

The Modi Exponentiated Exponential Distribution

Antoine Dieudonné Ndayisaba1*, Leo Odiwuor Odongo2, Anthony Ngunyi3

¹Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation (PAUSTI), Nairobi, Kenya

²Department of Mathematics and Actuarial Science, Kenyatta University (KU), Nairobi, Kenya

³Department of Mathematics, Statistics and Actuarial Sciences, Dedan Kimathi University of Technology (De KUT), Nairobi, Kenya

Email: *antoine.ndayisaba@students.jkuat.ac.ke

How to cite this paper: Ndayisaba, A.D., Odongo, L.O. and Ngunyi, A. (2023) The Modi Exponentiated Exponential Distribution. *Journal of Data Analysis and Information Processing*, **11**, 341-359. https://doi.org/10.4236/jdaip.2023.114017

Received: August 5, 2023 Accepted: September 16, 2023 Published: September 19, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

In this study, a new four-parameter distribution called the Modi Exponentiated Exponential distribution was proposed and studied. The new distribution has three shape and one scale parameters. Its mathematical and statistical properties were investigated. The parameters of the new model were estimated using the method of Maximum Likelihood Estimation. Monte Carlo simulation was used to evaluate the performance of the MLEs through average bias and RMSE. The flexibility and goodness-of-fit of the proposed distribution were demonstrated by applying it to two real data sets and comparing it with some existing distributions.

Keywords

Modi Family, Exponentiated Exponential Distribution, Maximum Likelihood Estimation, Order Statistics, Moments, Quantile Function

1. Introduction

Probability distributions are essential tools in a variety of disciplines as they provide statistical interpretations that help make sense of data. However, the existing distributions such as the exponential distribution have limitations making them insufficient in modeling a variety of data.

The exponential distribution is a continuous probability distribution commonly used in statistics and probability theory. It is frequently used to describe the time between events in a process where events occur independently at a fixed average rate [1].

To overcome these limitations, many authors have proposed modifications of

existing distributions. Gupta and Kundu [2] modified the exponential distribution by adding a shape parameter. However, their model is not flexible enough to control both skewness and kurtosis and also accommodate non-monotonic hazard rate shapes.

In the theory of probability distributions, choosing a particular probability distribution to represent real-life phenomena may be motivated by two factors such as tractability and flexibility [3]. Although a probability distribution's tractability may be advantageous in theory since it is simple to use, particularly when simulating random samples, practitioners and other stakeholders may be more interested in a distribution's flexibility. In reality, it is preferable to employ probability distributions that best fit the given data set rather than transforming the data set because doing so may compromise the data set's original features. Due to this, numerous attempts have been made in recent years to ensure that the current standard theoretical distributions are improved and extended [4] [5] [6]. Marshall and Olkin [7] proposed and studied a new method for adding a parameter to a family of distributions. Margaretha, et al. [8], in a conversation about employing Bayesian methodology to estimate parameters for the Exponentiated Exponential distribution in the context of left-censored data, in their work, Mahdavi and Kundu [9] introduced a new method for generating distributions and applied it to the exponential distribution. This generator was reported to control the skewness in the distribution which is not normal distributed. Many Researchers have made significant advancements in developing and extending various distributions.

Singh, et al. [10] and Niyoyunguruza, et al. [11] used the Marshall-Olkin generator method to extend distributions. They subsequently conducted parameter estimation using Maximum Likelihood Estimation (MLE) technique. By applying these extended distributions to a range of datasets, they demonstrated that these new distributions provided a better fit to the data when compared to the baseline distributions. Yahaya and Ieren [12] proposed the odd generalized exponential Gumbel distribution (OGEGD) for modeling lifetime data, Salem and selim [13] proposed Generalized Weibull-Exponential Distribution (GWED), the power Exponentiated family was proposed by Modi [14], Uwadi, et al. [15] introduced and studied the Exponentiated Gumbel Exponential (EGuE) distribution, and studied its mathematical and statistical properties. Modi, et al. [16], in their work proposed and studied a new family of distributions called Modi Exponential distribution, and applied it to two Real Datasets. In this study, we introduce the Modi Exponentiated Exponential distribution, which comprises four parameters. This distribution can be effectively used for fitting and analyzing data in various fields.

This paper is structured as follows: Section 2 provides a definition of the Modi generator, while Section 3 introduces the Exponentiated Exponential distribution. In Section 4, we present the Modi Exponentiated Exponential distribution, and develop its cumulative distribution function (CDF), probability density function (PDF), hazard rate function, and survival function. We also derive some statistical properties and mathematical properties of the proposed distribution. The estimation of the model parameters is given in Section 5. Simulation of study, Application to real data set, and Conclusion were given respectively in Sections 6, 7, and 8.

2. Modi Family

Modi, *et al.* [16] proposed and studied Modi family of distributions which is flexible and can be used to model a wide range of phenomena in various fields, including engineering, economics, and finance. The CDF, F(x) and PDF, f(x) of the Modi family are respectively, given by:

$$F(x) = \frac{\left(1 + \alpha^{\beta}\right)S(x)}{\alpha^{\beta} + S(x)}$$
(2.1)

$$f(x) = \frac{\left(1 + \alpha^{\beta}\right)\left(\alpha^{\beta}s(x)\right)}{\left[\alpha^{\beta} + S(x)\right]^{2}}$$
(2.2)

for all x > 0, where $\alpha > 0$, $\beta > 0$ and S(x) is the CDF of the existing distribution, and s(x) is the PDF of the existing distribution.

3. Exponentiated Exponential Distribution

The practice of adding a new parameter to an existing family of distribution functions is a common one in statistical distribution theory.

