Application of the Modified Exponentiated Exponential (MEEXP) Distribution in Survival Analysis

Arun Kumar Chaudhary
Department of Statistics, Nepal Commerce Campus, Tribhuvan University, Nepal
Email: akchaudhary1@yahoo.com

Lal Babu Sah Telee
Department of Statistics, Nepal Commerce Campus, Tribhuvan University, Nepal
Email: lalbabu3131@gmail.com (Corresponding Author)

Vijay Kumar
Department of Mathematics and Statistics, DDU Gorakhpur University, Gorakhpur,India.
Email: vkgkp@rediffmail.com

https://doi.org/10.3126/pravaha.v30i1.76883

Abstract

The exponential distribution is widely used in statistical modeling, particularly for time-to-event analysis in various fields. However, it often fails to capture complex real-world data with non-constant hazard rates. To address this limitation, we studied the Modified Exponentiated Exponential (MEEXP) Distribution, an extension of the traditional exponential distribution that introduces additional shape parameters for greater flexibility. This model accommodates diverse hazard rate behaviors such as increasing, decreasing, or bathtub-shaped hazard functions. We use Maximum Likelihood Estimation (MLE) to estimate the parameters of the MEEXP distribution and apply it to a real-world survival analysis dataset from the COVID-19 pandemic in Mexico. Our findings indicate that the MEEXP model offers a superior fit compared to traditional models such as the exponential and normal distributions. Through goodness-of-fit tests, residual analysis, and model comparison, we demonstrate the superior performance of the MEEXP model in capturing the underlying patterns in the data. Furthermore, sensitivity analysis highlights the impact of the model's parameters on its predictions, making it a valuable tool for reliability and survival analysis.

Keywords: *Exponentiated Exponential (EEXP) Distribution, Exponential Distribution, Failure rate function, Quantile function, Maximum likelihood estimation and Survival Analysis.*

Cite this paper
Chaudhary, A. K., Sah Telee, L. B., & Kumar, V. (2024). Application of the Modified Exponentiated Exponential (MEEXP) Distribution in Survival Analysis. *Pravaha, 30*(1), 1-10.

Introduction

For many years, classical probability models have been utilized to analyze real-world data across diverse fields such as risk assessment, hydrology, biology, climatology, geology, engineering, , life testing and finance. Despite their widespread application, these models often fail to provide a precise representation of complex data. Consequently, there is a growing need to improve existing distributions to better address the challenges encountered in these domains. By modifying the baseline models or introducing new parameters, researchers can create more adaptable distributions that outperform traditional classical models. This process involves enriching the foundational models with additional parameters, such as shape or scale parameters, to better capture the intricacies of real-world phenomena. These improvements enable the models to better match the observed data. Developing advanced probability models by integrating additional parameters marks a notable progression in the

realm of statistical analysis.
The exponential distribution is a fundamental and commonly utilized probability distribution in statistical modeling, especially for analyzing time-to-event data. It belongs to the family of continuous probability distributions and is often applied in scenarios involving reliability analysis, queuing theory, and life testing. The exponential distribution, defined by the rate parameter λ>0, is used to represent the time intervals between events in a Poisson process, where events occur independently and at a constant average rate (Ross, 2014).

The exponential distribution's probability density function (PDF) is expressed as:

$$f(x,\lambda) = \lambda e^{-\lambda x}; x > 0, \lambda > 0 \tag{1}$$

The exponential distribution's corresponding probability distribution function (or CDF) is

$$F(x,\lambda) = 1 - e^{-\lambda x}; x > 0, \lambda > 0 \tag{2}$$

Despite its simplicity and mathematical tractability, the exponential distribution has limitations in capturing more complex real-world data patterns, such as varying hazard rates (i.e., increasing, decreasing, or non-monotonic hazards). These limitations have motivated researchers to extend thebaseline exponential distribution by introducing additional shapeparameters to improve its flexibility and applicability (Gupta & Kundu, 1999). Such modifications often result in generalized or extended versions of the exponential distribution, which can better accommodate diverse datasets. One notable extension is the exponentiated exponential distribution, which introduces a shape parameter to allow for greater flexibility in modeling data with non-constant hazard rates (Nadarajah & Kotz, 2011).
Other modifications of the exponential distribution are beta exponential (Nadarajah & Kotz, 2006), Lindley-Exponential (LE) distribution (Bhati et al., 2015), An Extended Weighted Exponential Distribution (Mahdavi & Jabari, 2017), Exponentiated Gamma Exponential Distribution (Jabeen, & Para, 2018), Logistic Inverse Exponential Distribution (Chaudhary &Kumar, 2020), Lomax exponential distribution (Ijaz et al., 2019) and Arctan exponential extension distribution (Chaudhary & Kumar, 2021).
In similar way, more modifications of the distributions are Modified generalized exponential distribution (Telee & Kumar, 2023),Inverse Exponentiated Odd Lomax Exponential Distribution(Chaudhary et al., 2022), Lindley Generalized Inverted Exponential Distribution(Telee & Kumar, 2021), On Beta Exponentiated Lomax-Exponential Distribution (Sóyínká, 2023), half-Cauchy modified exponential distribution (Chaudhary & Kumar, 2022), Extended Kumaraswamy Exponential Distribution (Chaudhary et al., 2023),The Modi Exponentiated Exponential Distribution (Ndayisaba, 2023) ,The T-Exponentiated Exponential {Frechet} Family of Distributions (Chinazom et al., 2023), half -Cauchy Exponential Distribution (Jayakumar & Fasna, 2023),Uniformly Shifted Exponential Distribution (Alzaid,& Qarmalah, 2024) and New Extended Kumaraswamy Exponential Distribution(Chaudhary et al., 2024).
This study introduces a new and versatile probability model named the Modified Exponentiated Exponential (MEEXP) Distribution. This model serves as an extension of the classical Exponential distribution and is utilized to analyze a real-world survival analysis dataset.
This article is structured as follows: First, we present the Modified Exponentiated Exponential (MEEXP) Distribution and describe its fundamental statistical properties. We then utilize the MLE approach to determine the model parameters. To validate the model, we perform statistical tests to assess its goodness of fit. Following this, we apply the proposed model to a real-world dataset to demonstrate its practical utility. Finally, we provide concluding remarks.

