Pak.j.stat.oper.res. Vol.18 No.4 2022 pp 817-836 DOI: https://dx.doi.org/10.18187/pjsor.v18i4.3699

Pakistan Journal of Statistics and Operation Research

The Topp-Leone Extended Exponential Distribution: Estimation Methods and Applications

Aidi Khaoula[1*], N. Seddik-Ameur[2],
Abd EL-Baset A. Ahmad[3], Mundher A. Khaleel[4]

[*] Corresponding author

1. Laboratory of probability and statistics LaPS, University Badji Mokhtar- Annaba- Algeria, khaoula.aidi@yahoo.fr
2. Department of Mathematics, Assuit University, Assuit 71515, Egypt., abdelbasetsalem@yahoo.com
3. Laboratory of probability and statistics LaPS, University Badji Mokhtar- Annaba- Algeria , naciraseddik@yahoo.fr
4. Department of Mathematics, Tikrit University, Tikrit, Iraq, mun880088@tu.edu.iq .

Abstract

Al-Shomrani et al. (2016) introduced a new family of distributions $(TL-G)$ based on the Topp-Leone distribution (TL) by replacing the variable x by any cumulative distribution function $G(t)$. With only one extra parameter which control the skewness, this family is a good competitor to several generalized distributions used in statistical analysis. In this work, we consider the extended exponential as the baseline distribution G to obtain a new model called the Topp-Leone extended exponential distribution $TL-EE$. After studying mathematical and statistical properties of this model, we propose different estimation methods such as maximum likelihood estimation, method of ordinary and weighted least squares, method of percentile, method of maximum product of spacing, method of Cramer Von-Mises, modified least squares estimators and chi-square minimum method for estimating the unknown parameters. In addition to the classical criteria for model selection, we develop for this distribution a goodness-of-fit statistic test based on a modification of Pearson statistic. The performances of the methods used are demonstrated by an extensive simulation study. With applications to covid-19 data and waiting times for bank service, a comparison evaluation shows that the proposed model describes data better than several competing distributions.

Key Words: Maximum likelihood estimation, method of percentile, method of maximum product of spacing, method of Cramer Von-Mises, modified least squares estimators.

Mathematical Subject Classification: 62E10- 62F10- 62F12- 62G30.

1. Introduction

As it's a good competitor to beta distribution, the Topp-Leone distribution (TL) introduced by Topp and Leone (1955) and revisited by Nadarajah and Kotz (2003) attracts the attention of many authors. Ghitany et al. (2005) studied its hazard rate, mean residual life, reversed hazard rate, expected inactivity time, and its stochastic orderings, Dorp and Kotz (2006) showed its applications in financial engineering, Zhou et al. (2006) derived the distribution of some combinaisons of Topp-Leone variables, Kotz and Seier (2007) studied the behaviour of the kurtosis and Genç (2012) derived the moment of the order statistics. Despite its advantages, this J-shaped model support is bounded on $(0, 1)$ which reduces its applications, this is why some authors have proposed to extend its support. Among these genaralizations, we are interesting in a new family of distributions so-called the Topp-Leone family distributions $(TL-G)$ proposed recently by Al-Shomrani et al. (2016). With only one extra parameter which can control the skewness, these

new distributions are able to model different data because they can have both heavy and light tails which makes their scope wider than many of the models used in the analyzes. By replacing the Topp-Leone variable x by any cumulative distribution function $G(t)$, the proposed models will no longer be limited to data bounded between 0 and 1. After studying its statistical properties, the authors gave an example where $G(t)$ is the famous exponential distribution. Recently Rezaei et al. (2017) proposed the use of $[G(t)]^{\theta}$ instead of $G(t)$ to obtain an other generalization called TL-generated (TLG) distribution. The authors studied the maximum likelihood estimators and some special cases of this family, whereas Arshad and Jamal (2019) investigated the estimation of the scale parameter, the shape parameter and the reliability function based on recorded data. Also bayesian estimation of the unknown parameters is developed. Till now, researchers propose different forms of generalizations of Topp-Leone G family, such as Toppe-Leone power series distribution, type II Topp-Leone generated family, type II generalized Topp-Leone family of distributions and many others which are cited in Bantan et al. (2020). Along the same lines, we consider a new model called the Topp-Leone-extended exponential (TL-EE) distribution.
The extended exponential distribution introduced by Nadarajah and Haghighi (2010) is a good alternative to Weibull, exponentiated exponential and gamma distributions which can have serious limitation in the case where the pdf is monotonically decreasing and the hazard rate function is increasing. Despite its flexibility, this model has closed forms of the survival and rate functions which can be increasing, decreasing or constant. Thus motiveted us to use the extended exponential distribution (EE) as the baseline of $TL-G$ family. After the presentation of the TL-EE model, we develop its mathematical characteristics and statistical properties as the expansions of the cdf and pdf, the probability weighted moments and the order statistics, we propose different techniques to estimate the unknown parameters namely maximum likelihood method, method of ordinary and weighted least squares, method of percentile, method of maximum product of spacing, method of Cramer Von-Mises and modified least squares estimators. In addition to the classical criteria for model selection, we develop for this distribution a goodness-of-fit statistic test based on a modification of the chi-square Pearson statistic. An extensive simulation study is conducted to show the performances of the methods used. With an application to covid-19 data and waiting times for bank service confirm the usefullness of the proposed model.

2. Topp-Leone Extended Exponential Distribution (TL-EE)

As we can see in the statistical literature, several generalizations of classical distributions were proposed to better describe the observed data which become numerous and complex. Among these generalizations, the extended exponential distribution (EE), introduced by Nadarajah and Haghighi (2010), with parameters α and β given by its cumulative distribution function (cdf)

$$G(t,\alpha,\beta) = 1- \exp\left\{1-(1+\alpha t)^{\beta}\right\}, \quad t>0\ \alpha>0, \beta>0 \tag{1}$$

and its probability density function (pdf)

$$g(t,\alpha,\beta) = \alpha\beta\,(1+\alpha t)^{\beta-1}\exp\left\{1-(1+\alpha t)^{\beta}\right\}, \quad t>0, \alpha>0, \beta>0 \tag{2}$$

provides better fits than many classical models.
Also, Al-Shomrani et al. (2016) proposed a very interesting new family of distributions called the Topp-Leone-G distributions $(TL-G)$ based on the famous Topp-Leone model by replacing the variable t by any cumulative distribution function $G(t)$, called the baseline distribution. The cdf of the obtained model is given as follow:

$$F(t,\lambda) = \left\{1-[1-G(t)]^{2}\right\}^{\lambda}, \quad t\in\mathbb{R}, \lambda>0 \tag{3}$$

In this case the variable t is no longer limited between 0 and 1, which makes it possible to describe more data than the Topp-Leone distribution (TL) and the baseline distribution $G(t)$ also. Its corresponding pdf is:

$$f(t,\lambda) = 2\lambda g(t)\left\{1-[1-G(t)]\right\}^{\lambda-1}, \quad t\in\mathbb{R}, \lambda>0 \tag{4}$$

where λ is the extra shape parameter. By substituting Equations (1) and (2) into Equation (3), we derive the cdf of the

Topp-Leone Extended Exponential distribution $TL-EE$ as follows:

$$F(t,\alpha,\beta,\lambda)=\left\{1-e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\right\}^{\lambda},\quad t>0\ \alpha>0,\beta>0,\lambda>0 \tag{5}$$

So, Its corresponding pdf is:

$$f(t,\alpha,\beta,\lambda)=2\alpha\beta\lambda\left(1+\alpha t\right)^{\beta-1}e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\left\{1-e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\right\}^{\lambda-1} \tag{6}$$

The expressions for the survival function and the failure rate of the $TL-EE$ distribution become respectively:

$$S(t,\alpha,\beta,\lambda)=1-\left\{1-e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\right\}^{\lambda},\quad t>0\ \alpha>0,\beta>0,\lambda>0 \tag{7}$$

$$h(t,\alpha,\beta,\lambda)=\frac{2\alpha\beta\lambda\left(1+\alpha t\right)^{\beta-1}e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\left\{1-e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\right\}^{\lambda-1}}{1-\left\{1-e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\right\}^{\lambda}} \tag{8}$$

2.1. Quantile Function

The quantile function for this new model is obtained from the inverse of the cdf

$$Q(t,\theta)=q=F^{-1}(t,\theta)$$

as follows:

$$t=\frac{1}{\alpha}-\frac{1}{\alpha}\left[1-\frac{\ln(1-q^{1/\lambda})}{2}\right]^{1/\beta} \tag{9}$$

As shown in $Fig(1-2)$, the proposed model pdf can be decreasing, unimodal and skewed while the hazard function is increasing, decreasing, bathtub and J shapes.

