

Research Article

Modelling the COVID-19 Mortality Rate with a New Versatile Modification of the Log-Logistic Distribution

Abdisalam Hassan Muse ¹, Ahlam H. Tolba ², Eman Fayad,³ Ola A. Abu Ali,⁴ M. Nagy,^{5,6} and M. Yusuf ⁷

¹Department of Mathematics (Statistics Option) Programme, Pan African University, Institute of Basic Science, Technology and Innovation (PAUSTI), Nairobi 6200-00200, Kenya

²Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

³Department of Biotechnology, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁴Department of Chemistry, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁵Department of Statistics and Operation Research, Faculty of Science, King Saud University, Riyadh, Saudi Arabia

⁶Department of Mathematics, Faculty of Science, Fayoum University, Fayoum, Egypt

⁷Department of Mathematics, Faculty of Science, Helwan University, Cairo, Egypt

Correspondence should be addressed to Abdisalam Hassan Muse; muse.abdisalam@students.jkuat.ac.ke

Received 2 September 2021; Accepted 5 October 2021; Published 13 November 2021

Academic Editor: Ahmed Mostafa Khalil

Copyright © 2021 Abdisalam Hassan Muse et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The goal of this paper is to develop an optimal statistical model to analyze COVID-19 data in order to model and analyze the COVID-19 mortality rates in Somalia. Combining the log-logistic distribution and the tangent function yields the flexible extension log-logistic tangent (LLT) distribution, a new two-parameter distribution. This new distribution has a number of excellent statistical and mathematical properties, including a simple failure rate function, reliability function, and cumulative distribution function. Maximum likelihood estimation (MLE) is used to estimate the unknown parameters of the proposed distribution. A numerical and visual result of the Monte Carlo simulation is obtained to evaluate the use of the MLE method. In addition, the LLT model is compared to the well-known two-parameter, three-parameter, and four-parameter competitors. Gompertz, log-logistic, kappa, exponentiated log-logistic, Marshall–Olkin log-logistic, Kumaraswamy log-logistic, and beta log-logistic are among the competing models. Different goodness-of-fit measures are used to determine whether the LLT distribution is more useful than the competing models in COVID-19 data of mortality rate analysis.

1. Introduction

Models are at the heart of almost all statistical work. A statistical model is a classification of probability distributions. The distribution family can be parametric, semi-parametric, or nonparametric. Parametric models produce more efficient estimates with lower standard errors than nonparametric and semiparametric models [1], more specifically, if the distributional assumption is correct. In general, probability distributions have been widely used to model lifetime data in a variety of fields, particularly biomedical sciences and engineering. Because of the variability

of the data, the statistical models chosen have a significant impact on the quality of the modelling in terms of providing the best description of the phenomenon under consideration.

The generated family of distributions has a large influence on the quality of statistical analysis procedures, and much effort has gone into developing new statistical models. There are, however, a number of significant issues with actual data that do not fit into any of the commonly used mathematical models. As a result, the technique of extending a family of distributions by introducing new parameters is acknowledged in the statistical literature.

The log-logistic distribution is an excellent choice for analyzing data with unimodal or decreasing failure rates. However, in a number of situations where data behave monotonically, such as increasing failure rates, or non-monotonically, such as bathtub- or modified bathtub-shaped failure rates, the log-logistic model is not a good candidate model to use [2–5]. As a result, new extensions and generalizations to existing models are required for accurate and precise data modelling. As a result, numerous statistical techniques are intended to discover new modifications to classical models in order to achieve a better fit to the data of interest. Other techniques, on the other hand, such as the Sin-G family, Cos-G family, Tan-G family, and Sec-G family, provide a versatile generalization of the existing probability distribution without adding any extra parameters [6–10]. The techniques that are effectively used to extend the classical distributions without adding extra parameters are termed as “new trigonometric classes of probability distribution” [11].

A random variable X is said to have a log-logistic distribution with shape parameter $\beta > 0$ and scale parameter $\alpha > 0$, denoted by $X \sim \text{LL}(\alpha, \beta)$, if its cumulative distribution function (cdf) is defined by the following equation:

$$F(x; \varphi) = \frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta}, \quad x \geq 0, \varphi > 0. \quad (1)$$

The probability density function (pdf) is given by

$$f(x; \varphi) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2}, \quad x \geq 0, \varphi > 0. \quad (2)$$

The reliability (survival) function is given by

$$S(x; \varphi) = 1 - F(x; \varphi) = \frac{(x/\alpha)^{-\beta}}{1 + (x/\alpha)^{-\beta}}, \quad x \geq 0, \varphi > 0. \quad (3)$$

The failure (hazard) rate function is given by

$$\begin{aligned} h(x; \varphi) &= \frac{f(x; \varphi)}{S(x; \varphi)} = \frac{f(x; \varphi)}{1 - F(x; \varphi)} \\ &= \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{1 + (x/\alpha)^\beta}, \quad x \geq 0, \varphi > 0. \end{aligned} \quad (4)$$

The reversed hazard rate function (also known as the retro function) is given by

$$\begin{aligned} r(x; \varphi) &= \frac{f(x; \varphi)}{F(x; \varphi)} \frac{\left[(\beta/\alpha)(x/\alpha)^{\beta-1} / (1 + (x/\alpha)^\beta)^2 \right]}{\left[(x/\alpha)^\beta / (1 + (x/\alpha)^\beta) \right]} \\ &= \frac{(\theta/\alpha)(x/\alpha)^{-1}}{1 + (x/\alpha)^\beta}; \quad x \geq 0, \varphi > 0. \end{aligned} \quad (5)$$

The cumulative hazard rate function is given by

$$H(x) = -\log S(x; \varphi) = -\log \left(\frac{(x/\alpha)^{-\beta}}{1 + (x/\alpha)^{-\beta}} \right), \quad x \geq 0, \varphi > 0, \quad (6)$$

where $\phi = (\alpha, \beta)'$ is a vector of parameters.

It is well understood that the classical log-logistic distribution fails to capture the accurate phenomenon under investigation in many cases. As a result, several modifications have been proposed and researched. Inducting one or more parameters to the classical log-logistic distribution yields a modified form of the log-logistic distribution. When compared to the classic log-logistic distribution, several of these modified distributions have been found to be more flexible and adaptable of modelling real-life data. An up-to-date survey of recent modifications of the log-logistic distribution can be found in [12].

Most modifications of the log-logistic model in the statistical literature have been derived by adding extra parameters to control the shape of the skewness (or asymmetry) and the kurtosis of the distribution; see the exponentiated LL distribution [13], beta LL distribution [14], gamma LL distribution [15], Marshall–Olkin LL distribution [16], transmuted LL distribution [17], cubic transmuted LL distribution [18], McDonald LL distribution [19], and alpha power transformed LL distribution [20, 21]. Other generalizations and modifications of the log-logistic distribution developed recently can be seen in [22, 23].

The majority of techniques for extending the classical LL model produce a heavy-tailed distribution. Unfortunately, the abovementioned generalizations of the log-logistic distribution have some limitations. For instance, (i) adding extra parameter(s) to the distribution enhances flexibility, but even so, such practices usually result in reparameterization issues; (ii) the number of model parameters is increased, causing difficulty in estimating the model's parameters; (iii) some extending techniques reduce the tractability of the cdf, making manual computation of statistical properties more difficult; and (iv) other generalization techniques complicate the pdf, resulting in computational issues; incorporating new extra parameters into existing models increases flexibility, which is a desirable feature. On the other hand, it makes inferences more difficult [24, 25]. As a result, in this study, we modified the classical log-logistic distribution with a two-parameter continuous distribution referred to the log-logistic tangent distribution (LLT).

The primary goal of this current study is it to introduce and investigate a versatile modification of the log-logistic distribution using the Tan-G family generator method. As previously stated, the proposed two-parameter distribution may be more flexible than other popular log-logistic model extensions. It would eventually become clear that the proposed distribution family accommodates both monotone and nonmonotone hazard rates.