Modified Exponentiated Exponential (MEEXP) Distribution

In this study, we have applied Modified exponentiated exponential (MEEXP) model to analyze a survival real data set. The model MEEXP is defined by (Telee et al., 2024) to evaluate the impact of multi parameter space on reliability and validity study of the probability models. The model MEEXP has three parameters, α, β, and λ.

Equations (3) and (4) illustrate the cumulative distribution function (CDF) and the probability density function (PDF) of the MEEXP, respectively.

$$F(x,\alpha,\beta,\lambda)=1-e^{-\lambda x^{\alpha}e^{x\beta}};x>0,(\alpha,\lambda,\beta)>0 \tag{3}$$

$$f(x,\alpha,\beta,\lambda)=\lambda e^{\beta x}e^{-\lambda x^{\alpha}e^{x\beta}}\left(\alpha x^{\alpha-1}+x^{\alpha}\right);x>0,(\alpha,\lambda,\beta)>0 \tag{4}$$

Survival function

The MEEXP distribution's reliability function is

$$R(x,\alpha,\beta,\lambda)=e^{-\lambda x^{\alpha}e^{x\beta}};x>0,(\alpha,\lambda,\beta)>0 \tag{5}$$

Hazard rate function (HRF)

The MEEXPdistribution'sHRF is given by (6).

$$h(x)=\frac{f(x)}{R(x)}=\lambda e^{\beta x}\left(\alpha x^{\alpha-1}+x^{\alpha}\right);x>0,(\alpha,\lambda,\beta)>0 \tag{6}$$

The Reversed hazard function (RHR)

The reversed hazard rate function is expressed in Equation (7).

$$RHR(x)=\frac{f(x)}{F(x)}=\frac{\lambda e^{\beta x}e^{-\lambda x^{\alpha}e^{x\beta}}\left(\alpha x^{\alpha-1}+x^{\alpha}\right)}{1-e^{-\lambda x^{\alpha}e^{x\beta}}};x>0,(\alpha,\lambda,\beta)>0 \tag{7}$$

Cumulative hazard function (CHF)

The cumulative hazard function of the proposed model is given in Equation (8).

$$H(x)=-\log[R(x)]=-\log\left[e^{-\lambda x^{\alpha}e^{x\beta}}\right];x>0,(\alpha,\lambda,\beta)>0 \tag{8}$$

Quantile function (QF)

The quantile function of the MEEXP distribution is given in eq. (9),

$$Q_x(p)=\lambda x^{\alpha}e^{\lambda x}+\log(1-p)=0;\quad 0<p<1 \tag{9}$$

The generation of random deviates for the proposed model, as outlined in equation (9), is given by

$$Q_x(u)=\lambda x^{\alpha}e^{\lambda x}+\log(1-u)=0;\quad 0<u<1 \tag{10}$$

Where u denotes U (0, 1)'s uniform random variable.

Skewness and Kurtosis of MHCC distribution

The Bowley's coefficient of skewness is determined using quantiles and is calculated as follows:

$$S=\frac{Q_3-2Q_2+Q_1}{Q_3-Q_1}$$

The Moors' (1988) coefficient of kurtosis is defined in terms of octiles as

$$K_u(M)=\frac{Q(0.875)-Q(0.125)+Q(0.375)-Q(0.625)}{Q(3/4)-Q(1/4)}.$$

The pdf plots and hazard rate plots for different values of parameters are shown in figure 1. Hazard rate function of MEEXP is shown in figure 1. It is seen that there is more flexibility in pdf curves. The pdf curve of MEEXP is decreasing as well as positively skewed, approximately normal as well as negatively skewed in shape while the hazard rate curves are bathtub, inverted j and j shaped.

