Pak.j.stat.oper.res. Vol.18 No.4 2022 pp 883-900 DOI: https://dx.doi.org/10.18187/pjsor.v18i4.4149

Pakistan Journal of Statistics and Operation Research

Odd Lomax Generalized Exponential Distribution: Application to Engineering and COVID-19 data

Laxmi Prasad Sapkota[1*], Vijay Kumar[2]

[*] Corresponding author

1. Department of Mathematics and Statistics, DDU Gorakhpur University, India, laxmisapkota75@gmail.com
2. Department of Mathematics and Statistics, DDU Gorakhpur University, India, vkgkp@rediffmail.com

Abstract

This paper proposes the 4-parameter odd Lomax generalized exponential distribution for the study of engineering and COVID-19 data. The statistical and mathematical properties of this distribution such as a linear representation of the probability density function, survival function, hazard rate function, moments, quantile function, order statistics, entropy, mean deviation, characteristic function and average residual life function are established. The estimates of parameters of the proposed distribution are obtained using maximum likelihood estimation (MLE), Maximum product spacings (MPS), least-square estimation (LSE), and Cramer-Von-Mises estimation (CVME) methods. A Monte-Carlo simulation experiment is carried out to study the performance of MLEs. The applicability of the proposed distribution is evaluated using two real datasets related to engineering and COVID-19 death cases. All the computational work was performed in R programming software.

Key Words: Characteristic function; COVID-19; Generalized exponential distribution; Hazard function; Odd Lomax family.

Mathematical Subject Classification: 62F10, 62F40, 62E10, 62E15.

1. Introduction

Statistical models have a crucial role in modeling real-life datasets related to engineering, medicine, life sciences, etc. However classical probability models are not sufficient to fit the various varieties of the real datasets. Hence we need a flexible modified model that can address the deficiencies of the classical distributions. Generating a more versatile model however depends on the choice of base distribution that we want to modify (Lee et al., 2013). In this study, we have chosen generalized exponential distribution (Gupta and Kundu, 2001) as a base distribution. This model can be applied quite efficiently to study the positive lifetime data, especially in the replacement of two-parameter gamma and two-parameter Weibull distributions. It also can be used in the study of the lifetime of components that are used in parallel combinations (Gupta and Kundu, 2001). Suppose there are n components in a parallel system i.e. system will work at least one of the n parts of the system work. The lifetime distribution of n component is

$$F\left(t\right) = \left(1 - e^{-\lambda t}\right)^{n} ; t > 0$$

which is the cumulative distribution function of GED for $\alpha = n$. Further, the shape of the hazard function of GED is increasing for shape parameter > 1, decreasing for shape parameter < 1 and constant for shape parameter $= 1$. In the literature we can observe that many authors have created new distribution using this GED distribution such as beta generalized exponential distribution (Barreto-Souza et al., 2010), exponential extension distribution (Kumar, 2010), Nadarajah and Haghighi (2011) has introduced an extension of the exponential distribution, exponentiated NH distribution (Lemonte, 2013), generalized inverted exponential (Singh et al., 2013), odds generalized exponential-exponential distribution (Maiti and Pramanik, 2015), Weibull generalized exponential distribution (Mustafa et al., 2016) has introduced by computing extended exponential distribution (Gómez et al., 2014) and GED. A 4-parameter modified exponential distribution was introduced (Rasekhi et al., 2017), alpha power transformed generalized exponential distribution (Nadarajah and Okorie, 2018), truncated Cauchy power–exponential distribution (Chaudhary et al., 2020a), truncated Cauchy power–inverse exponential distribution (Chaudhary et al., 2020b), logistic-exponential power distribution (Joshi et al., 2020), and odd generalized exponential Lomax distribution (Mohamed et al., 2022).

The newly generated probability models are also useful to model daily death due to COVID-19 infection such as arcsine modified Weibull distribution is used by (Liu et al., 2021) to model the survival time of fifty-three COVID-19 infected patients taken from a hospital in China. Similarly Nagy et al. (2021) have studied the COVID-19 mortality data in Kingdom of Saudi Arabia and Latvia using new distribution discrete extended odd Weibull exponential distribution. Alsuhabi et al. (2022) have presented a new four-parameter continuous lifetime model with four parameters by compounding the Lomax and the Weibull distributions called extended odd Weibull Lomax distribution and used it to model COVID-19 dataset. Furthur an extension of Gumbel distribution has used to fit the two real data sets of COVID-19 (Hossam et al., 2022) and also applied different estimation methods. In this work we have also used the proposed distribution to study the COVID-19 death cases worldwide during January 21 to March 27, 2020.
The probability density function (pdf) and cumulative distribution function (cdf) of GED defined by (Gupta and Kundu, 2001) having shape parameter a and scale parameter δ are

$$g(x\,; a, \delta) = a\delta e^{-\delta x}\left(1 - e^{-\delta x}\right)^{a-1}; x > 0. \tag{1}$$

$$G(x\,; a, \delta) = \left(1 - e^{-\delta x}\right)^{a}; x > 0. \tag{2}$$

To extend the GED we have used the odd Lomax-G family of distribution defined by (Cordeiro et al., 2019) having cdf and pdf as

$$F\left(x;\delta,\theta\right) = 1 - \delta^{\theta}\left[\delta + \frac{G\left(x\right)}{1 - G\left(x\right)}\right]^{-\theta}; x > 0, \delta > 0, \theta > 0. \tag{3}$$

$$f\left(x;\delta,\theta\right) = \frac{\theta\delta^{\theta}g(x)}{\left[1 - G\left(x\right)\right]^{2}}\left[\delta + \frac{G\left(x\right)}{1 - G\left(x\right)}\right]^{-\theta-1}; x > 0. \tag{4}$$

Ogunsanya et al. (2019) have also defined another format of the odd Lomax-G family and defined the new 3-parameter model called odd Lomax exponential distribution. We feel that there is a lack of study using the odd Lomax-G family defined by (Cordeiro et al., 2019) to the best of our knowledge so we are attracted to this family of distribution. The different sections of the proposed work are presented as follows. The new odd Lomax generalized exponential distribution is introduced in Section 2. In section 3 we present the properties of the proposed model. To estimate the parameters of the model under study, we have employed four estimation techniques namely maximum likelihood estimation (MLE), Maximum product spacings (MPS), least-square estimation (LSE), and Cramer-Von-Mises (CVM) methods in Section 4. A simulation study for MLEs is carried out in section 5. In Section 6 we have considered two real datasets to analyze and explore the applications of the proposed distribution. Finally, Section 7 ends up with some general concluding remarks.

2. The odd Lomax generalized exponential (OLGE) distribution

Consider$X$ be a positive continuous random variable. If the random variable X follows OLGE distribution then its cdf must take the form like (inserting Equation (2) in Equation (3))

$$F\left(x;\Phi\right) = 1 - \delta^{\theta}\left[\delta + \frac{\left(1 - e^{-bx}\right)^{a}}{1 - \left(1 - e^{-bx}\right)^{a}}\right]^{-\theta}; x > 0, \Phi > 0. \tag{5}$$

The pdf corresponding to Equation (5) is

$$f(x;\Phi)=\frac{ab\theta\delta^{\theta}e^{-bx}\left(1-e^{-bx}\right)^{a-1}}{\left[1-\left(1-e^{-bx}\right)^{a}\right]^{2}}\left[\delta+\frac{\left(1-e^{-bx}\right)^{a}}{1-\left(1-e^{-bx}\right)^{a}}\right]^{-\theta-1};x>0. \tag{6}$$

where $\Phi=(a,b,\delta,\theta)$ is a vector of parameter space of OLGE distribution.

2.1. Reliability function

The reliability function of OLEG distribution is

$$R(x;\Phi)=\delta^{\theta}\left[\delta+\frac{\left(1-e^{-bx}\right)^{a}}{1-\left(1-e^{-bx}\right)^{a}}\right]^{-\theta};x>0,\Phi>0. \tag{7}$$

2.2. Hazard function

The hazard function of X reduces to

$$h(x;\Phi)=\frac{ab\theta e^{-bx}\left(1-e^{-bx}\right)^{a-1}}{\left[1-\left(1-e^{-bx}\right)^{a}\right]^{2}}\left[\delta+\frac{\left(1-e^{-bx}\right)^{a}}{1-\left(1-e^{-bx}\right)^{a}}\right]^{-1};x>0. \tag{8}$$

2.3. Reversed hazard rate function

$$rh(x;\Phi)=\frac{ab\theta\delta^{\theta}e^{-bx}\left(1-e^{-bx}\right)^{a-1}}{\left\{1-\delta^{\theta}\left[A(x)\right]^{-\theta}\right\}\left[1-\left(1-e^{-bx}\right)^{a}\right]^{2}}\left[A(x)\right]^{-\theta-1};x>0 \tag{9}$$

here $A(x)=\delta+\frac{\left(1-e^{-bx}\right)^{a}}{1-\left(1-e^{-bx}\right)^{a}}$. We have shown the various shapes of the pdf, reverse hrf, and hazard rate function (hrf) of OLGE distribution for some values of $\Phi=(a,b,\delta,\theta)$ in Figure 1. The hrf can have constant, increasing, decreasing and increasing and decreasing hazard functions.