The rest of the paper is organized as follows. The proposed family is discussed in Section 2. Section 3 presents the proposed distribution. Some mathematical and statistical

properties of the proposed distribution are discussed in Section 4. Section 5 presents the estimation of the unknown parameters of the model. Section 6 discusses a Monte Carlo simulation for the proposed model. In Section 7, an application to a real-life dataset is presented and discussed. The comparison of some of the parametric probability distributions and their submodels is presented. Finally, concluding remarks and a work summary are presented in Section 8.

2. The Proposed Family

Several continuous probability distributions involving trigonometric functions have been developed in the statistics literature by many researchers, with the tangent distribution being notable due to its wide range of applications in many real-life datasets from various disciplines. A useful survey in this context can be found in [11, 24–27]. In this study, we concentrate on the trigonometric classes of continuous probability distributions, which are described by a cdf concerning trigonometric functions (tangent, cosine, sine, secant, and different mixtures of them). The Tan-G family developed in [11, 26–29] is the fundamental work.

As indicated by its name, Souza et al. [28] introduced a new method for extending the classical probability distributions, leading to greater flexibility in analyzing and modelling different data types. They looked at a parent distribution, which is an arbitrary continuous probability distribution with a cdf and a corresponding pdf. The cumulative distribution function (cdf) is defined by the associated tangent family of distributions.

$$F_{\text{Tan}}(x) = \int_0^{\pi/4G(x)} \sec^2(t) dt. \quad (7)$$

Equation (7) can be written as

$$F_{\text{Tan}}(x) = \tan\left(\frac{\pi}{4} G(x)\right), \quad (8)$$

where $G(x)$ is the cdf of the parent (baseline) distribution, and if $G(x)$ has pdf $g(x)$, the pdf of the class is given by

$$f_{\text{Tan}}(x) = \frac{\pi}{4} g(x) \sec^2\left(\frac{\pi}{4} G(x)\right). \quad (9)$$

The survival function is obtained by

$$S_{\text{Tan}}(x) = 1 - \tan\left(\frac{\pi}{4} G(x)\right). \quad (10)$$

The failure rate function is given by

$$h_{\text{Tan}}(x) = \frac{f_{\text{Tan}}(x)}{S_{\text{Tan}}(x)} = \frac{\pi/4 g(x) \sec^2(\pi/4G(x))}{1 - \tan(\pi/4 G(x))}. \quad (11)$$

The reversed hazard rate function is given by

$$r_{\text{Tan}}(x) = \frac{f_{\text{Tan}}(x)}{F_{\text{Tan}}(x)} = \frac{\pi/4 g(x) \sec^2(\pi/4G(x))}{\tan(\pi/4 G(x))}. \quad (12)$$

Also, the cumulative hazard rate function is given by

$$H_{\text{Tan}}(x) = -\log S_{\text{Tan}}(x) = -\log\left\{1 - \tan\left(\frac{\pi}{4} G(x)\right)\right\}. \quad (13)$$

Recently, Souza et al. [28] proposed and studied the Burr-XII tangent distribution which serves as a potential lifetime model. Ampadu [29] introduced and studied the Weibull tangent distribution with applications to health science data. Hence, the Tan-G family is a special family for extending the classical well-known lifetime distributions without adding an extra parameter.

3. The Log-Logistic Tangent Distribution

In this section, the LLT distribution has been examined, considering that $G(x)$ is the cdf of the log-logistic distribution.

The distribution function (cdf) of the log-logistic tangent distribution, for $x > 0$, can be expressed as

$$F_{\text{Tan-LL}}(x; \phi) = \tan\left\{\frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta}\right)\right\}. \quad (14)$$

The corresponding probability density function to the abovementioned cdf is given by

$$f_{\text{Tan-LL}}(x; \phi) = \frac{\pi}{4} \left\{ \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2} \right\} \sec^2\left\{\frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta}\right)\right\}. \quad (15)$$

The survival function is expressed as follows:

$$S_{\text{Tan-LL}}(x; \phi) = 1 - \tan\left\{\frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta}\right)\right\}. \quad (16)$$

The hazard rate function is obtained by

$$h_{\text{Tan-LL}}(x; \phi) = \frac{\pi/4 \left\{ (\beta/\alpha)(x/\alpha)^{\beta-1} / (1 + (x/\alpha)^\beta)^2 \right\} \sec^2\left\{ \pi/4 \left((x/\alpha)^\beta / (1 + (x/\alpha)^\beta) \right) \right\}}{1 - \tan\left\{ \pi/4 \left((x/\alpha)^\beta / (1 + (x/\alpha)^\beta) \right) \right\}}. \quad (17)$$

The reversed hazard rate function is expressed as follows:

$$r_{\text{Tan-LL}}(x; \phi) = \frac{\pi/4 \left\{ (\beta/\alpha) (x/\alpha)^{\beta-1} / (1 + (x/\alpha)^\beta)^2 \right\} \sec^2 \left\{ \pi/4 \left((x/\alpha)^\beta / 1 + (x/\alpha)^\beta \right) \right\}}{\tan \left\{ \pi/4 \left((x/\alpha)^\beta / 1 + (x/\alpha)^\beta \right) \right\}}. \quad (18)$$

The cumulative hazard function can be given as follows:

$$H_{\text{Tan-LL}}(x) = -\log S_{\text{Tan-LL}}(x) = -\log \left[1 - \tan \left\{ \frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta} \right) \right\} \right]. \quad (19)$$

With $x \geq 0$, $\alpha, \beta > 0$, and $\varphi = (\alpha, \beta)$ is the vector parameter in all of the abovementioned equations, respectively.

Some possible shapes of the pdf, cdf, hazard function, survival function, and the reversed hazard functions of the log-logistic tangent distribution are displayed in Figures 1–7.

The hazard rate function of the proposed distribution can accommodate for both monotone (increasing and decreasing) and nonmonotone (i.e., unimodal) hazard rates, as it can be seen in Figures 4–6, and Figure 7 represents the reversed hazard function of the LLT distribution.

4. Some Statistical Properties of the Proposed Distribution

In this section, we use numerical examples to derive some mathematical properties of the log-logistic tangent distribution, such as the quantile function, skewness and kurtosis, moments, and residual and reverse residual life.

4.1. Quantile Function. For this model, the quantile function of the LLT distribution is used in theoretical aspects of distribution theory such as statistical simulations and applications. To generate random samples, the simulation algorithm used a quantile function by following the steps in Algorithm 1.

Let X be the Tan-G-distributed random variable. The quantiles can be utilized to obtain data of the distribution according to

$$x = Q(u) = F^{-1}(u) = G^{-1} \left[\frac{4}{\pi} \arctan(u) \right]. \quad (20)$$

The LLT distribution's quantile function is given by

$$Q_{\text{LLT}}(u) = \alpha \left(\frac{4/\pi \arctan(u)}{1 - 4/\pi \arctan(u)} \right)^{\frac{1}{\beta}}. \quad (21)$$

The lower quartile (Q1), median (Q2), and upper quartile (Q3) of LLT distribution can be derived from the equation of the quantile function by setting $u = 1/4, 1/2$, and $3/4$, respectively.

The quantiles of the LLT distribution for some parameter values are shown in Table 1.

4.2. Skewness and Kurtosis. Some of the properties of the continuous distribution can be studied through its asymmetry and kurtosis. The mathematical form of the Moors

Kurtosis and Galton asymmetry (or skewness) of the LLT model with two parameters is defined by the following relationship:

$$S_K = \frac{Q(3/4) + Q(1/4) - 2Q(2/4)}{Q(3/4) - Q(1/4)}, \quad (22)$$

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

where Q describes different quartile values.

The equations above can be solved as functions of the LLT quantile function. These measures have the advantage of being less sensitive in the presence of outliers and existing even when the distribution is devoid of moments.

4.3. Moments. Moments are essential in statistical modeling, particularly in role in applications. The LLT distribution's r th moment is defined as

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x; \alpha, \beta) dx. \quad (23)$$

In fact, we have

$$\mu'_r = \int_{-\infty}^{\infty} x^r \frac{\pi}{4} \left\{ \frac{(\beta/\alpha) (x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2} \right\} \sec^2 \left\{ \frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta} \right) \right\} dx. \quad (24)$$

The first five moments followed by the standard deviation, coefficient of variance, skewness, and kurtosis for some parameter values are shown in Table 2.