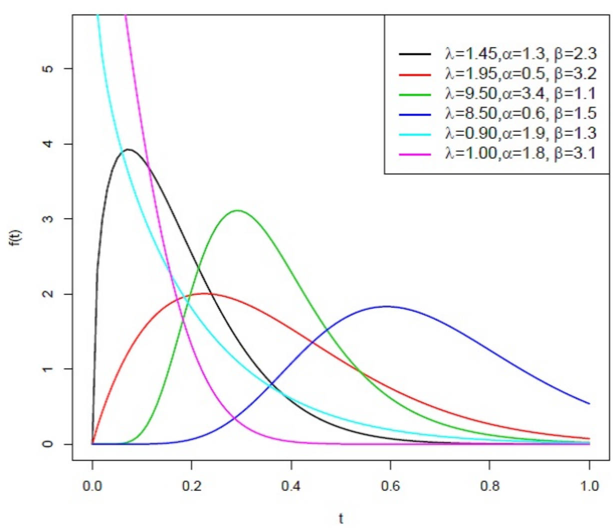

Fig 1. Probability density function of TL-EE distribution

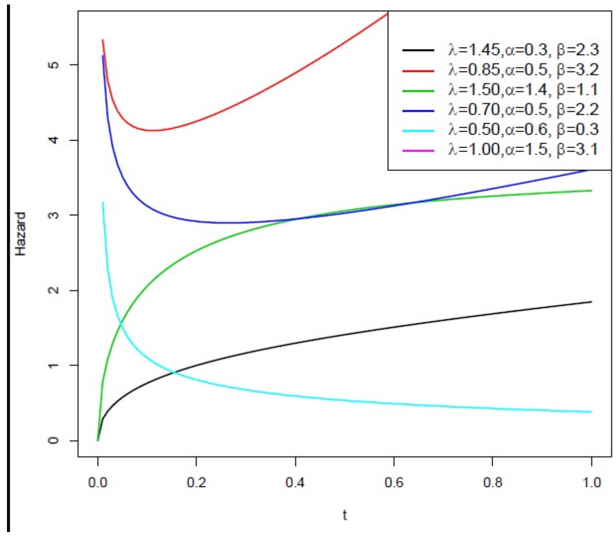

Fig 2. Hazard function of TL-EE distribution

3. Mathematical and Statistical Properties

In this section, we obtain useful mathematical representation for $TL-EE$ cdf and pdf forms, which will be used to obtain probability weighted moments, moments, moment generating function and order statistics of this distribution.

3.1. Expansions of the cdf and pdf of TL-EED

From the forms of the cdf and pdf of $TL-EE$ distribution in (5) and (6), we can find $F^s f$, as follows

$$F^s(t)f(t) = 2\alpha\beta\lambda\,(1+\alpha t)^{\beta-1}\, e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\left\{1-e^{2\left\{1-(1+\alpha t)^{\beta}\right\}}\right\}^{(s+1)\lambda-1}, \tag{10}$$

by using the generalized binomial expansion, (10) can be written in the form

$$F^s(t)f(t) = 2\alpha\beta\lambda(1+\alpha t)^{\beta-1}\sum_{j=0}^{\infty}(-1)^j A(j)\binom{(s+1)\lambda-1}{j}e^{-2(j+1)(1+\alpha t)^{\beta}}, \tag{11}$$

where $A(j) = e^{2(j+1)}$. So, the pdf (6), can be written in another form, by putting $s=0$ in (11), as follows

$$f(t) = 2\alpha\beta\lambda(1+\alpha t)^{\beta-1}\sum_{j=0}^{\infty}(-1)^j A(j)\binom{\lambda-1}{j}e^{-2(j+1)(1+\alpha t)^{\beta}}. \tag{12}$$

3.2. Probability weighted moments

Making use of the definition of probability weighted moments (PWMs) for a random variable X, denoted by $M_{r,s,m}$, see Greenwood et al.(1979), we have

$$M_{r,s,m} = E\left[X^r F^s(X)\bar{F}^m(X)\right] = \int_0^{\infty} x^r f(x)F^s(x)\bar{F}^m(x)dx. \tag{13}$$

We consider the PWM quantity $M_{r,s,0}\equiv\mu_{r,s}$, and $M_{r,0,0}\equiv\mu_r$ that represents rth non-central moments. So, putting $m=0$ in (13), and use (11) we obtain

$$\mu_{r,s} = 2\alpha\beta\lambda\sum_{j=0}^{\infty}(-1)^j A(j)\binom{(s+1)\lambda-1}{j}\int_0^{\infty} x^r(1+\alpha x)^{\beta-1}e^{-2(j+1)(1+\alpha t)^{\beta}}dx.$$

Making use of the transformation $(1+\alpha t)^{\beta}=z$, then after algebraic calculations we obtain

$$\mu_{r,s}=\frac{2\lambda}{\alpha^{r}}\sum_{j=0}^{\infty}\sum_{i=0}^{r}(-1)^{i+j+r}A(j)\binom{(s+1)\lambda-1}{j}\binom{r}{i}\int_{1}^{\infty}z^{i/\beta}e^{-2(j+1)z}dz,$$

considering the use of the formula $\int_{x}^{\infty}z^{k-1}e^{-Bz}dz=\frac{(k-1)!}{B^{k}}\sum_{\ell=0}^{k-1}\frac{(Bx)^{\ell}}{\ell!}e^{-Bx}$, with the use of $[i/\beta]$ (integer part of the value i/β) instead of i/β, we can obtain

$$\mu_{r,s}=\frac{\lambda}{\alpha^{r}}\sum_{j=0}^{\infty}\sum_{i=0}^{r}\sum_{\ell=0}^{[i/\beta]}\frac{(-1)^{i+j+r}\binom{r}{i}\binom{(s+1)\lambda-1}{j}[i/\beta]!}{2^{\frac{i}{\beta}-\ell}(j+1)^{\frac{i}{\beta}-\ell+1}\ell!}. \tag{14}$$

The r^{th} non-central moments, μ_r', of the $TL-EE$ distribution can be obtained from (14), setting $s=0$, by

$$\mu_r'=\frac{\lambda}{\alpha^{r}}\sum_{j=0}^{\infty}\sum_{i=0}^{r}\sum_{\ell=0}^{[i/\beta]}\frac{(-1)^{i+j+r}\binom{r}{i}\binom{\lambda-1}{j}[i/\beta]!}{2^{\frac{i}{\beta}-\ell}(j+1)^{\frac{i}{\beta}-\ell+1}\ell!}. \tag{15}$$

Also, $M_{1,s,0}\equiv\mu_{1,s}=E\left[XF^{s}(X)\right]$, L-moment, can be given from (14), setting $r=1$, by

$$\mu_{1,s}\frac{\lambda}{\alpha}\sum_{j=0}^{\infty}\sum_{i=0}^{1}\sum_{\ell=0}^{[i/\beta]}\frac{(-1)^{i+j+1}\binom{(s+1)\lambda-1}{j}[i/\beta]!}{2^{\frac{i}{\beta}-\ell}(j+1)^{\frac{i}{\beta}-\ell+1}\ell!}.$$

The moment generating function, $M_X(t)$, using the Maclaurin series expansion of the function $\exp(-tx)$ in terms of μ_r', can be given by the form

$$M_X(t)=\sum_{r=0}^{\infty}\frac{t^{r}}{r!}\mu_r',$$

which, using (15), gives

$$M_X(t)=\sum_{r,j=0}^{\infty}\sum_{i=0}^{r}\sum_{\ell=0}^{[i/\beta]}\frac{t^{r}}{r!}\frac{(-1)^{i+j+r}\lambda\binom{r}{i}\binom{\lambda-1}{j}[i/\beta]!}{2^{\frac{i}{\beta}-\ell}\alpha^{r}(j+1)^{\frac{i}{\beta}-\ell+1}\ell!}$$

3.3. Order statistics

The order statistics has a fundamental role in non-parametric statistics and statistical inference. So, we derive distributions of order statistics for the proposed distribution and establish their moments. If $X_1,X_2,\ldots,X_n$ is a random sample from a population with cdf $F(x)$ and pdf $f(x)$, then the corresponding order statistics are given by $X_{1:n},X_{2:n},\ldots,X_{n:n}$. The pdf of the d^{th} order statistic $X_{d:n}$, is given by

$$f_{d:n}(x)=\frac{n!}{(d-1)!(n-d)!}\left[F(x)\right]^{d-1}\left[\bar{F}(x)\right]^{n-d}f(x),$$

which can be written as

$$f_{d:n}(x)=\frac{n!}{(d-1)!}\sum_{w=0}^{n-d}\frac{(-1)^{w}}{w!(n-d-w)!}f(x)\left[F(x)\right]^{d+w-1}. \tag{16}$$

Substituting from (13), replacing s by $d+w-1$, into (16), we obtain

$$f_{d:n}(x)=2\alpha\beta\lambda\frac{n!}{(d-1)!}\sum_{j=0}^{\infty}\sum_{w=0}^{n-d}\frac{(-1)^{j+w}A(j)\binom{(w+d)\lambda-1}{j}}{w!(n-d-w)!}(1+\alpha x)^{\beta-1}e^{-2(j+1)(1+\alpha x)^{\beta}} \tag{17}$$

The r^{th} moment of the d^{th} order statistic can be given, from (17) and the same as deriving $\mu_{r,s}$, by

$$\mu_{d:n}^{(r)} = \frac{\lambda n!}{\alpha^r (d-1)!} \sum_{j=0}^{\infty} \sum_{w=0}^{n-d} \sum_{i=0}^{r} \sum_{\ell=0}^{[i/\beta]} \frac{(-1)^{i+j+w+r} 2^{\ell-\frac{i}{\beta}} \binom{r}{i} \binom{(w+d)\lambda-1}{j} [i/\beta]!}{w!\ell!(n-d-w)!(j+1)^{\frac{i}{\beta}-\ell+1}}.$$

4. Estimation methods

As it's well known, the properties of the maximum likelihood estimators are not always verified for small samples, this is why in the recent years, classical and new estimation methods have been developed. One purpose of this work is to investigate different methods to estimate the unknown parameters of this new model such as maximum likelihood estimation, ordinary least square, weighted least square methods and some methods based on the empirical function distribution.

