+\alpha_2)+\alpha_1-1)}\frac{(-1)^k}{(k+1)}\binom{j(\alpha_1+\alpha_2)+\alpha_1-1}{k} - \sum_{l=0}^{(\alpha_1(j+1)-1)}\frac{(-1)^l}{(l+1)}\binom{\alpha_1(j+1)-1}{l}\right]$$

4. Statistical Inference

4.1. Maximum Likelihood Estimation

Let $y_1, y_2, ..., y_n$ be a random sample from NEE distribution, then the logarithm of the likelihood function is given as follows:

$$l = nlog(log2^{\alpha}) + nlog\theta - \theta\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} log(1-e^{-\theta y_i})^{(\alpha-1)} + \sum_{i=1}^{n}(1-e^{-\theta y_i})^{\alpha} log2 \quad (5)$$

For getting maximum likelihood estimates of α and θ, we partially differentiate (5) with respect to the corresponding parameters and equating the derivatives to zero. This yeilds the following equations:

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} log(1-e^{-\theta y_i})\left(1+\sum_{i=1}^{n}(1-e^{-\theta y_i})^{\alpha} log2\right) = 0$$

$$\frac{\partial l}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^{n} y_i + (\alpha-1)\sum_{i=1}^{n}\left(\frac{y_i e^{-\theta y_i}}{1-e^{-\theta y_i}}\right) + \alpha log2\sum_{i=1}^{n} y_i e^{-\theta y_i}(1-e^{-\theta y_i})^{(\alpha-1)} = 0$$

Since the partial derivatives of the log-likelihood equation are not in explicit form, the maximum likelihood estimates cannot be obtained analytically. Instead, numerical methods such as optimization algorithms need to be employed to maximize the log-likelihood function and obtain the estimates. In this study, the maximum likelihood estimates were obtained using the optimization functions available in the R software package, which provides efficient tools for statistical analysis and estimation.

Theorem 2:*If the parameter θ is known ,then the MLE of α exists and is unique.*

Proof: Since,

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} log(1-e^{-\theta y_i})\left(1+\sum_{i=1}^{n}(1-e^{-\theta y_i})^{\alpha} log2\right)$$

$$\lim_{\alpha\to 0}\frac{\partial l}{\partial \alpha} = \infty + \sum_{i=1}^{n} log(1-e^{-\theta y_i}) = \infty$$

Also

$$\lim_{\alpha\to\infty}\frac{\partial l}{\partial\alpha}=0+\sum_{i=1}^{n}log(1-e^{-\theta y_i})<0$$

Therefore, there exists atleast one root say $\hat{\alpha}(0,\infty)$, such that $\frac{\partial l}{\partial\alpha}=0$
For uniqueness of root, we have

$$\frac{\partial^2 l}{\partial\alpha^2}=-\frac{n}{\alpha^2}+log2\sum_{i=1}^{n}\left(log(1-e^{-\theta y_i})\right)^2(1-e^{-\theta y_i})<0$$

whenever,

$$\frac{n}{\alpha^2}>log2\sum_{i=1}^{n}\left(log(1-e^{-\theta y_i})\right)^2(1-e^{-\theta y_i})$$

Theorem 3: *If the parameter α is known, then the MLE of θ exists and is unique.*
Proof: Since,

$$\frac{\partial l}{\partial\theta}=\frac{n}{\theta}-\sum_{i=1}^{n}y_i+(\alpha-1)\sum_{i=1}^{n}\left(\frac{y_ie^{-\theta y_i}}{1-e^{-\theta y_i}}\right)+\alpha log2\sum_{i=1}^{n}y_ie^{-\theta y_i}(1-e^{-\theta y_i})^{(\alpha-1)}$$

$$\lim_{\theta\to0}\frac{\partial l}{\partial\theta}=\infty-\sum_{i=1}^{n}y_i=\infty$$

Also

$$\lim_{\theta\to\infty}\frac{\partial l}{\partial\theta}=0-\sum_{i=1}^{n}y_i<0$$