4.4. Residual Life and Reverse Residual Life. The residual lifetime function (rlf) has broader uses in risk management and survival analysis. The rlf of the LLT random variable (r.v.) can be expressed as follows:

$$R_{(t)}(x) = \frac{S(x+t)}{S(t)},$$

$$R_{(t)}(x) = \frac{1 - \tan \left\{ \pi/4 \left((x+t/\alpha)^\beta / (1 + ((x+t)/\alpha)^\beta) \right) \right\}}{1 - \tan \left\{ \pi/4 \left((x/\alpha)^\beta / (1 + (x/\alpha)^\beta) \right) \right\}}. \quad (25)$$

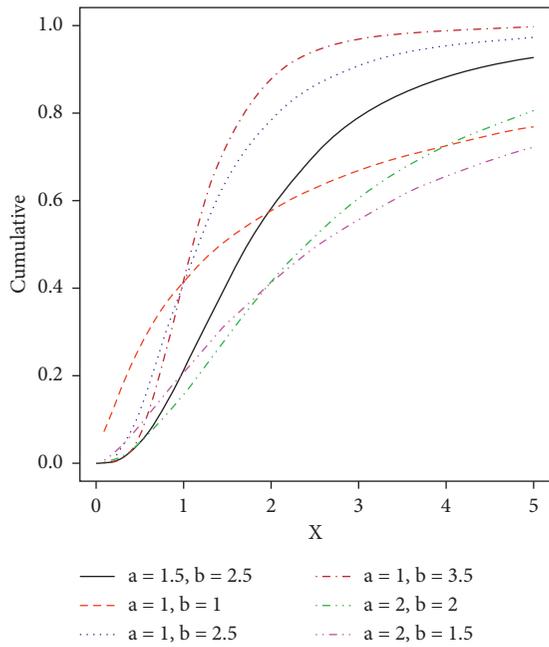


FIGURE 1: Visual display plots of the cdf of the LLT distribution.

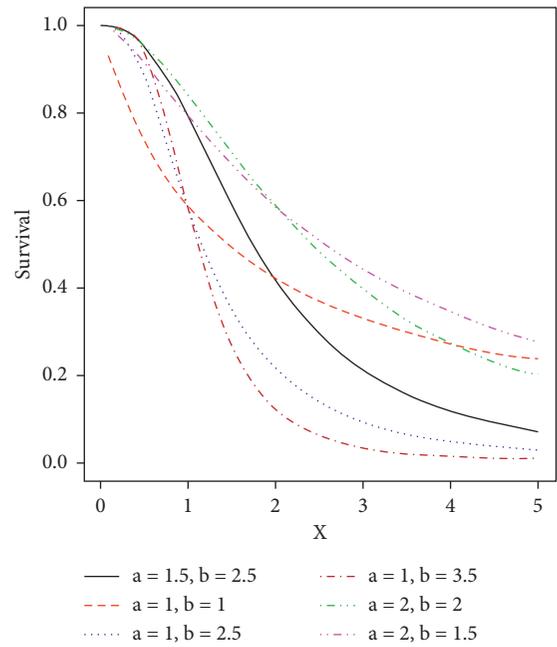


FIGURE 3: Plots of the survival function of the LLT distribution.

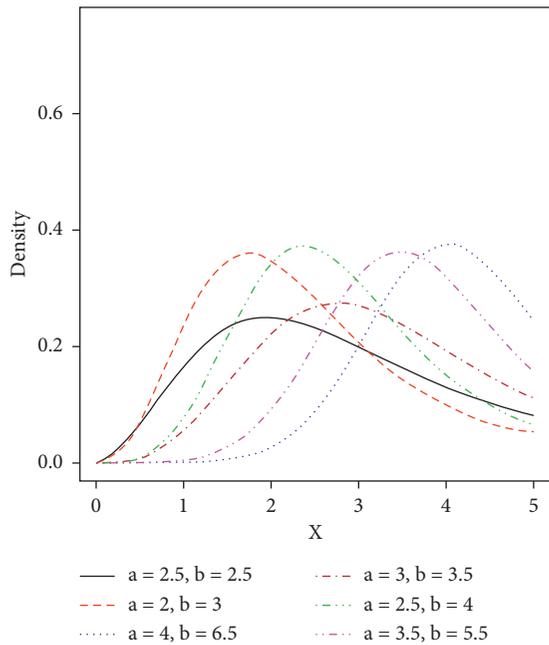


FIGURE 2: Visual display plots of the pdf of the LLT distribution.

Furthermore, the reverse residual lifetime function (rrlf) of the LLT r.v. can be obtained as follows:

$$\widehat{R}_{(t)}(x) = \frac{S(x+t)}{S(t)},$$

$$\widehat{R}_{(t)}(x) = \frac{1 - \tan\left\{\frac{\pi}{4} \left(\frac{(x-t)/\alpha^\beta}{1 + ((x-t)/\alpha)^\beta} \right)\right\}}{1 - \tan\left\{\frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1 + (x/\alpha)^\beta} \right)\right\}}. \tag{26}$$

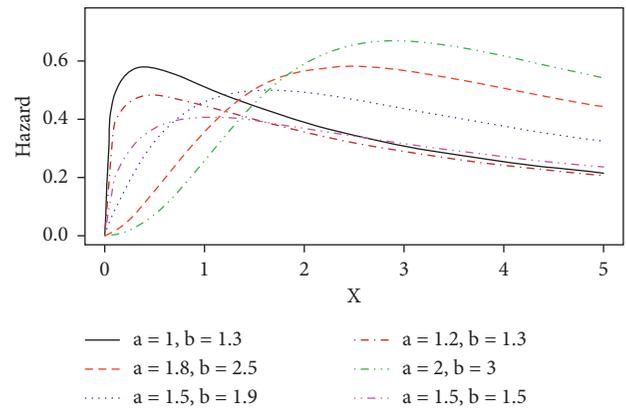


FIGURE 4: Unimodal hazard rate of the LLT distribution.

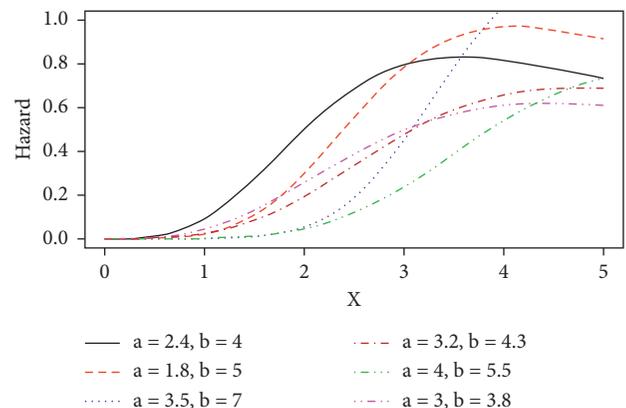


FIGURE 5: Plots of the LLT increasing hazard rate.

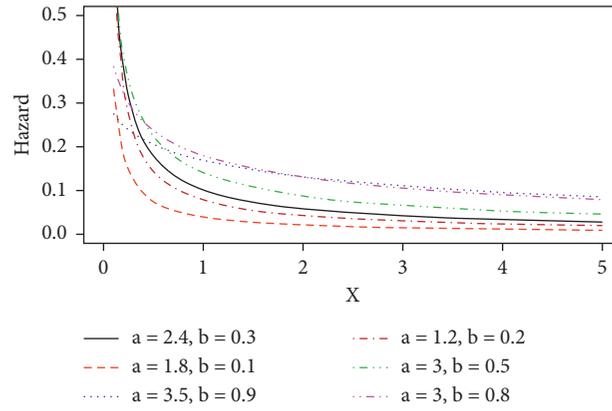


FIGURE 6: Plots of the LLT decreasing hazard rate.