Gupta and Kundu [2] proposed Exponentiated Exponential (EE) distribution. The CDF and PDF of EE are given respectively by:

$$F(x,\delta,\lambda) = \left(1 - e^{-\lambda x}\right)^{\delta}$$
(3.1)

where $\delta, \lambda, x > 0$;

$$f(x,\delta,\lambda) = \delta\lambda \left(1 - e^{-\lambda x}\right)^{\delta-1} e^{-\lambda x}$$
(3.2)

where λ is scale parameter, and δ is a shape parameter.

4. Modi Exponentiated Exponential (MEE) Distribution

In the field of statistical distribution theory, it is common to enhance the flexibility of a class of distribution functions by introducing an additional parameter. This practice can be very useful for data analysis, as it allows for greater versatility in modeling various types of data.

4.1. Cumulative Distribution Function and Survival Function

From the Equation 2.1 and Equation 3.1 the cumulative distribution F(x) of the MEE distribution is given as:

$$F(x,\delta,\lambda,\beta,\alpha) = \frac{\left(1+\alpha^{\beta}\right)\left(1-e^{-\lambda x}\right)^{\delta}}{\alpha^{\beta}+\left(1-e^{-\lambda x}\right)^{\delta}}$$
(4.1)

for all x > 0 and $\delta, \lambda, \beta, \alpha > 0$,

and the survival function is derived from Equations (2.2) as follows:

$$S(x,\delta,\lambda,\beta,\alpha) = 1 - F(x,\delta,\lambda,\beta,\alpha)$$

$$= 1 - \frac{(1+\alpha^{\beta})(1-e^{-\lambda x})^{\delta}}{\alpha^{\beta} + (1-e^{-\lambda x})^{\delta}}$$

$$= \frac{\alpha^{\beta} \left[1 - (1-e^{-\lambda x})^{\delta}\right]}{\alpha^{\beta} + (1-e^{-\lambda x})^{\delta}}$$
(4.2)

for all x > 0 and $\delta, \lambda, \beta, \alpha > 0$.

4.2. Probability Density Function and Hazard Rate Function

The PDF plays a crucial role in the modeling and analysis of continuous random variables.

The PDF of MEE is obtained from Equation 4.1 as follows:

$$f(x,\delta,\lambda,\beta,\alpha) = \frac{d}{dx}F(x,\delta,\lambda,\beta,\alpha)$$

$$= \frac{d}{dx}\frac{(1+\alpha^{\beta})(1-e^{-\lambda x})^{\delta}}{\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}}$$

$$= \frac{\lambda\delta\alpha^{\beta}(\alpha^{\beta}+1)(1-e^{-\lambda x})^{\delta}}{(e^{\lambda x}-1)[\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}]^{2}}$$
(4.3)

for all x > 0 and $\delta, \lambda, \beta, \alpha > 0$.

Figure 1 depicts several potential forms of the PDF for the MEE distribution across different parameter values. The MEE PDF has the potential to exhibit symmetry or unimodal as well as a reversed-*J* shape.

Parameter values were derived through the technique of Sensitivity Analysis, a valuable technique that enhances the understanding of input-variable-to-modeloutput connections. It offers insights into uncertainties and empowers decision-makers to make well-informed and resilient choices.

The hazard rate function is obtained by using Equations (4.2) and (4.3).

$$h(x,\delta,\lambda,\beta,\alpha) = \frac{f(x,\delta,\lambda,\beta,\alpha)}{S(x,\delta,\lambda,\beta,\alpha)}$$

$$= \frac{\lambda\delta(\alpha^{\beta}+1)(1-e^{-\lambda x})^{\delta}}{(e^{\lambda x}-1)[\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}][1-(1-e^{-\lambda x})^{\delta}]}$$
(4.4)

where λ is a scale parameter and α, β, δ are shapes parameters, for all x > 0, and $\alpha, \beta, \delta, \lambda > 0$.

The shapes of hazard rate function for the MEE are depicted in **Figure 2**. It demonstrates that the hazard rate function can take various shapes such as increasing, decreasing, constant, inverted bathtub, etc.



Figure 1. Plot of PDF of the MEE for various values of δ, λ, β and α .



Figure 2. Plot of hazard rate function of the MEE for various values of δ, λ, β and α .

4.3. Statisical and Mathematical Properties of the MEE Distribution

4.3.1. Quantile Function

The quantile function is crucial in generating random samples from a specified distribution.

Suppose that X is the random variable where $X \sim \text{MEE}(\lambda, \delta, \alpha, \beta)$, so to obtain the quantile function, one can solve an equation for x by using the Cumulative Distribution Function (CDF) of the Modi Exponentiated Exponential (MEE) as defined in Equation (4.1):

Then,

Let F(x) = u

$$u = \frac{\left(1 + \alpha^{\beta}\right)\left(1 - e^{-\lambda x}\right)^{\delta}}{\alpha^{\beta} + \left(1 - e^{-\lambda x}\right)^{\delta}},$$

$$\Rightarrow \frac{u}{1 + \alpha^{\beta}} = \frac{\left(1 - e^{-\lambda x}\right)^{\delta}}{\alpha^{\beta} + \left(1 - e^{-\lambda x}\right)^{\delta}},$$

$$(4.5)$$

$$\Rightarrow Q(u) = x_{u} = F^{-1}(u; \lambda; \delta; \alpha; \beta) = \frac{-\log\left[1 - \left(\frac{u\alpha^{\beta}}{1 + \alpha^{\beta} - u}\right)^{\frac{1}{\delta}}\right]}{\lambda}$$

where $F^{-1}(.)$ is the inverse of the distribution function of MEE and $0 \le u \le 1$. To obtain the values of the lower quartile, median, and upper quartile, one can use the quantile function by replacing *u* with 1/4, 1/2, and 3/4, respectively, giving:

The lower quartile as:

$$=\frac{-\log\left[1-\left(\frac{\alpha^{\beta}}{4\alpha^{\beta}+3}\right)^{\frac{1}{\delta}}\right]}{\lambda}$$
(4.6)

The median as:

$$=\frac{-\log\left[1-\left(\frac{\alpha^{\beta}}{2\alpha^{\beta}+1}\right)^{\frac{1}{\beta}}\right]}{\lambda}$$
(4.7)

The upper quartile as:

$$=\frac{-\log\left[1-\left(\frac{3\alpha^{\beta}}{4\alpha^{\beta}+1}\right)^{\frac{1}{\delta}}\right]}{\lambda}$$
(4.8)

4.3.2. Skewness and Kurtosis

Mathematically, the Moors Kurtosis and Galton skewness of the MEE distribution are stated as:

$$K_{M} = \frac{Q(7/8) + Q(3/8) - Q(5/8) - Q(1/8)}{Q(6/8) - Q(2/8)}$$

$$S_{K} = \frac{Q(3/4) + Q(1/4) - 2Q(2/4)}{Q(3/4) - Q(1/4)}$$
(4.9)

where the quartiles and octiles value is indicated by Q(.)

Table 1 provides information about how quantiles of the MEE distribution vary for different combinations of parameter values. It can be used to look up specific quantile values based on the desired parameters, which can be useful for statistical analysis and modeling.

4.3.3. The rth Moments of MEE Distribution

Calculation of the moments of a distribution is crucial for statistical analysis, especially in practical applications. Moments are used to determine various statistical measures such as measures of central tendency, dispersion, and shape.

The mathematical expression for the r^{th} moment is given by

$$\mu_r' = E\left(X^r\right) = \int_0^\infty x^r f\left(x\right) \mathrm{d}x \tag{4.10}$$

where f(x) is the PDF of the distribution.

By substituting (4.3) in (4.10)we get

$$\mu_r' = E\left(X^r\right) = \int_0^\infty x^r \frac{\lambda \delta \alpha^\beta \left(\alpha^\beta + 1\right) \left(1 - e^{-\lambda x}\right)^\delta}{\left(e^{\lambda x} - 1\right) \left[\alpha^\beta + \left(1 - e^{-\lambda x}\right)^\delta\right]^2} dx$$

By using binomial expansion, we get

$$\mu_r' = -\delta\alpha^{\beta} \left(\alpha^{\beta} + 1\right) \sum_{n=0}^{\infty} {\delta \choose n} (-1)^n \sum_{k=0}^{\infty} {-1 \choose k} (-1)^k \frac{1}{n-k} \sum_{p=0}^{\infty} {-2 \choose p} \left(\alpha^{\beta}\right)^{-2-p} \times \sum_{j=0}^{\infty} (-1)^j {\delta p \choose j} \int_0^\infty x^r \left(e^{-\lambda x}\right)^j dx$$

By using gamma function the integral becomes:

$$\int_0^\infty x^r \left(e^{-\lambda x} \right)^j dx = \frac{\Gamma(r+1)}{(\lambda n)^{r+1}},$$

since

$$\int_0^\infty t^{p-1} \mathrm{e}^{-qt} \mathrm{d}t = \frac{\Gamma p}{q^p}$$

Comparing with this, we can identify *t* as *x*, *p* as (r + 1), and *q* as λj . Finally,

$$\mu_{r}' = E\left(X^{r}\right) = -\delta\alpha^{\beta}\left(\alpha^{\beta}+1\right)\sum_{n=0}^{\infty} {\delta \choose n} (-1)^{n} \sum_{k=0}^{\infty} {-1 \choose k} (-1)^{k} \frac{1}{n-k} \sum_{p=0}^{\infty} {-2 \choose p} \left(\alpha^{\beta}\right)^{-2-p} \times \sum_{j=0}^{\infty} (-1)^{j} {\delta p \choose j} \frac{\Gamma\left(r+1\right)}{\left(\lambda j\right)^{r+1}}$$

$$(4.11)$$

Table 2 illustrates the adaptability of the MEE distribution concerning its mean and variance. The Coefficient of Skewness (CS) values indicate its potential to exhibit right, or near symmetrical skewness. Similarly, the Coefficient of Kurtosis (CK) values suggest that the MEE distribution can display mesokurtic, leptokurtic, or platykurtic traits. These attributes highlight the MEE distribution's versatility, making it an attractive choice for modeling needs.

$(\lambda,\delta,lpha,eta)$							
Quantiles	(1.5, 0.5, 0.5, 0.5)	(5, 1.5, 1.5, 1)	(4, 4, 2.5, 2)	(3, 2, 3, 1.5)	(5, 3, 2, 1)		
0.1	0.006668031	0.1428966	0.1233738	0.2899888	0.1146931		
0.2	0.029508056	0.2471598	0.1809695	0.4076942	0.1788438		
0.3	0.074244174	0.3524963	0.2336858	0.5133350	0.2398289		
0.4	0.149625137	0.4665985	0.2872349	0.6197005	0.3034492		
0.5	0.269490678	0.5963168	0.3451618	0.7343052	0.3738184		
0.6	0.457717346	0.7509293	0.4114854	0.8653468	0.4559811		
0.7	0.760043256	0.9468080	0.4926415	1.0257089	0.5582364		
0.8	1.280845704	1.2197961	0.6021814	1.2426781	0.6986824		
0.9	2.357441470	1.6833751	0.7828494	1.6016888	0.9341384		

Table 1. Quantiles of the MEE distribution for some parameter values.

Table 2. The first five moments, skewness, and kurtosis of the MEE distribution for different parameter values.