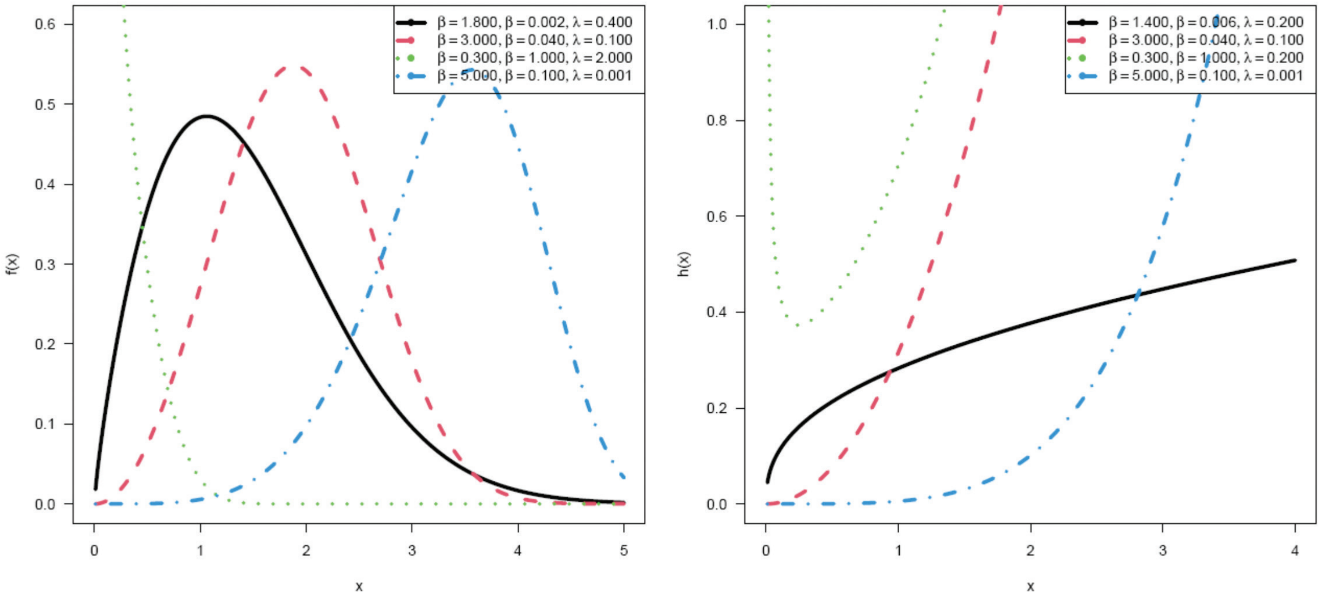

Fig 1.PDF and Hazard rate plots for MEEXP

Parameter estimation

In this part of the study, we performed parameter estimation for the proposed models. Common methods for parameter estimation include the Cramér-von Mises approach, Maximum Likelihood Estimation (MLE), and Least Squares Estimation. For this study, we have chosen to use the MLE approach exclusively.

Maximum Likelihood Estimation

The ML estimators (MLEs) for the MEEXP model, which are determined by the MLE technique, are presented below. For a random sample $\underline{x} = \left(x_1,\ldots,x_n\right)$ drawn from the MEEXP (α,β,λ), the estimators are obtained through the log-likelihood function.

$$l(\alpha,\beta,\lambda\,|\,\underline{x}) = n\ln\lambda + \beta\sum_{i=1}^{n} x - \lambda\sum_{i=1}^{n} \mathrm{x}^{\alpha} e^{x\beta} + \sum_{i=1}^{n} \ln(\alpha x^{\alpha-1} + x^{\alpha}) \tag{11}$$

In this case, the log-likelihood function is differentiated with respect to the unknown parameters α, β, and λ, and the resulting equations are solved to obtain the parameter estimates. Since manually estimating these parameters can be complex and time-consuming, we utilized R programming for the estimation process.

Real Data Analysis

A real dataset was used to evaluate the applicability of the proposed model. According to Bantan et al. (2023), the dataset includes the fatality rates of 106 patients in Mexico during the COVID-19 pandemic, which lasted from March 4, 2020, to July 20, 2020. For convenience, the rates have been scaled by dividing them by five. The dataset is as follows:

"1.7652, 1.2210, 1.8782, 2.9942, 2.0766, 1.4534, 2.6440, 3.2996, 2.3330, 1.2030, 2.1710, 1.2244, 1.3312, 0.6880, 1.1708, 2.1370, 2.0070, 1.0484, 0.8668, 1.0286, 1.5260, 2.9208, 1.5806, 1.2740, 0.7074, 1.2654, 0.9460, 0.6430, 1.8568, 2.5756, 1.7626, 2.0086, 1.4520, 1.1970, 1.2824, 0.6790, 0.8848, 1.9870, 1.5680, 1.9100, 0.6998, 0.7502, 1.3936, 0.6572, 2.0316, 1.6216, 1.3394, 1.4302, 1.3120, 0.4154, 0.7556, 0.5976, 0.6672, 1.3628, 1.5708, 1.6650, 1.7120, 0.6456, 1.4972, 1.3250, 1.2280, 0.9818, 0.9322, 1.0784, 2.4084, 1.7392, 0.3630, 0.6654,1.0812, 1.2364, 0.2082, 0.3600, 0.9898, 0.8178, 0.6718, 0.4140, 0.6596, 1.0634, 1.0884, 0.9114, 0.8584, 0.5000, 1.3070, 0.9296, 0.9394, 1.0918, 0.8240, 0.7884, 0.6438, 0.2804, 0.4876, 0.6514, 0.7264, 0.6466, 0.6054, 0.4704, 0.2410, 0.6436, 0.5852, 0.5202, 0.4130, 0.6058, 0.4116, 0.4652, 0.5052, 0.3846.”

