Journal of Econometrics and Statistics
Vol. 4, Issue 1, 2024, pp. 63-78
https://doi.org/10.47509/JES.2024.v04i01.04

Modified Arctan Exponential distribution with application to COVID-19 Second Wave data in Nepal

Arun Kumar Chaudhary[a] ,Lal Babu Sah Telee[b] and Vijay Kumar[c]

[a] *Department of Statistics, Nepal Commerce campus, Tribhuvan University* [b]*Department of Statistics, Nepal Commerce campus, Tribhuvan University and* [c]*Department of Mathematics and Statistics, DDU Gorakhpur University, Gorakhpur*

ABSTRACT
In this study, we introduce a novel trigonometric model known as the Modified Arctan Exponential distribution. This model is created by compounding the Cauchy family of distributions with the exponential distribution serving as the baseline distribution. Our aim is to use this model for analyzing lifetime data. We have derived mathematical expressions for various statistical functions, including the probability density function, distribution function, survival function, quantile function, hazard rate function, reversed hazard rate function, cumulative hazard rate function, skewness, and kurtosis. In addition, we have provided visual representations of the probability density and hazard rate curves. The COVID-19 second wave data in Nepal were collected from May 1, 2021, to September 30, 2021, as provided by Worldometer, World Health Organization (WHO). To assess the effectiveness of our proposed model, we applied it to a dataset concerning the second wave of COVID-19 cases in Nepal. We estimated the model parameters using three distinct techniques: maximum likelihood, least squares, and Cramer's-von Mises. To validate the model, we employed a range of statistical criteria, including Akaike's Information Criterion, Bayesian Information Criterion, Corrected Akaike's Information Criterion, and Hannan-Quinn Information Criterion. We also used P-P and Q-Q plots for validation purposes. To gauge the goodness of fit of our model to the data, we conducted the Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises tests. These tests were carried out to assess whether our model is suitable for analyzing the provided data. Our empirical findings demonstrate that, when compared to alternative lifetime distributions, our suggested distribution not only provides a better fit but also offers increased flexibility for the analysis of lifetime data. All numerical calculations were made using the R programming language.

KEYWORDS
Cauchy family of distribution, COVID-19, Exponential distribution, hazard rate function, Maximum Likelihood Estimation, Second wave.

1. Introduction

Lifetime distribution, alternatively referred to as survival analysis or time-to-event analysis, is a statistical method employed to examine the duration until a specific event of interest takes place. This event could be anything that has duration, such

Contatct[a]. Arun Kumar Chaudhary.Email:akchaudhary1@yahoo.com

Article History
Received : 04 February 2024; Revised : 06 March 2024; Accepted : 12 March 2024; Published : 15 April 2024

To cite this paper

Arun Kumar Chaudhary, Lal Babu Sah Telee & Vijay Kumar (2023). Modified Arctan Exponential distribution with application to COVID-19 Second Wave data in Nepal. *Journal of Econometrics and Statistics.* 4(1), 63-78.

as the failure of a machine, the onset of a disease, the time until a customer makes a purchase, or even the lifespan of a living organism. Lifetime distribution provides valuable insights into the probability of an event happening at a specific time and helps in making predictions and informed decisions in various fields, including engineering, life sciences, medicine, biology, insurance, healthcare, finance, and social sciences. Studying continuous probability distributions, such as the exponential, Cauchy, Gamma and Weibull distributions, is a common practice in statistical literature for the analysis of lifetime data. These probability distributions play a crucial role in understanding and modeling the variability in lifetimes, making them essential tools in fields like reliability engineering, survival analysis, and actuarial science. By examining these distributions, researchers and analysts can gain valuable insights into the behavior of data points over time, allowing them to make informed decisions and predictions in various real-world scenarios. In recent years, considerable attention from researchers has been directed towards the exponential distribution due to its efficacy in modelling lifetime data. Its favourable attributes stem from closed-form solutions available for numerous survival analyses. Nevertheless, although the exponential distribution is frequently applicable when assuming a consistent failure rate, real-world failure rates often display fluctuations. These variations in real-world failure rates can arise due to a multitude of factors, such as changing environmental conditions, aging components, or manufacturing inconsistencies. Consequently, relying solely on the assumption of a constant failure rate from the exponential distribution may not accurately model the complexities of many practical scenarios. Consequently, relying on the exponential lifetime model in a random manner can be both inadequate and unrealistic. In recent times, novel classes of models have emerged, focusing on adaptations and enhancements to classical probability models. Notably, Marshall and Olkin (2007) have introduced some of these modifications. Researchers have also made considerable efforts to develop new distributions that extend established ones and offer greater flexibility for handling encountered data. Some approaches involve introducing additional parameters to existing distributions to create broader families of models. Numerous models in the field of statistics have been introduced, featuring additional parameters designed to produce unique probability distributions. These additional parameters have been incorporated into the models to enhance their ability to capture complex patterns and variations in data. For example, Rinne (2009) and Pham and Lai (2007) have put forth models that include extra parameters aimed at generating novel statistical models. Over the past few decades, the Exponential distribution has often served as the parent distribution for creating novel distributions families. Numerous researchers have introduced various adaptations of the Exponential distribution. These modifications have paved the way for the development of diverse probability distributions, each tailored to address specific real-world scenarios and applications. Some of the well-known life time models found in the literature by the modification of the Exponential distribution are Exponential power distribution (Smith & Bain, 1975), Generalized exponential distribution (Gupta & Kundu ,1999), Exponential-Weibull distribution (Cordeiro et al., 2014), Beta GE (Barreto- Souza et al., 2010), Modified exponential distribution (Rasekhi et al.,2017), Transmuted GE distribution (Khan et al., 2017), Alpha power transformed extended exponential distribution (Hassan et al.,2018), A new extension of exponential distribution(Almarashi et al. ,2019), Truncated Cauchy power–exponential distribution (Chaudhary et al.,2020), Ranked set sampling with application of modified kies exponential distribution(Aljohani et al.,2021), Truncated exponentiated-exponential distribution(Ribeiro-Reis,2022), Inverse Exponentiated Odd Lomax Exponential Distribution (Chaudhary et al.,2022), Power-modified kies-exponential distribution(Afify et al.,2022) and the inverse exponential power distribution (Chaudhary et al., 2023). In the context of a non-negative random variable X, the exponential distribution with parameter δ is employed when its cumulative distribution function (CDF) can be formulated as follows:

$$G(x) = 1 - e^{-\delta x}; x \geq 0, \delta > 0 \tag{1.1}$$

An extension of the exponential distribution introduced by (Nadarajah & Haghighi,2011) allows for a wider range of probability distributions and is a generalization of the traditional exponential distribution, enabling more versatile modelling in various statistical applications. This extension allows for a more nuanced analysis and application of exponential-like phenomena in various fields. The generalization consistently places its mode at zero, while also accommodating the possibility of increasing, decreasing, and constant hazard rates. We have chosen the exponential distribution as the base distribution because of its extensive use, simplicity, and mathematical manageability. Its mathematical flexibility empowers researchers and practitioners to effectively model and analyze diverse phenomena within fields like statistics, physics, and engineering, making it a valuable tool for advancing knowledge and innovation. In this study, we present a trigonometric model. Incorporating a wide range of real-world applications, this trigonometric model is designed to be adaptable and user-friendly, making it a valuable tool for students, educators, and professionals alike. Throughout this study, we will delve into the theory behind our model, its practical implications, and examples of its successful application in various scenarios. Gómez-Déniz and Calderín-Ojeda (2015) defined to develop Pareto ArcTan (PAT) distribution by choosing the classical Pareto survival function as the parent distribution and incorporating the inverse of the circular tangent function to model Norwegian fire insurance data. Chaudhary et al. (2021) introduced the Arc tan generalized exponential distribution, which exhibits a flexible hazard rate. The flexibility of the hazard rate opens up opportunities for its application in various statistical and modelling contexts. Researchers can utilize this distribution to better understand and analyze data with non-constant hazard rates, contributing to advancements in statistical science. Chaudhary et al. (2021) also suggested Arc tan exponential extension distribution using the arc-tan-G family of distribution by choosing exponential extension distribution as parent distribution. The Arctan-X family of probability distributions, proposed by (Alkhairy et al., 2021), has garnered significant attention in the field of statistics and data analysis due to its unique properties and potential applications. These distributions are derived from the inverse trigonometric function known as the arctangent function. These distributions have demonstrated superior goodness-of-fit when applied to various types of data, including actuarial data, financial data, and related datasets within these fields. In this research, we have derived a new distribution using Cauchy family of distribution. In literature different probability models are available that are formulated using Cauchy family of distribution as Half-Cauchy modified exponential distribution (Chaudhary et al.,2022) and Half- Cauchy exponential extension distribution (Telee & Kumar, 2022). Consider Cauchy family of distribution on a non-negative random variable X such that x > 0 and $\theta > 0$ which is defined by

$$H(x) = \frac{2}{\pi} \arctan\left\{ -\frac{1}{\theta} \log\left\{1 - G(x)\right\} \right\} \tag{1.2}$$

where G(x) is distribution function of the base line distribution. The objective of this study is to create a more adaptable probability distribution and conduct data analysis for the second wave of COVID-19 in Nepal using this distribution. In pursuit of this objective, we aim to establish a probability distribution model that offers greater flexibility, allowing for a more comprehensive and insightful analysis of the COVID-19 second wave in Nepal. This model will serve as a valuable tool in understanding the dynamics and trends associated with the pandemic within the Nepalese context, ultimately aiding in the development of more effective public health strategies and interventions. The subsequent structure is employed to present the various sections of this study. Section 2 will introduce the Modified Arctan Exponential distribution (MATE) while elucidating its mathematical and statistical properties. Moving forward to Section 3, we will explore estimation techniques, including discussions on least-squares (LSE), Cramer-Von-Mises (CVME), and maximum likelihood (MLE). In Section 4, our focus will shift towards providing model parameter estimates using data from the COVID-19 second wave in Nepal. Additionally, we will showcase examples of different criteria used to assess the goodness of fit of the proposed model. All numerical calculations were caried out using the R programming language. In the concluding Section 5, this study has aimed to offer valuable insights into the field of statistical analysis and modeling. We hope that the information presented in this paper serves as a valuable resource for researchers, practitioners, and policymakers alike.

2. The Modified Arctan Exponential (MATE) Distribution

Using the exponential distribution as base distribution from equation (1.1) in Cauchy family of distribution of equation (1.2), the new derived distribution function is as follows:

$$H(x) = \frac{2}{\pi} \arctan\left\{ \frac{\delta x}{\theta} \right\} \tag{2.1}$$

Modifying above distribution function by adding two extra parameters α (scale parameter) and λ (shape parameter) to make it more applicable and flexible with distribution function as

$$F(x) = \left[\frac{2}{\pi} \arctan\left\{ \frac{\delta x e^{\alpha x}}{\theta} \right\} \right]^{\lambda} ; x \geq 0, \theta, \delta, \alpha, \lambda > 0 \tag{2.2}$$

For simplicity, we have assumed, $\frac{\delta}{\theta} = \beta$, hence

$$F(x) = \left[\frac{2}{\pi} \arctan\left\{ \beta x e^{\alpha x} \right\} \right]^{\lambda} ; x \geq 0, \alpha, \beta, \lambda > 0 \tag{2.3}$$

Expression (2.3) is the CDF of the proposed model which is denoted by MATE distribution and the associated probability density function (pdf) is given by (2.4).

$$f(x;\alpha,\beta,\lambda) = \lambda\beta\left(\frac{2}{\pi}\right)^{\lambda} e^{\alpha x}\left(1+\alpha x\right)\left\{\arctan\left(\beta x e^{\alpha x}\right)\right\}^{\lambda-1} \left\{1+\left(\beta x e^{\alpha x}\right)^2\right\}^{-1}; x \geq 0, \alpha,\beta,\lambda > 0 \tag{2.4}$$

In the following section, we delve into different key properties of the suggested model. These include the survival function, hazard rate function, reversed hazard rate function, cumulative hazard rate function and quantile function. Additionally, we investigate the model's behavior as it approaches its asymptotic limits.