		$(\lambda,\delta,lpha,eta$)		
Moments	(1.5, 1.5, 1.5, 0.3)	(1.5, 2.5, 0.6, 1)	(1.5, 1.5, 1.3, 2.3)	(1.5, 1.5, 1.5, 0.94)
$\mu_{\scriptscriptstyle 1}$	0.32011916	0.16205509	0.31214468	0.32319294
μ_{2}	0.18122137	0.06914166	0.17975234	0.18634641
$\mu_{_3}$	0.12208139	0.04086986	0.12274569	0.12687357
$\mu_{_4}$	0.09078026	0.02825694	0.09215417	0.09496380
μ_5	0.07180764	0.02135300	0.07339230	0.07544089
SD	0.28061557	0.20707441	0.28691119	0.28616907
CV	0.87659723	1.27780250	0.91916092	0.88544342
CS	0.61787869	1.77573677	0.64555447	0.38516391
СК	2.31891060	5.75953651	3.08746362	2.23680187

4.3.4. Order Statistics

Let's consider a finite random sample X_1, X_2, \dots, X_n drawn from a probability density function. From this sample, we define the f^{th} order statistic, X(j), where X(1) represents the smallest value in the sample, X(2) represents the second smallest value, and so on, up to X(n) representing the largest value. The Probability Density Function (PDF) of the f^{th} order statistic, with $1 \le j \le n$, can be expressed as follows:

$$g_{j:n}(x) = \frac{n!}{(j-1)!(n-j)!} f(x) F^{j-1}(x) (1-F(x))^{n-j}$$
(4.12)

by substituting Equations (4.3) and (4.2) in (5.8), the PDF of the f^{h} order statistic for the MEEE can be expressed as follows:

$$g_{j;n}(x) = \frac{n!\lambda\delta\alpha^{\beta}(\alpha^{\beta}+1)(1-e^{-\lambda x})^{\delta}}{(j-1)!(n-j)!(e^{\lambda x}-1)\left[\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}\right]^{2}} \times \left[\frac{(1+\alpha^{\beta})(1-e^{-\lambda x})^{\delta}}{\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}}\right]^{j-1} \left(1-\frac{(1+\alpha^{\beta})(1-e^{-\lambda x})^{\delta}}{\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}}\right)^{n-j} = \frac{n!\lambda\delta\alpha^{\beta}}{(j-1)!(n-j)!(e^{\lambda x}-1)\left[\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}\right]} \times \left(\frac{(1+\alpha^{\beta})(1-\alpha^{\beta})^{\delta}}{\alpha^{\beta}(1-(1-e^{-\lambda x})^{\delta})}\right)^{j} \left(\frac{\alpha^{\beta}(1-(1-e^{-\lambda x})^{\delta})}{\alpha^{\beta}+(1-e^{-\lambda x})^{\delta}}\right)^{n}$$

$$(4.13)$$

The PDF of smallest order statistic for the MEE occurs when the value of *j* is 1, and is given by

$$f_{1:n}(x) = \frac{n\lambda\delta(1+\alpha^{\beta})(1-\alpha^{\beta})^{\delta}}{\left(e^{\lambda x}-1\right)\left[\alpha^{\beta}+\left(1-e^{-\lambda x}\right)^{\delta}\right]\left(1-\left(1-e^{-\lambda x}\right)^{\delta}\right)}\left[\frac{\alpha^{\beta}\left(1-\left(1-e^{-\lambda x}\right)^{\delta}\right)}{\alpha^{\beta}+\left(1-e^{-\lambda x}\right)^{\delta}}\right]^{n} (4.14)$$

The PDF of largest order statistic for the MEE can be found when the index *j* is equal to *n*, and its PDF is given by

$$f_{n:n}(x) = \frac{n\lambda\delta\alpha^{\beta}}{\left(e^{\lambda x} - 1\right)\left[\alpha^{\beta} + \left(1 - e^{-\lambda x}\right)^{\delta}\right]} \left[\frac{\left(\alpha^{\beta} + 1\right)\left(1 - \alpha^{\beta}\right)^{\delta}}{\alpha^{\beta} + \left(1 - e^{-\lambda x}\right)^{\delta}}\right]^{n}$$
(4.15)

5. Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation (MLE) is a method of finding the values of the parameters that maximize the likelihood function. If we have *n* values x_1, x_2, \dots, x_n that are randomly selected from the MEE distribution, the log-likelihood function denoted by $l(\varphi)$ is given by:

$$l(\varphi) = \log L(\varphi) = \log \prod_{i=1}^{n} f(x_i; \varphi) = \sum_{i=1}^{n} \log f(x_i; \varphi)$$
(5.1)

where $\varphi = \alpha, \beta, \lambda, \delta$

Replacing Equation 4.3 in 5.1 to calculate the log-likelihood function that is associated with these values, we get

$$\sum_{i=1}^{n} \log f\left(x_{i};\lambda,\delta,\alpha,\beta\right) = \sum_{i=1}^{n} \log \frac{\lambda \delta \alpha^{\beta} \left(\alpha^{\beta}+1\right) \left(1-e^{-\lambda x_{i}}\right)^{\delta}}{\left(e^{\lambda x_{i}}-1\right) \left[\alpha^{\beta}+\left(1-e^{-\lambda x_{i}}\right)^{\delta}\right]^{2}}$$
(5.2)

$$\Rightarrow \log L(\varphi) = n \log \lambda + n \log \delta + n\beta \log \alpha + n \log (\alpha^{\beta} + 1) + \delta \sum_{i=1}^{n} \log (1 - e^{-\lambda x_i})$$

$$- \sum_{i=1}^{n} \log (e^{\lambda x_i} - 1) - 2 \sum_{i=1}^{n} \log \left[\alpha^{\beta} + (1 - e^{-\lambda x_i})^{\delta} \right]$$
(5.3)

