Marsall Olkin extended Topp–Leone distribution

AMEL A. TALAAT
Faculty of Commerce (Girls Campus)
Al-Azhar University, Cairo, Egypt

A b s t r a c t

Marshall and Olkin (1997) proposed a new method to establish more flexible new families of distributions by adding a parameter to a distribution. In this paper, Marshall-Olkin extended Topp-Leone (MOETL) distribution is introduced. Some statistical properties of the MOETL are studied. The performance of the maximum likelihood estimates is investigated by means of a Marckov chain Monte Carlo (MCMC) simulation study. Model parameters are obtained using maximum likelihood method of estimation. Life time data analysis is performed to demonstrate the models applicability and flexibility. Akaike and Bayesian information criteria (AIC, BIC) illustrate the extended distribution provides better fit compared to the original distribution.

Keywords : Marshall Olkin Extended, Topp–Leone Distribution, Maximum Likelihood Estimates. MCMC .

. 1. Introduction

Lifetime data plays an important role in a wide range of applications such as medicine, engineering, biological science, management, and public health. Statistical distributions are used to model the life of an item in order to study its important properties. Proper distribution may provide useful information that result in sound conclusions and decisions. When there is a need for more flexible distributions, many researchers are about to use the new one with more generalization. An excellent review of Lee *et al.* (2013) has provided thorough knowledge of several methods for generating families of continuous univariate distributions. According to their work, there are some general methods introduced prior to 1980, which were developed by the strategies based on differential equation, transformation and quantile function. In addition, they also put emphasis on the movement of those methods proposed since 1980s, which changed the momentum by adding extra parameters or combining existing distributions.

Topp-Leone(TL) distribution is a continuous unimodal distribution with bounded support. It is a two parametric family continuous distribution proposed by TL (1955). Such a distribution is useful for modelling lifetime phenomena, different aspect of this class of distributions have been studied
e.g. by Nadarajah and Kotz (2003). We say that a random variable X with range of values $(0, \lambda)$ has a TL distribution, and write $X \sim \mathrm{TL}(\alpha, \lambda)$, if the cumulative distribution function cdf is

$$F(\mathrm{x}) = \left(\frac{\mathrm{x}}{\lambda}\right)^{\theta} (2 - \frac{\mathrm{x}}{\lambda})^{\theta}, \quad 0 < \mathrm{x} < \lambda, \theta, \lambda > 0, \tag{1}$$

and the density function pdf is

$$f(x) = \frac{2\theta}{\lambda}\left(\frac{x}{\lambda}\right)^{\theta-1} (2 - \frac{x}{\lambda})^{\theta-1}, \quad 0 < x < \lambda, \theta, \lambda > 0 \tag{2}$$

Where α and λ are the scale and shape parameters respectively.
The TL distribution is known as the J-shaped distribution. This is due to the fact that $f(x) > 0$, $\acute{f}(x) < 0$, and $\acute{\acute{f}}(x) > 0$ for all $0 < x < \lambda$, where $\acute{f}$ and $\acute{\acute{f}}$ are the first and second derivatives of f respectively.

The univariate continuous TL distribution with bounded support was originally proposed by
TL (1955) and applied it as a model for some failure data. The probability density function pdf and cumulative distribution function cdf when $\lambda=0$ are respectively given, for $0 < \mathrm{x} < 1$ and $\theta > 0$, by

$$f(x) = \theta(2 - 2x)(2x - x^2)^{\theta-1}, \quad 0 < \mathrm{x} < 1, \ \theta > 0 \tag{3}$$

And

$$F(x) = (2x - x^2)^{\theta}, \quad 0 < \mathrm{x} < 1, \ \theta > 0 \tag{4}$$

In recent years, the Topp-Leone distribution has received a huge attention in the literature. For instance, Nadarajah and Kotz (2003) derived the structural properties of this distribution including explicit expressions for the moments, hazard rate function and characteristic function. Kotz and Van Dorp (2004) proposed a generalized version of the Topp-Leone distribution for modeling some financial data and studied its properties. Ghitany et al. (2005) discussed some reliability measures of the Topp-Leone distribution and their stochastic orderings. Kotz and Nadarajah (2006) gave a bivariate generalization of this distribution. Ghitany (2007) derived the asymptotic distribution of order statistics of this model. Vicari et al. (2008) introduced a two-sided generalized version of the distribution and discussed some of its properties. The moments of order statistics from this distribution were discussed by Gen¸c (2012) and MirMostafaee (2014). Gen.c (2013) considered the estimation of the stress-strength parameter for this distribution. Admissible minimax estimates for the shape of this distribution were derived by Bayoud (2016). Bayesian and non-Bayesian estimation of Topp-Leone distribution based lower record values were obtained by Mir Mostafaee et al. (2016).