3. Properties of OLGE distribution

3.1. Quantile function

The expression for quantile function can be calculated by inverting the cdf Equation (5) as

$$Q(p)=-\frac{1}{b}\left[\log\left\{1-\left(\frac{z}{1-z}\right)^{1/a}\right\}\right].$$

where $z=\left[\left\{(1-p)\,\delta^{-\theta}\right\}^{-1/\theta}-\delta\right]$ and $p\in U(0,1)$.

3.2. Linear form of OLGE distribution

Using the following two power series expansions we can express the density function of OLGE into a useful linear form as

$$(b+a)^{-n}=\sum_{i=0}^{\infty}(-1)^{i}\binom{n+i-1}{i}a^{n-i}b^{i}$$

$$(1-a)^{n}=\sum_{i=0}^{\infty}(-1)^{i}\binom{n}{i}a^{i}$$

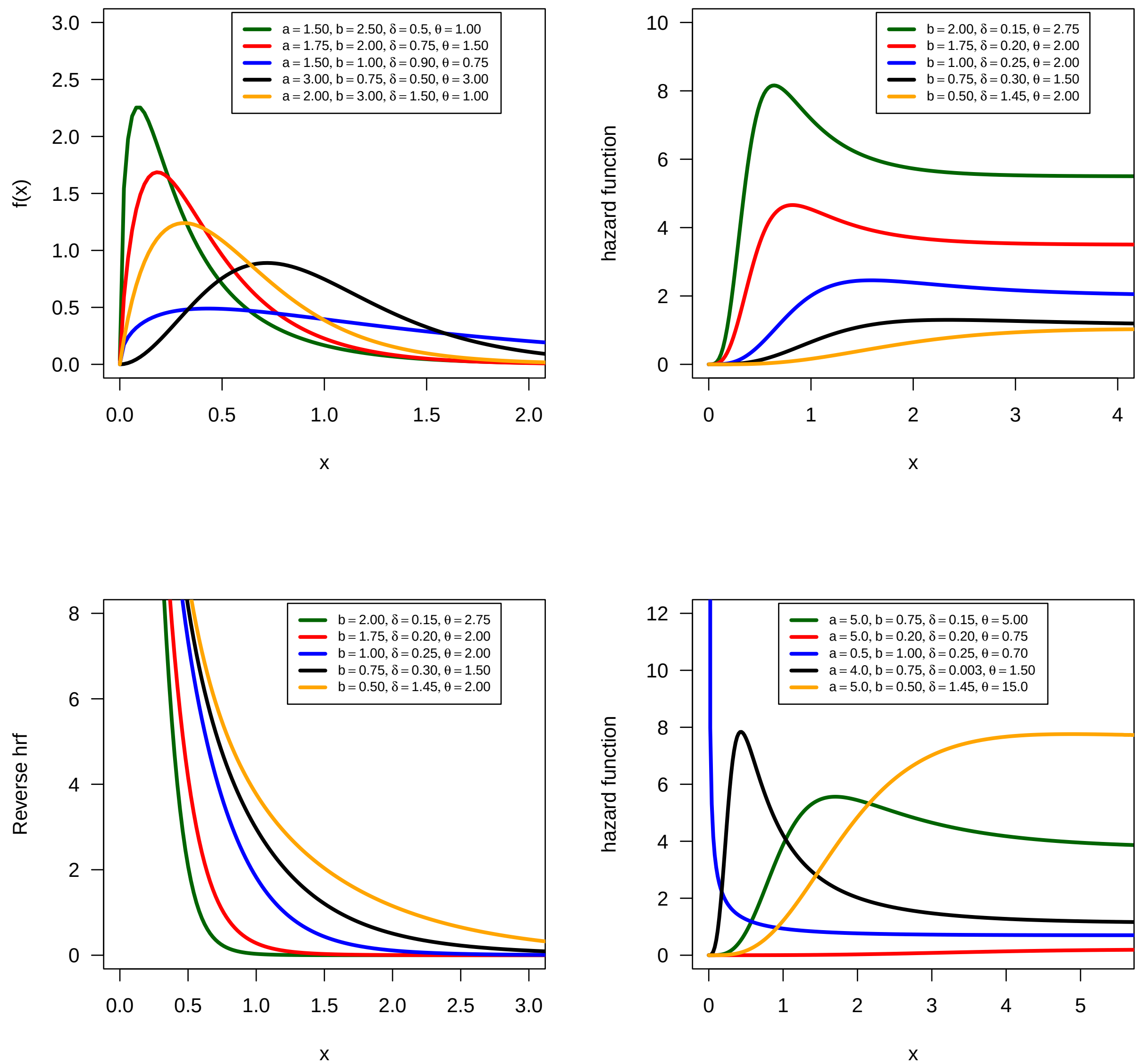

Figure 1: Graph of pdf and reverse hazard rate function (first column) and hazard functions (second column).

Hence the pdf define in Equation (6) reduces to

$$f\left(x;\Phi\right)=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}Z_{klm}\; e^{-(1+m)bx} \tag{10}$$

where $Z_{klm} = ab\theta\delta^{\theta+k}\left(-1\right)^{k+l+m}\begin{pmatrix}\theta+k\\ k\end{pmatrix}\begin{pmatrix}\theta+k-1\\ l\end{pmatrix}\begin{pmatrix}al-a\theta-ak-1\\ m\end{pmatrix}$. Now the pdf defined in Equation (10) can be expressed as the linear density of exponential distribution as

$$f\left(x;\Phi\right)=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}Z_{m}^{*}\, f^{*}(x;(1+m)b) \tag{11}$$

where $f^*(x;(1+m)b)$ is the density function of exponential distribution with shape parameter $\lambda = (1+m)b$ and $Z_m^* = a\theta \sum\limits_{k=0}^{\infty}\sum\limits_{l=0}^{\infty} \delta^{\theta+k}(-1)^{k+l+m}\begin{pmatrix}\theta+k\\k\end{pmatrix}\begin{pmatrix}\theta+k-1\\l\end{pmatrix}\begin{pmatrix}al-a\theta-ak-1\\m+1\end{pmatrix}$.

3.3. Moment

The moment of random variable X taken from origin i.e. raw moments can be calculated as

$$\begin{aligned}\mu_r^{'} &= \int\limits_0^{\infty} x^r f(x)\,dx \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} Z_{klm}\int\limits_0^{\infty} x^r e^{-(1+m)bx}dx \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} Z_{klm}\frac{\Gamma(r+1)}{[(1+m)b]^{r+1}}\end{aligned} \tag{12}$$

where $\Gamma(r+1)$ is the gamma function defined as $\frac{\Gamma(n+1)}{a^{n+1}} = \int\limits_0^{\infty} x^n \exp\{-ax\}\,dx$. Now r^{th} central moments can be calculated as

$$\mu_r = E\left(X-\mu_1^{'}\right)^r = \sum_{j=0}^{\infty}(-1)^j\begin{pmatrix}r\\j\end{pmatrix}\mu_1^{'}\mu_{r-j}^{'}.$$

3.4. Conditional Moment

It can be expressed for OLGE distribution as

$$\begin{aligned}\mu_{CM} = E(X^n/X>x) &= \frac{1}{R(x)}\int\limits_x^{\infty} t^n f(t)\,dt \\ &= \frac{1}{R(x)}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} Z_{klm}\frac{\Gamma(\{n+1,(1+m)bx\})}{[(1+m)b]^{n+1}}.\end{aligned} \tag{13}$$

here $R(x)$ is the reliability function defined in Equation (7).

