Journal of Statistical Theory and Applications
https://doi.org/10.1007/s44199-024-00076-w

RESEARCH ARTICLE

Topp-Leone Exponentiated Pareto Distribution: Properties and Application to Covid-19 Data

Fabio M. Correa[1] · **Braimah J. Odunayo**[1,2] · **Ibrahim Sule**[3] · **Olalekan A. Bello**[3]

Received: 23 October 2023 / Accepted: 21 April 2024

Abstract

This paper proposes a new Topp-Leone Exponentiated Pareto (TLEtP) distribution. The new distribution family is derived by expanding the Topp Leone-G family of distributions with additional positive shape parameters. The corresponding density and distribution functions are derived and shown. Some of the derived mathematical properties of the distribution include quantile function, ordinary and incomplete moments generating function (mgf), hazard function, survival function, odd function, probability weighted moment, and distribution of order statistic. The parameters of the distribution are estimated using Maximum Likelihood method. The proposed distribution's validity is demonstrated by fitting two sets of real data and comparing the results with two existing same-family distributions, the Topp-Leone Pareto type I(TLPI) and Pareto (P), with the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), respectively. The comparison of the proposed Topp-Leone Exponentiated Pareto (TLEtP) to the Topp-Leone Pareto type I(TLPI) and Pareto (P) distribution demonstrate that the TLEtP distribution offers a better fit for the data sets than the other two distributions.

Keywords Hazard function · Order statistics · Survival function

These authors contributed equally to this work.

Fabio M. Correa
mathiascorreaf@ufs.ac.za

[1] Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, Bloemfontein, South Africa

[2] Department of Mathematics and Statistics, Ambrose Alli University, Ekpoma, Nigeria

[3] Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

Published online: 08 May 2024

1 Introduction

Since the World Health Organization declared the Corona virus (COVID-19) pandemic outbreak a public health crisis and of international concern in 2020, approximately 704,241,673 cases and 7,006,177 deaths have been recorded worldwide as of 19 March 2024 [1]. Between 2020 and 2021, when the pandemic was at its peak, several efforts were made to develop and identify COVID-19 prevention methods. Despite global efforts to develop vaccines and cures, the primary strategy used to contain the spread was social distancing, isolation, personal and community hygiene, and mass testing. Prior to the advent of effective vaccines, a number of alternative methods were proposed to combat the pandemic. Most of these strategies were only able to reduce the spread of the disease through the use of personal hygiene measures such as face and nose masks, social distancing, patient quarantine, social isolation, mass testing and lockdown. Irrespective of the measures taken to limit the transmission of the pandemic, social distancing and isolation, as well as general lockdown measures, were all reported to be successful in reducing the spread of the virus [2–4].

The development and comparisons of life models (distributions) curves are always of key importance in modelling and predicting the outbreak of the disease. This is because if the data on the outbreak can be well modelled, the evolution of the disease can be traced, and then the spread can be prevented by studying the dynamics of the disease in the model. A good model in this resguard can be very helpful in decision-making by survival analysts and health authorities. The exponentiated-G class was first proposed by [5], who raised the cumulative distribution function (cdf) to a positive power parameter. Thus, becoming one of many methods proposed by statisticians where parameters are added to distributions. A number of techniques for creating new families of distributions have been investigated by [6]. New families of probability distributions that expand well-known distributions are frequently created to offer flexibility when modeling data in practice. In other words, distributions have been built by expanding popular families of continuous distributions. By enhancing the original model with at least one shape or scale parameter, these generalized distributions offer greater versatility. Additionally, statisticians create families of distributions so that distributions can be skewed to the right, left, unimodal, J or reversed J shape to provide consistently better fits, which equip other distributions with the same underlying model. In fact, a large number of families of continuous probability distributions have been developed with one or two parameters, such as the Marshall-Olkin-G [7], beta generalized-G [8], exponentiated generalized-G [9], exponentiated Gompertz using informative prior [10], Burr X-G distribution [11], generalized odd generalized exponential family [12], Lomax-G [13], mixture of Lomax distribution [14], Type I general exponential class of distributions by [15], transmuted Weibull-G family of distributions [16], and generalized odd generalized exponential family [17].

To obtain a better fit of the models, this article proposes the inclusion of new parameters and provides evidence that the use of two parameters provides a better fit compared to the original distributions. The Topp-Leone Exponentiated-G and Pareto distribution functions are used to generate a new distribution, the Topp-Leone Exponentiated Pareto (TLetP), and the mathematical properties are defined. The proposed Topp-Leone Exponentiated Pareto (TLEtP) distribution has an exceptional probability density function (pdf) due to its ability to model both positively and negatively skewed data with significant tails and symmetric data sets, making it a more flexible model. It can also be used to model increasing, decreasing and bathtub hazard and survival rates. These characteristics make the proposed model a novel distribution for modelling life data, including survival, reliability engineering and biomedical life testing. The proposed model is then applied to two real COVID-19 datasets to illustrate the properties and flexibility of the distribution. The first example is daily confirmed COVID-19 cases recorded in Pakistan and the results are compared with those obtained by [18] and the second example uses data from the COVID-19 pandemic in South Africa. The rest of the paper is organised as follows: Section 2 explains the procedure for generating the Topp Leonne Exponentiated Pareto (TLetP) distribution and develops its mathematical properties. Section 3 shows how the proposed distribution is applied to a real data set and section 4 contains the conclusion.