This paper will be organized as follows. In section 2, Marshall-Olkin extended top-leone distribution is introduced, some properties are studied theory and along with a graphical description.in section 3.in section 4, the maximum likelihood method is described. Simulation study and application to real data set performed in section 5 and 6.finally, concluding remarks are given in section 7

2. Marshall-Olkin Extended Topp-Leone (MOETL) Distribution:

Marshall and Olkin(1997) introduced a new family of distributions in an attempt to add a parameter to a family of distributions. Then, the Marshall-Olkin extended distribution has survival function
$\bar{F}(x) = 1 - F(x)$.Let $\bar{F}\,(x) = p(\times > x)$ is the survival function of a random variable X and $\alpha > 0$ be a parameter. Then

$$\bar{G} = \frac{\alpha \bar{F}\,(x)}{1-(1-\alpha)\bar{F}\,(x)}\,, -\infty < x < \infty, \alpha > 0 \qquad (5)$$

is a proper survival function. $\bar{G}\,(x, \alpha)$ is called Marshall-Olkin family of distributions.

The probability density function p.d.f .corresponding to (5) is given by

$$g(x, \alpha) = \frac{\alpha f(x)}{[1-(1-\alpha)\bar{F}\,(x)]^2} \quad , -\infty < x < \infty, \alpha > 0 \qquad (6)$$

where $f(x)$ is the p.d.f. corresponding to F(x). The hazard (failure) rate function is given by

$$h(x, \alpha) = \frac{r(x)}{1-(1-\alpha)\bar{F}\,(x)}, \text{where} \quad r(x) = \frac{f(x)}{\bar{F}\,(x)}. \qquad (7)$$

Similar models were considered, for example by Alice and Jose in (2003), (2005).

Now Let X follows TL (θ, ·) distribution, where $\theta > 0$. Then $\bar{F}\,(x)$ =$1 - (2x - x^2)^2$. Substituting in (5) we get a new distribution denoted by MOETL (θ , α) with survival function

$$\bar{G}(x,\alpha,\theta) = \frac{\alpha[1-(2-x^2)^{\theta}]}{1-(1-\propto)+(1-\propto)(2x-x^2)^{\theta}} \text{ , } 0 < x < 1, \alpha, \theta > 0 \quad (8)$$

The corresponding pdf and cdf are obtained respectively as

$$g(x,\alpha,\theta) = \frac{\alpha\theta[(2-2x)(2x-x^2)^{\theta-1}]}{[1-(1-\propto)(1-(2x-x^2)^{\theta})]^2}, \qquad 0 < x < 1, \alpha, \theta > 0 \quad (9)$$

And cdf is

$$G(x,\alpha,\theta) = \frac{(2-x^2)^{\theta}}{1-(1-\propto)+(1-\propto)(2x-x^2)^{\theta}}, \qquad 0 < x < 1, \alpha > 0 \quad (10)$$

The graphs of pdf. and distribution function cdf are drawn in Figures 1 and 3. The shape of the pdf $g(x,\alpha,\theta)$ Figure 1 depends on parameter α . Namely, if $\theta \epsilon$ (0,1), then the pdf is decreasing function on (0, θ) with g(0, α, θ) =θ and g(θ, α, θ) =$\alpha\theta[(2-2\theta)(2\theta-\theta^2)^{\theta-1}/[\alpha+(1-\alpha)(2\theta-\theta^2)^{\theta})]^2]$ Otherwise, if $\theta > 1$, then the p.d.f. is an decreasing function on (0, θ) with g(0, α, θ) =θ and g(θ, α, θ) =$\alpha\theta[(2-2\theta)(2\theta-\theta^2)^{\theta-1}/[\alpha+(1-\alpha)(2\theta-\theta^2)^{\theta})]^2$