3.5. Moment generating function

Let $M_x(t)$ denote moment generating function of random variable X of OLGE distribution, and then it can be expressed as

$$\begin{aligned}M_x(t) &= \sum_{n=0}^{\infty}\frac{t^n}{n!}\mu_r^{'} \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{t^n}{n!}Z_{klm}\frac{\Gamma(r+1)}{[(1+m)b]^{r+1}}.\end{aligned} \tag{14}$$

3.6. Characteristic Function

Let X be a random variable that follows OLGE distribution then its characteristic function can be defined as

$$\begin{aligned}\Omega_X(t) &= E\left(e^{itX}\right) = \int\limits_0^\infty e^{itx} f(x)\,dx \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} Z_{klm} \int\limits_0^\infty e^{-\{(1+m)b-it\}x} dx \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} \frac{Z_{klm}}{\{(1+m)b-it\}}\end{aligned} \tag{15}$$

here $i=\sqrt{-1}$ is a complex number.

3.7. Average Residual Life Function

Let $X-x$ is the expected remaining life of a component or an item given that it has survived up to time x can be calculated using pdf defined in Equation (10) as

$$\begin{aligned}\mu(x) &= \frac{\int\limits_0^\infty x\,f(x)\,dx}{1-F(x)} - x \\ &= \frac{\sum\limits_{k=0}^{\infty}\sum\limits_{l=0}^{\infty}\sum\limits_{m=0}^{\infty} Z_{klm} \frac{\Gamma(2)}{\{(1+m)b\}^2}}{1-F(x)} - x\end{aligned}. \tag{16}$$

3.8. Entropies

Renyi Entropy

Rényi (1961) has introduced the measures of entropy and it can be used to measure the variability of uncertainty and calculated as

$$I_\rho(X) = \frac{1}{1-\rho}\log\int\limits_{-\infty}^{+\infty}\{f(x)\}^\rho\,dx;\ \rho>0\ and\ \rho\neq 1. \tag{17}$$

After expanding $\{f(x)\}^\rho$ as we expand Equation (10) we get

$$\{f(x)\}^\rho = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} H_{k\,lm} e^{-\{(d+m)b\}x}. \tag{18}$$

Using Equation (18) in Equation (17) we get

$$\begin{aligned}I_\rho(X) &= \frac{1}{1-\rho}\log\left[\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} H_{k\,lm} \int\limits_0^{+\infty} e^{-\{(d+m)b\}x}dx\right] \\ &= \frac{1}{1-\rho}\log\left[\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} H_{k\,lm}\frac{1}{(d+m)b}\right]\end{aligned}$$

where $H_{k\,lm} = (ab\theta)^d\,\delta^{\theta d+k}\,(-1)^{k+l+m}\begin{pmatrix}\theta d+d+k-1\\ k\end{pmatrix}\begin{pmatrix}\theta d-d+k\\ l\end{pmatrix}\begin{pmatrix}al-ak-a\theta d-d\\ m\end{pmatrix}$.

q-entropy

For the random variable $X \sim OLGE\,(\Phi)$ distribution, the q-entropy is expressed as

$$M_q(X) = \frac{1}{1-q} \log \left[1 - \int_0^\infty \{f(x)\}^q \, dx \right]; \; q > 0 \; and \; q \neq 1. \tag{19}$$

Using the results of (18) by replacing ρ by q in (19) we can write

$$I_q(X) = \frac{1}{1-q} \log \left[1 - \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty H_{k\,lm} \frac{1}{(d+m)b} \right]. \tag{20}$$

3.9. Mean Deviation

The mean deviation taken from mean (μ) can be obtained as

$$\begin{aligned} Z\,(\mu) &= E\,|X-\mu| \\ &= 2\mu F\,(\mu) - 2\mu + \int_\mu^\infty x f\,(x)\,dx \\ &= 2\mu F\,(\mu) - 2\mu + \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty Z_{k\,lm} \frac{\Gamma\,\{2,(1+m)b\mu\}}{[(1+m)b]^2}. \end{aligned}$$

Similarly, the mean deviation from the median (M_d) is

$$\begin{aligned} Z\,(M_d) &= E\,|X-M_d| \\ &= M_d F\,(M_d) - M_d - \mu + 2\int_{M_d}^\infty x\; f\,(x)\,dx \\ &= M_d F\,(M_d) - M_d - \mu + 2\sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty Z_{k\,lm} \frac{\Gamma\,(\{2,(1+m)bM_d\})}{[(1+m)b]^2}. \end{aligned}$$

here μ and M_d are the values of the mean and median.

4. Estimation Methods

4.1. Maximum Likelihood Estimation (MLE) Method

To estimate the parameters of the OLGE distribution, we have used the MLE method. Let, $x_1, ..., x_n$ be a random sample from $OLGE\,(\Phi)$ and the likelihood function $L\,(a, b, \delta, \theta)$ is given by,

$$L\,(x\,;\Phi) = \left(ab\delta^\theta\theta\right)^n \prod_{i=1}^n \exp\,(-bx_i) \frac{\{1-\exp\,(-bx_i)\}^{a-1}}{\left[1-\{1-\exp\,(-bx_i)\}^a\right]^2} \left[\delta + \frac{\{1-\exp\,(-bx_i)\}^a}{1-\{1-\exp\,(-bx_i)\}^a}\right]^{-\theta-1} \tag{21}$$

Now log-likelihood density of Equation (21) is

$$\ell = n\log\left(ab\delta^\theta\theta\right) - b\sum_{i=1}^n x_i + \sum_{i=1}^n \log\{T(x_i)\}^{a-1} + 2\sum_{i=1}^n \log\{1-T(x_i)\} - (\theta+1)\log\left[\delta + \frac{\{T(x_i)\}^a}{1-\{T(x_i)\}^a}\right] \tag{22}$$

where $T(x_i) = 1 - \exp\,(-bx_i)$ To get the MLEs of the proposed distribution analytically we have to maximize Equation (22) with respect to model parameters.

4.2. Maximum Product of Spacings (MPS) estimation method

The MPS estimation method can be used as a substitute for the MLE method introduced by (Cheng and Amin, 1983) for the parameter estimation of continuous univariate probability models. Let $x_1 < ... < x_n$ be an ordered random sample from OLGE distribution then the uniform spacings of that random sample can be calculated as

$$\begin{aligned} D_i(\Phi) &= F\left(x_{(i)}, \Phi\right) - F\left(x_{(i-1)}, \Phi\right) \\ &= \delta^{\theta}\left[\delta + \frac{\left(1-e^{-bx_{(i-1)}}\right)^a}{1-\left(1-e^{-bx_{(i-1)}}\right)^a}\right]^{-\theta} - \delta^{\theta}\left[\delta + \frac{\left(1-e^{-bx_{(i)}}\right)^a}{1-\left(1-e^{-bx_{(i)}}\right)^a}\right]^{-\theta}, i = 1...(n+1) \end{aligned} \tag{23}$$

where $F\left(x_{0\,:n}\,|\Phi\right) = 0,\ F\left(x_{n+1\,:n}\,|\Phi\right) = 1$ and $\sum_{i=1}^{n+1} D_i(\Phi) = 1$. To obtain the MPS estimators we have to minimize the following function with respect to parameters as $K\left(\Phi\right) = \frac{1}{n+1}\sum_{i=1}^{n+1} \log\left[D_i(\Phi)\right]$.

4.3. Method of Least-Square Estimation (LSE)

The ordinary least square estimates for the OLGE distribution can be obtained by minimizing

$$K\left(X;\Phi\right) = \sum_{i=1}^{n}\left[F(X_{(i)};\Phi) - \frac{i}{n+1}\right]^2 \tag{24}$$

with respect to underlying parameters $\Phi = (a, b, \delta, \theta)$.
Consider the ordered random variables $X_{(1)} < \ldots < X_{(n)}$ where $\underline{X} = \{X_1, \ldots, X_n\}$ is a random sample of size n and $F(X_i)$ denotes the distribution function. The least-square estimators of $\Phi = (a, b, \delta, \theta)$ can be obtained by minimizing the Equation (25)

$$K\left(x;\Phi\right) = \sum_{i=1}^{n}\left[1 - \delta^{\theta}\left[\delta + \frac{\left(1-e^{-bx_{(i)}}\right)^a}{1-\left(1-e^{-bx_{(i)}}\right)^a}\right]^{-\theta} - \frac{i}{n+1}\right]^2 \tag{25}$$

4.4. Method of Cramer-Von-Mises estimation (CVME)

The Cramer-Von-Mises estimates $\Phi = (a, b, \delta, \theta)$ are obtained by minimizing the function

$$\begin{aligned} W\left(X;\Phi\right) &= \frac{1}{12n} + \sum_{i=1}^{n}\left[F\left(x_{i:n}|\Phi\right) - \frac{2i-1}{2n}\right]^2 \\ &= \frac{1}{12n} + \sum_{i=1}^{n}\left[1 - \delta^{\theta}\left[\delta + \frac{\left(1-e^{-bx_{(i)}}\right)^a}{1-\left(1-e^{-bx_{(i)}}\right)^a}\right]^{-\theta} - \frac{2i-1}{2n}\right]^2 \end{aligned} \tag{26}$$

5. Simulation Study

In this section, we have presented the Monte-Carlo simulation to study the performance of MLEs using the tools bias and mean square error (MSE). R programming language is used to perform the Monte-Carlo experiment. In this study, we have considered different sample sizes and different values of the parameter of OLGE distribution. By repeating 1000 times we have generated 20 samples from OLGE distribution of size $n = (50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000)$ and we have chosen two sets of actual parameter values as $set1 : (a = 1, b = 0.5, \delta = 2, \theta = 0.5)$ and $set2 : (a = 2, b = 0.25, \delta = 0.5, \theta = 1.5)$. The average values, biases and MSEs are reported in Appendix (A.1 to A.6) and also we have presented the results obtained via graphs (see Figures 2 and 3). We observed that the estimated values are nearly close to the actual parameter values as the size of the samples increases and also biases and MSEs are decreased as the sample size increases.