2 Methods

A four-parameter distribution, the Topp Leone Exponentiated-G Family of Distributions, is proposed by [19] with the probability density function (pdf) and cumulative distribution function (cdf) of the family are given as:

$$F(x;\alpha,\theta) = \{1 - [1 - (G(x,\psi))^{\alpha}]^{2}\}^{\theta} \tag{1}$$

and

$$f(x;\alpha,\theta) = 2\alpha\theta g(x;\psi)G(x;\psi)^{\alpha-1}\{1 - [G(x;\psi)]^{\alpha}\}\{1 - [1 - (G(x,\psi))^{\alpha}]^{2}\}^{\theta} \tag{2}$$

where ψ is the vector of parameters of the baseline distribution. The cdf and pdf of the baseline distribution correspond to the Pareto distribution.

$$G(x,\beta,\lambda) = 1 - \left(\frac{\beta}{x}\right)^{\lambda} \tag{3}$$

and

$$g(x;\beta,\lambda) = \lambda\beta^{\lambda}x^{-(\lambda+1)} \tag{4}$$

Some properties like moments, moment generating function, quantile, hazard function, survival function, odd function, distribution of order statistics, and estimate of the parameters using the method of maximum likelihood are described.

2.1 The Proposed Topp-Leone Exponentiated Pareto (TLEtP) distribution

In this sub-section, a new Topp-Leone exponentiated Pareto distribution is derived. To obtain the the new distribution, (3) is inserted into (1) to give the cumulative distribution function (cdf) of TLEtP as:

$$F(x;\alpha,\theta,\lambda,\beta)=\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta} \tag{5}$$

On differentiation Eq. (3) with respect x, we have the pdf given as:

$$\frac{\partial F(x;\alpha,\theta)}{\partial x}=2\alpha\theta\lambda\beta^{\lambda}x^{-(\lambda+1)}\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha-1}\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha}\right\}$$
$$\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta-1}$$

Therefore,

$$f(x;\alpha,\theta,\lambda,\beta)=2\alpha\theta\lambda\beta^{\lambda}x^{-(\lambda+1)}\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha-1}$$
$$\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha}\right\}\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta-1} \tag{6}$$

Where $x\geq 0$, $\beta>0$ is the scale parameters and $\alpha,\theta,\lambda>0$ are the shape parameters.

We can represent the distribution using binomial expansion on (6) as:

$$(1-Z)^{b-1}=\sum_{i=0}^{x}(-1)^{i}\binom{b-1}{i}Z^{i}$$

Then,

$$\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta-1}=\sum_{i=0}^{x}(-1)^{i}\binom{\theta-1}{i}\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2i}$$

$$\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2i+1}=\sum_{j=0}^{x}(-1)^{j}\binom{2i+1}{j}\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha j}$$

$$\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha(1+j)-1}=\sum_{k=0}^{x}(-1)^{k}\binom{\alpha(1+j)-1}{k}\left(\frac{\beta}{x}\right)^{\lambda k}$$

Therefore,

$$f(x;\alpha,\theta,\lambda,\beta) = 2\alpha\theta\lambda\beta^{(\lambda+\lambda k)} \sum_{i,j,k=0}^{\infty} (-1)^{i+j+k} \binom{\theta-1}{i} \binom{21+1}{j} \binom{\alpha(1+j)-1}{k} x^{-(\lambda+1+\lambda k)} \tag{7}$$

$$f(x;\alpha,\theta,\lambda,\beta) = \sum_{i,j,k=0}^{x} \gamma_i x^{-(\lambda+1+\lambda k)}$$

and

$$\gamma_i = 2\alpha\theta\lambda\beta^{(\lambda+\lambda k)} \sum_{i,j,k=0}^{\infty} (-1)^{i+j+k} \binom{\theta-1}{i} \binom{21+1}{j} \binom{\alpha(1+j)-1}{k}$$

We can also represent the distribution using binomial expansion on (5) as:

$$\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta h} = \sum_{m=0}^{\theta h}(-1)^{m}\binom{\theta h}{m}\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2m}$$

$$\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2m} = \sum_{q=0}^{\infty}(-1)^{q}\binom{2m}{q}\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha q}$$

$$\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha q} = \sum_{w=0}^{\infty}(-1)^{w}\binom{\alpha q}{w}\left(\frac{\beta}{x}\right)^{\lambda w} \tag{8}$$

$$F(x;\alpha,\theta,\lambda,\beta) = \beta^{\lambda w} \sum_{q,w=0}^{\infty} (-1)^{m+q+w} \binom{\theta h}{m}\binom{2m}{q}\binom{\alpha q}{w} x^{-\lambda}$$

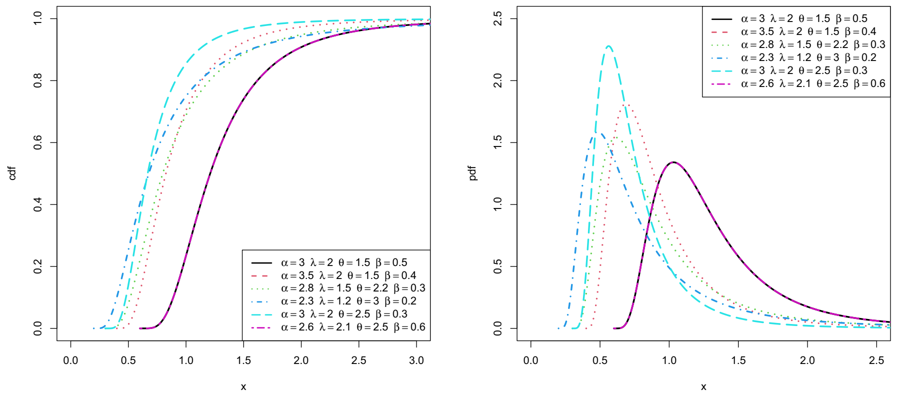
$$F(x;\alpha,\theta,\lambda,\beta) = \sum_{q,w=0}^{\infty} \psi_j x^{-\lambda}$$

and

$$\psi_j = \sum_{m=0}^{\theta h}\binom{\theta h}{m}\binom{2m}{q}\binom{\alpha q}{w}$$

Figure 1a and b show the cdf and pdf of the proposed distribution with different simulated parameters.