Table 1

Summary Statistics

Min.	Q1	Q2	Mean	Q3	Max.	SD	Skewness	Kurtosis
0.2082	0.6578	1.0559	1.1646	1.5188	3.2996	0.2082	0.9736	3.6675

The dataset demonstrates positive skewness and a lack of normality. To explore the nature of the model, we plotted the boxplot and TTT plot of the dataset, as shown in Figure 2.

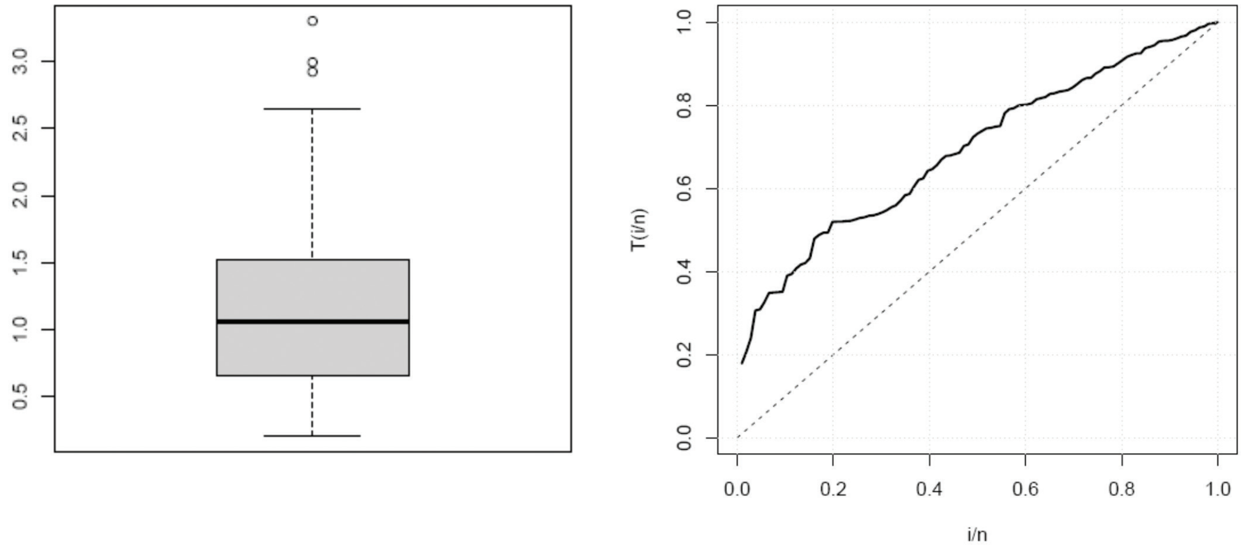

Figure 2: Boxplot (on the left) and TTT plot (on the right) for the analyzed dataset.

The box plot indicates that there are some outliers and that the data is positively skewed. Furthermore, the TTT plot, which is primarily located below the 45-degree line and seems almost convex, shows that the data follows a distribution with a decreasing hazard function.

Using R software's optim() function, the log-likelihood function described in Equation (11) was maximized to get the MLEs of the MEEXP model (R Core Team, 2023). Table 2 presents the estimated values.

Table 2

Estimated parameters using MLE and SE

Parameters	MLE	SE	95% C.I.
alpha	1.7215151	0.3431455	(1.671515, 1.771515)
beta	0.1222342	0.2445100	(0.112234, 0.132234)
lambda	0.5271803	0.1665411	(0.477180, 0.577180)

Figure 3 presents the Q-Q and P-P plots, indicating that the analyzed data provides a better fit to the MEEXP model.