4.1. Maximum Likelihood Estimation

Because of their properties the maximum likelihood estimators are preferred in providing the values of the unknown parameters. Consider $t_1, t_2, ...t_n$ a random sample distributed according to the $TL-EE$ distribution with parameters (α, β, λ), the likelihood function is

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

The log-likelihood function becomes

$$\log L = n\ln(2) + n\ln(\alpha\lambda\beta) + (\beta-1)\sum_{i=1}^{n}\ln(1+\alpha t_i) + 2n - 2\sum_{i=1}^{n}(1+\alpha t_i)^{\beta} + (\lambda-1)\sum_{i=1}^{n}\ln\left(1-e^{2\left\{1-(1+\alpha t_i)^{\beta}\right\}}\right)$$

the maximum likelihood estimators $\widehat{\alpha}$, $\hat{\beta}$ and $\widehat{\lambda}$ of the unknown parameters α, β and λ are derived from the nonlinear following score equations:

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} - 2\beta\sum_{i=1}^{n} t_i u_i^{\beta-1} + (\beta-1)\sum_{i=1}^{n}\frac{t_i}{1+\alpha t_i} + 2(\lambda-1)\beta\sum_{i=1}^{n}\frac{t_i u_i^{\beta-1} e^{2(1-u_i^{\beta})}}{1-e^{2(1-u_i^{\beta})}}$$

$$\frac{\partial L}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n}\ln u_i - 2\sum_{i=1}^{n} u_i^{\beta}\ln u_i + 2(\lambda-1)\sum_{i=1}^{n}\frac{u_i^{\beta}\ln(u_i)\, e^{2(1-u_i^{\beta})}}{1-e^{2(1-u_i^{\beta})}}$$

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n}\ln\left(1-e^{2\left\{1-(1+\alpha t_i)^{\beta}\right\}}\right)$$

where $u_i(\alpha, t_i) \equiv u_i = 1 + \alpha t_i$.

4.2. Method of least squares and weighted least squares

As the explicit forms of the maximum likelihood estimators cannot be obtained every time, so other methods are developed to overcome this problem. The least square (LS) and the weighted least square (WLS) are well known methods used for estimating the unknown parameters (Swain et al.,1988). Here, we consider the two methods to estimate the unknown parameters of the $TL-EE$ distribution. Let $t_{1,} t_{2.....}, t_n$ be the ordered observations obtained from a sample of size n from the $TL-EE$ distribution. By calculating the minimum of the function

$$S(\theta) = \sum_{i=1}^{n} \eta_i \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{i}{n+1} \right\}^2$$

with respect to α, λ and β respectively, the LS estimates $\hat{\alpha}_{LSE}$, $\hat{\lambda}_{LSE}$ and $\hat{\beta}_{LSE}$ can be obtained by setting $\eta_i = 1$, while we can obtain the WLS estimates $\hat{\alpha}_{WLS}$; $\hat{\lambda}_{WLS}$ and $\hat{\beta}_{WLS}$ by setting $\eta_i = \frac{(n+1)^2(n+2)}{i(n-i+1)}$. These estimates can also be obtained by solving the following equations:

$$\frac{\partial S(\theta)}{\partial \alpha} = \sum_{i=1}^{n} \eta_i \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{i}{n+1} \right\} \varphi_1(t_i, \theta) = 0$$

$$\frac{\partial S(\theta)}{\partial \beta} = \sum_{i=1}^{n} \eta_i \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{i}{n+1} \right\} \varphi_2(t_i, \theta) = 0$$

$$\frac{\partial S(\theta)}{\partial \lambda} = \sum_{i=1}^{n} \eta_i \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{i}{n+1} \right\} \varphi_3(t_i, \theta) = 0$$

where

$$\varphi_1(t_i, \theta) = 2\beta\lambda t_i u_i^{\beta-1} e^{2\left\{1-u_i^{\beta}\right\}} \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda-1}$$

$$\varphi_2(t_i, \theta) = 2\lambda u_i^{\beta} \ln u_i e^{2\left\{1-u_i^{\beta}\right\}} \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda-1}$$

$$\varphi_3(t_i, \theta) = \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} \ln \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}$$

where u_i, $i = 1, 2, ...n$ are the order observations of u_i as defined earlier.

4.3. Method of percentile

In this subsection, we estimate the unknown parameters of the $TL-EE$ distribution by the percentile method. This method was first introduced by Kao (1958) for estimating Weibull parameters. Let $p_i = \frac{i}{n+1}$ be the estimate of $F(t_i, \theta)$, then the percentile estimators $\hat{\alpha}_{PE}$, $\hat{\lambda}_{PE}$ and $\hat{\beta}_{PE}$ of the $TL-EE$ distribution parameters are the minimum with respect to α, λ and β of the function:

$$P(\theta) = \sum_{i=1}^{n} \left\{ t_i - \left[\frac{1}{\alpha} - \frac{1}{\alpha} \left[1 - \frac{\ln(1-p_i^{1/\lambda})}{2} \right]^{1/\beta} \right] \right\}^2$$

They are obtained as the solution of the system of equations of the first derivatives of the function above:

$$\frac{\partial P(\theta)}{\partial \alpha} = \sum_{i=1}^{n} \left\{ x_i - \left[\frac{1}{\alpha} - \frac{1}{\alpha} \left[1 - \frac{\ln(1-p_i^{1/\lambda})}{2} \right]^{1/\beta} \right] \right\} \varpi_1(t_i, \theta) = 0$$

$$\frac{\partial P(\theta)}{\partial \beta} = \sum_{i=1}^{n} \left\{ x_i - \left[\frac{1}{\alpha} - \frac{1}{\alpha} \left[1 - \frac{\ln(1-p_i^{1/\lambda})}{2} \right]^{1/\beta} \right] \right\} \varpi_2(t_i, \theta) = 0$$

$$\frac{\partial P(\theta)}{\partial \lambda} = \sum_{i=1}^{n} \left\{ x_i - \left[\frac{1}{\alpha} - \frac{1}{\alpha} \left[1 - \frac{\ln(1-p_i^{1/\lambda})}{2} \right]^{1/\beta} \right] \right\} \varpi_3(t_i, \theta) = 0$$

where $v_i(p_i, \lambda) \equiv v_i = 1 - \frac{\ln(1-p_i^{1/\lambda})}{2}$

$$\varpi_1(v_i,\theta) = -\frac{1}{\alpha^2} + \frac{1}{\alpha^2} v_i^{1/\beta}$$

$$\varpi_2(v_i,\theta) = \frac{1}{\alpha\beta^2} v_i^{1/\beta} \ln v_i$$

$$\varpi_3(v_i,\theta) = \frac{1}{2\alpha\beta\lambda^2} \frac{p_i^{1/\lambda} \ln(p_i) v_i^{\frac{1}{\beta}-1}}{1-p_i^{1/\lambda}}$$

4.4. Method of maximum product of spacing

According to Cheng and Amin (1983) and based on the idea of the differences between the values of the cdf at consecutive data points, the maximum product of spacing (MPS) is an estimation method as interesting as that of MLE. Moreover, Al-Mofleh and Afify (2019) concluded that the MPS estimator method outperforms all the other estimator methods.
Based on a random sample of size n from a distribution with cdf $F(t_i,\theta)$, the uniform spacings can be defined as follows

$$D_i(\theta) = F(t_i,\theta) - F(t_{i-1},\theta), \quad i = 1,2,...,n$$

where $F(t_0,\theta) = 0$ and $F(t_{n+1},\theta) = 1$. The MPS estimates denoted by $\hat{\alpha}_{MPS}, \hat{\lambda}_{MPS}$ and $\hat{\beta}_{MPS}$ maximizes the function $M(\theta)$ with respect to the unknown parameters α,λ and β :

$$\begin{aligned} M(\theta) &= \frac{1}{n+1}\sum_{i=1}^{n+1} \log D_i(\theta) \\ &= \frac{1}{n+1}\sum_{i=1}^{n+1} \log\left\{\left\{1-e^{2\left\{1-u_i^{\beta}\right\}}\right\}^{\lambda} - \left\{1-e^{2\left\{1-u_{i-1}^{\beta}\right\}}\right\}^{\lambda}\right\} \end{aligned}$$

or by solving the following equations

$$\frac{\partial M(\theta)}{\partial\alpha} = \frac{1}{n+1}\sum_{i=1}^{n+1} \frac{\varphi_1(t_i,\theta) - \varphi_1(t_{i-1},\theta)}{D_i} = 0$$

$$\frac{\partial M(\theta)}{\partial\beta} = \frac{1}{n+1}\sum_{i=1}^{n+1} \frac{\varphi_2(t_i,\theta) - \varphi_2(t_{i-1},\theta)}{D_i} = 0$$

$$\frac{\partial M(\theta)}{\partial\lambda} = \frac{1}{n+1}\sum_{i=1}^{n+1} \frac{\varphi_3(t_i,\theta) - \varphi_3(t_{i-1},\theta)}{D_i} = 0$$

where $\varphi_1(t_i,\theta)$, $\varphi_2(t_i,\theta)$ and $\varphi_3(t_i,\theta)$ are given earlier. For more details on this method, one can see Cheng and Stephens (1989).