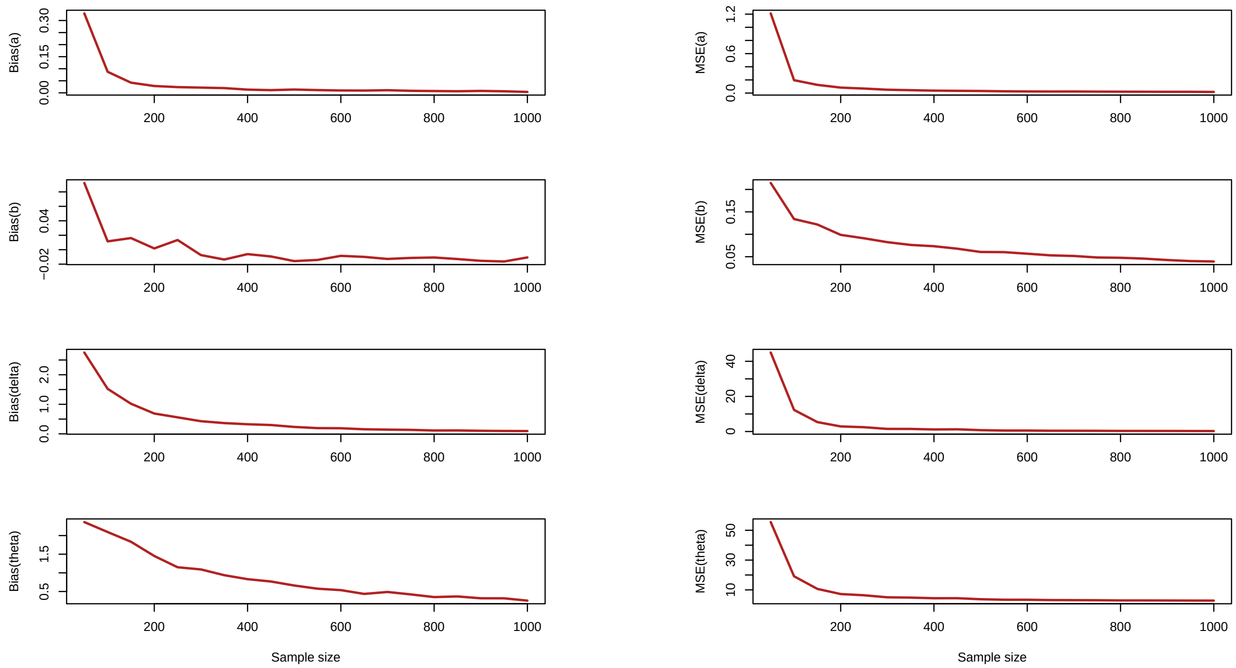

Figure 2: Bias and MSE for set 1

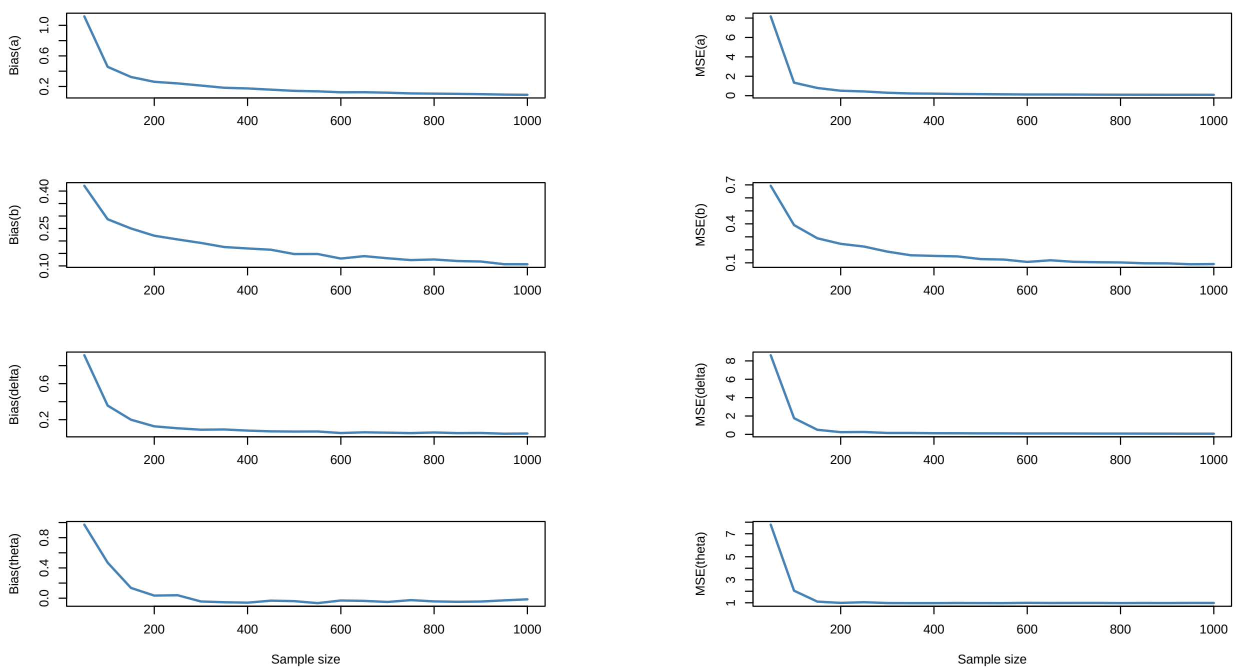

Figure 3: Bias and MSE for set 1

6. Illustration with Engineering and COVID-19 datasets

The capability of fit and flexibility of the proposed model is verified using two sets of real data by comparing it with some distributions that already exist in the literature.

6.1. Exploratory data analysis (EDA) of engineering dataset

The engineering dataset consists of $25(100cm)$ specimens of yarn, which were tested at a certain strain level, and it represents the number of cycles to failure (Lawless, 2011; Elshahhat and Elemary, 2021). The data are: $20, 15, 61, 38, 98,$

$42, 86, 76, 146, 121, 157, 149, 175, 180, 176, 180, 220, 198, 224, 264, 251, 282, 325, 321, 653$. The summary statistics of the dataset is presented in Table 1 and histogram and total time on test (TTT) plots are also displayed in Figure 4.