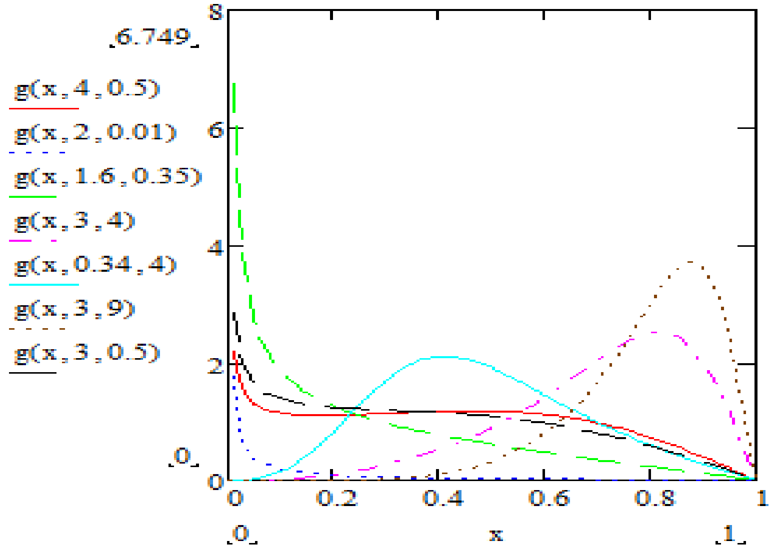

Figure 1: Graph of g(x) for θ = 4, 0.5, 0.01 and 5 for various values of α.

The hazard rate function of a random variable X with MOETL (α, θ) distribution is

$$h(x,\alpha,\theta) = \frac{\theta(2-2x)(2x-x^2)^{\theta-1}}{[1-(2x-x^2)^{\theta}][1-(1-\alpha)(1-(2x-x^2)^{\theta}]} \quad (11)$$

For $\theta \leq 0.7$ the hazard rate is initially constant and there exists an interval where it changes.
For $\theta > 0.7$ the hazard function is increasing. The graph of hazard rate function is drawn in Figure 2

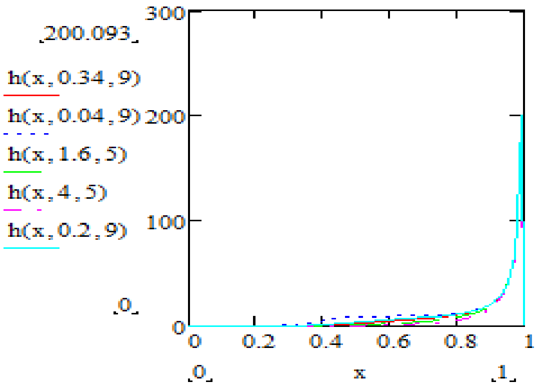

Figure 2: Graph of h(x) for θ = 9 and 5 for various values of α

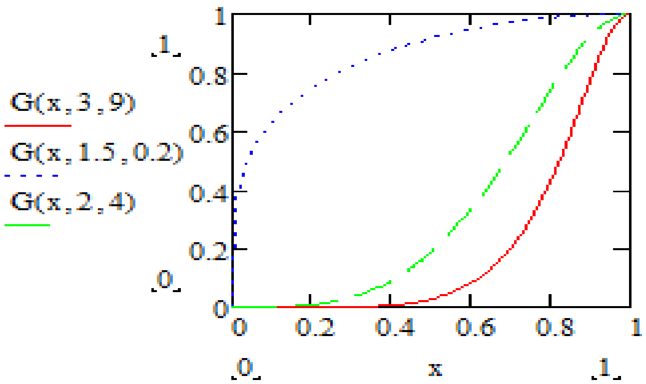

Figure 3: Graph of G(x) for θ = 4,9 and for various values of α

3. **Some properties of the Marshall-Olkin Extended Top-Leone distribution:**

In this section, some fundamental properties of MOETL distribution including mean, moments, variance, quantile and mode are presented

3.**1. Moments**

We consider a random variable X with MOETL (α, θ). The r-th non central moments of the MOETL distribution and variance are given, respectively, by

$$\grave{\mu}_r = \int_0^1 x^r \; \frac{\alpha\theta(2-2x)\left(2x-x^2\right)^{\theta-1}}{1-(1-\alpha)(2x-x^2)^{\theta}}\, dx \qquad (12)$$