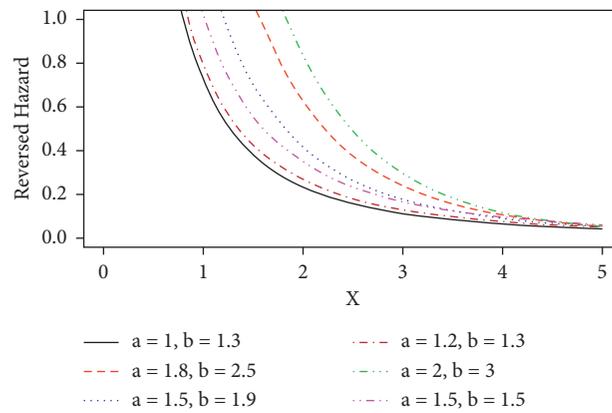


FIGURE 7: Plots of the retro hazard function of the LLT distribution.

(1) Generate $u \sim U(0, 1)$
 (2) Specify $G^{-1}(\cdot)$
 (3) Obtain an outcome of X by $X = Q(u)$

ALGORITHM 1: Random generator for the Tan-G class.

TABLE 1: The quantiles of the LLT distribution for some values of the parameters.

Quantiles	(α, β)				
	(0.5, 1.5)	(0.2, 0.9)	(0.3, 0.8)	(0.1, 1.2)	(0.05, 0.9)
0.1	0.1382	0.0235	0.0269	0.0201	0.0059
0.2	0.2415	0.0595	0.0767	0.0403	0.0149
0.3	0.3518	0.1113	0.1551	0.0644	0.0278
0.4	0.4797	0.1867	0.2776	0.0950	0.0467
0.5	0.6379	0.3001	0.4737	0.1356	0.0751
0.6	0.8473	0.4818	0.8065	0.1933	0.1204
0.7	1.1518	0.8036	1.4343	0.2838	0.2009
0.8	1.6688	1.4908	2.8745	0.4511	0.3727
0.9	2.8956	3.7351	8.0781	0.8983	0.9338

TABLE 2: First five moments, SD, CV, asymmetry, and kurtosis of the LLT distribution for different parameter values.

Moments	(α, β)				
	(1.1, 0.)	(1.2, 0.9)	(0.3, 0.8)	(0.1, 1.2)	(0.2, 0.4)
μ_1'	0.1383	0.1445	0.1727	0.1772	0.0920
μ_2'	0.0804	0.0862	0.0914	0.0744	0.0471
μ_3'	0.0562	0.0608	0.0611	0.0440	0.0314
μ_4'	0.0431	0.0469	0.0457	0.0306	0.0236
μ_5'	0.0349	0.0381	0.0364	0.0234	0.0188
SD	0.2476	0.2555	0.2482	0.2074	0.1965
CV	1.7906	1.7685	1.4377	1.1709	2.1388
CS	1.8534	1.7702	1.5728	1.7408	2.6374
CK	5.3540	4.9762	4.5195	5.6950	9.5119

5. Estimation of the Parameters

The maximum likelihood approach is used in this section to estimate the unknown parameters of the log-logistic tangent distribution based on a complete sample. Let $x = \{x_1, x_2, \dots, x_n\}^T$ represent n independent random variables drawn from the LLT distribution. The sample's likelihood function is defined as

$$L = \prod_{i=1}^n f(x_i, \alpha, \beta),$$

$$L(x; \alpha, \beta) = \prod_{i=1}^n \frac{\pi}{4} \left\{ \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1+(x/\alpha)^\beta)^2} \right\} \sec^2 \left\{ \frac{\pi}{4} \left(\frac{(x/\alpha)^\beta}{1+(x/\alpha)^\beta} \right) \right\}. \quad (27)$$

The log-likelihood function can be written as follows:

$$\begin{aligned} \ell = n \log \frac{\pi}{4} + n \log \beta - n\beta \log \alpha + (\beta - 1) \sum_{i=1}^n \log(x_i) \\ - 2 \sum_{i=1}^n \log \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right] \\ + 2 \sum_{i=1}^n \log \left[\sec \left\{ \frac{\pi}{4} \left(\left(\frac{x_i}{\alpha} \right)^\beta \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-1} \right) \right\} \right]. \end{aligned} \quad (28)$$

By taking the 1st derivatives of the log-likelihood function in equation (28) with respect to α and β parameters and setting the result to zero, we get

$$\begin{aligned} \frac{\partial \ell}{\partial \alpha} &= -\frac{n\beta}{\alpha} + \frac{2\beta}{\alpha} \sum_{i=1}^n \left(\frac{x_i}{\alpha} \right)^\beta \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-1} - \frac{\pi}{2} \sum_{i=1}^n \left(\frac{x_i}{\alpha} \right)^\beta \left(\frac{\beta}{\alpha} \right) \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-2} \tan \left\{ \frac{\pi}{4} \left(\left(\frac{x_i}{\alpha} \right)^\beta \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-1} \right) \right\}, \\ \frac{\partial \ell}{\partial \beta} &= \frac{n}{\beta} - n \log(\alpha) + \sum_{i=1}^n \log(x_i) - 2 \sum_{i=1}^n \left(\frac{x_i}{\alpha} \right)^\beta \log \left(\frac{x_i}{\alpha} \right) \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-1} \\ &\quad + \frac{\pi}{2} \sum_{i=1}^n \left(\frac{x_i}{\alpha} \right)^\beta \log \left(\frac{x_i}{\alpha} \right) \left(\frac{x_i}{\alpha} \right) \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-2} \tan \left\{ \frac{\pi}{4} \left(\left(\frac{x_i}{\alpha} \right)^\beta \left[1 + \left(\frac{x_i}{\alpha} \right)^\beta \right]^{-1} \right) \right\}. \end{aligned} \quad (29)$$

It is worth noting that the MLE's $\hat{\alpha}$ and $\hat{\beta}$ of α and β , respectively, can be achieved by equating the outcomes to zero and solving the system of nonlinear equations numerically. The well-known theory on MLE can be used under some standard regularity conditions, ensuring nice asymptotic properties [30].

6. Simulation Study

In this section, we use Monte Carlo (MC) simulation to evaluate the effectiveness of the Maximum Likelihood Estimation (MLE) method for estimating the log-logistic tangent distribution parameters. The simulation study is carried out to investigate the average bias (AB), mean square error (MSE), and root mean square error (RMSE) for the parameters of the proposed model. The simulation experiment was conducted by running a number of simulations with varying sample sizes and parameter values. The quantile function given in Eq. (21) was used to generate random samples for the LLT. The MC simulation study was iterated 750 times with sample sizes $n = 25, 50, \dots, 750$ and the parameter scenarios $\alpha = 2.0$ and $\beta = 1.5$ in set I and $\alpha = 3.0$ and $\beta = 2.0$ in set II.

The MLEs are ascertained for each item of simulated data, say $(\hat{\alpha}, \hat{\beta})$ for $i = 1, 2, \dots, 750$, and the AB, MSEs, and RMSEs of the parameters were calculated by

$$\begin{aligned} \text{AB} &= \frac{1}{N} \sum_{i=1}^N (\hat{\theta} - \theta), \\ \text{MSE} &= \frac{1}{N} \sum_{i=1}^N (\hat{\theta} - \theta)^2, \end{aligned} \quad (30)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\theta} - \theta)^2},$$

where $\theta = \alpha$ and β .

Table 3 shows the AB, MSE, and RMSE values of the parameters for various sample sizes. Based on these results, we draw the conclusion that MLEs do a good job of estimating parameters and that the estimates seem to be reasonably constant and closer to the true values for these sample sizes. In addition, Table 4 shows that the RMSE decreases as the sample size increases, as expected. Furthermore, as sample size increases, so does the AB. As a result, even if the sample size is small, the MLEs and their asymptotic results can be used to calculate confidence intervals for model parameters.

The simulation results for the aforementioned measures are depicted in Figures 8 and 9. These plots demonstrate that increasing the sample size n reduces the estimated biases. Furthermore, as the sample size n is increased, the estimated

TABLE 3: MC simulation outcomes for the proposed distribution using the values for MLEs, ABs, MSEs, and RMSEs.