5. Estimation method based on Goodness-of-fit statistics

The last years, some authors used the classical goodness-of-fit statistics such as Cramer-Von- Mises and Anderson-Darling statistics to derive the estimators of the unknown parameters. Bakouch et al. (2017) used these methods for the binomial exponential 2 distribution while Dey et al. (2020) evaluate the process capability index for normal distribution. Besides these methods and using this approach, we propose the use of the Pearson chi-square statistic X^2 where the estimators noted $CSMM$ are obtained by calculating the minimum of X^2 with respect to the unknown parameters.

5.1. Method of Cramer-Von-Mises

Mac-Donald (1971) and later Boos (1981) showed that the Cramer-von-Mises estimates ($CMEs$) based on the distance of Cramer-von-Mises goodness-of-fit statistics is the least biased compared to the other estimators. For calculation purposes, Boos (1981) gives the formula $C(\theta)$ for this statistic, where the estimators ensure its minimum with respect to the unknown paremeters

$$C(\theta) = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{2i-1}{2n} \right\}^2$$

They also can be obtained as the solution of the following equations:

$$\frac{\partial C(\theta)}{\partial \alpha} = \sum_{i=1}^{n} \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{2i-1}{2n} \right\} \varphi_1(t_i, \theta) = 0$$

$$\frac{\partial C(\theta)}{\partial \beta} = \sum_{i=1}^{n} \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{2i-1}{2n} \right\} \varphi_2(t_i, \theta) = 0$$

$$\frac{\partial C(\theta)}{\partial \lambda} = \sum_{i=1}^{n} \left\{ \left\{ 1 - e^{2\left\{1-u_i^{\beta}\right\}} \right\}^{\lambda} - \frac{2i-1}{2n} \right\} \varphi_3(t_i, \theta) = 0$$

5.2. Anderson-Darling Estimation method

The Anderson-Darling estimation method (ADE) is based on the classical goodness-of-fit statistic test (AD) proposed by Anderson and Darling (1952). Basically, this statistic is used to fit data to a theoretical hypothezised model F_0.When the parameters are unknown, they can be estimated by $\hat{\alpha}_{ADE}$, $\hat{\beta}_{ADE}$, $\hat{\lambda}_{ADE}$ which minimize the (AD) statistic given in the form:

$$ADE(\theta) = -n - n^{-1} \sum_{i=1}^{n} (2i-1) \left\{ \log F(t_i) + \log\left(1 - F(t_{-i+n+1})\right) \right\}$$

which is equivalent to cancel the first derivatives of this function with respect to α, β and λ. For more details on this method, one can consult Boos (1981). In their paper, Rodriguez et al. (2016) showed that the ADE estimation method gives the most efficient estimators. As known the Anderson-Darling statistic gives more weight for the tails of the distribution, so for right or left tailed distributions, right tail (ADE_{R-T}) and left tail (ADE_{L-T}) Anderson-Darling estimators are used in recent papers (Rodriguez et al.(2016), Dey et al. (2017), Ramadan et al. (2020). On an other hand, AKGÜL (2018) demonstrated that the minimum distance method estimation ADE is highly competitive method compared to ML estimation. Also, Al-Mofleh and Afify (2019) confirmed the superiority of this method.

5.3. The Right Tail and Left-Tail Anderson-Darling Estimation methods

As It 's known the Anderson-Darling statistic gives more weight for the tails of the distribution, so for right or left tailed distributions, right tail (ADE_{R-T}) and left tail (ADE_{L-T}) Anderson-Darling estimators are proposed to estimate the unknown parameters like in Rodriguez et al.(2016), Dey et al. (2017) and Ramadan et al. (2020).
The (ADE_{R-T}) $\hat{\alpha}_{ADE_{(R-T)}}, \hat{\beta}_{ADE_{(R-T)}}, \hat{\lambda}_{ADE_{(R-T)}}$ are obtained by minimizing

$$ADE_{(R-T)}(\theta) = \frac{n}{2} - 2\sum_{i=1}^{n} \log F(t_i) - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log\left\{1 - F(t_{-i+n+1})\right\}$$

And the (ADE_{L-T}) $\hat{\alpha}_{ADE_{(L-T)}}, \hat{\beta}_{ADE_{(L-T)}}, \hat{\lambda}_{ADE_{(L-T)}}$ are obtained by minimizing

$$ADE_{(L-T)}(\theta) = -\frac{3}{2}n + 2\sum_{i=1}^{n} F(t_i) - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log\left\{F(t_i)\right\}$$

with respect to the unknown parameters.

5.4. Chi-square minimum method

Based on the famous Pearson chi-square goodness-of-fit statistic X^2, the minimum chi-square estimation method $(CSMM)$ consist in minimizing this statistic with respect to the unknown parameters. Data are grouped into r classes I_j with υ_j numbers of observations and n is the sample size. The $CSMM$ estimators are obtained as $\widetilde{\theta} = \min_\theta X_n^2(\theta)$ where $X_n^2(\theta)$ is the Pearson statistic

$$X_n^2(\theta) = \frac{\sum_{j=1}^{r} (\upsilon_j - nF_j(\theta))^2}{nF_j(\theta)}$$

and $F(\theta)$ is the hypothezised theoretical distribution. For the $TL-EE$ distribution, the $CSMM$ estimators are obtained by canceling the first derivatives of the statistic $X_n^2(\theta)$ with respect to the unknown parameters and where

$$\frac{\partial X_n^2(\theta)}{\partial \alpha} = \left(-\frac{2r\beta\lambda t_i u_i^{\beta-1} e^{2\left\{1-u_i^\beta\right\}}(\upsilon_j - nF_j(\theta))}{1-e^{2\left\{1-u_i^\beta\right\}}}\right)\left(2\left\{1-e^{2\left\{1-u_i^\beta\right\}}\right\}^{\lambda-1} + \frac{\upsilon_j - nF_j(\theta)}{nF_j(\theta)}\right)$$

$$\frac{\partial X_n^2(\theta)}{\partial \beta} = \left(\frac{-2r\lambda u_i^\beta \ln u_i e^{2\left\{1-u_i^\beta\right\}}(\upsilon_j - nF_j(\theta))}{1-e^{2\left\{1-u_i^\beta\right\}}}\right)\left(2\left\{1-e^{2\left\{1-u_i^\beta\right\}}\right\}^{\lambda-1} + \frac{\upsilon_j - nF_j(\theta)}{nF_j(\theta)}\right)$$

$$\frac{\partial X_n^2(\theta)}{\partial \lambda} = \left(-r(\upsilon_j - nF_j(\theta))\ln\left\{1-e^{2\left\{1-u_i^\beta\right\}}\right\}\right)\left(2\left\{1-e^{2\left\{1-u_i^\beta\right\}}\right\}^{\lambda-1} + \frac{\upsilon_j - nF_j(\theta)}{nF_j(\theta)}\right)$$

6. Nikulin-Rao-Robson test statistic

The probability distribution used to describe any phenomenon is very important in the analysis. Since the twentieth century, researchers have not stopped developing techniques to validate the different models. In addition to the classical procedures, we propose in this work a criteria test statistic based on the Nikulin-Rao-Robson statistic (NRR) to fit data to the $TL-EE$ distribution. The great interest of the NRR statistic is that it recovers all the information lost while grouping data. Based on maximum likelihood estimation on initial data, this statistic Y^2 introduced by Nikulin (1975) and Rao and Robson (1974) is a modification of the wellknown chi-square Pearson statistic X^2 **which cannot be applied when the parameters of the distribution to be tested are unknown**. **The NRR statistic Y^2 follows a chi-square distribution with $r-1$ degrees of freedom where r is the number of the classes chosen (for more details one can see Voinov et al. (2013)).**
For testing the null hypothesis H_0 :

$$H_0 = P(T_i \leq t) = F(t,\theta), \qquad t \geq 0.$$

according to which a sample $T_1, T_2, ...T_n$ belongs to a parametric family $F(t,\theta)$ with unknown parameters, Y^2 consist in grouping data into r equiprobable intervals $I_1, I_2, ..., I_r$, where

$$I_j =]a_{j-1}, a_j]; \quad I_i \cap I_j = \phi \quad i \neq j; \qquad \bigcup_{j=1}^{r} I_j = R^1,$$

such as

$$p_j = \int_{a_{j-1}}^{a_j} f(t,\theta)dt = \frac{1}{r}, \;\; j = 1,2,..,r.$$

If $\upsilon = (\upsilon_1, \upsilon_2, ..., \upsilon_r)^T$ represents the number of observed T_i grouping into these intervals I_j, and the vector $X_n(\theta)$ is

$$X_n(\theta) = \left(\frac{\upsilon_1 - np_1(\theta)}{\sqrt{np_1(\theta)}}, \frac{\upsilon_2 - np_2(\theta)}{\sqrt{np_2(\theta)}}, ..., \frac{\upsilon_r - np_r(\theta)}{\sqrt{np_r(\theta)}}\right)^T.$$