2.1. Survival function

The survival function, labeled as S(x), shows the likelihood of surpassing a specific point x without encountering an event. It complements the cumulative distribution function (CDF). Equation (2.5) furnishes the survival function for the suggested model.

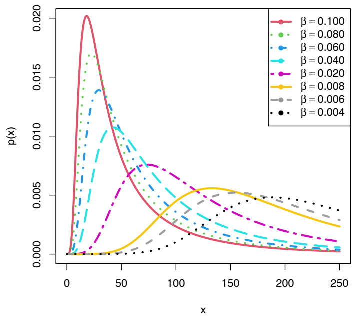
$$S(x) = 1 - \left[\frac{2}{\pi}\arctan\left\{\beta x e^{\alpha x}\right\}\right]^{\lambda}; x \geq 0, \alpha,\beta,\lambda > 0 \tag{2.5}$$

2.2. Hazard rate function

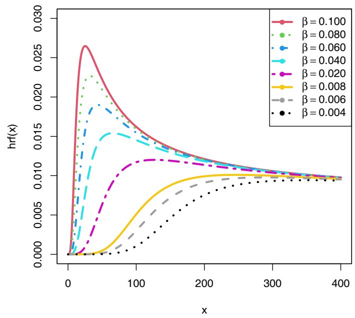
The hazard rate function, labeled as h(x), signifies how quickly failures happen at a given time point. It is computed by dividing the probability density function (pdf) by the survival function S(x) of the distribution. In the proposed model, equation (2.6) precisely defines h(x).

$$h(x) = \lambda\beta\left(\frac{2}{\pi}\right)^{\lambda} e^{\alpha x}\left(1+\alpha x\right)\left\{\arctan\left(\beta x e^{\alpha x}\right)\right\}^{\lambda-1} \left\{1+\left(\beta x e^{\alpha x}\right)^2\right\}^{-1}\left\{1-\left[\frac{2}{\pi}\arctan\left\{\beta x e^{\alpha x}\right\}\right]^{\lambda}\right\}^{-1}; x > 0 \tag{2.6}$$

Figure 1 presents probability density curves and hazard rate curves for fixed values of α=0.0073 and λ=5, at different values of β. The figure consists of two panels that illustrate important aspects of these curves for various parameter settings. In the left panel, we can observe the probability density curve, which demonstrates how it varies with changes in the parameters. This variability highlights the model's ability to adapt to different types of datasets, showcasing its flexibility. On the other hand, the right panel of Figure 1 displays hazard rate curves associated with specific parameter combinations. These hazard rate curves exhibit patterns that include both increasing and decreasing trends, as well as the distinctive shape resembling an inverted bathtub. This inverted bathtub shape in the hazard rate curves is a noteworthy feature. It suggests that the risk or hazard of an event can change over time in a non-linear manner. This graphical representation allows researchers and analysts to gain a comprehensive view of how these parameters influence the probability distribution and associated hazard rates, offering valuable information for making data-driven decisions in various contexts.



(a) Probability density



(b) Hazard rate function

Figure 1. Probability density curve and hazard rate curve.

The equation (2.7) defines the reversed hazard rate, denoted as hrex(x), for this model.

$$h_{rev}(x) = \lambda\beta\left(\frac{2}{\pi}\right)^{\lambda} e^{\alpha x}\left(1+\alpha x\right)\left\{\arctan\left(\beta x e^{\alpha x}\right)\right\}^{\lambda-1}\left\{1+\left(\beta x e^{\alpha x}\right)^{2}\right\}^{-1}\left[\frac{2}{\pi}\arctan\left\{\beta x e^{\alpha x}\right\}\right]^{-\lambda} \tag{2.7}$$

2.3. Cumulative Hazard rate function

Cumulative hazard rate function The equation (2.8) gives us the cumulative hazard rate function, denoted as H(x), for the suggested model.

$$H(x) = -\ln S(x) = -\ln\left\{1-\left\{\frac{2}{\pi}\arctan\left\{\beta x e^{\alpha x}\right\}\right\}^{\lambda}; x>0, \alpha, \beta, \lambda>0\right\} \tag{2.8}$$

2.4. Quantile function

The quantile function is defined for MATE by equation (2.9)

$$\beta x e^{\alpha x} - \tan\left(\frac{\pi}{2}p^{1/\lambda}\right) = 0;\;\; 0 \le p \le 1 \tag{2.9}$$

2.5. Asymptotic Behavior

We can examine the density function's behavior as it approaches zero and infinity by ensuring that $\lim_{x\to 0} f(x) = \lim_{x\to\infty} f(x)$. If the model follows these asymptotic properties, it will have a mode. This evaluation requires us to analyze the limits at both ends.