Table 1: Summary statistics of engineering dataset

Min.	First Quartile	median	mean	Third Quartile	Max.	Skewness	Kurtosis
15.0	86.0	175.0	178.3	224.0	653.0	1.627292	3.78314

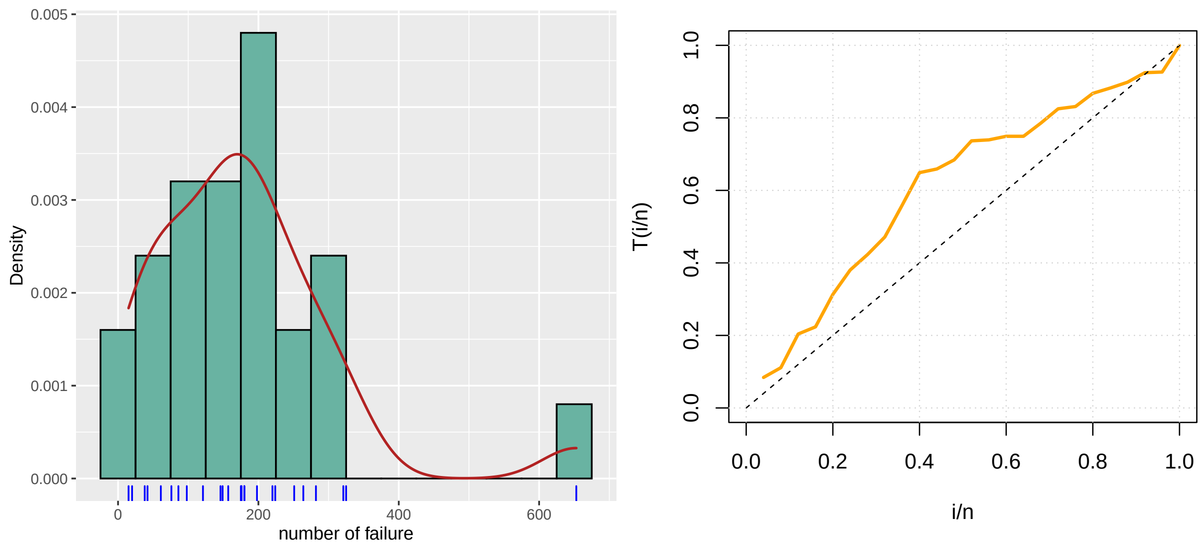

Figure 4: Histogram and TTT plots of engineering dataset.

6.2. EDA of COVID-19 dataset

The second dataset represents the daily death of COVID-19 across the world between January 21 to March 27, 2020, recorded from (Worldometers, 2020) for sixty-seven days. The data are $3, 3, 8, 16, 15, 24, 26, 26, 38, 43, 46, 45, 58, 64, 66, 73, 73, 86, 89, 97, 108, 97, 146, 122, 143, 143, 106, 98, 136, 117, 121, 113, 100, 158, 81, 64, 37, 58, 65, 54, 73, 67, 85, 83, 102, 107, 105, 228, 198, 271, 332, 353, 447, 405, 687, 642, 817, 972, 1079, 1356, 1625, 1629, 1873, 2381, 2388, 2791, 3271$.

The summary statistics of the dataset are presented in Table 2 and histogram and TTT plots are also displayed in Figure 5.

Table 2: Summary statistics of the COVID-19 dataset

Min.	First Quartile	median	mean	Third Quartile	Max.	Skewness	Kurtosis
3	64	102	408	301.5	3271	2.362177	4.858592

6.3. Estimation of the Unknown Parameters

In this study, four estimation methods are utilized to estimate the parameters of the proposed model and we have presented MLEs, MPS, LSEs, and CVMEs for OLGE distribution (see Table 3 and Table 4) for both datasets using the **maxLik()** function (Henningsen and Toomet, 2011) in R software (R Core Team, 2022) and (Lander, 2014). For both data sets, MLE performs better than MPS, LSE and CVME methods (see Figure 6).

6.4. Validity Test of the OLGE distribution

To test the validity of the proposed model we have presented the Kolmogorov-Smirnov (KS) test and found the values of test statistic $0.10788(0.9331)$ and $0.1213(0.2775)$ with $p - value$ in parenthesis respectively for both datasets also

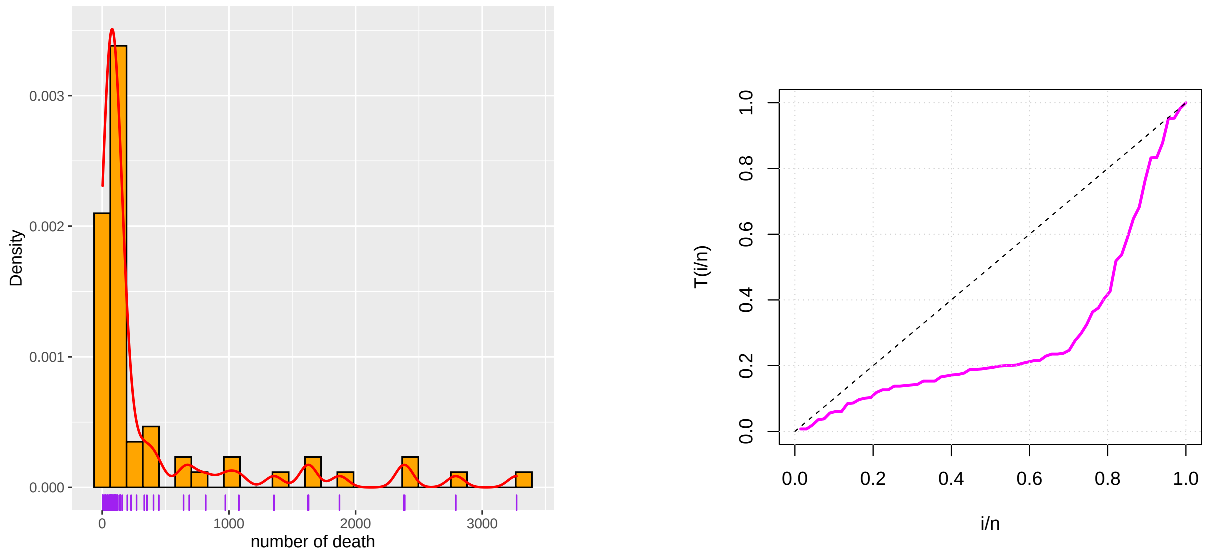

Figure 5: Histogram and TTT plots of COVID-19 dataset.

Table 3: Estimated values via MLE, MPS, LSE and CVME Methods (Engineering data)

Estimation Method	a	b	δ	θ	AIC	BIC
MLE	1.3396	0.0135	3.248	0.724	312.2311	317.1066
MPS	0.73	0.0156	4.9703	0.509	313.058	317.9335
LSE	2.0422	0.011	1.4127	0.7944	312.8834	317.7589
CVME	2.2581	0.0239	1.5752	0.2921	313.4652	318.3407

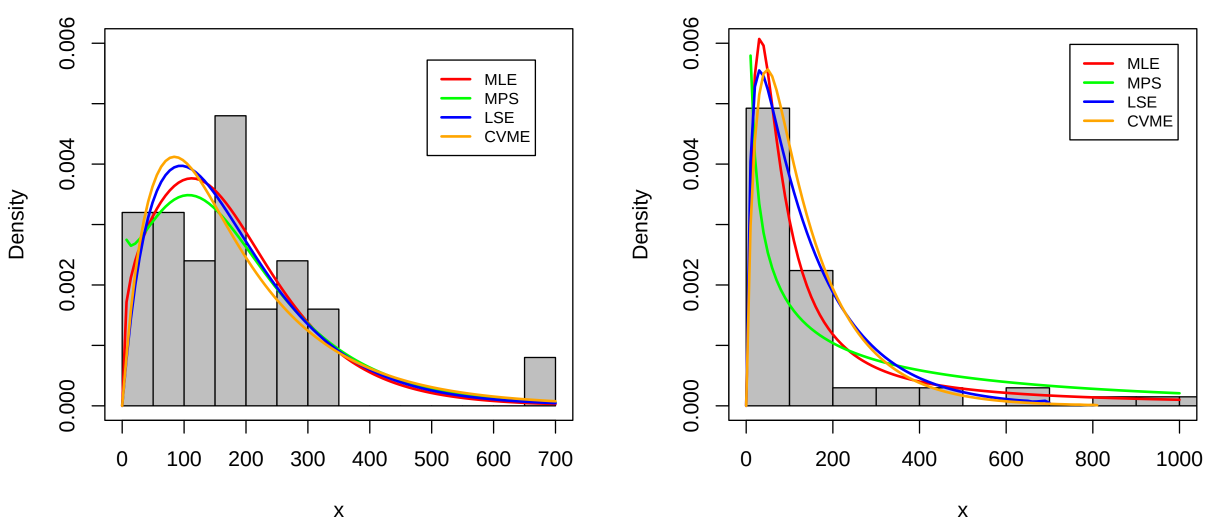

Figure 6: Comparison of the MLE, MPS, LSE and CVME Methods for engineering data (left panel) and COVID-19 data (right panel).

we have plotted KS plot and Q-Q plot (see Figures 7 and 8). These results support that the OLGE distribution fits data very well.