Parameters	n	I				II			
		MLE	AB	MSE	RMSE	MLE	AB	MSE	RMSE
α	25	2.059	0.059	0.273	0.522	3.060	0.060	0.368	0.607
	50	2.030	0.030	0.117	0.341	3.022	0.022	0.142	0.378
	100	2.010	0.010	0.066	0.256	3.017	0.017	0.086	0.294
	200	2.008	0.008	0.030	0.172	3.013	0.013	0.041	0.202
	400	2.004	0.004	0.015	0.124	3.008	0.008	0.019	0.137
	600	2.003	0.003	0.010	0.100	3.003	0.003	0.012	0.108
	700	2.002	0.002	0.007	0.086	3.002	0.002	0.009	0.098
β	25	1.584	0.084	0.079	0.282	2.124	0.124	0.156	0.395
	50	1.541	0.041	0.035	0.187	2.060	0.060	0.065	0.255
	100	1.518	0.018	0.017	0.130	2.023	0.023	0.029	0.173
	200	1.513	0.013	0.008	0.089	2.011	0.011	0.014	0.120
	400	1.505	0.005	0.004	0.062	2.007	0.007	0.007	0.085
	600	1.502	0.002	0.003	0.050	2.004	0.004	0.004	0.068
	700	1.501	0.001	0.002	0.044	2.003	0.003	0.003	0.058

TABLE 4: Descriptive measures of data I.

N	Min.	Median	Mode	Variance	Asymmetry	Kurtosis	Mean	Max.	CV	Geo. mean
51	2.857	5.882	5.5	5.962	0.450	-0.732	6.793	11.81	0.359	6.363

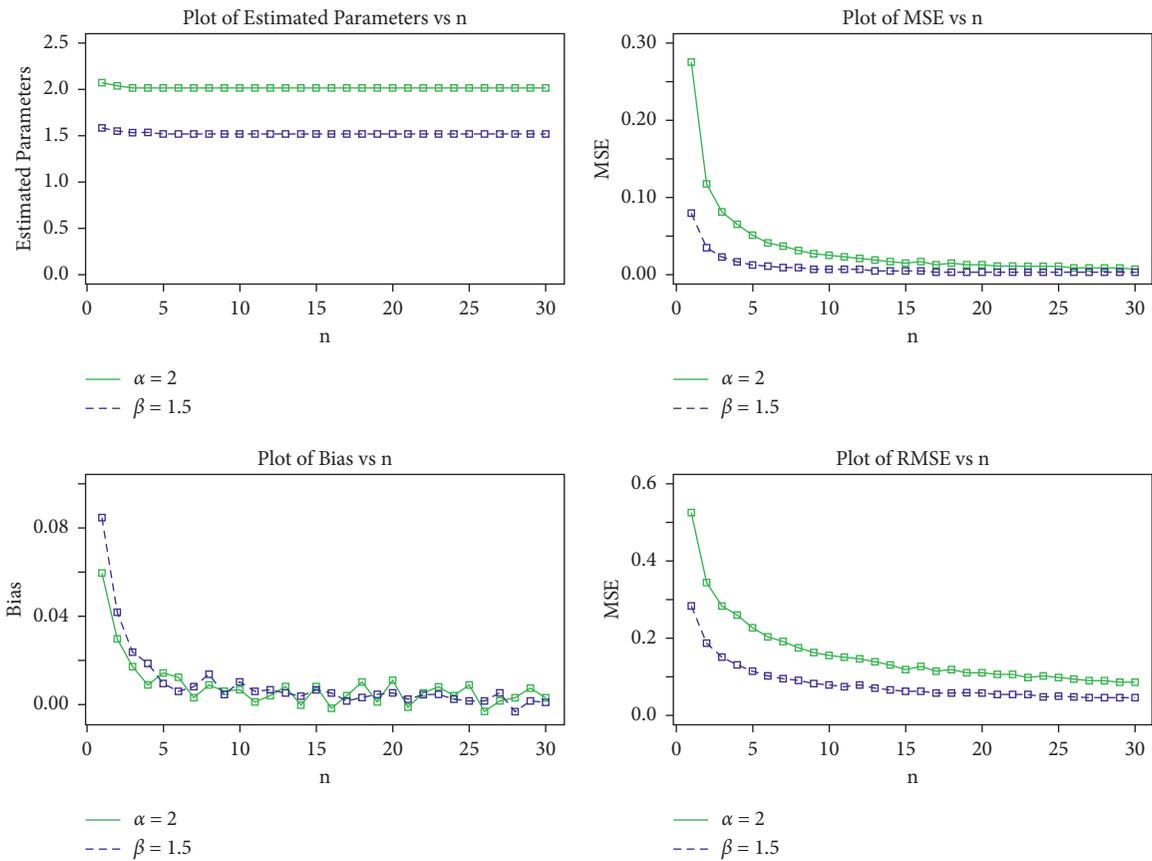


FIGURE 8: MLEs, MSEs, biases, and RMSE plots of the proposed distribution for set I values.

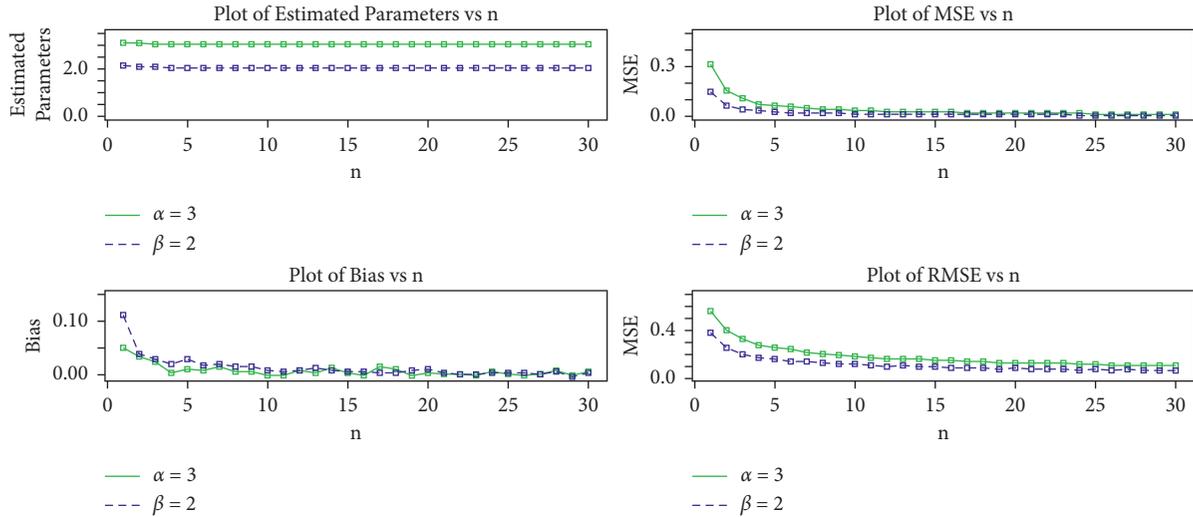


FIGURE 9: MLEs, MSEs, biases, and RMSE plots of the proposed distribution for set II values.

MSEs and RMSEs decay toward zero. These findings demonstrate the MLEs' efficiency as well as their consistency.

7. Applications to the COVID-19 Dataset

The log-logistic tangent distribution derivation's main interest is its application in data analysis goals, which makes it valuable in different disciplines, especially those associated with mortality data analysis. A number of proposed distributions for COVID-19 datasets have recently been proposed; for more information on these, please see [31–39].

The log-logistic tangent distribution was applied to a real-life COVID-19 mortality rate dataset from Somalia for demonstration purposes, and its performance was compared to the performance of other fitted models such as the log-logistic, Weibull, Gompertz, and kappa distributions. For the log-likelihood, the most common information criteria including Akaike Information Criterion (AIC), Hannan–Quinn Information Criterion (HQIC), Corrected Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC) values were used to select the most appropriate. The Anderson–Darling (A^*) statistic, the Cramer–von Mises (W^*) distance value, and the Kolmogorov–Smirnov ($K-S$) statistic, as well as the corresponding p value, are all recorded.

(1) Log-logistic:

$$f(x) = \frac{\beta/\alpha(x/\alpha)^{\beta-1}}{[1+(x/\alpha)^\beta]^2}, \quad \text{for } x \geq 0. \quad (31)$$

(2) Gompertz distribution:

$$f(x) = \alpha\beta \exp(\beta + \alpha x - \beta e^{\alpha x}), \quad \text{for } x \geq 0. \quad (32)$$

(3) Kappa distribution:

$$f(x) = \frac{\alpha}{\beta} \left[\alpha + \left(\frac{x}{\beta} \right)^\alpha \right]^{-(\alpha+1)/\alpha}, \quad \text{for } x \geq 0. \quad (33)$$

See also the other four recently modified log-logistic distributions with three parameters and four parameters [14, 16, 40, 41].