Using Eqs. (7) and (8), various properties of the TLEtP distribution are presented in section (2.2).



(a) cdf (b) pdf

Fig. 1 The cdf and pdf of TLEtP distribution with different parameter values

2.2 Mathematical properties

2.2.1 Moments

$$
\begin{aligned}
\mu_r &= E(X^r) = \int_{\beta}^{\infty} x^r f(x) dx \\
\mu_r &= \int_{\beta}^{\infty} x^r \sum_{i,j,k=0}^{\infty} \gamma_i x^{-(\lambda+1+\lambda k)} dx \\
&= \sum_{i,j,k=0}^{\infty} \gamma_i \int_{\beta}^{\infty} x^{-(\lambda+1+\lambda k)+r} dx \\
&= \sum_{i,j,k=0}^{\infty} \gamma_i \left[\frac{x^{-(\lambda+1+\lambda k)+r+1}}{(\lambda+1+\lambda k)+r+1} \right]_{\beta}^{\infty} \\
\mu_r &= \sum_{i,j,k=0}^{\infty} \gamma_i \left[\frac{\beta^{-(\lambda+1+\lambda k)+r+1}}{(\lambda+1+\lambda k)+r+1} \right]
\end{aligned}
\tag{9}
$$

2.3 Moment Generating Function (mgf)

The moment-generating function of X can be obtained using the following equation.

$$
M_x(t) = \sum_{i,j,k=0}^{\infty} \sum_{m=0}^{\infty} \frac{t^m}{m!} \gamma_i \left[\frac{\beta^{-(\lambda+1+\lambda k)+m+1}}{(\lambda+1+\lambda k)+m+1} \right]
$$

2.4 Quantile Function

The TLEtP distribution is simply approximated by inverting (5). If μ has a uniform $U \sim (0, 1)$ distribution, then the equation's solution is:

$$\begin{aligned}
\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta} &= \mu \\
\mu^{\frac{1}{\theta}} &= 1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2} \\
1-\mu^{\frac{1}{\theta}} &= \left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2} \\
\left(1-\mu^{\frac{1}{\theta}}\right)^{\frac{1}{2}} &= 1-\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha} \\
1-\left(1-\mu^{\frac{1}{\theta}}\right)^{\frac{1}{2}} &= \left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha} \\
\left[1-\left(1-\mu^{\frac{1}{\theta}}\right)^{\frac{1}{2}}\right]^{\frac{1}{\alpha}} &= 1-\left(\frac{\beta}{x}\right)^{\lambda} \\
1-\left[1-\left(1-\mu^{\frac{1}{\theta}}\right)^{\frac{1}{2}}\right]^{\frac{1}{\alpha}} &= \left(\frac{\beta}{x}\right)^{\lambda} \\
\left\{1-\left[1-\left(1-\mu^{\frac{1}{\theta}}\right)^{\frac{1}{2}}\right]^{\frac{1}{\alpha}}\right\}^{\frac{1}{\lambda}} &= \frac{\beta}{x} \\
x &= \frac{\beta}{\left\{1-\left[1-\left(1-\mu^{\frac{1}{\theta}}\right)^{\frac{1}{2}}\right]^{\frac{1}{\alpha}}\right\}^{\frac{1}{\lambda}}}
\end{aligned} \tag{10}$$

Equation (10) is the quantile function of the Topp-Leone exponentiated Pareto distribution and Eq. (11) is the median for the Topp-Leone exponential Pareto distribution.

$$x = \frac{\beta}{\left\{1-\left[1-\left(1-0.5^{\frac{1}{\theta}}\right)^{\frac{1}{2}}\right]^{\frac{1}{\alpha}}\right\}^{\frac{1}{\lambda}}} \tag{11}$$

2.5 Hazard Function

The hazard function is defined by:

$$H(x;\alpha,\theta,\beta,\lambda) = \frac{f(\alpha,\theta,\beta,\lambda)}{1 - F(\alpha,\theta,\beta,\lambda)}$$

Considering the definition above, the hazard function of TLEtP distribution is given as:

$$H(x;\alpha,\theta,\beta,\lambda) = \frac{2\alpha\theta\lambda\beta^{\lambda}x^{-(\lambda+1)}\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha-1}\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{\lambda}\right]^{\alpha}\right\}\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta-1}}{1-\left\{1-\left[1-\left(1-\left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta}} \tag{12}$$

Figure 2 shows the hazard function with different parameter values.

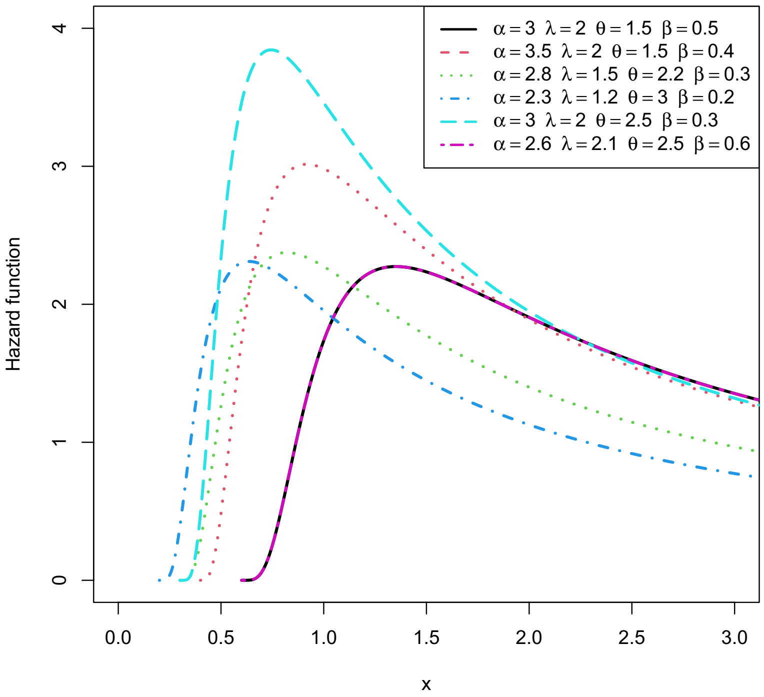


Fig. 2 Hazard function plot of TLEtP distribution with different parameter values