Table 4: Estimated values via MLE, MPS, LSE and CVME Methods (COVID-19 data)

Estimation Method	a	b	δ	θ	AIC	BIC
MLE	1.8458	0.0018	0.0121	0.4233	897.6426	906.4613
MPS	0.5805	0.0024	0.5634	0.5392	925.8073	934.6261
LSE	2.1056	0.0528	0.9203	0.1335	NA	NA
CVME	2.0497	0.0458	1.4566	0.1779	NA	NA

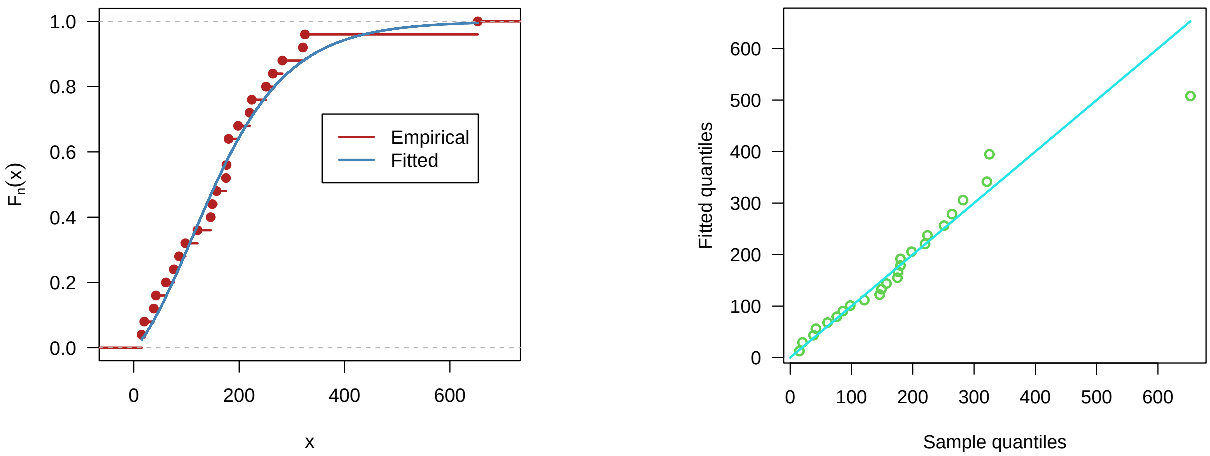

Figure 7: KS plot (left panel) and Q-Q plot (right panel) for the engineering dataset.

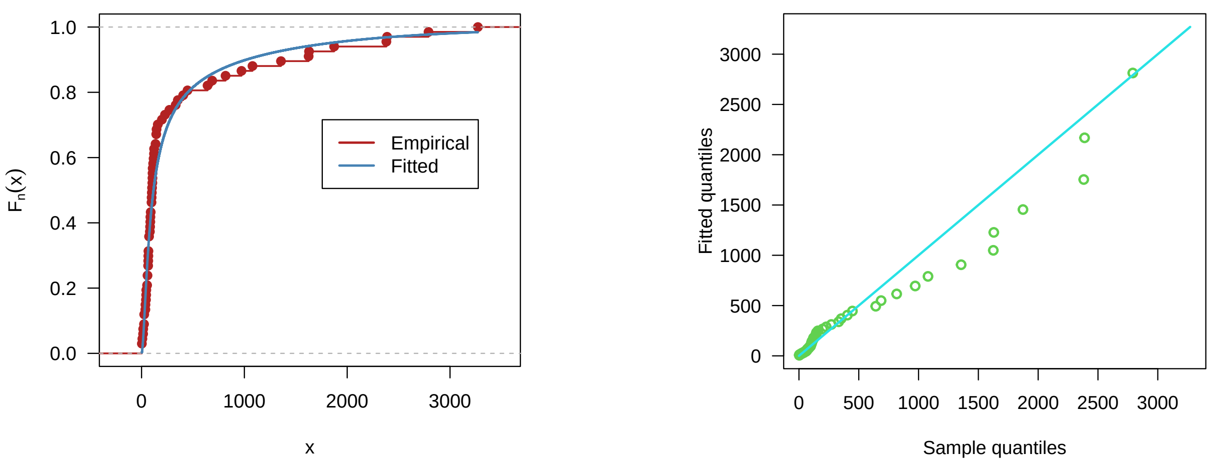

Figure 8: KS plot (left panel) and Q-Q plot (right panel) for the COVID-19 dataset.

6.5. Model Selection

To compare the performance of OLGE distribution with some modified form of exponential and generalized exponential distributions such as the odd generalized exponential Lomax (OGEL) distribution (Mohamed et al., 2022), generalized inverted exponential (GIE) distribution (Singh et al., 2013), odd Lomax exponential (OLE) distribution (Ogunsanya et al., 2019), exponential extension (EE) distribution (Kumar, 2010) and generalized exponential (GE) distribution (Gupta and Kundu2001). Some goodness of fit statistics such as maximum log-likelihood (LL), Kolmogorov-Smirnov (KS), Cramer-von Mises (W), and Anderson-Darling (A^2) tests are used for each model and results are presented in Table 5 and Table 6 for both datasets. It is observed that the OLGE distribution gets the minimum value of the test statistic and a higher p-value hence the proposed model is the best among five competing models for the both datasets. Further, this result has also been verified by the graphical method (see Figures 9 and 10).

Table 5: The goodness-of-fit statistics (engineering data)

Model	LL	KS(p-value)	W(p-value)	A^2(p-value)
OLGE	-152.1155	0.10788(0.9331)	0.2576(0.966)	0.0360(0.9556)
OGEL	-152.4016	0.1307(0.7866)	0.3344(0.9094)	0.0532(0.8605)
GIE	-158.091	0.2907(0.0293)	1.6802(0.1391)	0.3179(0.1201)
OLE	-154.6487	0.2633(0.0625)	1.1807(0.2745)	0.2244(0.2254)
EE	-153.2891	0.146(0.6609)	0.5608(0.6841)	0.0921(0.6292)
GE	-152.4905	0.1441(0.6773)	0.3877(0.8596)	0.0667(0.7765)

Table 6: The goodness-of-fit statistics (COVID-19 data)

Model	LL	KS(p-value)	W(p-value)	A^2(p-value)
OLGE	-444.8213	0.1213(0.2775)	0.9072(0.4095)	0.161(0.3584)
OGEL	-448.8046	0.2741(0.1602)	1.4377(0.1921)	0.1496(0.0999)
GIE	-460.3788	0.2971(0.00001)	3.5612(0.0144)	0.6634(0.0154)
OLE	-451.1612	0.3256(1.354e-06)	2.3541(0.0593)	0.4704(0.0471)
EE	-450.3552	0.3808(0.0810)	1.927(0.101)	0.1708(0.0402)
GE	-460.076	0.2644(0.0001)	4.6557(0.0042)	0.9857(0.0025)

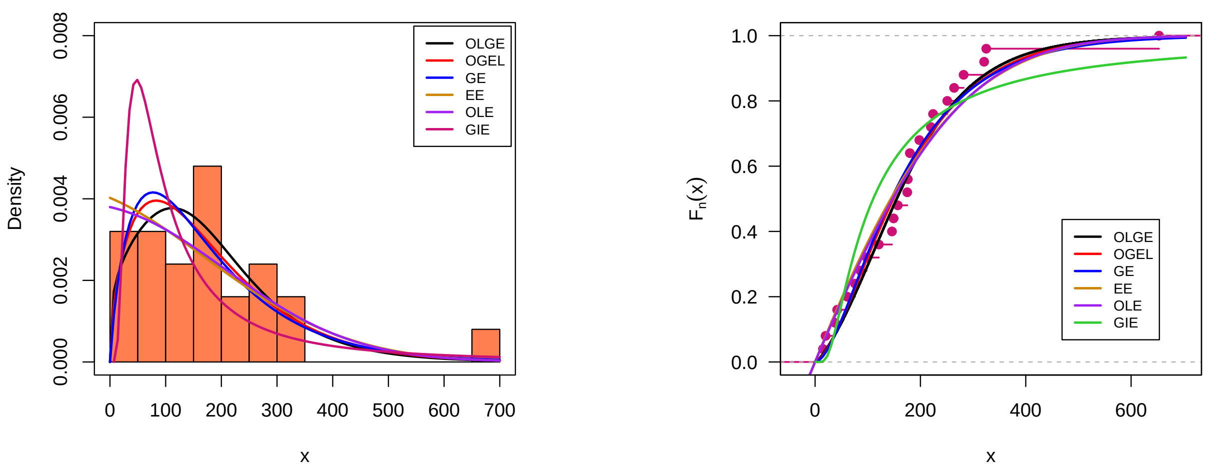

Figure 9: The Histogram and density function of fitted distributions and fitted CDF (for engineering data)

6.6. Prediction of Mortality Rates

The mortality rate of death cases in the world due to COVID-19 during the period January 21 to March 27, 2020, has been predicted based on MLEs using the OLGE distribution and reported in (Table 7).