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \lambda\beta\left(\frac{2}{\pi}\right)^{\lambda} e^{\alpha x}\left(1+\alpha x\right)\left\{\arctan\left(\beta x e^{\alpha x}\right)\right\}^{\lambda-1}\left\{1+\left(\beta x e^{\alpha x}\right)^{2}\right\}^{-1} = 0 \tag{2.10}$$

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \lambda\beta\left(\frac{2}{\pi}\right)^{\lambda} e^{\alpha x}\left(1+\alpha x\right)\left\{\arctan\left(\beta x e^{\alpha x}\right)\right\}^{\lambda-1}\Big\{1+\left(\beta x e^{\alpha x}\right)^{2}\Big\}^{-1} = 0 \tag{2.11}$$

Since $\lim_{x\to 0} f(x) = \lim_{x\to\infty} f(x)$ so we can say that the mode of the proposed model exists. This information helps us determine whether the distribution has a peak or mode, which can be important in various statistical and analytical contexts. Therefore, analyzing the asymptotic behavior of the density function is a fundamental step in understanding the behavior of the underlying data or model. Skewness characterizes the uniformity of the data. In this study, we utilized Bowley's skewness coefficient, as proposed by (Al-saiary et al., 2019), which is based on quantiles as follows:

$$\mathrm{SK}\,(\mathrm{B}) = \frac{\mathrm{Q}\,(0.75) + \mathrm{Q}\,(0.25)\,\text{-}2{*}\mathrm{Q}\,(0.50)}{\mathrm{Q}(0.75)\text{-}\mathrm{Q}(0.25)} \tag{2.12}$$

The Octiles Kurtosis coefficient, as described in studies by(Moors ,1998) and (Al-saiary et al. ,2019), can be determined using the following formula:

$$\mathrm{K_u}=\frac{\mathrm{Q}\,(0.375)\text{-}\mathrm{Q}(0.625)\text{-}\mathrm{Q}\,(0.125)+\mathrm{Q}\,(0.875)}{\mathrm{Q}\,(0.75)\text{-}\mathrm{Q}(0.25)} \tag{2.13}$$

3. Techniques of parameters estimation

Parameter estimation plays a crucial role in the process of fitting and formulating models. In practical applications, a variety of methods are available for estimating parameters. In this study, we employed three distinct approaches for parameter estimation: Maximum Likelihood Estimation (MLE), Least Square Methods of estimation (LSE), and Cramer Von Mises methods of estimation (CVM). Each of these estimation methods has its unique advantages and is applied in specific situations based on the nature of the data and the underlying statistical assumptions. MLE, for instance, is a widely used technique that aims to find parameter values that maximize the likelihood of observing the given data. On the other hand, LSE focuses on minimizing the sum of squared differences between observed and predicted values, making it suitable for certain regression models. CVM, a less commonly employed method, assesses the goodness of fit by comparing the empirical distribution of the data with the expected distribution defined by the estimated parameters. The choice of which estimation method to use can significantly impact the accuracy and reliability of the model, and it is essential to select the most appropriate approach for your specific research objectives and dataset characteristics. In this study, we explore the strengths and limitations of these three estimation techniques to better understand their suitability for the modeling task at hand.

3.1. Maximum Likelihood Estimation (MLE)

This estimation method depends on optimizing the model's log likelihood function. Imagine we have a random sample of 'n' items from MATE, which we'll represent as

.In this situation; the log likelihood function can be formulated as follows:

$$
\begin{aligned}
l(\alpha,\beta,\lambda|\underline{x}) = n\ln(\lambda\beta) + n\lambda\log(2/\pi) + \sum_{i=1}^{n} x_i + \sum_{i=1}^{n}(1+\alpha x_i) + \\
(\lambda-1)\sum_{i=1}^{n}\ln\left(\arctan\left(\beta x e^{\alpha x}\right)\right) - \sum_{i=1}^{n}\ln\left[1+\left(\beta x e^{\alpha x}\right)^2\right]
\end{aligned}
\tag{3.1}
$$

After obtaining the derivatives of equation (3.1) with respect to α, β, and λ, we can proceed to calculate the first-order and second-order partial derivatives of the log-likelihood function. To estimate the parameters of the proposed model, we equate the first-order derivatives to zero and solve for them. However, in practice, solving these first-order partial derivatives analytically can be complex or even impossible for certain models. This is particularly true for models with intricate or nonlinear relationships between parameters and the likelihood function. In such cases, resorting to computer programming and numerical optimization techniques becomes a practical and effective approach. These methods allow us to iteratively refine our parameter estimates until we converge to the values that optimize the likelihood function, ensuring our model fits the data as closely as possible.

3.2. Estimation using Least-Square (LSE)

We start by working with a set of ordered random variables, denoted as $\mathrm{X}_{(1)} < \mathrm{X}_{(2)} < \ldots < \mathrm{X}_{(\mathrm{n})}$. Afterward, we draw a random sample $\{X_1, X_2, \ldots, X_n\}$of size n from a distribution described by the function F(.). To create a function A, we use the cumulative distribution function (CDF) of these ordered statistics, which is represented as $F(X_{(i)})$and is explained in equation (3.2).

$$
A\left(x;\alpha,\beta,\lambda\right) = \sum_{i=1}^{n}\left[F(X_{(i)}) - \frac{i}{n+1}\right]^2 \tag{3.2}
$$

Once we've minimized function (3.2) with respect to the parameters, the next step involves solving for these parameter values, ensuring that they align with the intended specifications of the MATE model. This process is crucial for fine-tuning the model and making it suitable for its intended purpose.