2.6 Survival Function

The survival function, which is the probability of an item not failing prior to some time, can be defined as:

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

So, the survival function for the TLEtP distribution is given as:

$$S(x;\alpha,\theta,\beta,\lambda) = 1 - \left\{1 - \left[1 - \left(1 - \left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta}$$

And, the survival function plot can be seen in Fig. 3.

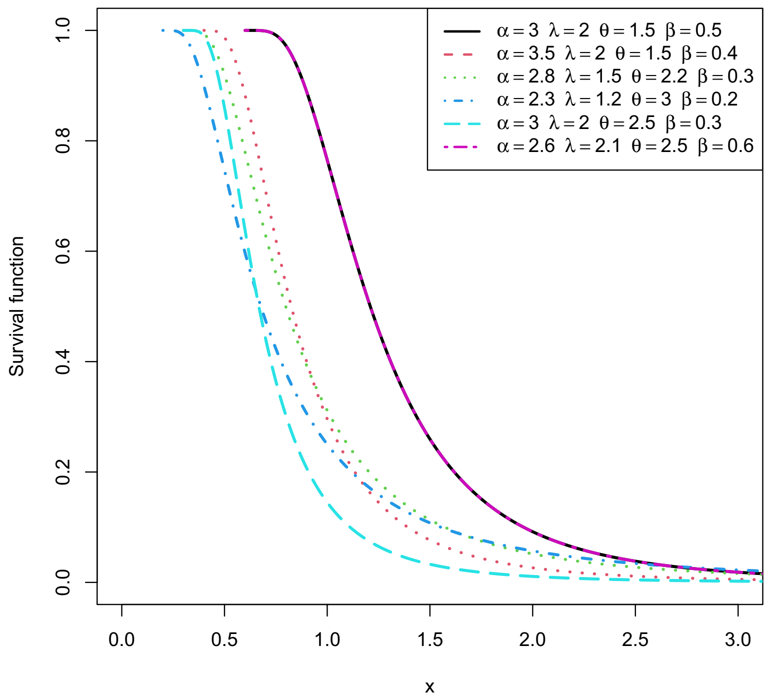


Fig. 3 Survival function plot of TLEtP distribution with different parameter values

2.7 Odds Function

The odds function is obtained using the relation:

$$O(x;\alpha,\theta,\beta,\lambda) = \frac{F(x;\alpha,\theta,\beta,\lambda)}{1 - F(x;\alpha,\theta,\beta,\lambda)}$$

The odds function of TLEtP distribution is given as:

$$O(x;\alpha,\theta,\beta,\lambda) = \frac{\left\{1 - \left[1 - \left(1 - \left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta}}{1 - \left\{1 - \left[1 - \left(1 - \left(\frac{\beta}{x}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}^{\theta}}$$

2.8 Distribution of Order Statistics

Let $X_1, X_2, \ldots, X_n$ be n independent random variable from the TLEtP distribution and let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ be their corresponding order statistics. Let $F_{r:n}(x)$ and $f_{r:n}(x)$, $r = 1, 2, 3, \ldots, n$ denote the cdf and pdf of the r^{th} order statistics $X_{r:n}$ respectively. The pdf of the r^{th} order statistics of $X_{r:n}$ is given as:

$$f_{r:n}(x;\alpha,\theta,\lambda,\beta) = \frac{1}{B(r, n-r+1)} \sum_{v=0}^{n-r} \sum_{i,j,k=0}^{\infty} \sum_{q,w=0}^{\infty} \gamma_i \psi_j x^{-(\lambda+1+\lambda k+\lambda d)} \quad (13)$$

Where

$$\gamma_i = 2\alpha\theta\lambda\beta^{(\lambda+\lambda k)} \binom{\theta-1}{i} \binom{2i+1}{j} \binom{\alpha(1+j)-1}{k}$$

and

$$\psi_j = \sum_{m=0}^{\theta(v+r-1)} \binom{\theta(v+r-1)}{m} \binom{2m}{q} \binom{\alpha q}{w}$$

The maximum order statistics is obtained by setting $r = n$ in Eq. (13) as:

$$f_{n:n} = (x;\alpha,\theta,\lambda,\beta) = n \sum_{i,j,k=o}^{\infty} \sum_{q,w=o}^{\infty} \gamma_i \psi_j x^{-(\lambda+1+\lambda k+\lambda d)} \quad (14)$$

And, the minimum order statistics is obtained by setting $r = 1$ in Eq. (13).

$$f_{1:n} = (x;\alpha,\theta,\lambda,\beta) = n\sum_{v=0}^{n-1}\sum_{i,j,k=o}^{\infty}\sum_{q,w=o}^{\infty}\gamma_i\psi_j x^{-(\lambda+1+\lambda k+\lambda d)} \tag{15}$$

Now

$$\psi_j = \sum_{m=0}^{\theta v}\binom{\theta v)}{m}\binom{2m}{q}\binom{\alpha q}{w}$$

2.9 Parameter Estimation

Since maximum likelihood estimators give the maximum information about the population parameters, this section presents the maximum likelihood estimates (MLEs) of the parameters that are inherent within the distribution function.