The NRR statistic Y^2 is defined by

$$Y_n^2\left(\widehat{\theta}\right) = X_n^2\left(\widehat{\theta}\right) + \frac{1}{n} l^T\left(\widehat{\theta}\right)\left(I\left(\widehat{\theta}\right) - J\left(\widehat{\theta}\right)\right)^{-1} l\left(\widehat{\theta}\right),$$

where $I\left(\widehat{\theta}\right)$ and $J\left(\widehat{\theta}\right)$ are the estimated information matrices on non-grouped and grouped data respectively, and $\widehat{\theta}$ is the vector of the maximum likelihood estimators on initial data. The components of the vector $l\left(\hat{\theta}\right) = \left(l_s\left(\hat{\theta}\right)\right)^T_{1\times s}$ and the matrix $J\left(\widehat{\theta}\right)$ are

$$l_s\left(\hat{\theta}\right) = \sum_{i=1}^{r} \frac{\nu_j}{p_j} \frac{\partial p_j\left(\hat{\theta}\right)}{\partial \hat{\theta}_s}\ ,\ J\left(\widehat{\theta}\right) = \left(\frac{1}{\sqrt{p_j}} \frac{\partial p_j\left(\hat{\theta}\right)}{\partial \theta_k}\right)^T_{r\times s} \left(\frac{1}{\sqrt{p_j}} \frac{\partial p_j\left(\hat{\theta}\right)}{\partial \theta_k}\right)_{r\times s},\ j = 1,2,..,r \text{ and } k = 1,..,s.$$

s represents the parameters number. The distribution of $Y^2\left(\widehat{\theta}\right)$ is a chi-square with $r-1$ degrees of freedom.

To construct the test statistic Y^2 corresponding to the $TL-EE(\theta)$, we calculate the maximum likelihood estimators $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})^T$, the limit intervals a_j and the derivatives $\frac{\partial p_j(\hat{\theta})}{\partial \hat{\theta}_k}$ for our model:

$$\frac{\partial p_j\left(\theta\right)}{\partial \hat{\alpha}} = \varphi_1(t_j,\theta) - \varphi_1(t_{j-1},\theta)$$

$$\frac{\partial p_j(\theta)}{\partial \hat{\beta}} = \varphi_2(t_j,\theta) - \varphi_2(t_{j-1},\theta)$$

$$\frac{\partial p_j(\theta)}{\partial \hat{\lambda}} = \varphi_3(t_j,\theta) - \varphi_3(t_{j-1},\theta)$$

Then, after computing the estimated information matrices $I\left(\widehat{\theta}\right)$ and $J\left(\widehat{\theta}\right)$, we obtain the statistic $Y^2\left(\widehat{\theta}\right)$ for the $TL-EE$ distribution. This method was used to construct goodness-of-fit test statistics for some generalized models, see Chouia and Seddik-Ameur (2014), Aidi and Seddik-Ameur (2016), Treidi and Seddik-Ameur (2016), Tilbi and Seddik-Ameur (2017).

7. Simulations

7.1. Estimation methods

In this section, we study the performance of the estimation methods used in this paper. At that end, we generated $10,000$ samples with different sizes ($n = 10$, $n = 20$, $n = 30$, $n = 80$, $n = 200$ and $n = 350$) from the $TL-EE$ model, and we compute the different estimators, their mean square errors (values in brakets), **and the estimated average widths (AW)** of the unknown parameters by using R statistical software. The results are given by sample size in Tables 1a, 1b, 1c, 1d, 1e and 1f.

$n=10$	$\alpha=1.5$	AW_α	$\beta=3$	AW_β	$\lambda=2$	AW_λ
MLE	1.5798(0.0259)	0.0507	2.9416(0.0181)	0.0354	2.0523(0.0193)	0.0378
WLS	1.4626(0.0194)	0.0380	3.0385(0.0149)	0.0292	2.0493(0.0171)	0.0335
PEs	1.5346(0.0178)	0.0349	2.9548(0.0140)	0.0274	2.0453(0.0164)	0.0321
MPS	1.4637(0.0209)	0.0409	3.0392(0.0167)	0.0327	2.0507(0.0190)	0.0372
$CMEs$	1.5382(0.0197)	0.0386	2.9524(0.0160)	0.0313	2.0473(0.0182)	0.0356
$OLSE$	1.5339(0.0171)	0.0335	2.9529(0.0128)	0.0251	2.0462(0.0153)	0.0299
$CSMM$	1.5496(0.0218)	0.0427	3.0409(0.0187)	0.0366	2.0451(0.0204)	0.0399
ADE	1.5399(0.0236)	0.0462	2.9648(0.0201)	0.0394	2.0483(0.0216)	0.0423
ADE_{L-R}	1.5454(0.0254)	0.0497	2.9631(0.0211)	0.0414	2.0495(0.0232)	0.0454
ADE_{L-T}	1.5477(0.0267)	0.0523	2.9615(0.0226)	0.0443	2.0514(0.0263)	0.0515

$Table$ 1a. Estimator values of α, β and λ with different estimation methods

$n=20$	$\alpha=1.5$	$\mathbf{AW}_\alpha$	$\beta=3$	$\mathbf{AW}_\beta$	$\lambda=2$	$\mathbf{AW}_\lambda$
MLE	1.5546(0.0216)	0.0423	2.9574(0.0156)	0.0305	2.0646(0.0174)	0.0341
WLS	1.4673(0.0179)	0.0345	3.0297(0.0124)	0.0243	2.0413(0.0152)	0.0297
PEs	1.5316(0.0163)	0.0319	2.9636(0.0119)	0.0233	2.0377(0.0145)	0.0284
MPS	1.4679(0.0195)	0.0382	3.0310(0.0142)	0.0278	2.0428(0.0171)	0.0335
$CMEs$	1.5347(0.0182)	0.0356	2.9603(0.0136)	0.0266	2.0397(0.0163)	0.0319
$OLSE$	1.5328(0.0156)	0.0305	2.9612(0.0103)	0.0201	2.0386(0.0134)	0.0262
$CSMM$	1.4613(0.0203)	0.0397	3.0327(0.0163)	0.0319	2.0382(0.0185)	0.0362
ADE	1.5374(0.0223)	0.0437	2.9732(0.0176)	0.0344	2.0407(0.0197)	0.0386
ADE_{L-R}	1.5423(0.0241)	0.0472	2.9713(0.0183)	0.0358	2.0413(0.0213)	0.0417
ADE_{L-T}	1.5456(0.0252)	0.0493	2.9697(0.0192)	0.0376	2.0436(0.0244)	0.0478

$Table$ 1b. Estimator values of α, β and λ with different estimation methods

$n=30$	$\alpha=1.5$	AW_α	$\beta=3$	AW_β	$\lambda=2$	AW_λ
MLE	1.5218(0.0118)	0.0231	2.9808(0.0094)	0.0184	2.0229(0.0112)	0.0219
WLS	1.4696(0.0141)	0.0276	3.0256(0.0123)	0.0241	2.0279(0.0148)	0.0290
PEs	1.5286(0.0148)	0.0290	2.9793(0.0106)	0.0207	2.0242(0.0129)	0.0252
MPS	1.4691(0.0168)	0.0329	3.0269(0.0130)	0.0254	2.0317(0.0153)	0.0299
$CMEs$	1.5288(0.0158)	0.0309	2.9776(0.0116)	0.0227	2.0272(0.0139)	0.0272
$OLSE$	1.5271(0.0137)	0.0268	2.9783(0.0109)	0.0213	2.0265(0.0134)	0.0262
$CSMM$	1.4637(0.0171)	0.0335	3.0318(0.0152)	0.0297	2.0361(0.0175)	0.0343
ADE	1.5352(0.0183)	0.0358	2.9751(0.0157)	0.0307	2.0368(0.0184)	0.0360
ADE_{L-R}	1.5363(0.0188)	0.0368	2.9743(0.0168)	0.0329	2.0379(0.0192)	0.0376
ADE_{L-T}	1.5417(0.0197)	0.0386	2.9709(0.0173)	0.0339	2.0435(0.0198)	0.0388