Then

$$\grave{\mu}_1 = \int_0^1 x \; \frac{\alpha\theta(2-2x)\left(2x-x^2\right)^{\theta-1}}{1-(1-\alpha)(2x-x^2)^{\theta}} dx \qquad (1٣)$$

And

$$\grave{\mu}_2 = \int_0^1 x^2 \; \frac{\alpha\theta(2-2x)\left(2x-x^2\right)^{\theta-1}}{1-(1-\alpha)(2x-x^2)^{\theta}} \qquad (14)$$

Hence

$$\text{Var(x)}= \int_0^1 x^2 \; \frac{\alpha\theta(2-2x)\left(2x-x^2\right)^{\theta-1}}{1-(1-\alpha)(2x-x^2)^{\theta}} dx - \left(\int_0^1 \boldsymbol{x} \; \frac{\alpha\theta(2-2x)\left(2x-x^2\right)^{\theta-1}}{1-(1-\alpha)(2x-x^2)^{\theta}} \; dx\right)^2 \qquad (15)$$

The value of the function in (12),(13) can be obtained by a numerical calculation

moments / parameters	μ_1	μ_2	μ_3	μ_4
α= 2 θ=4	**0.773**	**0.623**	**0.516**	**0.436**
α= 0.3 θ=5	**0.352**	**0.194**	**0.124**	**0.086**
α= 3 θ=0.5	**0.463**	**0.254**	**0.157**	**0.106**
α= 0.7 θ=0.9	**0.256**	**0.118**	**0.63**	0.044

Figure 4 some values of the moments for various values of α and θ

3.**2** The quantile x_q of order q for (α,θ) from MOETL is given by

If the quantile of X, say x_q follows easily by inverting $\bar{G}(x) = 1 - q$ where $G(x) = q$,$0 \leq q \leq 1$,and G as (10) then.

$$x_q = G^{-1}(q) = 1 - \sqrt{1 - \left({}^{q\alpha}/_{1-q(1-\alpha)}\right)^{1/\theta}} \;, q \in (0,1). \tag{16}$$

Where $G^{-1}(.)$ is inverse of cdf of MOETL distribution

Quantiles of interest can be obtained from (10) by substituting appropriate values for q. In particular, the median of MOETL distribution can be obtained by substituting q=0.5 in (15) which given

$$Median(x) = q(0.5) = 1 - \sqrt{1 - \left({}^{0.5\alpha}/_{1-0.5(1-\alpha)}\right)^{1/\theta}} \tag{17}$$

3.**3** The mode of the MOETL distribution can be found by solving d log $(x,,\theta)/\, d\, x$ = 0

d Log $(x,,\theta)/\, d\, x$=

$$8\alpha^2(1-\alpha)\theta^2\left[\left((2-x)^2(2x-x^2)^{2\theta-2}/\left(\alpha+(1-\alpha)(2x-x^2)^{\theta}\right)^3\right) - \left(2(\theta-1)(1-x)^2(2x-x^2)^{2\theta-2} - (2x-x^2)^{\theta-1}/\left(\alpha+(1-\alpha)(2x-x^2)^{\theta}\right)^2\right)\right]=0 \tag{18}$$

For this selected parameters values θ=6, α=0.8 the mode is 0.55 as shown in Figure 5 which indicates that the PDF is unimodal for the selected parameter values.

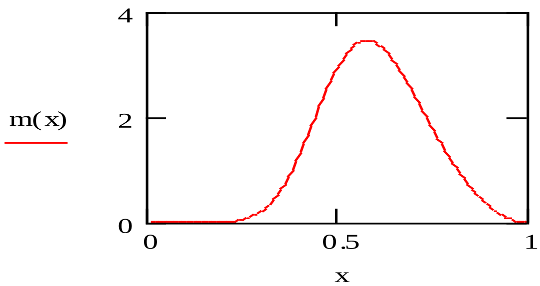

Figure 5: Graph of m(x) for θ = 6 and α=0.8

4. Maximum Likelihood Estimation of MOETL Distribution

We consider the estimation of the unknown parameters of the MOETL distribution by the method of maximum likelihood. Let $\underline{X}$= $(X_1, X_2, X_3, \dots X_n)$ be a random sample of size n from MOETL (α, θ) in (9). Let $\underline{\beta}$= $(\alpha, \theta)^T$ the vector of model parameters. Then the (ML) log-likelihood function l (β) can be written as