Let $X_1, X_2, \dots, X_n$ be random variables of the TLEtP distribution of size n. The log-likelihood function of the TLEtP distribution is obtained as:

$$l(\phi) = nlog(2) + nlog(\alpha) + nlog(\theta) + nlog(\lambda) + n\lambda log(\beta) - (\lambda+1)\sum_{i=1}^{n} log(x) + \sum_{i=1}^{n} log\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\} + (\theta-1)\sum_{i=1}^{n} log\left\{1-\left[1-\left(1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\} \tag{16}$$

Deriving the Eq. (16), we have:

$$\frac{\partial L(\phi)}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\frac{\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha} log\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]}{\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\}} - (\theta-1)\sum_{i=1}^{n}\frac{\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\}\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha} log\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]}{\left\{1-\left[1-\left(1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\}} \tag{17}$$

$$\frac{\partial L(\phi)}{\partial\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} log\left\{1-\left[1-\left(1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right)^{\alpha}\right]^{2}\right\} \tag{18}$$

$$\frac{\partial L(\phi)}{\partial\lambda} = \frac{n}{\lambda} + nlog(\beta) - \sum_{i=1}^{n} log(x_i)$$
$$- \sum_{i=1}^{n} \frac{\alpha\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha-1}\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]log\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]}{\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\}}$$
$$- (\theta-1)\sum_{i=1}^{n}\alpha\frac{\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\}\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha-1}\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]log\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]}{\left\{1-\left[1-\left(1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right)\right]^{\alpha}\right]^{2}\right\}} \tag{19}$$

$$\frac{\partial L(\phi)}{\partial\beta} = \frac{n\lambda}{\beta} + \sum_{i=1}^{n}\frac{\alpha\lambda\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha-1}\left(\frac{\beta}{x_i}\right)^{\lambda-1}\frac{1}{x_i}}{\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\}} -$$
$$(\theta-1)\sum_{i=1}^{n}\frac{\alpha\lambda\left\{1-\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha}\right\}\left[1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right]^{\alpha-1}\left(\frac{\beta}{x_i}\right)^{\lambda-1}\frac{1}{x_i}}{\left\{1-\left[1-\left(1-\left(\frac{\beta}{x_i}\right)^{\lambda}\right)\right]^{\alpha}\right]^{2}\right\}} \tag{20}$$

Equations (17), (18), and (20) do not have a simple form and are intractable. As a result, we have to resort to using interactive procedures to estimate the parameters.

3 Application Using Real-Life Data Sets

The application of the TLEtP distribution is demonstrated using two data sets. The performance of the distribution with regards to providing reasonable parametric fit to the data sets was compared to Topp-Leone Pareto type I(TLPI) and Pareto (P) distributions using the Akaike information criterion (AIC) and Bayesian information criterion (BIC), respectively. The model with a minimum value of AIC and BIC is chosen as the best model to fit the data set.

Table 1 Positive cases recorded in Pakistan from March 24 to April 24, 2020 (Al-Marzouki et al. 2020)

108	102	133	170	121	99	236	178	150
161	258	172	407	577	210	243	281	186
254	336	342	269	520	414	463	514	427
796	555	742	642	785	783	605	751	806

Table 2 The MLEs and information criteria of the models based on confirmed COVID-19 cases recorded in Pakistan from March 24 to April 28, 2020

Model	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\beta}$	-ll	AIC	BIC
TLEtP	129.1655	1.1920	0.7297	7.1520	243.8687	495.7374	502.0714
TLPI	231.1232	8.1379	0.8163	–	245.9008	497.8016	504.5522
P	0.4191	32.0879	–	–	273.2473	550.4946	553.6616

3.1 Example 1

The first data set as listed in the Table 1 represents the daily confirmed COVID-19 cases recorded in Pakistan from March 24 to April 28, 2020, previously used by Al-Marzouki, et al. (2020).

The TLEtP distribution provides a better fit for the data set compared to the TLPI model. From Table 2, the TLEtP distribution has the highest log-likelihood and the

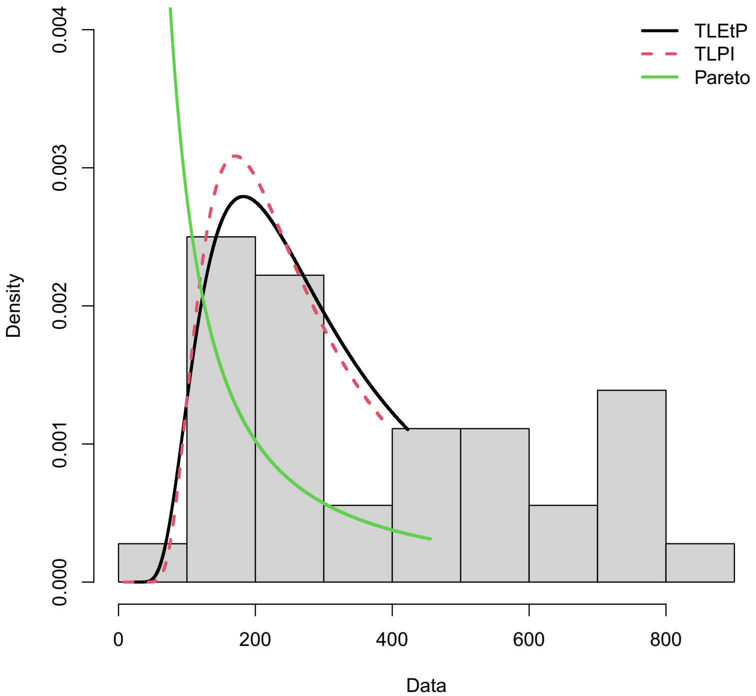


Fig. 4 Fitted pdfs for the TLEtP, TLPI, and Pareto distribution based on confirmed COVID-19 cases recorded in Pakistan from March 24 to April 28, 2020