Therefore, there exists atleast one root say $\hat{\theta}(0,\infty)$, such that $\frac{\partial l}{\partial\theta}=0$
For uniqueness of root, we have

$$\frac{\partial^2 l}{\partial\theta^2}=-\frac{n}{\theta^2}+\sum_{i=1}^{n}{y_i}^2e^{-\theta y_i}\left(\frac{(1-\alpha)}{(1-e^{-\theta y_i})^2}-\alpha log2(1-\alpha e^{-\theta yi})\right)<0$$

whenever,

$$\frac{n}{\alpha^2}>\sum_{i=1}^{n}{y_i}^2e^{-\theta y_i}\left(\frac{(1-\alpha)}{(1-e^{-\theta y_i})^2}-\alpha log2(1-\alpha e^{-\theta yi})\right)$$

4.2. Simulation Study

To assess the stability and consistency of the estimates obtained from the NEE distribution, a simulation exercise was conducted using R software. The simulation involved generating two sets of samples with sizes of 50 and 100, which were then replicated 1000 times for various combinations of the parameters $\alpha=(0.5,1,1.5,2,3)$ and $\theta=$

$(0.5, 1, 1.5, 2, 3, 5)$
For each parameter combination and sample size, the average values of the maximum likelihood estimates (MLEs) were computed, along with the empirical mean squared errors (MSEs) and bias. The results of the simulation exercise are presented in Table 1 for the sample size of 50 and Table 2 for the sample size of 100. The simulation results showed that as the sample size increased, the MSE and bias decreased in all cases. This indicates that the MLEs obtained from the NEE distribution are stable and consistent, and the estimation accuracy improves with larger sample sizes.

Table 1: Average values of MLEs their corresponding MSEs and Bias (n=50).

Parameter		MLEs		MSE		Bias	
α	θ	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$
0.5	0.5	0.57134	0.52205	0.08633	0.00652	0.07134	0.02205
	1	0.55059	1.03528	0.08402	0.02709	0.05059	0.03528
	1.5	0.54335	1.55186	0.07599	0.06009	0.04335	0.05186
	2	0.57550	2.08942	0.09867	0.09950	0.07550	0.08942
	3	0.55956	3.10780	0.08314	0.23320	0.05956	0.10780
	5	0.50185	5.23818	0.06469	1.34852	0.00185	0.34878
1	0.5	1.06282	0.51665	0.06197	0.00816	0.06282	0.01665
	1	1.05083	1.03376	0.05703	0.03266	0.05083	0.03376
	1.5	1.06133	1.57631	0.05132	0.08446	0.06133	0.07631
	2	1.06030	2.06750	0.05578	0.13775	0.06030	0.06750
	3	0.55738	3.14291	0.09749	0.48546	0.05738	0.14291
	5	1.05850	5.22939	0.05975	0.83857	0.05850	0.22939
1.5	0.5	1.62057	0.52156	0.17097	0.00814	0.12057	0.02156
	1	1.60750	1.04112	0.16592	0.03129	0.10750	0.04112
	1.5	1.61132	1.56079	0.16248	0.06684	0.11132	0.06079
	2	1.61974	2.10354	0.16130	0.12921	0.11974	0.10354
	3	1.61806	3.10803	0.15985	0.23010	0.11806	0.10803
	5	1.63724	5.24881	0.19127	0.70677	0.13724	0.24881
2	0.5	2.17301	0.51720	0.35351	0.00650	0.17301	0.01720
	1	2.17243	1.04111	0.34631	0.02540	0.17243	0.04111
	1.5	2.18124	1.56159	0.33782	0.05801	0.18124	0.06159
	2	2.19453	2.08430	0.35438	0.11187	0.19453	0.08430
	3	2.20274	3.14401	0.42489	0.27736	0.20274	0.14401
	5	2.16861	5.18609	0.36986	0.68634	0.16861	0.18609
3	0.5	3.23995	0.51141	0.91498	0.00580	0.23995	0.01141
	1	3.30667	1.03846	0.91361	0.02381	0.30667	0.03846
	1.5	3.32444	1.57016	0.99342	0.05924	0.32444	0.07016
	2	3.30723	2.05501	1.01406	0.09144	0.30723	0.05501
	3	3.24981	3.09117	0.90045	0.20010	0.24981	0.09117
	5	3.23030	5.11994	0.80221	0.53437	0.23030	0.11994

Table 2: Average values of MLEs their corresponding MSEs and Bias (n=100).