The AIC is

$$\text{AIC} = 2k - 2l. \quad (34)$$

The BIC is

$$\text{BIC} = k \ln(n) - 2l. \quad (35)$$

The CAIC is

$$\text{CAIC} = \frac{2nk}{n-k-1} - 2l. \quad (36)$$

The HQIC is

$$\text{HQIC} = 2k \ln(\ln(n)) - 2l, \quad (37)$$

where the log-likelihood function is represented by l , the sample size by n , and the number of model parameters by k . The following goodness-of-fit measures are being considered.

The Anderson–Darling (A^*) test statistic is given by

$$A^* = -n - \frac{1}{n} \sum_{i=1}^n (2l - 1) \times [\ln G(X_i) + \ln\{1 - G(X_{n-i+1})\}], \quad (38)$$

The Cramer–von Mises (W^*) test statistics is given by

$$W^* = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{2i-1}{2n} + G(X_i) \right]^2, \quad (39)$$

where x_i is the i th observation in the sample and n is the sample size. When the data are sorted in ascending order, x_i is calculated.

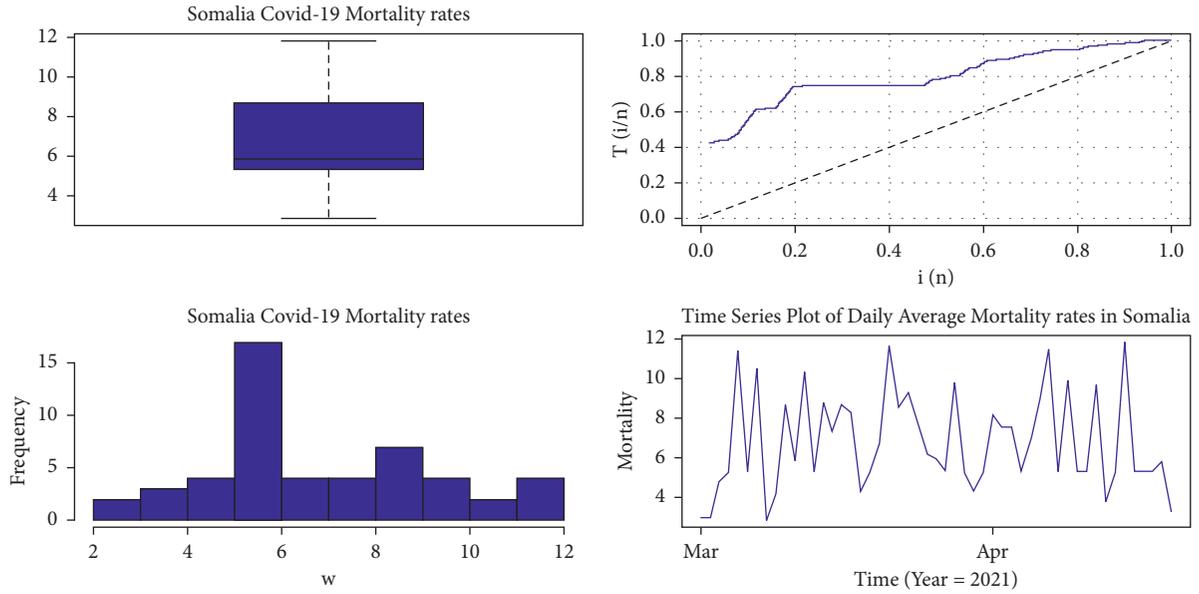


FIGURE 10: Box plot, TTT plot, histogram, and time-series plot for dataset I.

TABLE 5: Likelihood and the information criteria for data I.

Distribution	ℓ	AIC	BIC	CAIC	HQIC
LLT	-116.839	236.677	240.541	237.927	238.154
LL	-118.198	238.997	241.860	239.246	239.473
Gompertz	-120.495	244.991	248.854	245.8241	246.467
Kappa	-127.196	258.391	262.255	258.641	259.867
Expo LL	-116.995	239.991	245.786	240.501	242.205
Beta LL	-115.559	239.119	246.846	239.988	242.071
Kumw LL	-115.970	239.940	247.667	240.810	242.892
Mo-LL	-116.998	239.996	245.792	240.507	242.212

The best model has the lowest AIC, CAIC, BIC, and HQIC, as well as the lowest A^* , W^* , and $K-S$ values. In addition, the best model is chosen as the one with the highest log-likelihood function value, and p values for the $K-S$ statistics are applied to compare the competitive distributions.

7.1. Data I: Somalia COVID-19 Mortality Rate Dataset. The dataset contains COVID-19 mortality rate from Somalia during the time between 1st March 2021 to 20th April 2021 (see <https://covid19.who.int/>). These data are made up of a rough mortality rate. The data frame contains 51 observations, which are formed by using the daily cumulative cases (DCCs) and daily new deaths (DNDs). The data are as follows:

3.008, 2.963, 4.762, 5.263, 11.382, 5.330, 10.471, 2.857, 4.274, 8.633, 5.882, 10.280, 5.330, 8.730, 7.377, 8.696, 8.257, 4.294, 5.330, 6.769, 11.627, 8.547, 9.302, 7.742, 6.173, 6.015, 5.330, 9.770, 5.330, 4.294, 5.330, 8.120, 7.547, 7.563, 5.330, 6.875, 8.800, 11.429, 5.330, 9.898, 5.330, 5.330, 9.630, 3.750, 5.330, 11.808, 5.330, 5.330, 5.330, 5.785, and 3.265.

7.2. Exploratory Data Analysis. The primary goal of data analysis is to extract information from the data. We used five different techniques to perform exploratory data analysis in this work: (1) data descriptive statistics; (2) Box plot; (3) TTT plot; (4) histogram; and (5) time-series plot. Figure 10 displays the Box, TTT, histogram, and times-series plots for our dataset.

The total-time-on-test (TTT) plot is a graphical method for determining the failure rate function's shape. There is qualitative information about the shape of the failure rate function in many real-world applications that can aid in the selection of a particular distribution. The TTT plot for our dataset in this work is shown, and it shows an increasing failure rate shape.

Table 4 shows us the descriptive statistics of the COVID-19 data for mortality rate between 1st March to 20 April 2021 by computing specific aspects of the data (central tendency and spread).

The proposed LLT distribution has the lowest AIC, CAIC, BIC, and HQIC values and the highest log-likelihood values in Table 5, so it is selected as the most suitable model among the competing distributions considered in this work.

TABLE 6: MLE model parameter estimators and goodness-of-fit tests for the COVID-19 data.

Distributions	Estimates	W^*	A^*	$K - S$ (P value)
LLT (δ, θ)	$\delta = 5.904$ $\theta = 4.814$	0.199	1.067	0.169 (0.206)
llog (α, β)	$\alpha = 6.388$ $\beta = 4.633$	0.203	1.089	0.169 (0.109)
Gompertz (α, β)	$\alpha = 0.365$ $\beta = 0.021$	0.221	1.266	0.181 (0.016)
Kappa (α, β)	$\alpha = 58.227$ $\beta = 11.071$	0.322	1.856	0.267 (0.002)
Exp LL (α, β, γ)	$\alpha = 1.050$ $\beta = 4.561$ $\gamma = 6.285$	0.210	1.085	0.197 (0.113)
Beta LL ($\alpha, \beta, \gamma, \delta$)	$\alpha = 5.431$ $\beta = 20.368$ $\gamma = 1.360$ $\delta = 17.705$	0.214	1.075	0.178 (0.104)
Kum LL ($\alpha, \beta, \gamma, \delta$)	$\alpha = 2.319$ $\beta = 5.749$ $\gamma = 2.182$ $\delta = 7.903$	0.201	1.071	0.161 (0.101)
Mo-LL (α, β, γ)	$\alpha = 6.064$ $\beta = 4.633$ $\gamma = 4.330$	0.203	1.089	0.169 (0.109)