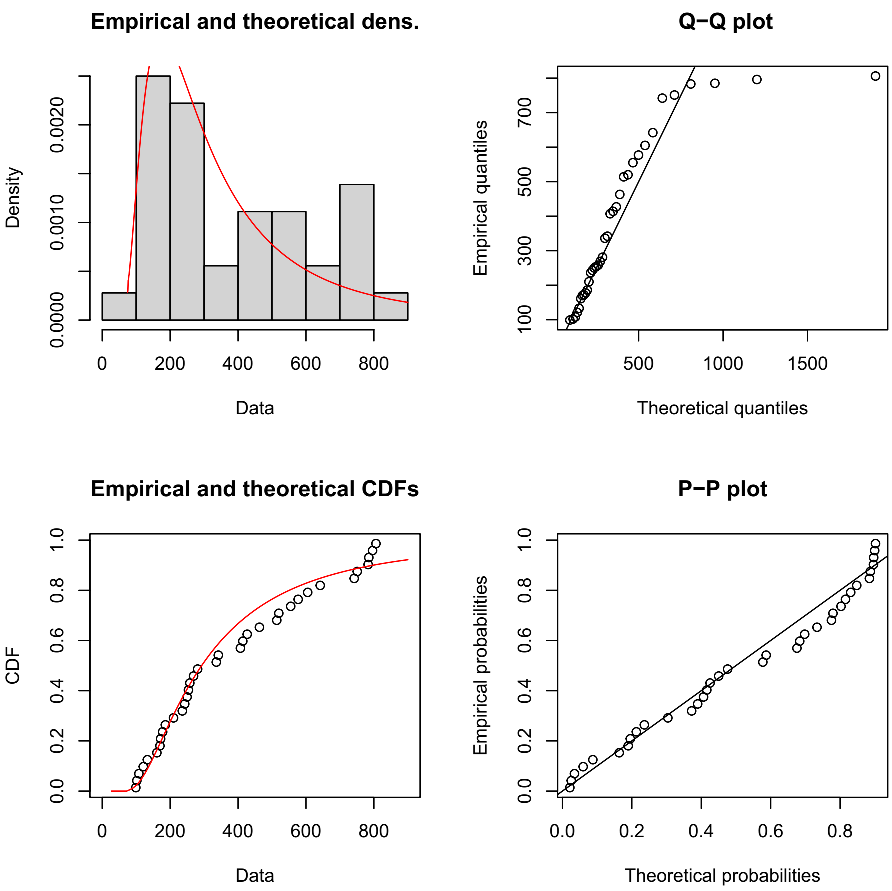


Fig. 5 Fitted cdf, pdf, Q-Q and P-P plots of confirmed COVID-19 cases recorded in Pakistan from March 24 to April 28, 2020 for the TLEtP distribution

smallest AIC and BIC values compared to the other fitted models. This fit can also be seen in Figs. 4 and 5.

3.1.1 Example 2

The second data set as listed in the Table 3 represents the active COVID-19 cases per day in South Africa from March 11 to August 31, 2020.

Table 3 Active COVID-19 cases per day in South Africa from March 11 to August 31, 2021

6	3	8	14	23	1	23	31	34	52	38	34
52	38	34	128	152	155	218	243	17	93	46	8
82	43	80	70	31	63	96	89	69	25	145	99
143	91	99	178	251	124	142	165	170	318	267	141
185	247	203	354	297	304	385	447	437	352	236	424
663	525	595	637	698	724	665	785	831	1160	918	767
803	1134	988	1218	1240	1032	649	1673	1466	1837	1727	1716
1674	1455	1713	3267	2642	2539	2312	2594	2112	2430	3147	3359
3809	4302	3495	2801	4078	3478	3825	4966	4621	4288	4518	5688
6579	6215	7210	6334	6130	6945	8124	8728	9063	10,853	8773	8971
10,134	8810	13,734	12,288	13,497	12,058	11,554	10,496	12,757	13,172	13,373	13,285
13,449	9300	8170	13,150	13,104	13,944	12,204	11,233	7096	7232	11,362	11,046
11,014	10,107	8195	5377	4456	8559	8307	7292	7712	6670	3740	2511
2810	3946	6275	4513	3692	2541	2258	3916	3880	3398	3707	2728
1677	1567	2684	2585	1846	2419	2505	1985				

Table 4 The MLEs and information criteria of the models based on confirmed COVID-19 cases per day in South Africa from March 11 to August 31, 2020

Model	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\beta}$	-ll	AIC	BIC
TLEtP	11.1481	0.2743	3.4610	0.0128	1595.0150	3198.0300	3210.644
TLPI	73.8834	0.0124	0.2106	–	1608.6940	3223.3890	3232.849
Pareto	0.2941	36.5483	–	–	1603.8520	3211.7040	3221.8560

The analysis of the data from South Africa showed similar results that of Pakistan, i.e., the model with the TLEtP distribution showed a better fit compared to other models. From Table 4, the TLEtP distribution has the highest log-likelihood and the smallest AIC and BIC. This fit can be seen in Figs. 6 and 7.