Parameter		MLEs		MSE		Bias	
α	θ	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$
0.5	0.5	0.51387	0.50797	0.03745	0.00312	0.01387	0.00797
	1	0.53577	1.02252	0.03693	0.01068	0.03577	0.02252
	1.5	0.51881	1.53341	0.03299	0.02613	0.01881	0.03341
	2	0.51376	2.03174	0.03739	0.05001	0.01376	0.03174
	3	0.52029	3.06187	0.03304	0.10124	0.02029	0.06187
	5	0.51707	5.14111	0.03603	0.60966	0.01707	0.14111
1	0.5	1.03448	0.51070	0.02512	0.00366	0.03448	0.01070
	1	1.02819	1.01506	0.02224	0.01562	0.02819	0.01506
	1.5	1.02463	1.52034	0.02356	0.03403	0.02463	0.02034
	2	1.05308	2.02560	0.12496	0.16825	0.05307	0.02560
	3	1.08331	3.02208	0.16356	0.42361	0.08331	0.02208
	5	1.02143	5.26282	0.13960	1.19519	0.02143	0.13960
1.5	0.5	1.56550	0.50785	0.07016	0.00304	0.06550	0.00785
	1	1.55509	1.02565	0.06599	0.01410	0.05509	0.02565
	1.5	1.57412	1.53681	0.07259	0.03316	0.07412	0.03681
	2	1.57002	2.03858	0.07667	0.05358	0.07002	0.03858
	3	1.57157	3.08789	0.06444	0.12156	0.07157	0.08789
	5	1.55026	5.10185	0.07385	0.36036	0.05025	0.10185
2	0.5	2.06975	0.50704	0.12767	0.00291	0.06975	0.00704
	1	2.05223	1.01001	0.11809	0.01091	0.05223	0.01001
	1.5	2.07430	1.52753	0.13023	0.02794	0.07430	0.02753
	2	2.08165	2.03086	0.13014	0.04544	0.08165	0.03086
	3	2.09106	3.04667	0.14016	0.10894	0.09106	0.04667
	5	2.06878	5.08603	0.11650	0.30130	0.06878	0.08603
3	0.5	3.12878	0.50681	0.39204	0.00269	0.12878	0.00681
	1	3.11934	1.01162	0.34365	0.01003	0.11934	0.01162
	1.5	3.11678	1.52036	0.34228	0.02314	0.11678	0.02036
	2	3.12212	2.02530	0.35640	0.039001	0.12212	0.02530
	3	3.17807	3.07972	0.38750	0.10216	0.17807	0.07972
	5	3.10338	5.06260	0.30608	0.21289	0.10338	0.06260

5. Application

To justify the validity and applicability of the NEE distribution two real data sets have been used. The data set I consists of 63 observations of the gauge lengths of 10mm. The data was taken from Kundu and Raqab (2009) The data set II reported by Bader and Priest (1982) on failure stresses (in GPa) of 65 single carbon fibers of lengths 50 mm.
For data set I, we compared the proposed NEE distribution with several other models namely, Marshall Olkin exponential (MOE) Marshall and Olkin (1997) alpha power exponential (APE) Mahdavi and Kundu (2017) exponentiated exponential (EE) Gupta and Kundu (2001) novel alpha power transformed exponential (NAPTE) Ijaz et al. (2021) Weibull (W), gamma (G), Rayleigh (R), Exponential (E), Weibull exponential (WE) Oguntunde et al. (2015) and exponentiated transmuted exponential (ETE) Al-Kadim and Mahdi (2018) distributions. For data set II, the proposed model is compared with alpha power exponential (APE) Mahdavi and Kundu (2017) exponentiated exponential (EE) Sarhan and Zaindin (2009) novel alpha power transformed exponential (NAPTE) Ijaz et al. (2021) Rayleigh (R), Exponential (E), Weibull exponential (WE) Oguntunde et al. (2015) and exponentiated transmuted exponential (ETE) Al-Kadim and Mahdi (2018) distributions.

Table 3: MLEs (standard errors in parentheses), K-S Statistic, and p-values for the data set I.