$Table$ 1c. Estimator values of α, β and λ with different estimation methods

$n = 80$	$\alpha = 1.5$	AW_α	$\beta = 3$	AW_β	$\lambda = 2$	AW_λ
MLE	1.5133(0.0089)	0.0174	2.9873(0.0069)	0.0135	2.0141(0.0077)	0.0151
WLS	1.4784(0.0112)	0.0219	3.0194(0.0098)	0.0192	2.0193(0.0112)	0.0219
PEs	1.5195(0.0101)	0.0197	2.9864(0.0081)	0.0158	2.0152(0.0093)	0.0182
MPS	1.4779(0.0125)	0.0245	3.0207(0.0105)	0.0205	2.0219(0.0114)	0.0223
$CMEs$	1.5203(0.0109)	0.0213	2.9837(0.0091)	0.0178	2.0182(0.0106)	0.0207
$OLSE$	1.5186(0.0104)	0.0203	2.9848(0.0084)	0.0164	2.0172(0.0095)	0.0186
$CSMM$	1.4725(0.0143)	0.0280	3.0253(0.0125)	0.0245	2.0275(0.0134)	0.0262
ADE	1.5267(0.0151)	0.0295	2.9816(0.0133)	0.0260	2.0282(0.0143)	0.0280
ADE_{L-R}	1.5278(0.0159)	0.0311	2.9805(0.0138)	0.0270	2.0293(0.0159)	0.0311
ADE_{L-T}	1.5332(0.0167)	0.0307	2.9773(0.0148)	0.0290	2.0349(0.0168)	0.0329

$Table$ 1d. Estimator values of α, β and λ with different estimation methods

$n = 200$	$\alpha = 1.5$	AW_α	$\beta = 3$	AW_β	$\lambda = 2$	AW_λ
MLE	1.5084(0.0062)	0.0121	2.9922(0.0047)	0.0092	2.0089(0.0051)	0.0099
WLS	1.4832(0.0085)	0.0166	3.0149(0.0074)	0.0145	2.0147(0.0080)	0.0156
PEs	1.5146(0.0073)	0.0143	2.9913(0.0056)	0.0109	2.0096(0.0064)	0.0125
MPS	1.4827(0.0098)	0.0192	3.0162(0.0081)	0.0158	2.0163(0.0085)	0.0166
$CMEs$	1.5152(0.0082)	0.0160	2.9886(0.0069)	0.0135	2.0123(0.0075)	0.0147
$OLSE$	1.5146(0.0076)	0.0148	2.9889(0.0065)	0.0127	2.0114(0.0069)	0.0135
$CSMM$	1.4773(0.0113)	0.0221	3.0208(0.0103)	0.0201	2.0217(0.0108)	0.0211
ADE	1.5218(0.0124)	0.0243	2.9865(0.0108)	0.0211	2.0226(0.0117)	0.0229
ADE_{L-R}	1.5238(0.0132)	0.0258	2.9854(0.0114)	0.0223	2.0237(0.0122)	0.0239
ADE_{L-T}	1.5283(0.0139)	0.0272	2.9819(0.0122)	0.0239	2.0293(0.0132)	0.0258

$Table$ 1e. Estimator values of α, β and λ with different estimation methods

$n = 350$	$\alpha = 1.5$	AW_α	$\beta = 3$	AW_β	$\lambda = 2$	AW_λ
MLE	1.5052(0.0043)	0.0084	2.9965(0.0030)	0.0058	2.0054(0.0032)	0.0062
WLS	1.4866(0.0069)	0.0135	3.0106(0.0057)	0.0111	2.0113(0.0061)	0.0119
PEs	1.5084(0.0054)	0.0105	2.9954(0.0039)	0.0076	2.0062(0.0045)	0.0088
MPS	1.4841(0.0082)	0.0160	3.0117(0.0063)	0.0123	2.0126(0.0065)	0.0127
$CMEs$	1.5106(0.0065)	0.0127	2.9928(0.0052)	0.0101	2.0086(0.0054)	0.0105
$OLSE$	1.5095(0.0057)	0.0111	2.9931(0.0050)	0.0098	2.0077(0.0050)	0.0098
$CSMM$	1.4809(0.0097)	0.0190	3.0159(0.0086)	0.0168	2.0183(0.0089)	0.0174
ADE	1.5172(0.0108)	0.0211	2.9908(0.0091)	0.0178	2.0192(0.0098)	0.0192
ADE_{L-R}	1.5190(0.0115)	0.0225	2.9897(0.0097)	0.0190	2.0203(0.0104)	0.0203
ADE_{L-T}	1.5246(0.0126)	0.0246	2.9862(0.0105)	0.0205	2.0259(0.0111)	0.0217

$Table$ 1f. Estimator values of α, β and λ with different estimation methods

Considering the obtained values of the different estimators, we can draw the following conclusions:

1. For little sample sizes ($n \prec 30$), the percentile PEs , the maximum product of spacing MPS, the least squares $OLSE$ and the CMS estimation methods for α; and the MPS and the weighted least squares WLS and the Anderson-Darling estimation (ADE) estimation methods are the most efficient for estimating β, but for λ the results are the same for all the methods used.

2. When $n \geq 30$, and as It was expected the maximum likelihood estimation method gives the best results. It can also be noted that the PEs and the CMS methods give very good results for both large and small samples.

7.2. The Y^2 statistic

To verify the null hypothesis H_0 for the $TL-EE$ distribution, samples of respective sizes $n = 30, 80, 200, 350, 500$ and 1000, are generated $N = 10,000$ times from this model. The Y^2 values of the proposed NRR test criterion are computed for all samples and the different empirical levels of rejection of the null hypothesis H_0, for $Y^2 > \chi^2_\varepsilon(r-1)$ are compared to their levels of theoretical significance ε ($\varepsilon = 0,01$, $0,05$, $0,10$). The results are given in Table 2.

$N = 10.000$	$n_1 = 30$	$n_2 = 80$	$n_3 = 200$	$n_4 = 350$	$n_5 = 500$	$n_6 = 1000$
$\varepsilon = 0.01$	0.0058	0.0063	0.0078	0.0086	0.0092	0.0099
$\varepsilon = 0.05$	0.0456	0.0466	0.0472	0.0482	0.0490	0.0498
$\varepsilon = 0.10$	0.0949	0.0954	0.0968	0.0977	0.0982	0.0995

$Table 2$.Comparison of theoretical risks and corresponding empirical risks
$\varepsilon = 0.01; 0.05; 0.10$

The simulation shows, taking into account the simulation errors, that the levels simulated for the statistic Y^2 and those corresponding to the theoretical levels of the chi-square distribution with $(r-1)$ degrees of freedom, are close to each other. Consequently, we can say that the statistic test proposed in this work, can suitably fit data to the $TL-EE$ model.

8. Applications

The analysis of three real data sets is proposed to show the usefullness of the proposed distribution in modeling different phenomenon and the performances of the methods used to determine the unknown parameters. In addition to the classical methods of model selection, we calculate Y^2 statistic to prove that the proposed model fits data better than the commun used alternative distributions such us Beta extended exponential $(Be-EE)$, Weibll extented exponential $(W-EE)$, extented exponential (EE), Lomax and Burr XII distributions.

Covid-19 data

As the covid-19 pandemic occupies all the news, we propose to study the contamination phenomenon in one of the countries most affected, India. Data (10^2) consist in number of contaminations in this countrie relating to the period from 1 may to 14 june 2020 and taken from the siteweb (Coronavirus Update (Live): 7,114,524 Cases and 406,552 Deaths from COVID-19 Virus Pandemic - Worldometer). For calculation purpose, we consider data $(\times 10^{-2})$:
2394, 2442, 2806, 3932, 2963, 3587, 3364, 3344, 3113, 4353, 3607, 3524, 3764, 3942, 3787, 4864, 5050, 4630, 6147, 5553, 6198, 6568, 6629, 7113, 6414, 5843, 7293, 7300, 8105, 8336, 8782, 7761, 8821, 9633, 9889, 9471, 10438, 10864, 8442, 8852, 12375, 11128, 11320, 12023, 11157.
Table 3 represents the values of the parameter estimators for the hypothezised distribution $TL-EE$ obtained by the different methods and the corresponding p-values. Since $n = 45 \succ 30$ and as It was established by the simulation study the MLE and PEs methods give the best results.

$Method$	α	λ	β	$-NLL$	KS	$p-value$
MLE	6.18025	0.01583	1.11818	214.0236	0.12303	0.46673
WLS	6.16346	0.01267	1.13076	214.9967	0.12764	0.45983
PEs	6.23146	0.01476	1.12463	214.3497	0.12412	0.46467
MPS	6.25317	0.01736	1.14138	215.1746	0.12889	0.45768
$CMEs$	6.17067	0.01376	1.11013	214.8787	0.12633	0.46156
$CSMM$	6.13469	0.01794	1.11356	215.2334	0.12983	0.45618
LSE	6.24767	0.01602	1.10346	214.6184	0.12567	0.46213
ADE	6.27413	0.01316	1.14796	215.3412	0.13146	0.45562
ADE_{L-R}	6.15934	0.01896	1.12103	215.5613	0.13213	0.45434
ADE_{L-T}	6.29794	0.01403	1.09946	215.7364	0.13264	0.45219

$Table 3$.Different parameter estimators for covid-19 data

As the maximum likelihood parameter estimators for the competing distributions are needed to compute their corresponding NRR test statistic Y^2, the MLE values are calculated and given in Table 4.