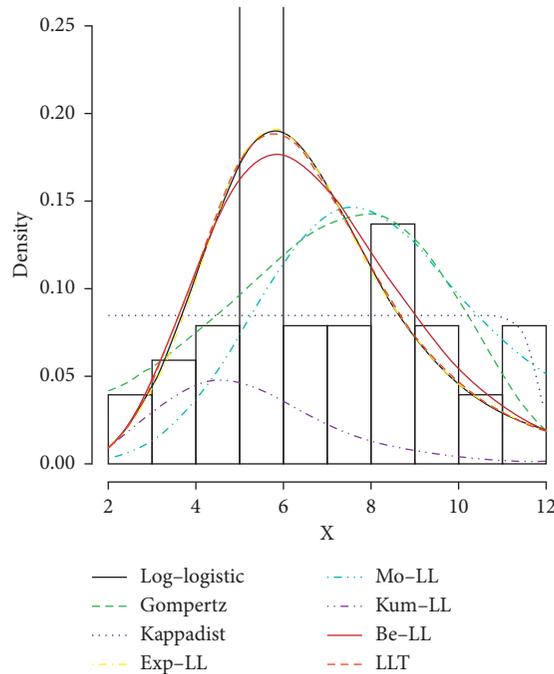


FIGURE 11: Some fitted pdf's of the fitted distributions in dataset I.

Table 6 shows the parameter estimate and p value for the Cramer-von Mises (W^*), Anderson-Darling (A^*), and Kolmogorov-Smirnov ($K-S$) tests for all competing distributions using abovementioned dataset I. According to Table 6, the proposed LLT distribution has the lowest A^* , W^* , and $K-S$ tests, as well as the highest p value. As a result, among the competing distributions considered in this study,

the proposed LLT distribution is chosen as the most appropriate model.

The estimated pdf for the competing models is shown in Figure 11, the estimated cdf for the competing models is shown in Figure 12, and the estimated cdf, Kaplan-Meier, pdf, and PP plots for the proposed model are shown in Figure 13.

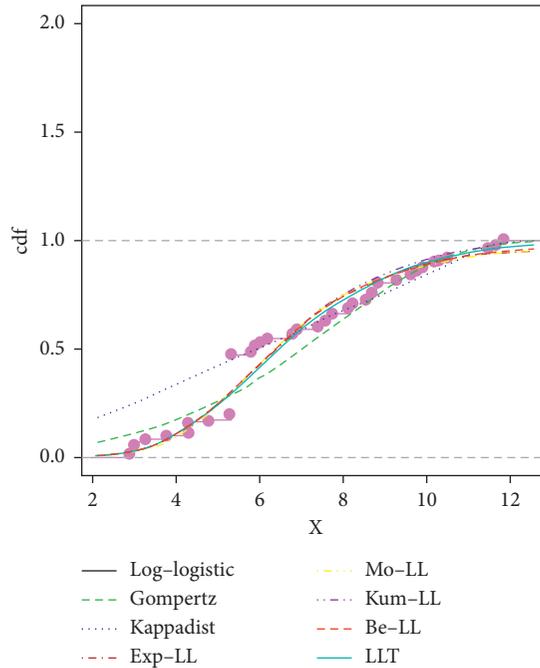


FIGURE 12: Some fitted cdf's of the fitted distributions in dataset I.

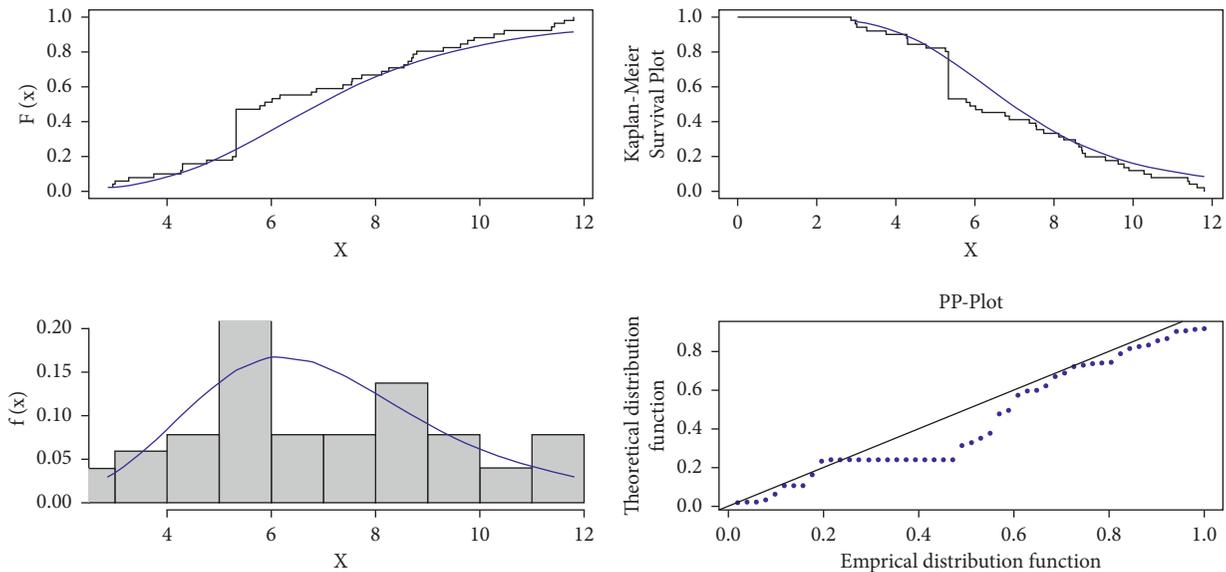


FIGURE 13: The estimated cdf, Kaplan-Meier, estimated pdf, and PP plots of the log-logistic tangent distribution for dataset I.

8. Conclusions

The two-parameter log-logistic model has found widespread use in statistical sciences, especially survival analysis, biomedical sciences, engineering, and actuarial sciences. In this paper, we proposed a new probability distribution that derives from a combination of log-logistic distribution and a Tan-G trigonometric class of distribution. This model was named log-logistic tangent distribution “LLT distribution” and has been successfully extended in this work. Reliability function, pdf, cdf, failure rate function, reversed failure rate

function, and the cumulative failure rate function were carefully derived, and expressions for the basic statistical properties for the log-logistic tangent distribution have been developed. The log-logistic tangent distribution was applied to the COVID-19 mortality rate dataset and provided a better fit than the known two-parameter, three-parameter, and four-parameter competitors. Gompertz, log-logistic, kappa, exponentiated log-logistic, Marshall-Olkin log-logistic, Kumaraswamy log-logistic, and beta log-logistic based on goodness-of-fit tests, selection criteria such as AIC, BIC, CAIC, and HQIC values, and the log-likelihood value were

applied. We, therefore, conclude that the log-logistic tangent distribution is the most adequate model among the ones considered, as well as a very competitive model for describing different datasets in different areas of application.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Taif University Researchers Supporting Project, no. TURSP-2020/220, Taif University, Taif, Saudi Arabia.