	Estimates			Statistics	
Model	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	K-S	p-value
NEE	255.22906 (137.17114)	2.07686 (0.20096)	-	0.084193	0.76320
MOE	5276.27857 (4826.29410)	2.83399 (0.29582)	-	0.09427	0.63000
APE	6.68161 (1.67772)	1.04561 (4.81109)	-	0.18082	0.03250
EE	218.23178 (111.017903)	1.94584 (0.19274)	-	0.08880	0.70310
NAPTE	5.56317 (1.18632)	1.04092 (4.70195)	-	0.16544	0.06358
W	4.85846 (0.24854)	0.00337 (0.00105)	-	0.11335	0.39320
G	25.59175 (0.26389)	8.36522 (0.03032)	-	0.08980	0.70120
R	2.20666 (0.13900)	-	-	0.36072	1.517e-07
E	0.32687 (0.04118)	-	-	0.48600	2.378e-13
WE	0.00881 (0.00405)	0.94623 (0.97387)	1.48174 (1.57994)	0.13208	0.22180
ETE	218.44687 (2009.70045)	0.00100 (9.33959)	1.94583 (0.19836)	0.08880	0.70310

Table 4: $-2l(\hat{\theta})$, AIC, AICC, BIC for the data set I.

Model	$-2l(\hat{\theta})$	AIC	AICC	BIC
NEE	112.7573	116.7573	116.9573	121.0436
MOE	118.6353	122.6353	122.8353	126.9216
APE	147.2311	151.2311	151.4311	155.5173
EE	113.0324	117.0324	117.2324	121.3187
NAPTE	144.4528	148.4528	148.6528	152.7391
W	124.9071	128.9071	129.1071	133.1934
G	113.7575	117.7575	118.1643	118.1643
R	187.0399	187.0399	189.1055	191.1830
E	266.8915	268.8915	268.9571	271.0347
WE	138.4807	144.4807	144.8875	150.9102
ETE	113.0324	119.0324	119.4392	125.4619

Table 5: MLEs (standard errors in parentheses), K-S Statistic, and p-values for the data set II.

	Estimates			Statistics	
Model	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	K-S	p-value
NEE	211.69341 (102.44910)	2.73363 (0.24746)	-	0.09439	0.6087
APE	7.37702 (2.37265)	1.42395 (6.43925)	-	0.19828	0.0120
EE	176.34593 (80.65879)	2.54031 (0.23657)	-	0.098335	0.5559
NAPTE	2.72827 (1.18633)	1.39546 (6.29347)	-	0.20394	0.0089
R	1.61335 (0.10005)	-	-	0.35092	2.231e-07
E	0.445624 (0.00116)	-	-	0.46779	8.837e-13
WE	0.00399 (0.00116)	1.19794 (0.24405)	1.90580 (1.73691)	0.09358	0.6196
ETE	176.51866 (400.00971)	0.00100 (2.25642)	2.54031 (0.23695)	0.09833	0.5559

Table 6: $-2l(\hat{\theta})$, AIC, AICC, BIC for the data set II.

Model	$-2l(\hat{\theta})$	AIC	AICC	BIC
NEE	75.27530	79.27530	79.46885	83.62407
APE	110.2652	114.2652	114.4587	118.6140
EE	76.72420	80.72420	80.91775	85.07298
NAPTE	108.9861	112.9861	113.1797	117.3349
R	151.5756	153.5756	153.6391	155.7500
E	235.0765	237.0765	237.1400	239.2509
WE	78.12431	84.12431	84.51775	90.64747
ETE	76.72430	82.72430	83.11775	89.24746

From table (3),(4), (5) and (6) it is evident that the NEE distribution consistently yields the smallest values of the criteria -2l, AIC, AICC, and BIC, along with the maximum p-value when compared to all other distributions. Therefore, we can conclude that the proposed model provides the best fit for both data sets I and II

Pak.j.stat.oper.res. Vol.20 No.2 2024 pp 217-231 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.3845