$l(\beta) = n\,log(\alpha) + nlog\theta + \sum_{i=1}^{n} log(2-2x_i) + (\theta-1)\sum_{i=1}^{n}\log(2x - x_i{}^2) - 2\sum_{i=0}^{1}\log[\alpha + (1 - \alpha)(2x - x^2)^\theta]$ (19)

Equation (19) can be maximized with respect to α and θ.

That is, solve the following two non- linear equation using iterative procedure.

$$\frac{\partial l}{\partial\hat{\alpha}} = \frac{n}{\hat{\alpha}} - 2\sum_{i=1}^{n}\frac{[1-(2x-x^2)^{\hat{\theta}}]}{[\hat{\alpha}+(1-\hat{\alpha})(2x-x^2)^{\hat{\theta}}]} = 0 \qquad (٢٠)$$

$$\frac{\partial l}{\partial\hat{\theta}} = \frac{n}{\hat{\theta}} + \sum_{i=1}^{n}\log(2x_i - x_i^2) - 2\sum_{i=1}^{n}\frac{[(1-\hat{\alpha})(2x-x_i{}^2)^{\theta}\log(2x-x^2)]}{[\hat{\alpha}+(1-\hat{\alpha})(2x-x^2)^{\hat{\theta}}]} = 0 \qquad (21)$$

It is clear that, the system of nonlinear equation (20-21) does not have an analytic solution in α and θ, therefore a numerical method is needed to obtain the solution.

The approximate confidence interval of the parameter $\beta(\alpha,\theta)^T$ can be obtained based on asymptotic distribution of the ML estimates of β. Then the 100(1-γ/2) % approximate confidence of the parameters $\beta(\alpha,\theta)^T$ are

$$\hat{\alpha}_{ML} \pm z\alpha_{/2}\sqrt{var(\hat{\alpha}_{ML})} \qquad (٢٢)$$

And

$$\hat{\theta}_{ML} \pm z\alpha_{/2}\sqrt{var(\hat{\theta}_{ML})} \qquad (٢٣)$$

5. **Simulation study**

In this section, some simulation results are provided to observe the behavior of the MLEs in terms of the sample size n. The importance of the MOETL distribution using real data is also illustrated.

Simulation study has been performed to investigate the performance of the ML estimates. This simulation was conducted for different sample sizes and different parameter values to study the effect of the new parameter α in two cases, the case of α < 1 and the case of α > 1.

The biases and mean squared errors of the MLE, and relative bias of MLE estimators of tow parameter have been investigated in a small simulation experiment. This experiment was conducted using the MATHCAD software .The

ML estimator was obtained by solving equation (20) and (21) using the MATHCAD. The Monte Carlo simulation involved 10000 replications, and the

MOETL random were generated by the distribution function. That is, $x = 1 - \sqrt{1 - \left(\frac{u\alpha}{1-u(1-\alpha)}\right)^{\frac{1}{\theta}}}$, where u is drawn from a U(0,1) distribution (Nadarajah and Kotz, 2003, p.317).

Table 1 reports the ML estimates of the unknown parameters along with the R bias,the mean square error (MSE), and Confidence of Interval(CI) of the parameter.

Table 1: simulation results mean, Rbias, MSE, and CI

α	θ	Sample size	parameter	Mean	R bias	MSE	CI Lower	CI upper
0.35	**1.6**	20	**α**	0.327	0.09	0.078	0.0001	0.871
			θ	3.385	0.128	1.303	1.281	5.488
		50	**α**	0.313	0.045	0.025	0.0007	0.669
			θ	3.137	0.046	0.412	1.908	4366
		100	**α**	0.306	0.019	0.011	0.11	0.511
			θ	3.073	0.024	0.195	2.22	3.95
0.4	**2**	20	**α**	0.327	0.09	0.78	0.0001	0.871
			θ	4.513	0.128	2.31	1.709	7.313
		50	**α**	0.313	0.043	0.025	0.0001	0.619
			θ	4.182	0.046	0.732	1.544	5.821
		100	**α**	0.305	0.018	0.011	0.102	0.509
			θ	4.097	0.024	0.342	1.967	5.228