Table 7: Prediction of Deaths

No. of Death	0-500	500-1000	1000-1500	1500-2000	2000-2500	2500-3000	3000 and above
Probability	0.8153	0.0829	0.0372	0.0219	0.0141	0.0093	0.0193

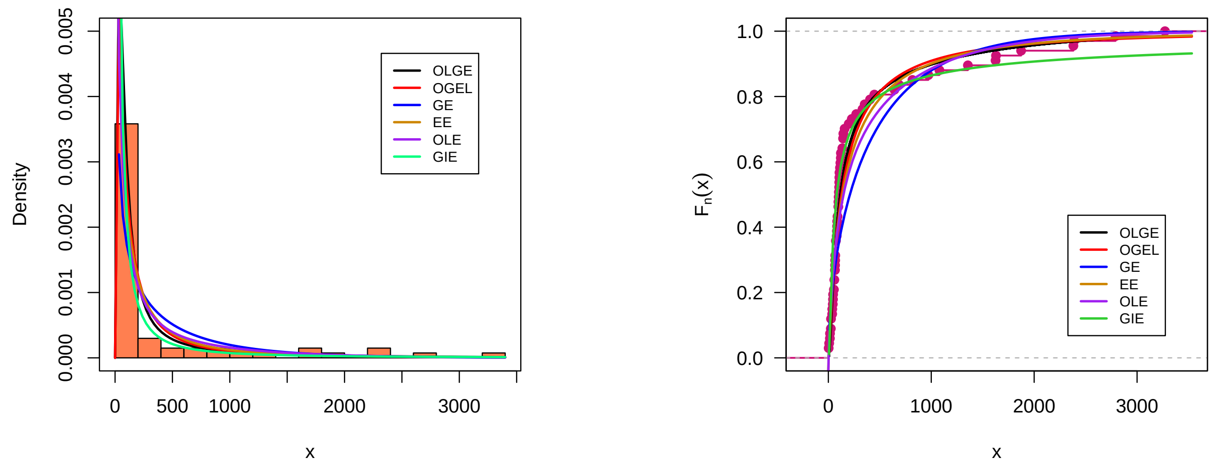

Figure 10: The Histogram and density function of fitted distributions and fitted CDF (for CVOID-19 data)

7. Conclusion

We suggest a univariate continuous probability distribution having four parameters named odd Lomax generalized exponential distribution. Some significant statistical properties of the OLGE distribution are discussed such as a linear mixture of exponential densities, moment and moment generating function, mean deviation, entropy and characteristic function. The parameters of the proposed distribution are estimated by using maximum likelihood estimation (MLE), Maximum product spacings (MPS), least-square estimation (LSE), and Cramer-Von-Mises estimation (CVME) methods. A Monte-Carlo simulation experiment is carried out to study the behavior of MLEs and we found that bias and MSEs decrease as the sample size increases this indicates that MLEs are asymptotically unbiased and consistent. The two real datasets related to engineering and death cases of COVID-19 are analyzed to explore the potentiality and suitability of the proposed model and observed that the presented model is quite better than five other existing models taken for illustration. We expect this model can be used in the field of survival analysis, engineering, applied statistics, and probability theory. Bayesian Analysis and time series regression analysis can be performed for further study using this model.

References

1. Alsuhabi, H., Alkhairy, I., Almetwally, E. M., Almongy, H. M., Gemeay, A. M., Hafez, E., Aldallal, R., and Sabry, M. (2022). A superior extension for the lomax distribution with application to covid-19 infections real data. *Alexandria Engineering Journal*, 61(12):11077–11090.
2. Barreto-Souza, W., Santos, A. H., and Cordeiro, G. M. (2010). The beta generalized exponential distribution. *Journal of statistical Computation and Simulation*, 80(2):159–172.
3. Chaudhary, A. K., Sapkota, L. P., and Kumar, V. (2020a). Truncated cauchy power–exponential distribution with theory and applications. *IOSR Journal of Mathematics (IOSR-JM)*, 16(6):44–52.
4. Chaudhary, A. K., Sapkota, L. P., and Kumar, V. (2020b). Truncated cauchy power–inverse exponential distribution: Theory and applications. *IOSR Journal of Mathematics (IOSR-JM)*, 16(4):12–23.
5. Cheng, R. and Amin, N. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(3):394–403.
6. Cordeiro, G. M., Afify, A. Z., Ortega, E. M., Suzuki, A. K., and Mead, M. E. (2019). The odd lomax generator of distributions: Properties, estimation and applications. *Journal of Computational and Applied Mathematics*, 347:222–237.
7. Elshahhat, A. and Elemary, B. R. (2021). Analysis for xgamma parameters of life under type-ii adaptive progressively hybrid censoring with applications in engineering and chemistry. *Symmetry*, 13(11):2112.
8. Gómez, Y. M., Bolfarine, H., and Gómez, H. W. (2014). A new extension of the exponential distribution. *Revista Colombiana de Estadística*, 37(1):25–34.
9. Gupta, R. D. and Kundu, D. (2001). Generalized exponential distribution: different method of estimations.

Journal of Statistical Computation and Simulation, 69(4):315–337.
10. Henningsen, A. and Toomet, O. (2011). maxlik: A package for maximum likelihood estimation in r. *Computational Statistics*, 26(3):443–458.
11. Hossam, E., Abdulrahman, A. T., Gemeay, A. M., Alshammari, N., Alshawarbeh, E., and Mashaqbah, N. K. (2022). A novel extension of gumbel distribution: statistical inference with covid-19 application. *Alexandria Engineering Journal*, 61(11):8823–8842.
12. Joshi, R. K., Sapkota, L. P., and Kumar, V. (2020). The logistic-exponential power distribution with statistical properties and applications. *International Journal of Emerging Technologies and Innovative Research*, 7(12):629–641.
13. Kumar, V. (2010). Bayesian analysis of exponential extension model. *J. Nat. Acad. Math*, 24:109–128.
14. Lander, J. P. (2014). *R for everyone: Advanced analytics and graphics*. Pearson Education.
15. Lawless, J. F. (2011). *Statistical models and methods for lifetime data*. John Wiley & Sons.
16. Lee, C., Famoye, F., and Alzaatreh, A. Y. (2013). Methods for generating families of univariate continuous distributions in the recent decades. *Wiley Interdisciplinary Reviews: Computational Statistics*, 5(3):219–238.
17. 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.
18. Liu, X., Ahmad, Z., Gemeay, A. M., Abdulrahman, A. T., Hafez, E., and Khalil, N. (2021). Modeling the survival times of the covid-19 patients with a new statistical model: A case study from china. *Plos one*, 16(7):e0254999.
19. Maiti, S. S. and Pramanik, S. (2015). Odds generalized exponential-exponential distribution. *Journal of data science*, 13(4):733–753.
20. Mohamed, H., Mousa, S. A., Abo-Hussien, A. E., and Ismail, M. M. (2022). Estimation of the daily recovery cases in egypt for covid-19 using power odd generalized exponential lomax distribution. *Annals of Data Science*, 9(1):71–99.
21. Mustafa, A., El-Desouky, B. S., and AL-Garash, S. (2016). Weibull generalized exponential distribution. *arXiv preprint arXiv:1606.07378*.
22. Nadarajah, S. and Haghighi, F. (2011). An extension of the exponential distribution. *Statistics*, 45(6):543–558.
23. Nadarajah, S. and Okorie, I. E. (2018). On the moments of the alpha power transformed generalized exponential distribution. *Ozone: Science & Engineering*, 40(4):330–335.
24. Nagy, M., Almetwally, E. M., Gemeay, A. M., Mohammed, H. S., Jawa, T. M., Sayed-Ahmed, N., and Muse, A. H. (2021). The new novel discrete distribution with application on covid-19 mortality numbers in kingdom of saudi arabia and latvia. *Complexity*, 2021.
25. Ogunsanya, A., Sanni, O., and Yahya, W. (2019). Exploring some properties of odd lomax-exponential distribution. *Annals of Statistical Theory and Applications (ASTA)*, 1:21–30.
26. R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
27. Rasekhi, M., Alizadeh, M., Altun, E., Hamedani, G., Afify, A. Z., and Ahmad, M. (2017). The modified exponential distribution with applications. *Pakistan Journal of Statistics*, 33(5).
28. Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 1. Berkeley, California, USA.
29. Singh, S. K., Singh, U., and Kumar, M. (2013). Estimation of parameters of generalized inverted exponential distribution for progressive type-ii censored sample with binomial removals. *Journal of Probability and Statistics*, 2013.
30. Worldometers (2020). Reported death cases 2020 worldwide. https://www.worldometers.info/coronavirus/about/.