Differentiating 5.3 partially with respect to each parameter and equating to

zero gives

$$\frac{\partial}{\partial \alpha} \log L(\varphi) = \frac{n\beta}{\alpha} + \frac{n\beta\alpha^{\beta-1}}{\alpha^{\beta} + 1} - 2\sum_{i=1}^{n} \frac{\beta n\alpha^{\beta-1}}{\alpha^{\beta} + (1 - e^{-\lambda x_i})^{\delta}} = 0$$
(5.4)

$$\frac{\partial}{\partial\beta}\log L(\varphi) = n\log\alpha + \alpha + \frac{\beta\alpha^{\beta-1}}{\alpha^{\beta}+1} - 2\sum_{i=1}^{n} \frac{2\beta\alpha^{\beta-1}}{\alpha^{\beta} + (1 - e^{-\lambda x_i})^{\delta}} = 0$$
(5.5)

$$\frac{\partial}{\partial\lambda}\log L(\varphi) = \frac{n}{\lambda} + \delta \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} - \sum_{i=1}^{n} \frac{x_i e^{\lambda - x_i}}{1 - e^{-\lambda x_i}} - \sum_{i=1}^{n} \frac{x_i e^{\lambda x_i}}{e^{\lambda x_i} - 1} - 2\sum_{i=1}^{n} \frac{\delta \left(1 - e^{-\lambda x_i}\right)^{\delta^{-1}} \lambda e^{-\lambda x_i}}{\alpha^{\beta} + \left(1 - e^{\lambda x_i}\right)^{\delta}} = 0$$
(5.6)

$$\frac{\partial}{\partial\delta}\log L(\varphi) = \frac{n}{\delta} + \sum_{i=1}^{n}\log\left(1 - e^{-\lambda x_{i}}\right) - 2\sum_{i=1}^{n}\frac{\delta\left(1 - e^{-\lambda x_{i}}\right)^{\delta-1}\lambda e^{-\lambda x_{i}}}{\alpha^{\beta} + \left(1 - e^{-\lambda x_{i}}\right)^{\delta}} = 0 \quad (5.7)$$

Based on our observations, it is evident that Equations 5.4 - 5.7 can not be solved analytically, meaning there is no direct mathematical solution for them. Consequently, we need to resort to a numerical optimization technique to find their solutions. In this study, the Broyden-Fletcher-Goldfarb-Shannon (BFGS) algorithm was employed to estimate the parameters of the MEE distribution.

6. Simulation Study

This section describes a Monte Carlo simulation study conducted to examine the behavior of Maximum Likelihood Estimators (MLEs) for the parameters of the MEE model. A simulation study was conducted to examine the accuracy of the Maximum Likelihood Estimators (MLEs) for four model parameters (α , β , δ , and λ) in terms of their average biases (ABs) and root mean squared errors (RMSEs). In order to obtain random samples from the MEE, the inverse of the CDF presented in Equation 4.5 was utilized. To accomplish this, we generated 1000 replications for each of the sample sizes

n = 50,75,100,125,200,275,600,625,650, and n = 700 using Equation 4.5 for various combinations of parameter values (α , β , λ , δ).

The parameter values were provided in two different sets.

Set I: $\lambda = 0.3$, $\delta = 0.5$, $\alpha = 1.0$, $\beta = 0.07$

Set II: $\lambda = 0.52$, $\delta = 0.47$, $\alpha = 0.97$, $\beta = 0.065$

The average biases were computed by:

$$AB_{(\eta)} = \frac{1}{R} \sum_{i=1}^{R} (\hat{\eta}_i - \eta)$$
(6.1)

and the root mean square errors by:

$$RMSE_{(\eta)} = \sqrt{\frac{1}{R} \sum_{i=1}^{R} (\hat{\eta}_{i} - \eta)^{2}}$$
(6.2)

where η is the parameter in question, while $\hat{\eta}_i$ is its estimated value at the *t*^h replication at each sample size, and *R* is the total number of replication.

The following table presents the MLEs, ABs, and RMSEs values for different

sample sizes, corresponding to the parameters, λ , δ , α , β .

Table 3 presents the simulation results for Maximum Likelihood Estimators (MLEs), Average Biases (ABs), and Root Mean Square Errors (RMSEs) for the parameter values in Set I and set II. Analyzing **Table 3** in set I, it can be observed that as the sample size increases, the MLEs approach the true parameter values for the MEE distribution. Generally, the ABs, and RMSEs of the parameter estimators decrease as the sample size increases.

Table 3, on the other hand, displays the simulation results for the parameter values in Set II.

Indeed, the simulation results for the MEE distribution in **Table 3** for set II demonstrate a similar pattern. As the sample size increases, the MLEs become closer to the true values of the parameters. Additionally, the Average Biases (ABs), and Root Mean Square Errors (RMSEs) for the parameter estimators tend to decrease, indicating improved accuracy, and precision as the sample size increases.

Sensitivity Analysis methodology was used to determine parameter values, and parameter estimations were carried out through Monte Carlo simulation using R software.

7. Application to Real Data Set

In this section, we fitted the Modi Exponentiated Exponential distribution to two real data sets and observed its flexibility compared to other well-known distributions. The analysis was conducted using R software. We calculated values of various information criteria such as Akaike Information Criterion (AIC), Hannan Quin Information Criterion (HQIC), Bayesian Information Criterion (BIC), and Consistent Akaike Information Criterion (CAIC). Additionally, we performed Kolmogorov-Smirnov (K-S), Cramér-von Mises test (W*), and Anderson-Darling test (A*) to assess the goodness of fit for the considered distributions. The distribution with the highest log-likelihood and the highest p-value for the K-S test and the lowest AIC, BIC, HQIC, CAIC, W*, A*, K-S was considered the best.

The PDFs of the existing distributions compared with the Modi Exponentiated Exponential distribution are presented in Table 4.