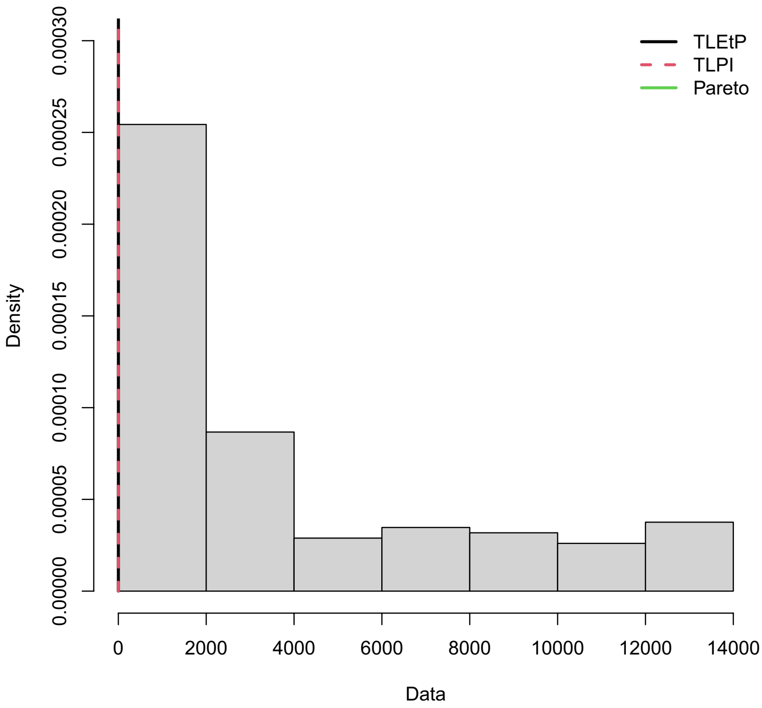


Fig. 6 Fitted pdfs for the TLEtP, TLPI, and Pareto distribution based on confirmed COVID-19 cases per day in South Africa from March 11 to August 31, 2020

4 Conclusion

Recently, statisticians have increasingly created extended distributions from families of existing distributions. We have successfully created the novel Topp-Leone Exponentiated Pareto (TLEtP) family of distributions. Furthermore, some mathematical characteristics of the newly developed distribution are derived. By analyzing the COVID-19 outbreak in Pakistan and South Africa and comparing the data with two members of the Topp-Leone Pareto type I (TLPI) and Pareto (P) families of distributions, we show the adaptability and applicability of the new distributions. Compared to (TLPI) and (P) distributions, the proposed distribution (TLEtP) provides a better fit for the data set. As a result, we anticipate that the proposed distributions (TLEtP) will be useful in a variety of domains, particularly in reliability studies when the hazard rate is increasing.

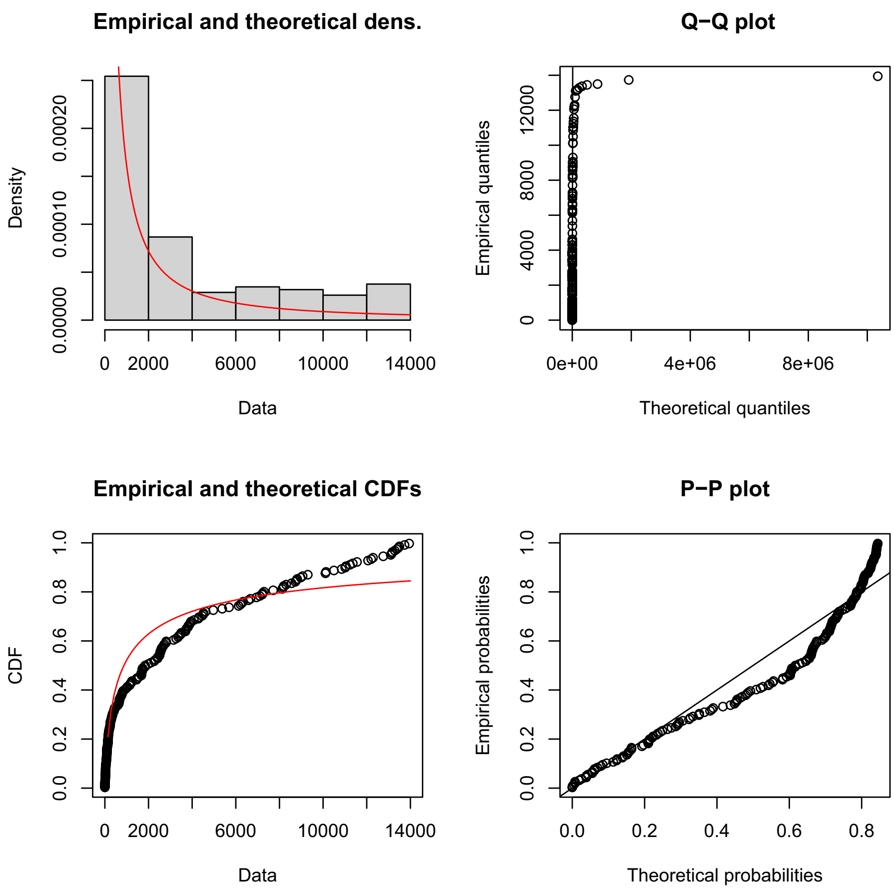


Fig. 7 Fitted cdf, pdf, Q-Q, and P-P plots of confirmed COVID-19 cases per day in South Africa from March 11 to August 31, 2020 for the TLEtP distribution

Acknowledgements The authors are grateful to the anonymous referee for a careful review and helpful comments, which improved this paper. We would like to thank the University of the Free State for paying the APC.

Author Contributions Conceptualization of the study was done by FMC, BJO, IS, and OAB, and all participated in its design and coordination. FMC, BJO, and IS were involved in the formal investigation, and FMC and OAB performed the statistical analysis. This was concluded with review and editing by FMC, BJO, IS, and OAB, and all authors read and approved the manuscript.