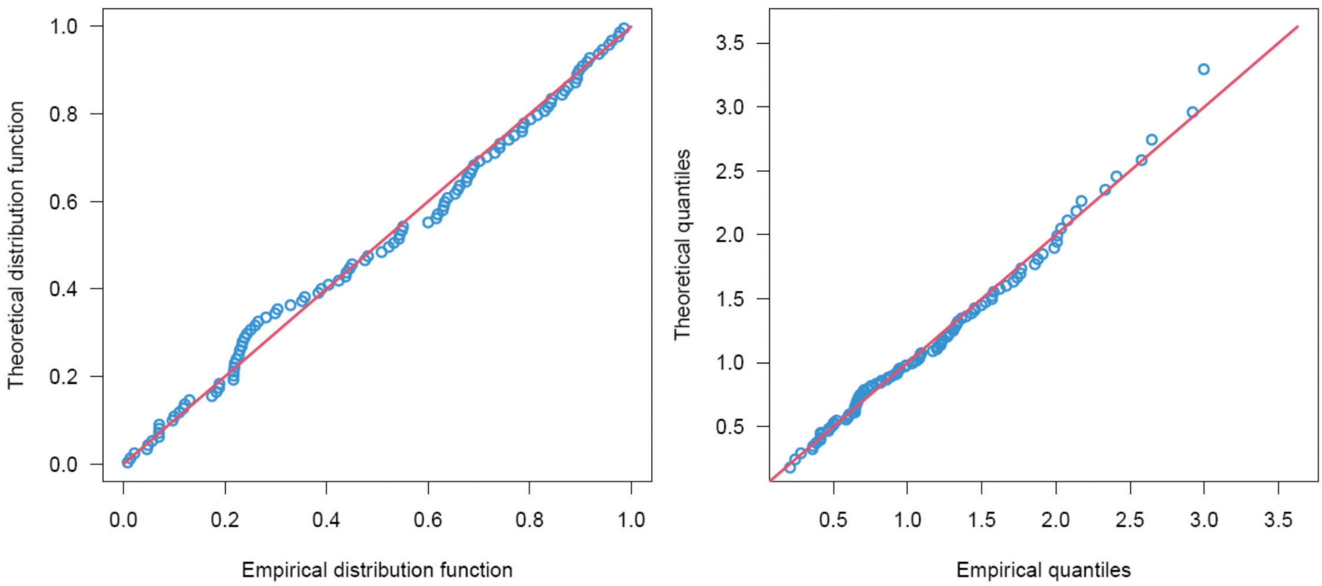

Figure 3: The P-P plot (left) & Q-Q plot (right) of the MHCC Model.

Table 3 presents the estimated parameter values of the MHCC model along with the log-likelihood, AIC, BIC, CAIC, and HQIC criteria. The corresponding values for these criteria and the model parameters are summarized in Table 3.

Table 3

Estimated parameters along with log-likelihood (LL) and information criteria values.

-LL	AIC	BIC	CAIC	HQIC	KS	W	A2
93.715	193.430	201.420	193.665	196.669	0.062 (0.806)	0.089 (0.640)	0.715 (0.546)

Figure 4 displays the histogram versus fitted pdf plot and ECDF versus cdf plots.

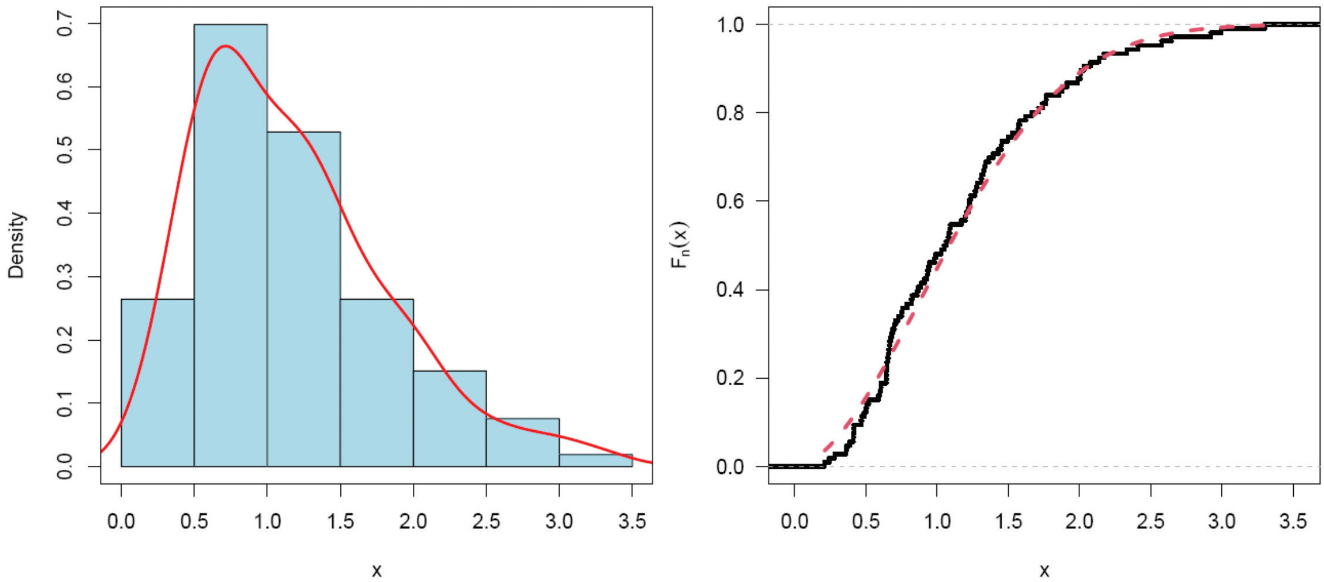

Figure 4: Histogram versus fitted pdf (Left) and ECDF versus fitted cdf (Right)

Figure 5 displays the residual plots for the data sets showing that the model has captured the underlying data structure approximately well. Figure 5 also displays normal Q-Q plot.