$Distributions$	$Estimates$
$TLEE$	$\alpha = 6.1802, \lambda = 0.0158, \beta = 1.1181$
$Bet - EE$	$\alpha = 4.6433, \beta = 1.4169, \lambda = 0.0162, s = 1.3723$
$W - EE$	$\alpha = 6.1434, \beta = 0.5619, \lambda = 0.2661, s = 0.1474$
EE	$\alpha = 0.01745, \beta = 0.89775$
$Lomax$	$\alpha = 0.01596, \beta = 1.05082$
$BurrXII$	$\alpha = 0.09777, \beta = 2.50348$

$Table$ 4. ML parameter estimates for the alternative distributions

To distinguish between the proposed model and its alternatives, we use classical criteria for model selection and the NRR statistic with the corresponding values summarized in Table 5.

$Distributions$	Y^2	W	A	$K - S$	$p - value$
$TLEE$	7.8296	0.08895	0.62356	0.12303	0.4667372
$Bet - EE$	7.9345	0.10527	0.69173	0.12303	0.4667372
$W - EE$	8.2637	0.09445	0.64255	0.12777	0.4195012
EE	8.9134	0.10532	0.68837	0.30802	0.00027397
$Lomax$	9.8462	0.12711	0.79317	0.31798	0.00015165
$BurrXII$	10.4936	0.13976	0.84959	0.54036	7.91589$e - 13$

$Distributions$	$-NLL$	AIC	$CAIC$	BIC	$HQIC$
$TLEE$	214.0236	434.0472	434.6326	439.4672	436.0677
$Bet - EE$	213.8315	435.6629	436.6629	442.8896	438.3569
$W - EE$	213.4943	434.9886	435.9886	442.2152	437.6826
EE	234.8657	473.7444	474.0302	477.3578	475.0914
$Lomax$	248.1394	500.2863	500.572	503.9896	501.6333
$BurrXII$	292.2254	588.4508	588.7366	592.0642	589.7979

$Table5$. values of criteria statistics for model selection

The p-values indicates that the null hypothesis H_0 cannot be rejected for $TL - EE$, $Bet - EE$ and $W - EE$ distributions, however the newly developed $TL - EE$ displays a very good potential in Table 5 as it has the lowest values for th $Y^2, W, A, AIC, CAIC, BIC, HQIC$,and $K - S$ statistics. We conclude that the phenomenon of the contamination by covid-19 virus can be described by the $TL - EE$ distribution in a satisfactory manner and better than all the alternatives. The pdf graphs given in Figure 3 and 4 show that the histogram of these data are very close to the $TL - EE$ pdf curve.

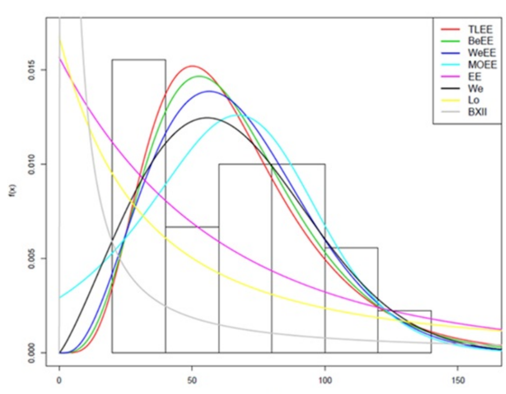

Fig3.Histogram plot of the dataset with the compared distribution

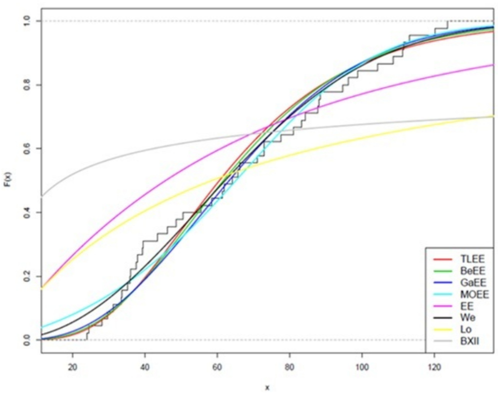

Fig4. Empirical cdf of the datasetwith the compared distributions

Waiting times

This application relates to the waiting times (in minutes) before service of 100 bank customers:
0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5.

Since n is large, we calculate only the ML estimators of the unknown parameters and the values of the different statistics for model selection in order to choose whether the proposed model suited to these data better than the alternative distributions.

$Distributions$	Y^2	NLL	AIC	$CAIC$	BIC	$HQIC$	KS	$p-value$
$TL-EE$	7.8235	317.0435	640.0869	640.3369	647.9024	643.25	0.0369	0.9991
$BeEE$	7.9462	317.0439	642.0877	642.5088	652.5084	646.3051	0.0366	0.9993
$GaEE$	8.2264	322.1976	650.3952	650.6452	658.2107	653.5583	0.1138	0.1493
$MOEE$	8.1292	319.1373	644.2745	644.5245	652.09	647.4376	0.0471	0.9794
EE	8.3178	327.2185	658.4584	658.5821	663.6687	660.5671	0.1626	0.0100

$Table$ 6.Model selection criteria values for competing distribution for waiting times in service bank

Considering the obtained results in Table 6, this new distribution models these data as well as possible and this is confirmed by the pdf plots of Figures 5 and 6.

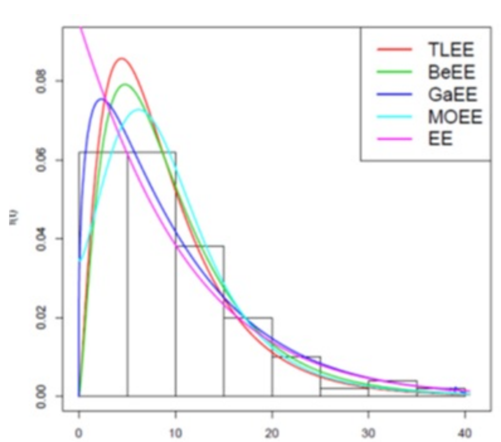

Fig5.Histogram plot of the dataset with the compared distribution

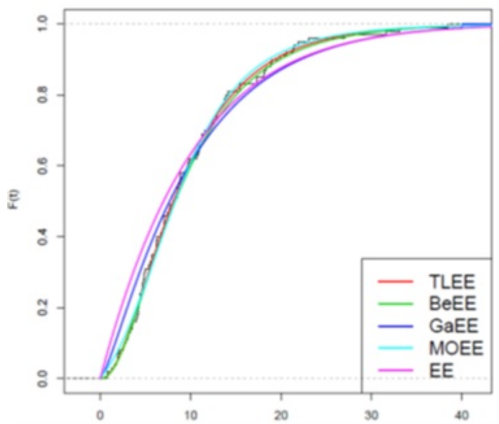

Fig6.Empirical cdf of the dataset with the compared distributions

Remission times

The third example concerned the remission times (in months) of a random sample of 128 bladder cancer patients reported in Lee and Wang (2003):

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 12.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

To show that this model can also describe this type of data, we use all classical criteria for model selection and the modified chi-square NRR. All the corresponding values are computed and given in Table 7.

$Distributions$	Y^2	NLL	AIC	$CAIC$	BIC	$HQIC$	KS	$p-value$
$TL-EE$	8.9462	410.397	826.795	826.989	835.351	830.272	0.0429	0.9721
$BeEE$	9.2341	410.195	828.390	828.715	839.798	833.025	0.0418	0.9781
$GaEE$	9.6945	411.926	829.852	830.046	838.408	833.329	0.0631	0.6875
$MOEE$	9.4569	410.825	827.651	827.845	836.208	831.128	0.0511	0.8907
EE	9.3351	414.228	832.456	832.552	838.161	834.774	0.0930	0.2179

$Table7.values\ of\ different$ criteria for model selection for remission times of bladder cancer patients

From Table 7, we see that the smallest values of the different model criteria selection are obtained for the $TL-EE$ distribution which confirm that the proposed model describes these data better than all the alternatives.