Appendix

Table A.1: Estimates for set1

n	$a = 1$	$b = 0.5$	$\delta = 2$	$\theta = 0.5$
50	1.3295814	0.5923795	4.753318	2.8617164
100	1.0871675	0.5115287	3.5234137	2.594559
150	1.0421402	0.5161087	3.0194602	2.333846
200	1.0282517	0.5017459	2.6876562	1.9522008
250	1.0237234	0.5134218	2.5586245	1.6499123
300	1.0216291	0.4925625	2.428145	1.592046
350	1.0196297	0.4863902	2.3642415	1.4371739
400	1.013246	0.493893	2.325024	1.33077
450	1.0111406	0.4906614	2.2988432	1.2682909
500	1.013646	0.484208	2.234005	1.160491
550	1.0113526	0.4858464	2.1943205	1.0754713
600	1.00978	0.4914759	2.1894946	1.0369218
650	1.0093922	0.4900785	2.1547149	0.935341
700	1.0109245	0.4872083	2.1430901	0.9873087
750	1.0083527	0.4886213	2.1351356	0.922056
800	1.0074945	0.4892342	2.1132166	0.8503103
850	1.0065793	0.4870848	2.1148698	0.8674024
900	1.0078942	0.4846366	2.1034206	0.8182404
950	1.0063818	0.4836907	2.0971905	0.8168285
1000	1.0037205	0.4893072	2.0946992	0.7556323

Table A.2: Bias for set1

n	$a = 1$	$b = 0.5$	$\delta = 2$	$\theta = 0.5$
50	0.3295814	0.0923795	2.753318	2.3617165
100	0.0871675	0.0115287	1.5234137	2.094559
150	0.0421403	0.0161087	1.0194602	1.833846
200	0.0282517	0.0017459	0.6876562	1.452201
250	0.0237234	0.0134218	0.5586245	1.1499123
300	0.0216291	-0.0074375	0.428145	1.092046
350	0.0196297	-0.0136098	0.3642415	0.9371739
400	0.0132461	-0.006107	0.3250241	0.8307703
450	0.0111406	-0.0093386	0.2988432	0.7682909
500	0.0136463	-0.015792	0.2340046	0.660491
550	0.0113526	-0.0141536	0.1943206	0.5754713
600	0.00978	-0.0085241	0.1894946	0.5369218
650	0.0093922	-0.0099215	0.1547149	0.435341
700	0.0109245	-0.0127917	0.1430901	0.4873087
750	0.0083527	-0.0113787	0.1351356	0.422056
800	0.0074945	-0.0107658	0.1132166	0.3503103
850	0.0065793	-0.0129152	0.1148698	0.3674024
900	0.0078942	-0.0153634	0.1034206	0.3182404
950	0.0063818	-0.0163093	0.0971905	0.3168285
1000	0.0037205	-0.0106928	0.0946992	0.2556323

Table A.3: Mean square errors set1

n	$a=1$	$b=0.5$	$\delta=2$	$\theta=0.5$
50	1.211596	0.2144661	45.009518	55.519472
100	0.1957614	0.1340271	12.308519	19.12876
150	0.1246451	0.1217517	5.3600766	10.668457
200	0.0830171	0.0987142	2.9023194	7.215288
250	0.0684513	0.0910644	2.4594024	6.3852757
300	0.0507718	0.0824469	1.4892253	5.0236602
350	0.0440997	0.0763008	1.477688	4.8204124
400	0.0368833	0.0733858	1.1771981	4.4022705
450	0.0339244	0.0678648	1.2616456	4.4081753
500	0.0317175	0.0605511	0.781347	3.7333608
550	0.0268312	0.0602028	0.5539991	3.3869608
600	0.0245975	0.0566879	0.5580211	3.376505
650	0.0234139	0.0530465	0.4503567	3.1645013
700	0.0239104	0.0516685	0.4329501	3.1122205
750	0.0219878	0.0482927	0.4060388	3.0614456
800	0.02078	0.0476411	0.3375802	2.9272299
850	0.0201333	0.0457505	0.3299239	2.9245333
900	0.0189145	0.0426188	0.3207862	2.881048
950	0.0188624	0.0402541	0.2908654	2.832437
1000	0.0171926	0.0393464	0.2690358	2.803133

A.4: Estimates set2

n	$a=2$	$b=0.25$	$\delta=0.5$	$\theta=1.5$
50	3.1177063	0.6711956	1.4157631	2.4721277
100	2.4570414	0.5370403	0.8565649	1.970193
150	2.3242213	0.5000795	0.6994618	1.6357556
200	2.2610228	0.470992	0.6259053	1.5343338
250	2.2403475	0.456084	0.6043913	1.5393505
300	2.212038	0.4420504	0.5884175	1.4565349
350	2.1828094	0.4256901	0.5911937	1.4462613
400	2.1744221	0.4198937	0.5785921	1.441012
450	2.1582356	0.4147477	0.5703443	1.4670874
500	2.1424863	0.3975438	0.5678462	1.461286
550	2.1364442	0.397737	0.5687254	1.4348304
600	2.1240121	0.3794616	0.5516165	1.4695137
650	2.12497	0.3891445	0.5595599	1.4644239
700	2.1186386	0.3806773	0.5554503	1.4501559
750	2.1091821	0.3732648	0.5508524	1.4745983
800	2.1063297	0.3757564	0.5576492	1.4576827
850	2.1030752	0.3695668	0.5507369	1.4522757
900	2.0995089	0.3675717	0.5521869	1.4559166
950	2.0936721	0.3568325	0.5441705	1.4710771
1000	2.0907246	0.3565701	0.5461962	1.4852708

Table A.5: Bias set2

n	$a = 2$	$b = 0.25$	$\delta = 0.5$	$\theta = 1.5$
50	1.1177063	0.4211956	0.9157631	0.972127
100	0.4570414	0.2870403	0.3565649	0.470193
150	0.3242213	0.2500795	0.1994618	0.1357556
200	0.2610228	0.220992	0.1259053	0.0343337
250	0.2403475	0.206084	0.1043913	0.0393505
300	0.212038	0.1920504	0.0884175	-0.0434651
350	0.1828094	0.1756901	0.0911937	-0.0537387
400	0.1744221	0.1698937	0.0785921	-0.0589881
450	0.1582356	0.1647477	0.0703443	-0.0329126
500	0.1424863	0.1475438	0.0678462	-0.038714
550	0.1364442	0.147737	0.0687254	-0.0651696
600	0.1240121	0.1294616	0.0516165	-0.0304863
650	0.12497	0.1391445	0.0595599	-0.0355761
700	0.1186386	0.1306773	0.0554503	-0.0498441
750	0.1091821	0.1232648	0.0508524	-0.0254018
800	0.1063297	0.1257564	0.0576492	-0.0423173
850	0.1030752	0.1195668	0.0507369	-0.0477243
900	0.0995089	0.1175717	0.0521869	-0.0440834
950	0.0936722	0.1068325	0.0441705	-0.0289229
1000	0.0907246	0.1065701	0.0461962	-0.0147293

Table A.6: Mean square errors set2

n	$a = 2$	$b = 0.25$	$\delta = 0.5$	$\theta = 1.5$
50	8.1772258	0.6929987	8.6196629	7.7881367
100	1.3400216	0.3904067	1.7621869	2.0490572
150	0.7920442	0.2888463	0.4983121	1.0993885
200	0.5092755	0.24603	0.2406515	0.9888409
250	0.4355471	0.2251577	0.2580189	1.0492362
300	0.2990486	0.1861827	0.1530715	0.9762365
350	0.2271321	0.1582706	0.1524354	0.970048
400	0.2071816	0.1531205	0.1263929	0.9692087
450	0.1750197	0.1496214	0.1207365	0.980048
500	0.1596171	0.1288027	0.1086715	0.9729792
550	0.1400643	0.124998	0.1080503	0.970599
600	0.1204872	0.1068673	0.0968028	0.9935698
650	0.1208916	0.1194807	0.0985731	0.9794532
700	0.1146175	0.107546	0.0952925	0.9843919
750	0.1045939	0.1046591	0.0873012	0.9855964
800	0.0973783	0.1028302	0.0864718	0.9711735
850	0.0941655	0.096507	0.0823667	0.9808929
900	0.089781	0.0957085	0.0790988	0.974725
950	0.0914647	0.0888862	0.0737782	0.9854371
1000	0.0869548	0.0904649	0.0735407	0.9811482