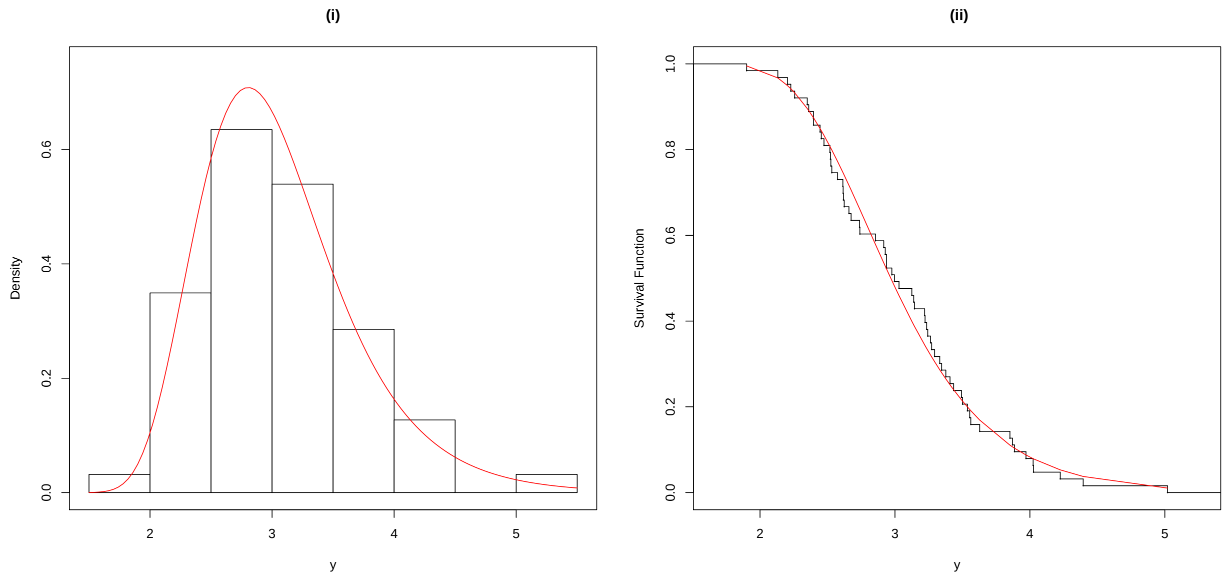

Figure 2: (i) The relative histogram and the fitted NEE distribution. (ii) The empirical survival function and fitted NEE survival function for data set I.

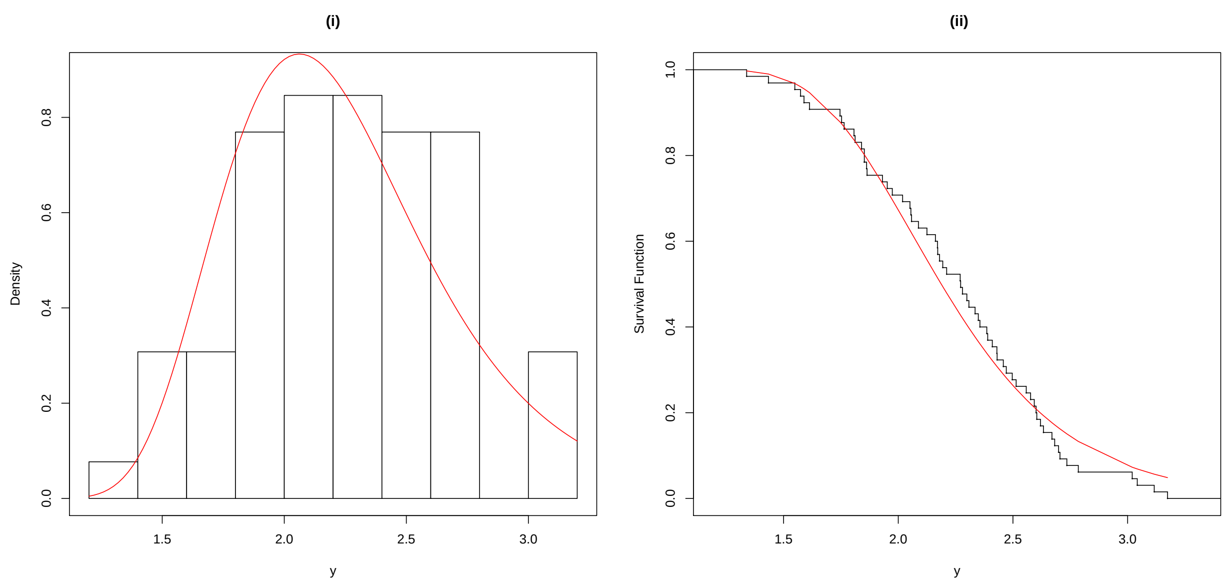

Figure 3: (i) The relative histogram and the fitted NEE distribution. (ii) The empirical survival function and fitted NEE survival function for data set II.