Data set I: This data set presented in [17] contains information on the remission times, measured in months, for a group of 128 individuals diagnosed with bladder cancer. 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.5, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.00, 3.36, 6.93, 8.65, 12.63, and 22.69.

	$(\lambda, \delta, \alpha, \beta) = (0.3, 0.5, 1.0, 0.07)$ $(\lambda, \delta, \alpha, \beta) = (0.52, 0.47, 0.95, 0.12)$					2,0.47,0.95,0.0	65)	
	Set I					Se	et II	
	MLE					Ν	ILE	
п	â	$\hat{\delta}$	â	β	â	$\hat{\delta}$	â	β
50	0.3196859	0.5409758	3.4886202	0.9259900	0.5440354	0.5126671	2.8060344	0.7634655
75	0.3112706	0.5293138	2.7584619	0.6064322	0.5318997	0.5023858	1.9331472	0.4791677
100	0.3055520	0.5267967	2.0916017	0.4039657	0.5240320	0.4933882	1.4257057	0.33972273
125	0.3026006	0.5183084	1.8401000	0.3124900	0.0072818	0.0022034	5.2892570	0.3734995
200	0.3012543	0.5139124	1.3534001	0.1925341	0.5176685	0.4800144	1.0008279	0.1463059
275	0.3003015	0.5090554	1.1222609	0.1374888	0.5176876	0.4748563	0.9802502	0.1052886
600	0.2995486	0.5019489	1.0040428	0.0774720	0.5199701	0.4703388	0.9772399	0.0673495
625	0.2994533	0.5017486	1.0015491	0.0765860	0.5199667	0.4700723	0.9693667	0.0655118
650	0.2992176	0.5014539	0.9831785	0.0765069	0.5198862	0.4700318	0.958920	0.0654146
700	0.2992179	0.5012098	0.9885469	0.0740976	0.519817	0.4700174	0.9496249	0.0653499
		A	В			1	AB	
n	λ	$\hat{\delta}$	\hat{lpha}	\hat{eta}	λ	$\hat{\delta}$	\hat{lpha}	β
50	0.0196859	0.0409758	2.4886202	0.8559900	0.0240354	0.0426671	1.8560344	0.6984655
75	0.0112706	0.0293138	1.7584619	0.5364322	0.0118997	0.0323858	0.9831472	0.4141677
100	0.0055520	0.0267967	1.0916017	0.3339657	0.0040320	0.0233882	0.4757057	0.2747227
125	0.0026006	0.0183084	0.8401000	0.2424900	0.0017034	0.0167820	0.3818860	0.1997395
200	0.0012543	0.0139124	0.3534000	0.1225341	0.5176685	0.4800144	1.0008279	0.1463059
275	0.0003015	0.0090554	0.1222609	0.0674888	0.5176876	0.4748563	0.9802502	0.1052886
600	-0.0004514	0.0019489	0.0040428	0.0074728	-0.0000299	0.0003388	-0.0027601	0.0023495
625	-0.0004671	0.0017486	0.0015491	0.0065860	-0.0000333	0.0000723	-0.0046333	0.0005118
650	-0.0007824	0.0014539	-0.0108215	0.0065069	-0.0000414	0.0000132	-0.0050796	0.0004146
700	-0.0008826	0.0012098	-0.0114531	0.0040976	-0.0000674	0.0000047	-0.0073751	0.0003499
		RM	ISE			RI	MSE	
п	â	$\hat{\delta}$	â	\hat{eta}	â	$\hat{\delta}$	â	β
50	0.1284746	0.1122043	7.5995099	1.9944528	0.2077717	0.0998491	5.7092375	1.7251820
75	0.0896363	0.0872377	6.2949962	1.4750860	0.1327728	0.0783012	3.9018322	1.1582821
100	0.0761222	0.0739686	4.5863408	0.9805357	0.0109710	0.0036614	4.7063140	0.8058986
125	0.0624115	0.0586510	4.0503263	0.7705012	0.0853335	0.0469405	2.2998384	0.6111460
200	0.0412558	0.0443379	2.8825240	0.4568065	0.0455862	0.0338128	0.7020969	0.225840
275	0.0313782	0.0295924	1.1509773	0.2422337	0.0316109	0.0205347	0.3285671	0.1467336
600	0.0093709	0.0112332	0.3535583	0.0395004	0.0021541	0.0045906	0.0362593	0.0307185
625	0.0090843	0.0105629	0.3110703	0.0335318	0.0010524	0.0022868	0.0270256	0.0161850
650	0.0088504	0.0028265	0.2978253	0.0325252	0.0005547	0.0021469	0.0243816	0.0146910
700	0.0063642	0.0016918	0.1926575	0.0240313	0.0005490	0.0014986	0.0118613	0.0110662

Table 3. The outcomes of a Monte Carlo simulation, investigation for the parameters in set I and set II.

DOI: 10.4236/jdaip.2023.114017

352

Journal of Data Analysis and Information Processing

Distributions	PDFs
EE	$f(x) = \delta \lambda \left(1 - e^{-\lambda x} \right)^{\delta - 1} e^{-\lambda x}, x > 0$
ME	$f(x) = \frac{\lambda \alpha^{\beta} e^{-\lambda x}}{\left(1 + \alpha^{\beta}\right) \left(1 - \frac{e^{-\lambda x}}{1 + \alpha^{\beta}}\right)^{2}}, x > 0$
KE	$f(x) = \alpha \beta \lambda e^{-\lambda x} \left(1 - e^{-\lambda x}\right)^{\alpha - 1} \left(1 - \left[1 - e^{-\lambda x}\right]^{\alpha}\right)^{\beta - 1}, x > 0$
E	$f(x) = \lambda e^{-\lambda x}, x > 0$

 Table 4. The existing distributions compared with Modi Exponentiated Exponential distribution.