Funding No funding was received to conduct this study.

Data Availability The data is available in the article

Declarations

Consent to participate Not applicable

Ethics approval The research work in this article is new and unique and it contains no ethics issues.

Consent for publication Not applicable

References

1. COVID - Coronavirus Statistics - Worldometer. https://www.worldometers.info/coronavirus/ Accessed 2024-03-19
2. Anderson, R.M., Heesterbeek, H., Klinkenberg, D., Hollingsworth, T.D.: How will country-based mitigation measures influence the course of the COVID-19 epidemic? **395**(10228), 931–934 (2020) https://doi.org/10.1016/S0140-6736(20)30567-5 . Accessed 2024-03-19
3. Mitjà, O., Arenas, A., Rodó, X., Tobias, A., Brew, J., Benlloch, J.M.: Experts' request to the spanish government: move spain towards complete lockdown **395**(10231), 1193–1194 (2020) https://doi.org/10.1016/S0140-6736(20)30753-4 . Accessed 2024-03-19
4. Saez, M., Tobias, A., Varga, D., Barceló, M.A.: Effectiveness of the measures to flatten the epidemic curve of COVID-19. the case of spain **727**, 138761 (2020) https://doi.org/10.1016/j.scitotenv.2020.138761 . Accessed 2024-03-19
5. Gupta, R.C., Gupta, P.L., Gupta, R.D.: Modeling failure time data by lehman alternatives. Communications in Statistics - Theory and Methods **27**(4), 887–904 https://doi.org/10.1080/03610929808832134 . Accessed 2023-09-17
6. Alzaatreh, A., Lee, C., Famoye, F.: A new method for generating families of continuous distributions. METRON **71**(1), 63–79 https://doi.org/10.1007/s40300-013-0007-y . Accessed 2023-09-17
7. Marshall, A.W., Olkin, I.: A new method for adding a parameter to a family of distributions with applicatino to the exponental and weibull families. Biometrika **84**(3), 641–652
8. Eugene, N., Lee, C., Famoye, F.: BETA-NORMAL DISTRIBUTION AND ITS APPLICATIONS. Communications in Statistics - Theory and Methods **31**(4), 497–512 https://doi.org/10.1081/STA-120003130 . Accessed 2023-09-17
9. Cordeiro, G.M., Ortega, E.M.M., Cunha, D.C.C.D.: The exponentiated generalized class of distributions. Journal of Data Science **11**(1), 1–27 https://doi.org/10.6339/JDS.2013.11(1).1086 . Accessed 2023-09-17
10. Aslam, M., Afzaal, M., Ishaq Bhatti, M.: A study on exponentiated gompertz distribution under bayesian discipline using informative priors. Statistics in Transition New Series **22**(4), 101–119 (2021) https://doi.org/10.21307/stattrans-2021-040 . Accessed 2023-12-27
11. Yousof, H.M., Afify, A.Z., Hamedani, G.G., Aryal, G.: The burr x generator of distributions for lifetime data. Journal of Statistical Theory and Applications **16**(3), 288 https://doi.org/10.2991/jsta.2017.16.3.2 . Accessed 2023-09-17
12. Alizadeh, M., Ghosh, I., Yousof, H.M., Rasekhi, M., Hamedani, G.G.: The generalized odd generalized exponential family of distributions: Properties, characterizations and application.

Journal of Data Science **15**(3), 443–466 https://doi.org/10.6339/JDS.201707_15(3).0005 . Accessed 2023-09-17

13. Cordeiro, G.M., Ortega, E.M.M., Popović, B.V., Pescim, R.R.: The lomax generator of distributions: Properties, minification process and regression model. Applied Mathematics and Computation **247**, 465–486 https://doi.org/10.1016/j.amc.2014.09.004 . Accessed 2023-09-17
14. Younis, F., Aslam, M., Bhatti, M.I.: Preference of prior for two-component mixture of lomax distribution. Journal of Statistical Theory and Applications **20**(2), 407 (2021). https://doi.org/10.2991/jsta.d.210616.002. Accessed 2023-12-27
15. Hamedani, D.G.G., Yousof, H.M., Rasekhi, M., Alizadeh, M., Najibi, S.M.: Type i general exponential class of distributions. Pakistan Journal of Statistics and Operation Research **14**(1), 39 https://doi.org/10.18187/pjsor.v14i1.2193 . Accessed 2023-10-11
16. Mansour, M.M., Abd Elrazik, E.M., Afify, A.Z., Ahsanullah, M., Altun, E.: The transmuted transmuted-g family: properties and applications. Journal of Nonlinear Sciences and Applications **12**(4), 217–229 https://doi.org/10.22436/jnsa.012.04.03 . Accessed 2023-10-11
17. Alizadeh, M., Cordeiro, G.M., Pinho, L.G.B., Ghosh, I.: The Gompertz-G family of distributions. Journal of Statistical Theory and Practice **11**(1), 179–207 (2017). https://doi.org/10.1080/15598608.2016.1267668. Accessed 2023-04-19
18. Al-Marzouki, S., Jamal, F., Chesneau, C., Elgarhy, M.: Topp-leone odd frechet generated family of distributions with applicationsto COVID-19 data sets. Computer Modeling in Engineering & Sciences **125**(1), 437–458 https://doi.org/10.32604/cmes.2020.011521 . Accessed 2023-09-17
19. Ibrahim, S., Doguwa, S.I., Audu, I., Muhammad, J.H.: On the topp leone exponentiated-g family of distributions: Properties and applications. Asian Journal of Probability and Statistics, 1–15 https://doi.org/10.9734/ajpas/2020/v7i130170 . Accessed 2023-10-05