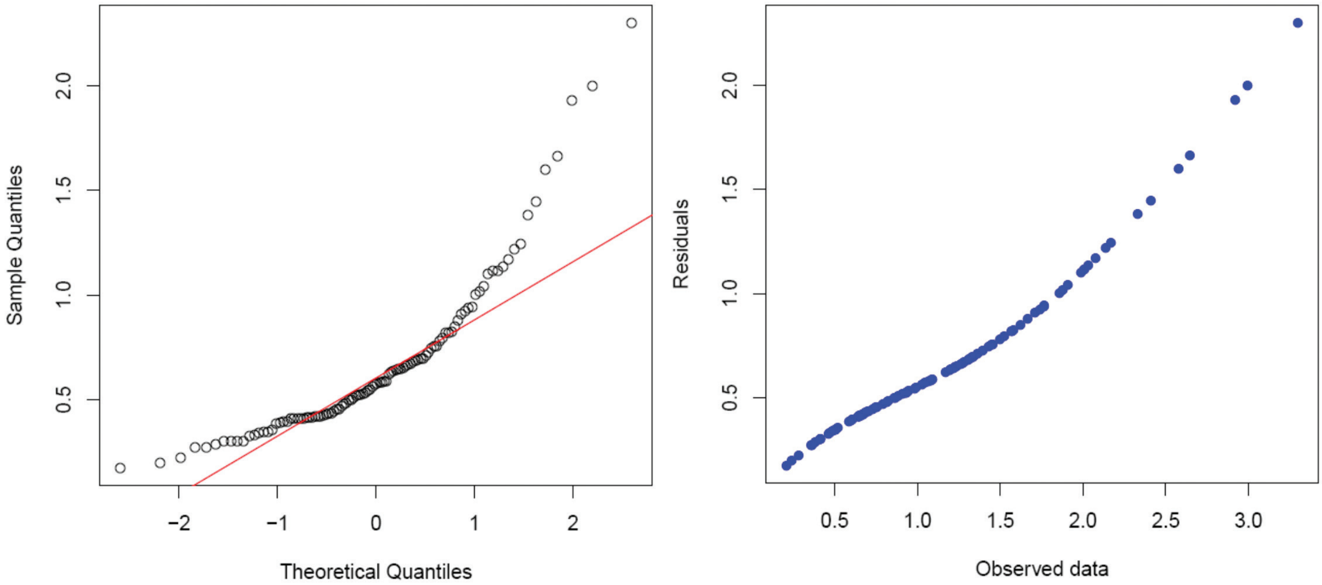

Figure 5: Residual Plot (Left) and Normal Q-Q plot (Right)

Model is compared to other model; we have chosen Normal distribution as well as exponential distribution and mentioned the AIC values of the models shown table 4. The models are exponential distribution, Weibull Extension (WE) distribution (Tang et al., 2003), Normal Distribution, Modified Weibull (MW) distribution (Lai et al.,2003), and Logistic Inverse Exponential (LIE) Distribution (Chaudhary & Kumar, 2020).

Table 4

AIC for MEEXP and some competitive models

Models	MEEXP	WE	MW	LIE	Normal	Exponential
AIC	193.4301	193.5786	193.7115	196.7763	212.4799	246.2955

Sensitivity analysis of the parameters

Sensitivity analysis involves evaluating how variations in input parameters affect the uncertainty in a model's output. For complex models, it helps identify which parameters have the greatest impact on the model's predictions and can guide decision-making, model improvement, and resource allocation. Figure 6 shows the sensitivity analysis for the parameters alpha, beta and lambda.