3.3. Cramer-Von-Mises (CVM) method

We can determine the values of α, β, and λ by using this method to minimize the function (3.3).

$$
Z\left(X;\alpha,\beta,\lambda\right) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F\left(x_{i:n}|\alpha,\beta,\lambda\right) - \frac{2i-1}{2n}\right]^2 \tag{3.3}
$$

To find both the first and second-order partial derivatives of function Z, we perform differentiation on equation (3.3) with respect to α, β, and λ. Solving these nonlinear equations enables us to determine the estimated parameters. This information is essential for understanding how changes in the input variables α, β, and λ impact the behavior of the function. It provides a deeper understanding of the relationships between these parameters and allows us to make more informed decisions or predictions based on the function's behavior. Parameter estimation and applicability testing

4. Parameter estimation and applicability testing

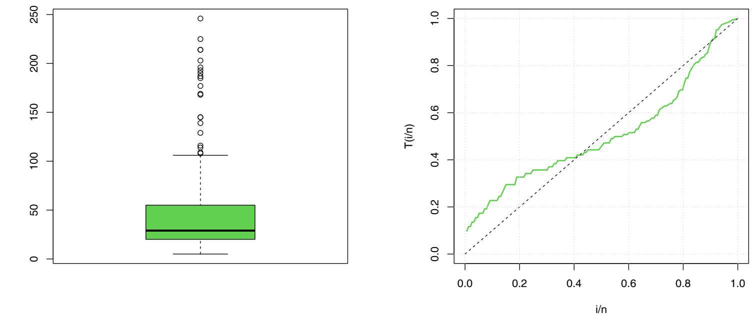
For testing the applicability of the model, a second wave of COVID-19 real data set is used. The second wave of the COVID-19 pandemic in Nepal commenced in early April 2021, with just one reported death on the first day of that month. As the days progressed, the daily infection rate surged significantly, culminating in a peak in May. On May 11, 2021, the country saw an average of 9,317 new cases reported daily, while daily fatalities reached approximately 225, as reported by the Ministry of Health and Population, Government of Nepal. Likewise, the maximum number of deaths occurred in Nepal during the second wave on May 19, 2021 was 246 out of 8064 new cases reported. In order to evaluate the suitability of the proposed model, we utilized a dataset encompassing the total number of COVID-19-related deaths in Nepal during the second wave, spanning from May 1, 2021, to September 30, 2021, as provided by (Worldometer, 2023). 19, 27, 37, 55, 58, 54, 50, 53, 88, 139, 225, 168, 214, 203, 187, 145, 214, 196, 246, 190, 177, 129, 193, 185, 169, 145, 106, 96, 116, 109, 114, 68, 101, 75, 101, 68, 99, 92, 108, 81, 59, 67, 61, 46, 53, 41, 52, 39, 44, 34, 51, 46, 41, 81, 24, 27, 30, 34, 42, 20, 41, 33, 34, 19, 27, 23, 15, 28, 29, 20, 22, 20, 18, 12, 18, 33, 43, 21, 23, 32, 25, 30, 24, 18, 16, 18, 25, 20, 33, 16, 27, 23, 18, 23, 24, 35, 37, 25, 19, 55, 22, 35, 30, 32, 27, 20, 33, 35, 27, 42, 33, 26, 30, 24, 24, 35, 44, 26, 25, 27, 24, 16, 20, 20, 16, 23, 9, 20, 20, 21, 10, 14, 21, 12, 13, 16, 19, 7, 11, 10, 16, 12, 8, 5, 6, 13, 9, 6, 9, 7, 12, 8, 12. Parameters are estimated using optim () function of R language programming (R Core Team, 2022). To investigate the exploratory characteristics of the curve, we computed summary statistics for the utilized data and created graphical plots, including Box plots and TTT plots. The summary statistics can be found in Table 1. This holistic approach provides a robust foundation for further analysis and interpretation of the curve's dynamics.

Table 1. Summary statistics of the data.

Min.	Q1	Q2	Q3	Mean	Max.	S.D.	Skewness	Kurtosis
5.00	20.00	29.00	55.00	51.35	246.00	53.91	1.88	5.66

The data exhibits positive skewness and is not normally distributed. It has a minimum value of 5 and a maximum value of 246, resulting in a range of 241.In the second wave of COVID-19 in Nepal; it was found that an average of 51 people per day lost their lives. Figure 2 presents both the box plot and the TTT plot for this dataset. The box plot displays the interquartile range, median, and any potential outliers, while the TTT plot can provide additional insights into the distribution shape and symmetry. Analyzing these plots and the characteristics of the data can help researchers and analysts make informed decisions about how to handle and interpret this non-normally distributed dataset

Parameters estimated using MLE is mentioned in table 2. Table 2 also contains the



(a) Boxplot. (b) TTT plot.

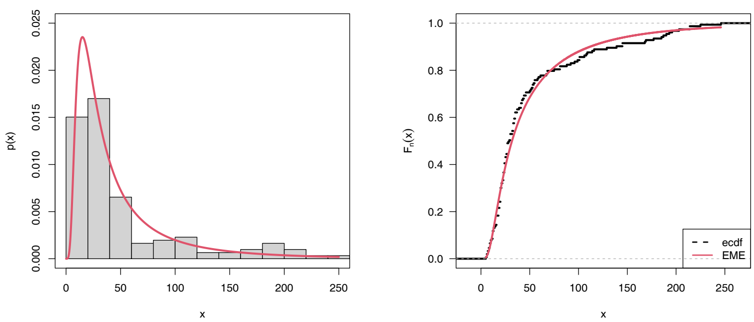
Figure 2. Boxplot and TTT plot of the data.

standard error of estimates of the parameters. In Figure 3, we plot and display his-

Table 2. Maximum likelihood estimates and standard error of estimates(SE).

Parameters	MLE	SE
α	0.0073	0.0020
β	0.1279	0.0972
λ	5.3110	4.0442

tograms alongside the probability density curve, as well as the empirical cumulative distribution function (CDF) compared to the fitted CDF. Comparing the histogram to the probability density curve and the empirical cumulative distribution function (CDF) to the fitted CDF in Figure 3 allows us to assess the model's goodness of fit to the considered data.