Table 5 presents the summary characteristics of data set I. The data is skewed to the right, as indicated by a skewness coefficient of 3.325, and it exhibits significant tailing in its distribution, with a kurtosis coefficient of 16.15.

The parameter known as the Kurtosis coefficient was obtained from the data set using R software through descriptive statistics.

In **Figure 3**, the histogram of the data displays a right-skewed and the presence of outliers is effectively revealed by the box plot.

In **Figure 4**, the TTT plot of the data set, it can be observed that the hazard rate function is an inverted bathtub shape while the violin plot highlights that the majority of values are concentrated around the median.

Table 6 and **Table 7** provide the AIC, HQIC, BIC, and CAIC values, along with the K-S, W*, and A* tests. Based on these results, the MEE model emerges as the most favorable choice because it has the lowest values for AIC, HQIC, BIC, CAIC, W*, K-S and A*, indicating better goodness-of-fit. Additionally, it exhibits the highest p-value for the K-S statistic and log-likelihood function value, further supporting its superiority.

Figure 5 illustrates a plot of fitted densities, comparing the MEE distribution to its sub-models using the bladder cancer data set. The plot reveals that the MEE distribution demonstrates a favorable and encouraging fit when compared to the existing distributions.

Data set II: This data set consists of the waiting times (in minutes) of one hundred bank customers before they receive service. This data set has been previously analyzed by Ghitany, *et al.* [18]. They have fitted both the Lindley distribution and the exponential distribution to this data. The data set is provided below:

0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11, 11, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2,, 18.4, 18.9, 19, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27, 31.6, 33.1, 38.5.

Table 8 presents the summary characteristics of data set II. The data is skewed to the right, as indicated by a skewness coefficient of 1.47277, and it can be



Table 5. Comprehensive overview of the bladder cancer data set: Descriptive analysis.

Figure 4. TTT and violin plots of bladder cancer data set.

platykurtic based on kurtosis coeficient of 2.54029. The parameter known as the Kurtosis coefficient was obtained from the data set using descriptive statistics in the R software.

In the **Figure 6**, the TTT plot of the data set shows that the hazard rate function is on the rise, indicating an increasing shape and the histogram illustrates right skewed of the data.

Distributions	Estimates	Log-Likelihood	K-S	p-Value	A*	W*
	$\hat{\lambda} = 0.07099569$	-410.34	0.033238	0.9989	0.2364337	0.03395614
MEE	$\hat{\delta}$ = 1.64959239					
MEE	$\hat{\alpha} = 0.02692251$					
	$\hat{\beta} = 0.34715153$					
	$\hat{\lambda} = 0.121140$	-413.07	0.072453	0.5124	0.6706381	0.1116195
EE	$\hat{\delta} = 1.217358$					
	$\hat{\lambda} = 0.105736$	-414.34	0.22927	0.4642	0.5345996	0.08818205
ME	$\hat{\alpha} = 3.592623$					
	$\hat{\beta} = 3.116176$					
	$\hat{\theta} = 1.4189689$	-412.4623	0.06904403	0.5751033	0.5594285	0.09164695
KE	$\hat{\alpha} = 3.592623$					
	$\hat{\gamma} = 3.116176$					
Е	$\hat{\lambda} = 0.1067831$	-417.83	0.082926	0.3422	0.7123287	0.1186696

Table 6. Maximum Likelihood Estimates and goodness-of-fit tests for data set I.

 Table 7. A summary of the results from the information criteria analysis conducted on data set I.

Distributions	AIC	BIC	CAIC	HQIC
MEE	828.6901	834.0982	829.0153	831.3253
EE	830.1486	835.8526	830.2446	832.4662
ME	834.689	843.2451	834.8825	838.1654
KE	830.9246	839.4807	831.1182	834.401
E	835.6646	839.5166	836.6963	839.8234



Figure 5. The densities of the bladder cancer data set estimated using different distribution models.



Table 8. Investigation of customer waiting duration for bank services: A descriptive analysis.

Figure 6. TTT and histogram plots of waiting time data set.

In the **Figure 7**, the violin plot emphasizes that most values are centered around the median. The box plot effectively identifies the presence of outliers in the data.

The MEE model is considered the best model based on the information provided in **Table 9** and **Table 10**. This is because it has the lowest values for AIC, HQIC, BIC, W*, A* and CAIC, indicating better model fit. Additionally, it has the highest p-value for the K-S statistic and log-likelihood function value, further supporting its superiority compared to other models.

Figure 8 illustrates a plot of fitted densities, comparing the MEE distribution to its sub-models using the waiting time data set. The plot reveals that the MEE distribution demonstrates a favorable and encouraging fit when compared to the existing distributions.

8. Conclusion

In this paper, we introduced a new four-parameter model called the Modi Exponentiated Exponential (MEE) distribution, and applied to two real data sets. We have examined the mathematical and statistical properties of this proposed distribution. We derived expressions for its t^{th} moment, survival function, hazard rate function, cumulative distribution, and quantile function. Furthermore, through various plots, we have observed that the MEE distribution exhibits different shapes, indicating its versatility in fitting data sets with diverse distributions. We have also obtained the Probability Density Function (PDF) of its minimum and maximum order statistics.