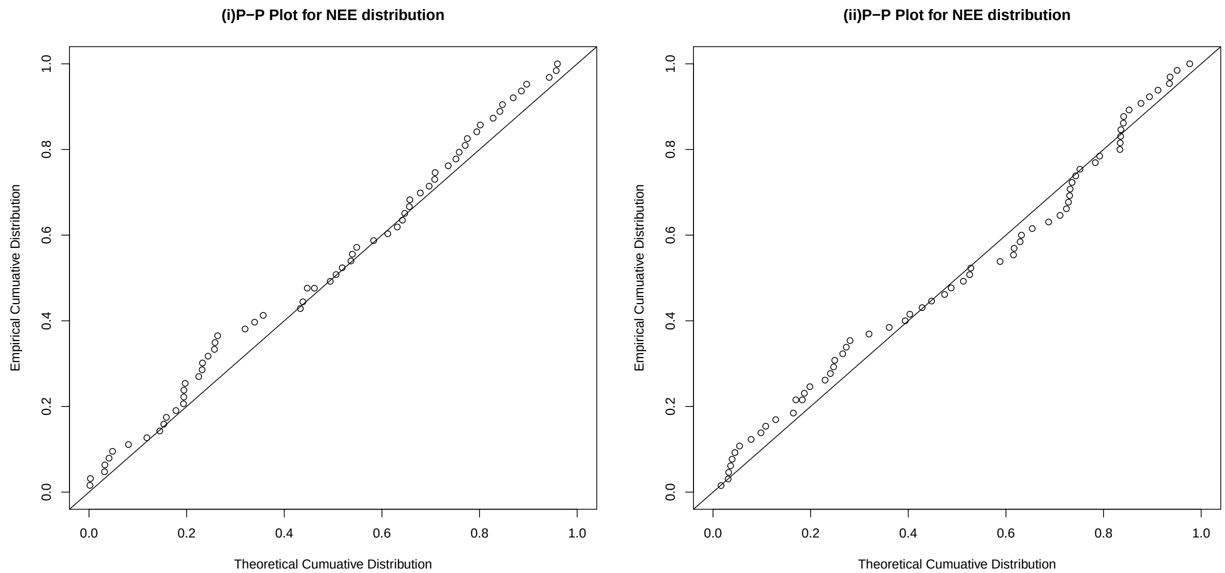

Figure 4: P-P plot for the NEE distribution for data set I and data set II

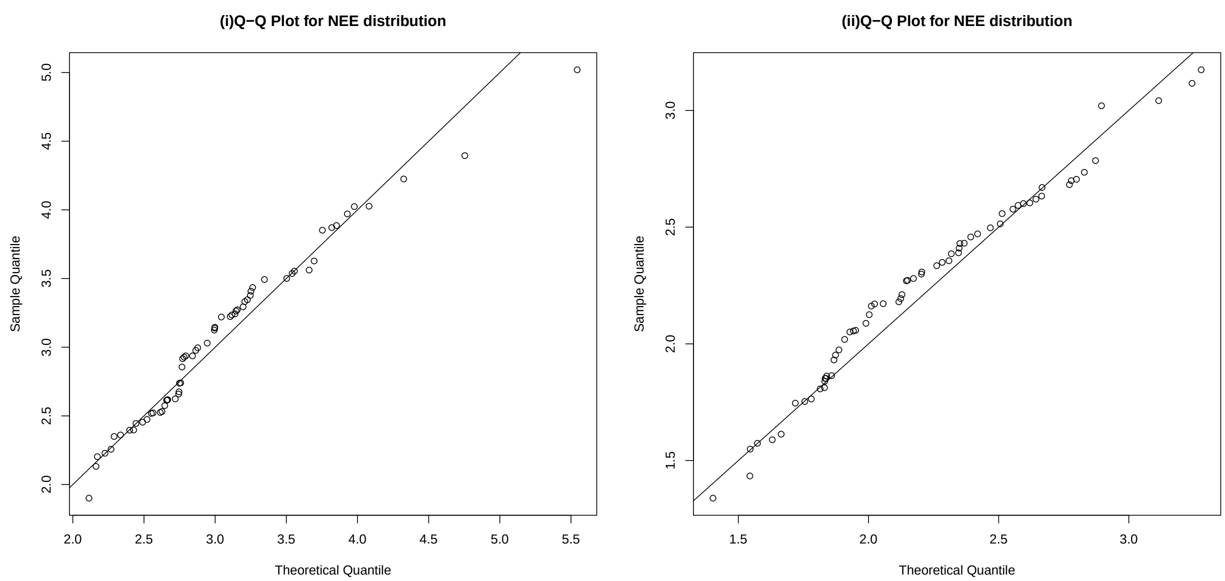

Figure 5: Q-Q plot for the NEE distribution for data set I and data set II

Figure 2(i) and 3(i) display relative histograms for data set I and II respectively. Also, the Figure 2(ii) and 3(ii) shows the plots of the fitted NEE survival function and empirical survival function of the data set I and II, respectively. The Q-Q and P-P plots are presented by 4(i) and (ii) and 5(i) and (ii) for data set I and II, respectively, which permits us to compare the empirical distribution of the data with the NEE distribution, these graphical measures also support the results provided in tables (3), (4), (5) and (6)