0.2	**0.23**	20	α	2.218	0.891	5.579	0.0001	6.828
			θ	0.678	0.643	7.934	0.0001	6.188
		50	α	3.154	0.154	4.019	0.0001	7.071
			θ	2.346	0.173	1.065	0.441	4.252
		100	α	3.073	0.024	1.788	0.493	5.669
			θ	2.177	0.089	0.449	0.911	3.44
0.5	**0.4**	20	α	2.385	0.193	22.70	0.0001	11.69
			θ	0.674	0.348	0.173	0.0001	1.413
		50	α	2.099	0.05	1.548	0.0001	4.53
			θ	0.566	0.132	0.047	0.163	0.969
		100	α	1.105	0.035	1.086	0.501	5.64
			θ	0.163	0.081	0.039	0.0001	5.294
0.01	**1.31**	20	α	3.26	0.913	13.85	0.0001	10.536
			θ	2..975	0.512	14.315	0.0001	10.273
		50	α	3.151	0.905	4.154	0.0001	7.135
			θ	2.352	0.824	4.894	0.0001	6.653
		100	α	2.05	0.874	0.632	0.497	3.608
			θ	2.119	0.841	0.648	0.558	3.679
0.2	**0.23**	20	α	2.218	0.891	5.579	0.0001	6.828
			θ	0.679	0.643	7.934	0.0001	6.188
		50	α	2.094	0.953	1.485	0.0001	11.95
			θ	0.056	0.872	3.821	0.0001	5.492
		100	α	2.090	0.970	0.687	0.0001	7.22
			θ	0.53	0.0.03	3.02	0.421	0.805

It is clear from the results that the MSE for the estimates of α and θ are smaller than their corresponding R bias of ML estimates of MOETL. The results reveal that the estimates are stable and quite close to the true parameter values for the large sample size. Furthermore, as the sample size increases the MSEs decreases in all cases, but for special case the MSE for the value of parameter α is very large as parameter is large.

6. Application

In this section we have given a numerical application of the MOETL distribution. We have used data from Nigm et al. (2003) and are about ordered failure of 20 components. The data is given below

0.0009, 0.004, 0.0142, 0.0221, 0.0261, 0.0418, 0.0473, 0.0834, 0.1091, 0.1252, 0.1404, 0.1498, 0.175, 0.2031, 0.2099, 0.2168, 0.2918, 0.3465, 0.4035, 0.6143.

First, it was checked whether the MOETL distribution can be used or cannot to analyze this data set. The Kolmogorov–Smirnov goodness of fit statistic has a value of 0.1869with a p-value of 0.606. This indicates that the data is adequately fitted by MOETL distribution. The plot of data with MOETL distribution is given below. The ML of MOETL distribution of α is 0.352 and the ML of θ is 0.835.the data is highly positively skewed with coefficient of skewness of i.1.439. Figure 5 display the MOETL distribution with new extracted parameters

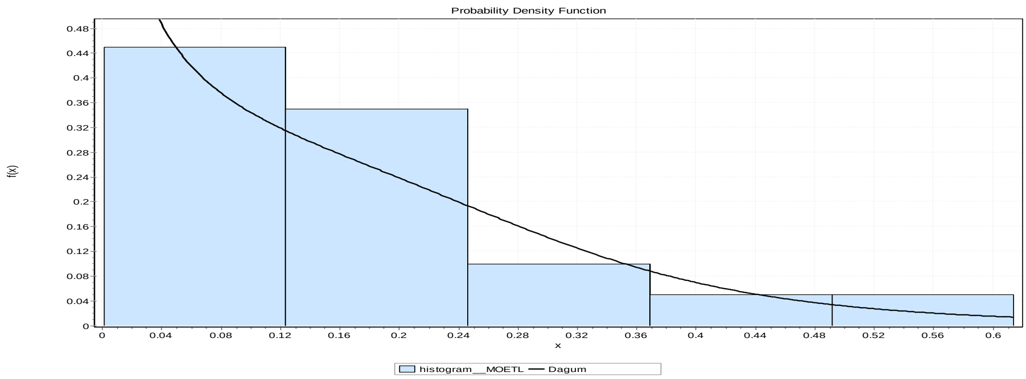

Figure 5: Histogram of Data and Fitted MOETL Distribution

The graph also shows that the data is adequately fitted by MOETLE distribution

The flexibility of the MOETL distribution is measured by comparing the fitted model with the original distribution TL (α) distribution by calculating the (AIC) and (BIC) as presented in Table 2. It can be seen that the TL distribution has higher AIC and BIC values compared to the MOETL distribution, which confirms that MOETL is more suitable for this
data set.