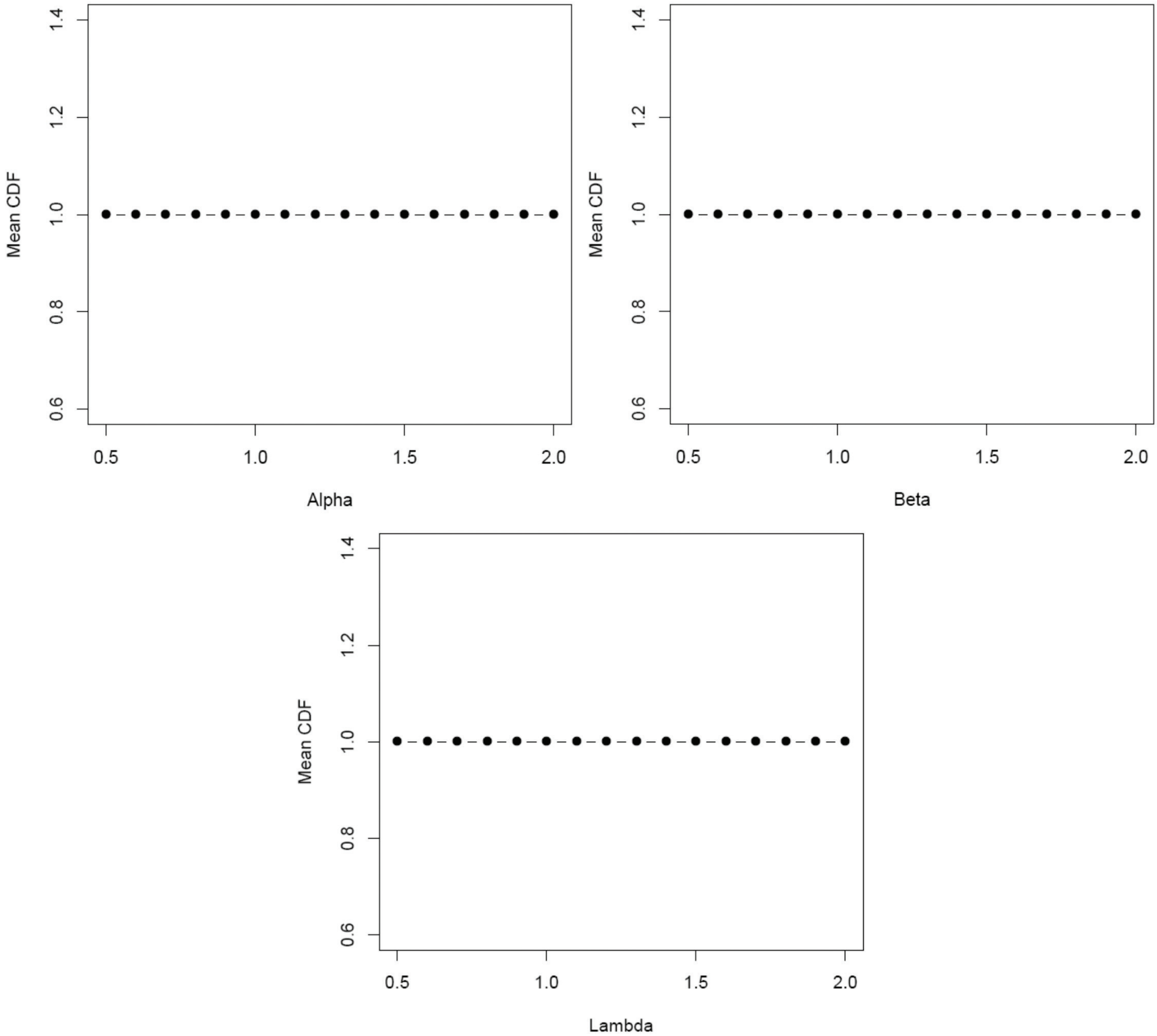

Figure 6: Sensitivity plots for parameters alpha, beta and lambda.

The mean value for the cross validation of 0.4284 is obtained indicates the average performance of the model across all the cross-validation folds. Cross-validation is typically used to evaluate the performance of a model by splitting the data into subsets (folds), training the model on a subset, and testing it on another. The average result across all the folds is then computed

Conclusion

This article analyzes the Modified Exponentiated Exponential (MEEXP) Distribution as a more flexible and adaptive model for survival analysis and reliability studies. By incorporating additional parameters, the MEEXP distribution can capture a wide range of hazard rate shapes, including those with non-monotonic behaviors. The model's parameters were determined using the MLE technique, and its effectiveness was demonstrated through the analysis of a real-world dataset from the COVID-19 pandemic. The MEEXP distribution outperformed classical models such as the exponential and normal distributions, as evidenced by goodness-of-fit tests, residual plots, and information criteria metrics (AIC, BIC, etc.). Sensitivity analysis further confirmed the robustness of the model across varying parameter values. The MEEXP model's flexibility in capturing complex data patterns makes it a valuable tool for various applications, including epidemiology, engineering, finance, and risk assessment. Future research could explore further extensions and refinements to this model to enhance its applicability to even more diverse datasets.

References

Alzaid, A. A., & Qarmalah, N. (2024). Uniformly Shifted Exponential Distribution. *Axioms, 13*(6), 339. https://doi.org/10.3390/axioms13060339

Bantan, R. A., Ahmad, Z., Khan, F., Elgarhy, M., Almaspoor, Z., Hamedani, G. G., & Gemeay, A. M. (2023). Predictive modeling of the COVID-19 data using a new version of the flexible Weibull model and machine leaning techniques. *Mathematical Biosciences and Engineering, 20*(2), 2847-2873. http://dx.doi.org/10.3934/mbe.2023134

Bhati, D., Malik, M. A., & Vaman, H. J. (2015). Lindley–exponential distribution: properties and applications. *Metron, 73,* 335-357.https://doi.org/10.1007/s40300-015-0060-9

Chaudhary, A. K., & Kumar, V. (2020). The Logistic NHE Distribution with Properties and Applications. *International Journal for Research in Applied Science & Engineering Technology (IJRASET), 8*(12), 591-603. https://doi.org/10.22214/ijraset.2020.32565

Chaudhary, A. K., & Kumar, V. (2020). Logistic Inverse Exponential Distribution with Properties and Applications. *International Journal of Mathematics Trends and Technology-IJMTT, 66*(10),151-162. https://doi.org/10.14445/22315373/IJMTT-V66I10P518

Chaudhary, A. K., & Kumar, V. (2021). Arctan exponential extension distribution with properties and applications. *International Journal of Applied Research, 7*(1), 432-442. https://doi.org/10.22271/allresearch.2021.v7.i1f.8251