6. Concluding Remarks

In conclusion, the introduction of the New Exponentiated Transformation marks a significant advancement in statistical modeling. This new family of distribution functions offers enhanced flexibility without the addition of extra parameters, providing a versatile framework for capturing the complexities inherent in diverse data structures.The investigation conducted in this study demonstrates that the NEE distribution surpasses several reputed models mentioned in the manuscript in terms of flexibility and applicability. By eschewing the burden of an extra parameter, the NET family of distribution strikes a balance between simplicity and versatility, making it an attractive choice for a wide range of statistical applications. Through a thorough analysis using two real data sets, it is observed that the proposed NEE distribution outperforms all other competitive models considered in this study. Exhibiting superior fit and predictive accuracy across various metrics, the NEE distribution emerges as a promising tool for modeling complex data patterns with precision and efficiency.

References

1. Ahmad, A., Rather, A. A., Gemeay, A. M., Nagy, M., Sapkota, L. P., and Mansi, A. (2024). Novel sin-g class of distributions with an illustration of lomax distribution: Properties and data analysis. *AIP Advances*, 14(3).
2. Al-Kadim, K. A. and Mahdi, A. A. (2018). Exponentiated transmuted exponential distribution. *Journal of University of Babylon for Pure and Applied Sciences*, 26(2):78–90.
3. Bader, M. and Priest, A. (1982). Statistical aspects of fibre and bundle strength in hybrid composites. *Progress in science and engineering of composites*, pages 1129–1136.
4. Cordeiro, G. M., Ortega, E. M., and da Cunha, D. C. (2013). The exponentiated generalized class of distributions. *Journal of data science*, 11(1):1–27.
5. Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 75(371):667–672.
6. Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1):117–130.
7. Hassan, A., Dar, I., and Lone, M. (2021). A novel family of generating distributions based on trigonometric function with an application to exponential distribution. *Journal of scientific research*, 65(5):172–179.
8. Ijaz, M., Mashwani, W. K., Göktaş, A., and Unvan, Y. A. (2021). A novel alpha power transformed exponential distribution with real-life applications. *Journal of Applied Statistics*, pages 1–16.
9. Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of hydrology*, 46(1-2):79–88.
10. Kundu, D. and Raqab, M. Z. (2009). Estimation of r= p (y¡ x) for three-parameter weibull distribution. *Statistics & Probability Letters*, 79(17):1839–1846.
11. Lone, M. A., Dar, I. H., and Jan, T. (2022). An innovative method for generating distributions: Applied to weibull distribution. *Journal of Scientific Research*, 66(3).
12. Lone, M. A., Dar, I. H., and Jan, T. (2024). A new family of generalized distributions with an application to weibull distribution. *Thailand Statistician*, 22(1):1–16.
13. Mahdavi, A. and Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557.
14. Mahmood, Z. and Chesneau, C. (2019). A new sine-g family of distributions: properties and applications.
15. Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3):641–652.
16. Maurya, S., Kaushik, A., Singh, S., and Singh, U. (2017). A new class of distribution having decreasing, increasing, and bathtub-shaped failure rate. *Communications in Statistics-Theory and Methods*, 46(20):10359–10372.
17. Maurya, S., Kumar, D., Singh, S., and Singh, U. (2018). One parameter decreasing failure rate distribution. *International Journal of Statistics & Economics*, 19(1):120–138.
18. Maurya, S. K., Kaushik, A., Singh, R. K., Singh, S. K., and Singh, U. (2016). A new method of proposing distribution and its application to real data. *Imperial Journal of Interdisciplinary Research*, 2(6):1331–1338.
19. Oguntunde, P., Balogun, O., Okagbue, H., and Bishop, S. (2015). The weibull-exponential distribution: Its properties and applications. *Journal of Applied Sciences*, 15(11):1305–1311.
20. Sarhan, A. M. and Zaindin, M. (2009). Modified weibull distribution. *APPS. Applied Sciences*, 11:123–136.