(a) Histogram versus pdf (b) Empirical versus fitted cdf

Figure 3. Histogram versus pdf and Empirical versus fitted cdf

To assess the reliability of the estimated parameters, we employ two additional techniques: the Least Square Method (LSE) and Cramer's von Mises (CVM) methods of estimation in Table 3. The parameters obtained through these methods exhibit very similar results. Table 4 displays a comparative evaluation of estimation techniques using four distinct information criteria values and negative log-likelihood values.

Table 3. Estimated parameters using MLE, LSE and ,CVM.

Parameters	MLE	LSE	CVM
α	0.0073	0.0086	0.0091
β	0.1279	0.1277	0.1257
λ	5.3110	5.3104	5.3105

The examination discloses that the Maximum Likelihood Estimation (MLE) approach demonstrates the most favorable information criteria values in contrast to the other two techniques This robust preference for the Maximum Likelihood Estimation (MLE) method is indicative of its superior performance in modeling the COVID-19 second wave dataset. The MLE method outperforms the other approaches in terms of fitting the data and providing a more accurate representation of the underlying patterns. This finding underscores the significance of selecting the MLE method when dealing with similar datasets, as it consistently outshines alternative methods in estimating the parameters and delivering a more precise analysis.

Table 4. Information criteria for different methods of estimation.

Methods	LL	AIC	BIC	CAIC	HQIC
MLE	-732.3457	1470.691	1479.783	1470.852	1474.384
CVM	-732.7788	1471.558	1480.649	1471.719	1475.251
LSE	-732.6053	1471.211	1480.302	1471.372	1474.904

To check the model validation, P-P plot and Q-Q plots of the model taking the data are obtained and displayed in figure 4.

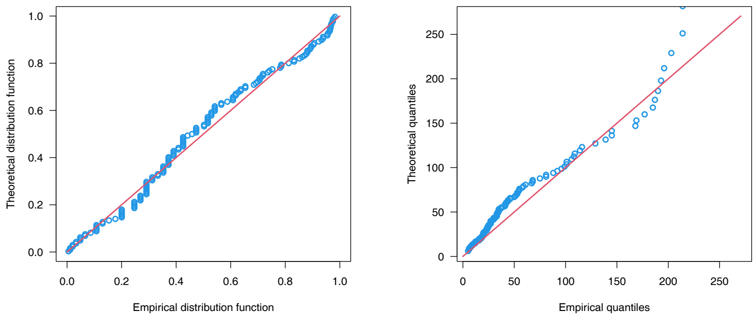
To assess and compare the suitability of the proposed model MATE alongside with methods of estimation, we calculate test statistics, including Kolmogorov-Smirnov (KS), Cramer's Von Mises(W), and Anderson-Darling(A^2). Table 5 conveniently presents the corresponding p-values for these statistics. These p-values serve as valuable indicators, aiding us in making informed assessments and comparisons of the methods of estimation performance and suitability in fitting the data. It also contains the corresponding p values. It shows that methods of estimations used here supports the goodness of fit under Kolmogorov-Smirnov, Cramer's Von Mises, and Anderson-Darling methods of testing goodness of fit.

Table 5. Goodness of fit statistics and p values for different methods of estimation.

Methods	KS(p-value)	W(p-value)	A^2(p-value)
MLE	0.0792(0.2918)	0.1694(0.3358)	0.9891(0.3629)
CVM	0.0693(0.4536)	0.1477(0.3970)	1.1693(0.2795)
LSE	0.0716(0.4130)	0.1482(0.3953)	1.1021(0.3078)

5. Model Comparison

This study assesses the proposed model by comparing it to six other models documented in existing literature. The six lifetime models under consideration include include the Generalized Exponential Extension (GEE) distribution (Lemonte, 2013),



(a) P-P plot (b) Q-Q plot

Figure 4. P-P and Q-Q plot

Lomax Exponentiated Weibull (LEW) distribution (Ansari & Nofal, 2020), Generalized Weibull Extension (GWE) (Sarhan and Apaloo, 2013), Odd Lomax Exponential (OLE) distribution (Ogunsanya et al., 2019), Weibull Extension (WE) distribution (Tang et al., 2003) and Half Logistic Nadarajah Haghighi (HLNHE) distribution (Joshi & Kumar,2020a).
Table 6 provides the estimated parameters for all these models using the given COVID-19 second wave real dataset in Nepal.

Table 6. Estimated parameters of competing models.

Methods	α	β	λ	θ
MATE	0.0073	0.1279	5.3110	-
GEE	0.2933	23.3805	5.3161	-
LEW	73.3614	0.6270	-	0.1537
GWE	375.6404	0.1155	0.4643	-
OLE	0.0796	0.0794	3.4232	-
WE	10.1098	0.4131	0.0143	-
HLNHE	0.0108	0.1779	19.3000	-

We compared the models by calculating different information criteria values for each of them and summarized the findings in Table 7. The results indicate that the proposed model has the lowest information criteria values, indicating that it is a more suitable fit for the dataset when compared to the other competing models. This superiority in information criteria values suggests that the suggested model provides a more precise representation of the data, surpassing the alternative models being considered. These findings highlight the robustness and effectiveness of the proposed model in capturing the underlying patterns and relationships within the dataset. Moreover, the superior performance of the proposed model underscores its potential in enhancing our understanding of the underlying dynamics within the dataset. This model's ability to outperform its competitors in terms of information criteria values signifies its suitability for making informed decisions and predictions based on the data.

To assess and compare the suitability of the proposed model MATE alongside competing models, we calculate test statistics, including Kolmogorov-Smirnov, Cramer's Von Mises, and Anderson-Darling. Table 8 conveniently presents the corresponding p-values for these statistics. These p-values serve as valuable indicators, aiding us in making informed assessments and comparisons of the model's performance and suitability in fitting the data.

Table 7. Information criteria values for MATE and competing models.