$Distributions$	$Estimates$
$TL-EE$	$\lambda = 1.6530, \alpha = 0.1864, \beta = 0.5523$
$BeEE$	$\alpha = 1.6130, \beta = 8.6106, \lambda = 0.0911, s = 0.2997$
$GaEE$	$\beta = 0.2624, \gamma = 0.5665, c = 0.9020$
$MOEE$	$\alpha = 29.7347, \beta = 7.9998, \theta = 0.3745$
EE	$\alpha = 0.1240, \lambda = 0.9116$

$Table$ 8. maximum likelihood estimates

$Method$	α	λ	β	NLL	KS	$p-value$
MLE	0.1864	1.6530	0.5523	410.3979	0.00429	0.9721
WLS	0.1963	1.3641	0.6166	410.4966	0.00492	0.9786
PEs	0.1533	1.6177	0.5414	410.4126	0.00441	0.9737
MPS	0.1724	1.6022	0.5289	410.4393	0.00456	0.9742
$CMEs$	0.1632	1.7266	0.5677	410.4675	0.00472	0.9759
$CSMM$	0.1912	1.7521	0.5723	410.4914	0.00489	0.9773
$MLSE$	0.1384	1.4161	0.6272	410.5133	0.00499	0.9792
$OLSE$	0.1589	1.6933	0.5518	410.4515	0.0466	0.9754
ADE	0.1833	1.7689	0.6319	410.5243	0.0501	0.9815
ADE_{L-R}	0.1767	1.7702	0.6334	410.5413	0.0523	0.9836
ADE_{L-T}	0.1798	1.7836	0.6462	410.5533	0.0554	0.9898

$Table$ 9.Different parameter estimators for remission times bladder cancer patients

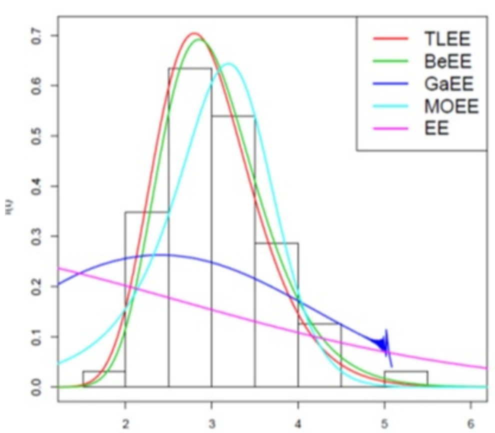

Fig7. Histogram plot of the dataset with the compared distribution

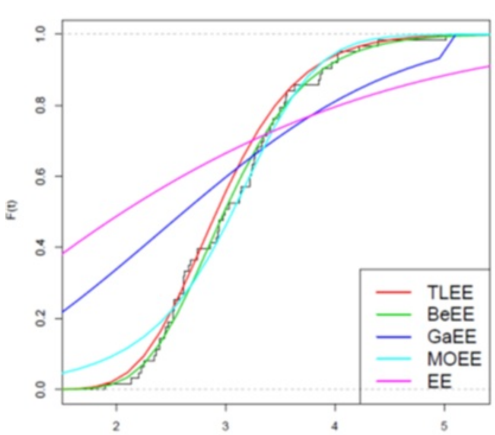

Fig8.Empirical cdf of the dataset with the compared distributions

As seen in $Fig8$ and $Fig9$, remission times of bladder cancer patients are well fitted by the $TL-EE$ model compared to the other distributions.

Conclusion:

In this work, we have proposed a new distribution named the Topp-Leone extended exponential $(TL-EE)$ distribution. Different methods of estimation were used to calculate the unkown parameters. The simulation study showed that the PSE and MPSE estimator methods outperform all the other estimator methods in the case of small samples. However, the MLE and PSE methods are the best ones for large samples. The usefullness of this model is demonstrated by mean of three applications from different fields. The data on covid-19 virus contaminations have been relatively well described by this model.

References

1. Al-Shomrani, A., Arif, O., Shawky, A., Hanif, S. and Shahbaz, M.Q. (2016). Topp–Leone family of distributions: some properties and application. Pakistan Journal of Statistics and Operation Research, 12, 443–451.
2. Abdel-Ghaly, A. A. , Al-Dayian, G.R. and Al-Kashkari, F.H. (1997). The use of Burr type XII distribution on software reliability growth modelling. Microelectronics Reliability, 37, 305-313.
3. Genc, A.I. (2012). Moments of order statistics of Topp–Leone distribution. Statistical Papers, 53, 117–131.
4. Aidi, K. and Seddik-Ameur, N. (2016). Chi-square tests for generalized exponential AFT distributions with censored data. Electronic journal of Applied Statistical Analysis, 09(2), 371-384.
5. Cheng, R.C.H. and Amin, N.A.K. (1983). Estimating Parameters in Continuous Univariate Distribution with a Shifted Origin. Journal of the Royal Statistical Society, 45, 394-403.
6. Chouia, S. and Seddik-Ameur, N. (2017). Un test du chi carré modifié pour le modèle de Bertholon avec des données censurées, les communications dans la simulation statistique et le calcul, 46(1), 593-602.
7. Boos, D.D. (1982). Minimum Anderson-Darling estimation. Communications in Statistics-Theory and Methods, 11, 2747–2774.
8. Fatma Gül AKGÜL (2018). Comparison of the Estimation Methods for the Parameters of Exponentiated Reduced Kies Distribution, Journal of Natural and Applied Sciences, 22(3), 1209-1216.
9. Ghitany, M.E., Kotz, S. and Xie, M. (2005). On some reliability measures and their stochastic orderings for the Topp–Leone distribution. J. Appl. Stat., 32, 715–722.
10. Giovani Carrara Rodrigues, Francisco Louzada and Pedro Luiz Ramos (2016). Poisson–exponential distribution: different methods of estimation. Journal of applied statistics. DOI: 10.1080/02664763.2016.1268571
11. Al-Mofleh, H. and Afify, A. Z. (2019). A generalization of Ramos-Louzada distribution: Properties and estimation, Journal Not Specified, pages 1 – 25.
12. Jamal, Q.A. and Arshad, M. (2019). Statistical inference for Topp–Leone-generated family of distributions based on records. Journal of Statistical Theory and Applications, 18, 65–78.
13. Swain, J. J., Venkatraman, S. and Wilson, J. R. (1988). Least-squares estimation of distribution functions in johnson's translation system. Journal of Statistical Computation and Simulation. 29(4), 271–297. DOI: 10.1080/00949658808811068
14. Aidi, K., Dey, S., Kumar, D. and Seddik-Ameur, N. (2021). Different cal methods of estimation and Chi-Squared classical goodness-of-fit test for unit generalized inverse Weibull distribution. Austrian Journal of Statistics, 50, 77–100.
15. Lee, E.T. and Wang, J.W. (2003). Statistical Methods For Survival Data Analysis. Third Edition, Wiley, New York. https://doi.org/10.1002/0471458546
16. Arshad, M. and Jamal, Q.A.(2019). Statistical inference for Topp–Leone-generated family of distributions based on records, J. Mod. Appl. Stat. Methods.18(1), 65 - 78.
17. Zhou, M., Yang, D., Wang, Y. and Nadarajah, S. (2006). Some J-shaped distributions: Sums, products and ratios. Reliability and Maintainability Symposium, RAMS'06 Annual, IEEE, 175–181.
18. Nadarajah, S. and Kotz, S. (2003). Moments of some J-shaped distributions. Journal od Applied Statistics. 30, 311–317.
19. Seddik-Ameur, N. and Treidi, W. (2018). On testing the fit of accelerated failure time and proportional hazard Weibull extension models. Journal of Statistical Theory and Practice. DOI: 10.1080/15598608.2017.1397568.
20. Nikulin, M. S. (1975). Chi-square test for normality. In: Proceedings of the International Vilnius Conference on Probability Theory and Mathematical Statistics. 2, 199–122
21. Bantan, R. A. R., Jamal, F., Chesneau, C. and Elgarhy,M. (2020). Type II Power Topp-Leone Generated Family of Distributions with Statistical Inference and Applications. Symmetry, 12(1), 75, https://doi.org/10.3390/sym12010075.

22. Rao, K. C. and Robson, D. S. (1974). A Chi-square statistic for goodness-of-fit tests within the exponential family. Communication in Statistics, 3, 1139-1153.
23. Cheng, R. and Stephens, M. (1989). A goodness-of-fit test using Moran's statistic with estimated parameters. Biometrika, 76(2), 385–392.
24. Rezaei, S., Sadr, B.B., Alizadeh, M. and Nadarajah, S. (2016). Topp-Leone generated family of distributions: Properties and applications. Commun. Stat. Theory Methods, 46, 2893–2909.
25. Dey, S., Saha, M., Yadav, A. S. and Ali, S. (2021). Confidence intervals of the index Cpk for normally distributed quality characteristics using Classical and Bayesian methods of estimation. Brazilian Journal of Probability and Statistics, 35(1), 138–157.
26. Dey, S., Kumar, D., Ramos, P.L. and Louzada, F. (2017). Exponentiated Chen distribution: Properties and estimation. Communications in Statistics - Simulation and Computation, 46(10), 8118–8139.
27. Topp, C.W. and Leone, F.C. (1955). A family of J-Shaped frequency functions. J. Am. Stat. Assoc. 50, 209–219.
28. Voinov, V., Nikulin, M. and Balakrishnan, N. (2013). Chi-Squared Goodness of Fit Tests with Applications, Academic Press, Elsevier.
29. ZeinEldin, R. A., Ahsan ul Haq, M., Sharqa, H. and Mahmoud, E. (2020). Alpha Power Transformed Inverse Lomax Distribution with Different Methods of Estimation and Applications, Complexity; Volume 2020. DOI:10.1155/2020/1860813.