References

- [1] A. H. Muse, S. Mwalili, O. Ngesa, S. J. Almalki, and G. A. Abd-Elmougod, "Bayesian and classical inference for the generalized log-logistic distribution with applications to survival data," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5820435, 24 pages, 2021.
- [2] S. Nadarajah, G. M. Cordeiro, and E. M. M. Ortega, "The exponentiated G geometric family of distributions," *Journal of Statistical Computation and Simulation*, vol. 85, no. 8, pp. 1634–1650, 2015.
- [3] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, vol. 99, p. 465, Hoboken, NJ, USA, 2002.
- [4] S. A. Khan and S. K. Khosa, "Generalized log-logistic proportional hazard model with applications in survival analysis," *Journal of Statistical Distributions and Applications*, vol. 3, no. 1, 2015.
- [5] S. A. Khan, "Exponentiated Weibull regression for time-to-event data," *Lifetime Data Analysis*, vol. 24, no. 2, pp. 328–354, 2018.
- [6] M. H. Tahir and G. M. Cordeiro, "Compounding of distributions: a survey and new generalized classes," *Journal of Statistical Distributions and Applications*, vol. 3, no. 1, 2015.
- [7] Z. Ahmad, G. G. Hamedani, and N. S. Butt, "Recent developments in distribution theory: a brief survey and some new generalized classes of distributions," *Pakistan Journal of Statistics and Operation Research*, vol. 15, no. 1, pp. 87–110, 2019.
- [8] M. H. Tahir and S. Nadarajah, "Parameter induction in continuous univariate distributions: well-established G families," *Anais da Academia Brasileira de Ciências*, vol. 87, no. 2, pp. 539–568, 2015.
- [9] L. Souza, W. R. O. Júnior, C. C. R. De Britoz, C. Chesneaux, T. A. E. Ferreiray, and L. G. M. Soaresy, "On the Sin-G class of distributions: theory, model and application," *Journal of Mathematical Modeling*, vol. 7, no. 3, pp. 357–379, 2019.
- [10] L. Souza, W. R. Junior, C. C. Brito, and C. Chesneau, "General properties for the cos-G class of distributions with applications," *Eurasia Journal of Mathematics*, vol. 2, no. 2, pp. 63–79, 2019.
- [11] L. Souza, "New trigonometric classes of probabilistic distributions," Thesis, Universidade Federal Rural de Pernambuco, Brazil, 2015.
- [12] A. H. Muse, S. M. Mwalili, and O. Ngesa, "On the log-logistic distribution and its generalizations: a survey," *International Journal of Statistics and Probability*, vol. 10, no. 3, p. 93, Apr. 2021.
- [13] G. S. Rao, R. R. L. Kantam, K. Rosaiah, and S. Prasad, "Reliability test plans for type-II exponentiated log-logistic distribution," vol. 5, no. 1, pp. 55–64, 2012.
- [14] A. J. Lemonte, "The beta log-logistic distribution," *Brazilian Journal of Probability and Statistics*, vol. 28, no. 3, pp. 313–332, 2014.
- [15] M. W. A. Ramos, "The zografos-balakrishnan log-logistic distribution," *Journal of Statistical Distributions and Applications*, vol. 12, no. 3, pp. 244–255, 2013.
- [16] W. Gui, "Marshall-olkin extended log-logistic distribution and its application in minification processes," *Applied Mathematical Sciences*, vol. 7, no. 77–80, pp. 3947–3961, 2013.
- [17] D. C. T. Granzotto and F. Louzada, "The transmuted log-logistic distribution: modeling, inference, and an application to a polled tabapua race time up to first calving data," *Communications in Statistics - Theory and Methods*, vol. 44, no. 16, pp. 3387–3402, 2015.
- [18] M. M. Rahman, B. Al-Zahrani, S. H. Shahbaz, and M. Q. Shahbaz, "Transmuted probability distributions: a review," *Pakistan Journal of Statistics and Operation Research*, vol. 16, no. 1, pp. 83–94, 2020.
- [19] M. H. Tahir, M. Mansoor, M. Zubair, and G. G. Hamedani, "McDonald log-logistic distribution with an application to breast cancer data," *Journal of Statistical Theory and Applications*, vol. 13, no. 1, p. 65, 2014.
- [20] A. S. Malik and S. P. Ahmad, "An extension of log-logistic distribution for analyzing survival data," *Pakistan Journal of Statistics and Operation Research*, vol. 16, no. 4, pp. 789–801, 2020.
- [21] M. A. Aldahlan, "Alpha power transformed log-logistic distribution with application to breaking stress data," *Advances in Theoretical and Mathematical Physics*, vol. 2020, Article ID 2193787, 9 pages, 2020.
- [22] M. M. Mansour, "A new log-logistic lifetime model with mathematical properties, copula, modified goodness-of-fit test for validation and real data modeling," *Mathematics*, vol. 8, no. 9, 2020.
- [23] N. M. Alfaer, A. M. Gemeay, H. M. Aljohani, and A. Z. Afify, "The extended log-logistic distribution: inference and actuarial applications," *Mathematics*, vol. 9, no. 12, p. 1386, 2021.
- [24] Y. Liang Tung, Z. Ahmad, and E. Mahmoudi, "The arcsine-X family of distributions with applications to financial sciences," *Computer Systems Science and Engineering*, vol. 39, no. 3, pp. 351–363, 2021.
- [25] W. He, Z. Ahmad, A. Z. Afify, and H. Goual, "The arcsine exponentiated-X family: validation and insurance application," *Complexity*, vol. 2020, Article ID 8394815, 18 pages, 2020.
- [26] A. L. P. dos Santos, "Method to generate growth and degrowth models obtained from combinations of existing models applied to agrarian sciences," *Biorxiv*, 2019, <https://www.biorxiv.org/content/10.1101/585224v1.full>.
- [27] T. U. Case, "Method for generating distributions and classes of probability distributions: method for generating distributions and classes of probability distributions," 2015, <https://arxiv.org/ftp/arxiv/papers/1504/1504.01062.pdf>.
- [28] L. Souza, W. R. d. O. Júnior, C. C. R. d. Brito, C. Chesneau, R. L. Fernandes, and T. A. E. Ferreira, "Tan-G class of trigonometric distributions and its applications," *Cubo (Temuco)*, vol. 23, no. 1, pp. 1–20, 2021.

- [29] C. B. Ampadu, "The tan-G family of distributions with illustration to data in the health sciences," *Physical Science & Biophysics Journal*, vol. 3, no. 3, pp. 4–6, 2019.
- [30] G. Casella and R. L. Berger, *Statistical Inference*, Cengage Learning, Boston, MA, USA, 2021.
- [31] J. Chu, "A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain," *PLoS One*, vol. 16, no. 3, Article ID e0249037, 2021.
- [32] M. Zuo, S. K. Khosa, Z. Ahmad, and Z. Almaspoor, "Comparison of COVID-19 pandemic dynamics in asian countries with statistical modeling," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 4296806, 16 pages, 2020.
- [33] E. M. Almetwally, "The odd weibull inverse topp-leone distribution with applications to covid-19 data," *Annals of Data Science*, 2021.
- [34] J. Zhao, Z. Ahmad, Z. Almaspoor, M. El-Morshedy, and A. Z. Afify, "Modeling COVID-9 pandemic dynamics in two asian countries," *Computers, Materials & Continua*, vol. 67, no. 1, pp. 965–977, 2021.
- [35] H. M. Almongy, E. M. Almetwally, H. M. Aljohani, A. S. Alghamdi, and E. H. Hafez, "A new extended Rayleigh distribution with applications of COVID-19 data," *Results in Physics*, vol. 23, Article ID 104012, 2021.
- [36] E. M. Almetwally, "Application of COVID-19 pandemic by using odd lomax-G inverse Weibull distribution," *Mathematical Sciences Letters*, vol. 10, no. 2, pp. 47–57, 2021.
- [37] M. M. Mansour, M. A. Farsi, S. M. Mohamed, and E. M. Abd Elrazik, "Modeling the COVID-19 pandemic dynamics in Egypt and Saudi Arabia," *Mathematics*, vol. 9, no. 8, pp. 1–13, 2021.
- [38] X. Liu, Z. Ahmad, A. M. Gemeay, A. T. Abdulrahman, E. H. Hafez, and N. Khalil, "Modeling the survival times of the COVID-19 patients with a new statistical model: A case study from China," *PLoS One*, vol. 16, no. 7, 2021.
- [39] O. O. Rasaki, L. A. Nafiu, H. M. Abdisalam, and S. B. Thierno, "Stochastic modelling of the dynamics of the SARS-CoV-2 epidemic: an Africa perspective," *American Journal of Mathematics and Statistics*, vol. 11, no. 2, pp. 41–48, 2021.
- [40] K. Rosaiah, R. R. L. Kantam, and S. Kumar, "Reliability test plans for exponentiated log-logistic distribution," *Economic Quality Control*, vol. 21, no. 2, pp. 279–289, 2006.
- [41] T. V. F. de Santana, E. M. M. Ortega, G. M. Cordeiro, and G. O. Silva, "The Kumaraswamy-log-logistic distribution," *Journal of Statistical Theory and Applications*, vol. 11, no. 3, pp. 265–291, 2012.