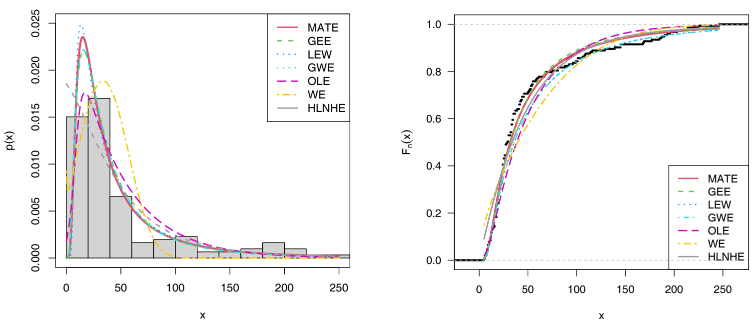
Methods	LL	AIC	BIC	CAIC	HQIC
MATE	-732.3457	1470.691	1479.783	1470.852	1474.384
GWE	-733.7814	1473.563	1482.654	1473.724	1477.256
GEE	-733.8717	1473.743	1482.835	1473.904	1477.436
LEW	-733.9826	1473.965	1483.056	1474.126	1477.658
OLE	742.4357	1490.871	1490.871	1491.032	1494.564
HLNHE	-753.5461	1513.092	1522.184	1513.253	1516.785
WE	-776.5519	1559.104	1568.195	1559.265	1562.797

Table 8. Goodness of fit statistics and p values for different methods of estimation .

Methods	KS(p-value)	W(p-value)	A^2(p-value)
MATE	0.0792(0.2918)	0.1694(0.3358)	0.9891(0.3629)
GEE	0.0848(0.2208)	0.1979(0.2718)	1.1902(0.2712)
LEW	0.0846(0.2234)	0.2077(0.2532)	1.2407(0.2525)
GWE	0.0834(0.2376)	0.1919(0.2840)	1.1624(0.2822)
OLE	0.1631(0.0006)	0.9614(0.0030)	4.8013(0.0036)
HLNHE	0.1270(0.0144)	0.7129(0.0120)	4.6682(0.0041)
WE	0.1601(0.0008)	1.5553(0.0001)	8.9759(0.0000)

When the test statistics values are lower and the p-values are higher for the proposed model in comparison to those of the competing model, it indicates that the proposed model provides a better fit to the dataset than the competing model. This suggests that the proposed model offers a stronger alignment with the dataset, demonstrating its superior performance when compared to the competing model.

To compare the proposed model with other models, we've generated probability density function (PDF) curves for all the models and a corresponding histogram, as shown in the left panel of Figure 5. Additionally, in the right panel of Figure 5, we've plotted empirical cumulative distribution function (CDF) curves for all the models, along with the fitted CDF curve. This graphical representation aids in our comparative evaluation of the models under consideration.



(a) Histogram versus pdfs (b) Fitted cdf versus emperical cdfs

Figure 5. Histogram versus pdfs and fitted cdf versus emperical cdfs

6. Conclusion

In this research, we introduce a novel trigonometric model called the Modified Arctan Exponential distribution (MATE). The MATE is constructed by combining the Cauchy family of distributions with the exponential distribution as the base model. This distribution exhibits a positively skewed and unimodal shape. We conducted a comprehensive analysis of various statistical properties associated with the MATE model, revealing its remarkable flexibility in accommodating both increasing and decreasing hazard functions, as well as an inverted bathtub-shaped hazard function. These findings were derived from a thorough graphical analysis of the Probability Density Function (PDF) and Hazard Rate Function (HRF) of the MATE. To estimate the model's parameters, we utilized three distinct methods: Cramer's-von Mises Estimation (CVME), Least Squares Estimation (LSE), and Maximum Likelihood Estimation (MLE). These approaches provided valuable insights into the accuracy of parameter estimation for our model. Furthermore, we assessed the performance of the MATE distribution by applying it to real-world data from the second wave of COVID-19 in Nepal. The results of this application demonstrated that the MATE distribution offers superior fitting performance compared to several other commonly used lifetime models. This highlights the potential utility of the MATE as a valuable tool for analyzing lifetime data, particularly in complex and dynamic scenarios such as the COVID-19 pandemic. Our study focused on analyzing the suitability of the MATE distribution for modeling and understanding complex and dynamic scenarios, such as the spread of infectious diseases. This is especially significant in the context of public health, where understanding the dynamics of disease spread and accurately modeling its impact is crucial for effective decision-making and resource allocation.

In sum up, our research introduces the Modified Arctan Exponential distribution as a valuable addition to the toolbox of statistical models for analyzing lifetime data. The practical demonstration of its effectiveness in modeling COVID-19 data highlights its potential for addressing complex, real-world challenges. As we continue to refine and adapt statistical tools for a rapidly changing world, the MATE distribution stands out as a versatile and powerful option for a wide range of applications. Through rigorous data analysis and the application of this distribution model, we aspire to provide valuable insights that can inform healthcare professionals, policymakers, and the general public. These insights may contribute to more informed decision-making, better resource allocation, and improved strategies for mitigating the effects of future waves of the COVID-19 pandemic in Nepal.

References

[1] Marshall, A. W., and Olkin, I. (2007). A new family of distributions for lifetimes. Journal of Statistical Planning and Inference, 137(10), 3285-3290.

[2] Gupta, R. D., & Kundu, D. (1999). Theory & methods: Generalized exponential distributions. Australian & New Zealand Journal of Statistics, 41(2), 173-188.

[3] Pham, H., & Lai, C. D. (2007). On recent generalizations of the Weibull distribution. IEEE transactions on reliability, 56(3), 454-458.

[4] Smith, R. M., & Bain, L. J. (1975). An exponential power life-testing distribution. Communications in Statistics-Theory and Methods, 4(5), 469-481.