Chaudhary, A. K., Telee, L. B. S., & Kumar, V. (2022). Inverse Exponentiated Odd Lomax Exponential Distribution: Properties and Applications. *Nepalese Journal of Statistics, 6*(01), 29–50. https://doi.org/10.3126/njs.v6i01.50801

Chaudhary, A. K., Sah Telee, L. B., & Kumar, V. (2023). Extended Kumaraswamy Exponential Distribution with Application to COVID-19 Data set. *Journal of Nepal Mathematical Society, 6*(1), 1–10. https://doi.org/10.3126/jnms.v6i1.57657

Chaudhary, A. K., Telee, L. B. S., Karki, M., & Kumar, V. (2024). Statistical analysis of air quality dataset of Kathmandu, Nepal, with a New Extended Kumaraswamy Exponential Distribution. *Environmental Science and Pollution Research, 31*(14), 21073-21088.https://doi.org/10.1007/s11356-024-32129-z

Chaudhary, A. K., and Kumar, V. (2022). Half Cauchy-Modified Exponential Distribution: Properties and Applications. *Nepal Journal of Mathematical Sciences (NJMS), 3*(1), 47-58.https://doi.org/10.3126/njmathsci.v3i1.44125

Chinazom, O. C., Chinaka, N. E., & Azubuike, I. M. (2023). The T-Exponentiated Exponential {Frechet} Family of Distributions: Theory and Applications. *Asian Journal of Probability and Statistics, 23*(4), 8-25. https://doi.org/10.9734/ajpas/2023/v23i4509

Gupta, R. D., & Kundu, D. (1999). Generalized exponential distributions. *Australian & New Zealand Journal of Statistics, 41*(2), 173–188. https://doi.org/10.1111/1467-842X.00068

Ijaz, M., Asim, S. M., & Alamgir. (2019). Lomax exponential distribution with an application to real-life data. *PloS one, 14*(12), e0225827.https://doi.org/10.1371/journal.pone.0225827

Jayakumar, K. and Fasna. K. (2023). Half Cauchy Exponential Distribution: Estimation and Applications. *Reliability: Theory & Applications, 18*(3 (74)), 387-401.

Jabeen, S., & Para, B. (2018). Exponentiated Gamma Exponential Distribution. *Sohag Journal of Mathematics: An International Journal, 5*(3), 79-84.http://dx.doi.org/10.18576/sjm/050301

Kharazmi, O., Mahdavi, A., &Fathizadeh, M. (2014). Generalized weighted exponential distribution. *Communications in Statistics-Simulation and Computation, 44*(6), 1557-1569. http://dx.doi.org/10.1080/0

3610918.2013.824092

Lai, C. D., Xie, M., & Murthy, D. N. P. (2003). A modified Weibull distribution. *IEEE Transactions on reliability, 52*(1), 33-37. https://doi.org/10.1109/TR.2002.805788

Mahdavi, A., & Jabari, L. (2017). An extended weighted exponential distribution. *Journal of Modern Applied Statistical Methods, 16,* 296-307.https://doi.org/10.56801/10.56801/v16.i.898

Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician), 37*(1), 25-32.https://doi.org/10.2307/2348376

Nadarajah, S., & Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering & System Safety, 91*(6), 689–697. https://doi.org/10.1016/j.ress.2005.05.008

Nadarajah, S., & Kotz, S. (2011). The exponentiated exponential distribution: A survey. *International Journal of Statistics, 6*(2), 150–168.https://doi.org/10.1007/s10182-011-0154-5

Ndayisaba, A. , Odongo, L. 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

R Core Team (2023). *R: A Language and environment for statistical computing.* (Version 4.1) [Computer software]. Retrieved fromhttps://cran.r-project.org. (R packages retrieved from CRAN snapshot 2023-04-07).

Ross, S. M. (2014). *Introduction to probability models (11th Ed.)*. Academic Press. https://doi.org/10.1016/B978-0-12-407948-9.00012-8,

Sóyínká, A. T. (2023). On Beta Exponentiated Lomax-Exponential Distribution with applications. *Scientific African, 20,* e01701. https://doi.org/10.1016/j.sciaf.2023.e01701

Tang, Y., Xie, M., & Goh, T. N. (2003). Statistical analysis of a Weibull extension model. *Communications in Statistics-Theory and Methods, 32*(5), 913-928. https://doi.org/10.1081/STA-120019952

Telee, L. B. S., & Kumar, V. (2021). Lindley Generalized Inverted Exponential Distribution: *Model and Applications. Pravaha, 27*(1), 61-72.https://doi.org/10.3126/pravaha.v27i1.50616

Telee, L. B. S., & Kumar, V. (2023). Modified generalized exponential distribution. *Nepal Journal of Mathematical Sciences, 4*(1), 21-32.https://doi.org/10.3126/njmathsci.v4i1.53154

Telee, L. B. S., Chaudhary, A. K., Karki, M., & Kumar, V. (2024). The Evolution of Probability Distributions: From Simple to Complex Parameter Spaces. *Eurasian Journal of Economics and Statistics, 1*(1), 1-14.