Table2: ML estimates of MOETL parameters, CI and along with the corresponding AIC, BIC statistics

<table>
<tr><th>Model</th><th colspan="5">ML</th></tr>
<tr><td rowspan="2">TL Distribution</td><td colspan="2">estimate</td><td>CI (95%)</td><td>AIC</td><td>BIC</td></tr>
<tr><td>θ</td><td>0.901</td><td>(0.886,0.916</td><td>-3.461</td><td>-4.327</td></tr>
<tr><td rowspan="2">MOETL</td><td>α</td><td>0.352</td><td>(0.336,0.367)</td><td rowspan="2">-10.09</td><td rowspan="2">-11.785</td></tr>
<tr><td>θ</td><td>0.835</td><td>(0.82,0.85)</td></tr>
</table>

7. Conclusion

In this paper, an extended model based on TL distribution is investigated. The method for generating the MOETL distribution is presented in Section 3 This MOE family of distributions has closed forms of cdf, pdf, survival function and hazard function. We apply the general properties to the MOETL distribution. Some statistical properties of this model are obtained. Through numerical simulation, the MLE of the parameters are calculated and discussed. Finally, real data set is fitted to this model and is shown to be appropriate.

References

[١] Alice, T., Jose, K.K. (2003). Marshall- Olkin Pareto Processes, Far East Journal of Theoretical Statistics, 9(2), 117-132.

[2] Alice, T., Jose, K.K.(2005). Marshall-Olkin logistic processes, Science, Technology and Arts Research (STAR) Journal 6(1) ,1-11 are an international.

[3] Al-Zahrani, B, (2012). Goodness-of-fit for the Topp-Leone distribution with unknown parameters. Applied Mathematical Sciences 6(128), 6355 – 6363.

[4] Bayoud, H. A., (2016).Admissible minimax estimators for the shape parameter of Topp-Leone distribution. Communications in Statistics – Theory and Methods, 45, 71-82.

[5] Genç, A. I. (2012). Moments of order statistics of Topp-Leone distribution. Statistical Papers, 53, 117- 131.

[6] Ghitany, M.E., (2007). Asymptotic distribution of order statistics from the Topp-Leone distribution. International Journal of Applied Mathematics, 20, 371–376.

[7] Kotz S, Van Dorp JR (2004). Beyond Beta: Other continuous families of distributions with bounded support and applications, Singapore, World Scientific, 290, 35.

[8] Mi rMostafaee, S.M.T.K.(2014). On the moments of order statistics coming from the Topp - Leone distribution. Statistics & Probability Letter, 95,85-91

[9] Mir Mostafaee, S.M.T.K., Mahdizadeh M., Aminzadeh, M. (2016).Bayesian inference for the Topp Leone distribution based on lower k-record values. Japan Journal of Industrial and Applied Mathematics. 33(3), 637-669.

[10] Nadarajah, S. and Kotz, S. (2003). Moments of some J-shaped distributions. Journal of Applied Statistics, 30, 311-317.

[11] Nadarajah, S., Kotz, S. (2004). The Beta Gumbel Distribution, Mathematical Problems in Engineering, 4, 323-332.

[۱2] Nigm, A. M. AL-Hussaini, E. K. and Jaheen, Z. F. (2003). Bayesian one-sample prediction of future observations under Pareto distribution. Statistics, 37, (6), 527–536.

[13] Topp, C. W. and Leone, F. C. (1955). A family of J-shaped frequency functions. Journal of the American Statistical Association. 50(269), 209-219.

[14] Vicaria, D., Dorpb, J. R. V., and Kotzb. S. (2008).Two-sided generalized Topp and Leone (TS- \lo.nmgGTL) distributions. Journal of Applied Statistics. 35(10), 1115-1129.