[5] Cordeiro, G. M., Ortega, E. M., & Lemonte, A. J. (2014). The exponential–Weibull life-

time distribution. Journal of Statistical Computation and simulation, 84(12), 2592-2606. https://doi.org/10.1080/00949655.2013.797982

[6] Barreto-Souza, W., Santos, A. H., & Cordeiro, G. M. (2010). The beta generalized exponential distribution. Journal of statistical Computation and Simulation, 80(2), 159-172.

[7] Rasekhi, M., Alizadeh, M., Altun, E., Hamedani, G. G., Afify, A. Z., & Ahmad, M. (2017). THE MODIFIED EXPONENTIAL DISTRIBUTION WITH APPLICATIONS. Pakistan Journal of Statistics, 33(5).

[8] Khan, M. S., King, R., & Hudson, I. L. (2017). Transmuted generalized exponential distribution: A generalization of the exponential distribution with applications to survival data. Communications in Statistics-Simulation and Computation, 46(6), 4377-4398. https://doi.org/10.1080/03610918.2015.1118503

[9] Hassan, A. S., Mohamd, R. E., Elgarhy, M., & Fayomi, A. (2018). Alpha power transformed extended exponential distribution: properties and applications. Journal of Nonlinear Sciences and Applications, 12(4), 62-67.

[10] Almarashi, A. M., Elgarhy, M., Elsehetry, M. M., Kibria, B. M., & Algarni, A. (2019). A new extension of exponential distribution with statistical properties and applications. Journal of Nonlinear Sciences & Applications (JNSA), 12(3).

[11] Chaudhary, A. K., Sapkota, L. P., & Kumar, V. (2020). Truncated Cauchy power–exponential distribution: Theory and Applications. IOSR Journal of Mathematics (IOSR-JM), 16(6), 44-52. DOI: 10.9790/5728-1606014452

[12] Almarashi, A. M., Elgarhy, M., Elsehetry, M. M., Kibria, B. M., & Algarni, A. (2019). A new extension of exponential distribution with statistical properties and applications. Journal of Nonlinear Sciences & Applications (JNSA), 12(3). http://dx.doi.org/10.22436/jnsa.012.03.01

[13] Ribeiro-Reis, L. D. (2022). Truncated exponentiated-exponential distribution: A distribution for unit interval. Journal of Statistics and Management Systems, 25(8), 2061-2072. https://doi.org/10.1080/09720510.2022.2060613

[14] 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. DOI: https://doi.org/10.3126/njs.v6i01.50801

[15] Afify, A. Z., Gemeay, A. M., Alfaer, N. M., Cordeiro, G. M., & Hafez, E. H. (2022). Power-modified kies-exponential distribution: properties, classical and Bayesian inference with an application to engineering data. Entropy, 24(7), 883. https://doi.org/10.3390/e24070883

[16] Chaudhary, A. K., Sapkota, L. P., & Kumar, V (2023). Inverse Exponential Power distribution: Theory and Applications. International Journal of Mathematics, Statistics and OperationsResearch,3(1),175-185. https://doi.org/10.47509/IJMSOR.2023.v03i01.11

[17] Nadarajah, S., & Haghighi, F. (2011). An extension of the exponential distribution. Statistics, 45(6), 543-558.https://doi.org/10.1080/02331881003678678

[18] Gómez-Déniz, E., & Calderín-Ojeda, E. (2015). Modelling insurance data with the Pareto ArcTan distribution. ASTIN Bulletin: The Journal of the IAA, 45(3), 639-660. https://doi.org/10.1017/asb.2015.9

[19] Chaudhary, A. K., Sapkota, L. P., & Kumar, V. (2021). Some properties and applications of arctan generalized exponential distribution. International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), 10(1), 456-468.

[20] Alkhairy, I., Nagy, M., Muse, A. H., & Hussam, E. (2021). The Arctan-X family of distributions: Properties, simulation, and applications to actuarial sciences. Complexity, 1-14. https://doi.org/10.1155/2021/4689010

[21] Chaudhary, A. K., & Kumar, V. (2022). Half Cauchy-modified exponential distribution: properties and applications. Nepal Journal of Mathematical Sciences, 3(1), 47-58. https://doi.org/10.3126/njmathsci.v3i1.44125

[22] Telee, L. B. S., & Kumar, V. (2022). Some properties and applications of Half-Cauchy exponential extension distribution. Int J Stat Appl Math ;7(6):91-101. https://doi.org/10.22271/maths.2022.v7.i6b.902

[23] Al-saiary, Z. A., Bakoban, R. A., & Al-zahrani, A. A. (2019). Characteriza-

tions of the Beta Kumaraswamy Exponential Distribution. Mathematics, 8(1), 23. https://doi.org/10.3390/math8010023

[24] 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

[25] Worldometer (2023). Second wave of COVID-19 in Nepal. Retrieved from https://covid.ourworldindata.org/data/owid-covid-data.xlsx

[26] R Core Team (2022). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2023-04-07).

[27] Lemonte, A. J. (2013). A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. Computational Statistics & Data Analysis, 62, 149-170. https://doi.org/10.1016/j.csda.2013.01.011

[28] Ansari, S. I., & Nofal, Z. M. (2021). The Lomax exponentiated Weibull model. Japanese Journal of Statistics and Data Science, 4, 21-39. https://doi.org/10.1007/s42081-020-00073-0

[29] Sarhan, A. M., & Apaloo, J. (2013). Exponentiated modified Weibull extension distribution. Reliability Engineering & System Safety, 112, 137-144.

[30] Ogunsanya, A. S., Sanni, O. O., & Yahya, W. B. (2019). Exploring some properties of odd Lomax-exponential distribution. Annals of Statistical Theory and Applications (ASTA), 1, 21-30.

[31] 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