Distributions	Estimates	Log-Likelihood	K-S	p-Value	A*	W*
	$\hat{\lambda} = 0.1475045$	-313.514	0.03596	0.9995	0.1245139	0.01697144
MEE	$\hat{\delta} = 2.3053274$					
MEE	$\hat{\alpha} = 2.2326181$					
	$\hat{\beta} = 1.4412144$					
	$\hat{\lambda} = 0.1591537$	-317.1	0.040245	0.9969	0.1427685	0.02074274
EE	$\hat{\delta} = 2.1834307$					
	$\hat{\lambda} = 0.1012398$	-329.02	0.28139	0.6532	0.1351024	0.01919179
ME	$\hat{\alpha} = 4.2890999$					
	$\hat{\beta} = 5.7491249$					
	$\hat{\theta} = 0.6011951$	-317.0035	0.0361401	0.9994518	0.1261128	0.0179404
KE	$\hat{\alpha} = 2.5249861$					
	$\hat{\gamma} = 0.2448366$					
E	$\hat{\lambda} = 0.1012478$	-332.5209	0.37302	0.005022	0.1790281	0.02703896

Table 0 Maximum	likalihaad	actimation	and goodn	and of fit a	analyzia fa	r data	oot II
Table 9. Maximum	iikeimood	estimation	ana gooan	ess-01-111 a	anaiysis io	r data	set II.

 Table 10. A summary of the results from the information criteria analysis conducted on data set II.

Distributions	AIC	BIC	CAIC	HQIC
MEE	635.0292	642.4499	637.4502	638.2466
EE	638.1906	643.4009	638.3143	640.2993
ME	664.0467	671.8622	664.0467	667.2097
KE	640.007	647.8225	640.257	643.1701
Е	669.0826	672.6469	665.0826	667.0961



Figure 8. Histogram and fitted densities of the waiting time data set for different distributions.

To estimate the parameters of the MEE distribution, we employed the method of maximum likelihood estimation. Monte Carlo simulation was used to assess the performance of MLEs. The study observed that MLEs demonstrate good accuracy and consistent estimating of model parameters. As the sample size increases, MLEs tend to approach the true values of the parameters, as indicated by the decreasing ABs. Additionally, RMSEs also decrease with increasing sample size. Our analysis demonstrates that MEE distribution outperforms the existing distributions, in modeling the two data sets considered in this study.

Acknowledgements

The authors would like to acknowledge and thank the Pan African University, Institute for Basic Sciences, Technology, and Innovation (PAUSTI) for their support in conducting this study.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Mansoor, M., Tahir, M.H., Cordeiro, G.M., Provost, S.B. and Alzaatreh, A. (2019) The Marshall-Olkin Logistic-Exponential Distribution. *Communications in Statistics-Theory and Methods*, 48, 220-234. <u>https://doi.org/10.1080/03610926.2017.1414254</u>
- 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, 117-130. https://doi.org/10.1002/1521-4036(200102)43:1<117::AID-BIMJ117>3.0.C

<u>O;2-R</u>

- [3] Oguntunde, P., Adejumo, A., Okagbue, H. and Rastogi, M. (2016) Statistical Properties and Applications of a New Lindley Exponential Distribution. *Gazi University Journal of Science*, 29, 831-838.
- [4] Merovci, F. (2013) Transmuted Exponentiated Exponential Distribution. *Mathematical Sciences and Applications E-Notes*, 1, 112-122.
- [5] Gupta, R.D. and Kundu, D. (2007) Generalized Exponential Distribution: Existing Results and Some Recent Developments. *Journal of Statistical Planning and Inference*, 137, 3537-3547. <u>https://doi.org/10.1016/j.jspi.2007.03.030</u>
- [6] Algamal, Z.Y. (2008) Exponentiated Exponential Distribution as a Failure Time Distribution. *IRAQI Journal of Statistical Science*, 14, 63-75. https://doi.org/10.33899/igjoss.2008.31434
- [7] 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, 641-652. <u>https://doi.org/10.1093/biomet/84.3.641</u>
- [8] Margaretha, M., Fithriani, I. and Abdullah, S. (2020) Parameter Estimation of Exponentiated Exponential Distribution for Left Censored Data Using Bayesian Method. *AIP Conference Proceedings*, 2242, Article ID: 030029. https://doi.org/10.1063/5.0007892
- [9] 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, 6543-6557. https://doi.org/10.1080/03610926.2015.1130839
- [10] Kr Singh, R., Yadav, A.S., Singh, S.K. and Singh, U. (2016) Marshallolkin Extended Exponential Distribution: Different Method of Estimations. *Journal of Advanced Computing*, 5, 12-28. <u>https://doi.org/10.7726/jac.2016.1002</u>
- [11] Niyoyunguruza, A., Odongo, L.O., Nyarige, E., Habineza, A. and Muse, A.H. (2023) Marshall-Olkin Exponentiated Frechet Distribution. *Journal of Data Analysis and Information Processing*, **11**, 262-292. <u>https://doi.org/10.4236/jdaip.2023.113014</u>
- [12] Yahaya, A. and Ieren, T.G. (2017) On Odd Generalized Exponential Gumbel Distribution with Its Applications to Survival Data. *Journal of the Nigerian Association* of Mathematical Physics, **39**, 149-158.
- [13] Salem, H.M. and Selim, M.A. (2014) The Generalized Weibull-Exponential Distribution: Properties and Applications. *International Journal of Statistics and Applications*, 4, 102-112.
- [14] Modi, K. (2021) Power Exponentiated Family of Distributions with Application on Two Real-Life Datasets. *Thailand Statistician*, **19**, 536-546.
- [15] Uwadi, U.U., Okereke, E.W. and Omekara, C.O. (2019) Exponentiated Gumbel Exponential Distribution: Properties and Applications. *American Journal of Applied Mathematics and Statistics*, 7, 178-186.
- [16] Modi, K., Kumar, D. and Singh, Y. (2020) A New Family of Distribution with Application on Two Real Datasets on Survival Problem. *Science & Technology Asia*, 25, 1-10.
- [17] Lee, E.T. and Wang, J. (2003) Statistical Methods for Survival Data Analysis. Volume 476, John Wiley & Sons, Hoboken.
- [18] Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008) Lindley Distribution and Its Application. *Mathematics and Computers in Simulation*, 78, 493-506. <u>https://doi.org/10.1016/j.matcom.2007.06.007</u>