International Journal of Mathematical Analysis and Modelling

(Formerly Journal of the Nigerian Society for Mathematical Biology)

Volume 6, Issue 1 (July), 2023

ISSN (Print): 2682 - 5694
ISSN (Online): 2682 - 5708

International Journal of Mathematical Analysis and Modelling
Volume 6, Issue 1, July 2023, pages 115 – 125

Type II Top-Leone generalized Lomax distribution and application to hydrological data

O. Fatoki[*†], N.F. Adeleye[*], and D.O. Omosigho[‡]

Abstract

In this paper, a new three-parameter extended Lomax distribution called the Type II Top-Leone Lomax distribution was developed. We demonstrated its flexibility and tractability over inverse Lomax distribution through various theoretical and practical approaches. The derived properties include the quantiles, moments, Renyi entropy, and order statistics. Finally, we consider a river data set to demonstrate the flexibility of the newly developed model.

Keywords: moments; entropy; Type II Top-Leone Lomax distribution; quantiles

1 Introduction

Introduction In the last decades, the Lomax distribution introduced by Lomax (1954), has been discovered to be very useful in several areas of applications most especially in applied sciences such as applications in flood, queue theory, internet traffic control, life testing, wind speed, sea waves and many others. The use of Lomax distribution is motivated by its heavy tail and simplicity in applications. It provides an alternative to the exponential-type distributions (Rayleigh, Weibull, exponential, Gompertz, pareto.), for further study see Bryson (1974). A random variable X is said to be distributed according to Lomax distribution, if it distribution function is given by

$$G(x) = 1 - \left(1 + \frac{x}{a}\right)^{-c} \tag{1}$$

with $a, c > 0.$ Thus, c is a shape parameter and a is a scale parameter. The corresponding probability density function (pdf) associated to (1) is given by

$$g(x) = \frac{c}{a}\left(1 + \frac{x}{a}\right)^{-c}. \tag{2}$$

It should be noted that the pdf of Lomax distribution is naturally a special case of some well-known distributions, and this includes Pareto type II, Pareto type IV, Feller-Pareto, and many others. However, the Lomax distribution has limited scope of applications and this includes: Lack of flexibility, poor fits etc., when used to model real life data which exhibits non-monotonic or bathtub

[*]Department of Statistics, Ogun State Institute of Technology, Igbesa, 0gun State, Nigeria.
[†]Corresponding author: fatoki_olayode@yahoo.com
[‡]Department of Mathematics and Statistics, Federal Polytechnic, Ado-Ekiti, Nigeria.

failure rate. Efforts have been made to address the problem aforementioned and this involved the generalization of the Lomax distribution in other to induce flexibility into the distribution and also improve its fits for a better modeling potentials. Among these are: Al-Zahrani and Sagor (2014) who developed and studied the Poisson Lomax distribution, the gamma Lomax distribution by Cordeiro and Ortega (2015), Marshall-Olkin Extended Lomax distribution by Ghitany et al. (2007), , the Exponential Lomax distribution was studied by El-Bassiouny et al. (2015), Kilany (2016) studied the weighted Lomax distribution, the Weibull Lomax distribution was studied by Tahir et al. (2015), Abdul-Moniem and Abdel-Hameed (2012) developed and studied the Exponentiated Lomax distribution, the transformed-transformer Lomax distribution was studied by Alzaghal and Hamed (2019), and the Power Lomax distribution was developed and studied by Rady et al. (2016).

Elgarhy et al. (2018) developed and studied a new family of distribution called Type II Topp-Leone (TIITL-G) family of distributions which add a new parameter to the baseline distribution to produce a more flexible distribution that represents a wide range of behavior of the hazard function (increasing, decreasing, non-monotonic, bathtub shapes). The new generalization is a suitable alternative for transmuted-G, exponentiated-G, odds generalized exponential-G, and beta-G, Kumaraswamy-G families, and many more others in modeling real-life data. The CDF for the TIITL-G family of distribution is given by

$$F(x;\lambda,\zeta) = 1 - [1 - G(x;\zeta)^2]^{\lambda}, \quad x \in \mathbb{R}, \tag{3}$$

where λ is a positive shape parameter and $G(x;\zeta)$ is a cdf of a baseline continuous distribution which sometimes may depend on a parameter vector ζ. The associated pdf corresponding to (3) is given by

$$f(x;\lambda,\zeta) = 2\lambda g(x;\zeta)G(x;\zeta)[1 - G(x;\zeta)^2]^{\lambda-1}, \quad x \in \mathbb{R}, \tag{4}$$

where $g(x;\zeta)$ is the pdf corresponding to $G(x;\zeta)$.

Elgarhy et al. (2018), described (TIITLG) family as a simplified version of the Kumaraswamy-G family. The added shape parameter λ is to possibly control the tail weight and skewness of the cdf of baseline distribution. The TIITL-G family has been used by several authors to extend several distributions which includes Type-II Topp-Leone generalized inverse Rayleigh distribution by Mohammed and Yahia (2019), Type-II Topp-Leone inverse Rayleigh distribution by Yahia and Mohammed (2019), Type-II Topp-Leone Inverted Kumaraswamy distribution by Zein Eldin et al. (2019), Type-II Topp-Leone Bur XII distribution by Ogunde and Oyebimpe (2022). In this work, we employ the same approach in extending the Lomax distribution to obtain a new extension of the L distribution called the Type-II Topp-Leone Lomax (TIITLL) distribution. The new distribution is more flexible with some desirable properties which enable its use in modeling real-life data of different shapes of the hazard function.

2 Type II Topp-Leone Lomax distribution

A random X is said to follow a Type-II Topp-Leone Lomax (TIITLL) distribution if its cdf is given by

$$F(x;a,c,\lambda) = 1 - \left[1 - \left(1 - \left(1 + \frac{x}{a}\right)^{-c}\right)^2\right]^{\lambda}, \quad x \in \mathbb{R}; \tag{5}$$

where a, c, λ are positive shape parameters. The corresponding pdf to (5) is given by

$$f(x;a,c,\lambda)=\frac{2\lambda}{a}c\left(1+\frac{x}{a}\right)^{-c-1}\left[1-\left(1+\frac{x}{a}\right)^{-c}\right]\left[1-\left(1-\left(1+\frac{x}{a}\right)^{-c}\right)^{2}\right]^{\lambda-1},\quad x\in\mathbb{R}, \tag{6}$$

where a, c, $\lambda > 0$, and are all shape parameters. The graph of the density function and distribution function are given in figures 1 and 2 presented below for different values of the parameters. The graphs show that the $TIITLL$ distribution is unimodal and right-skewed.

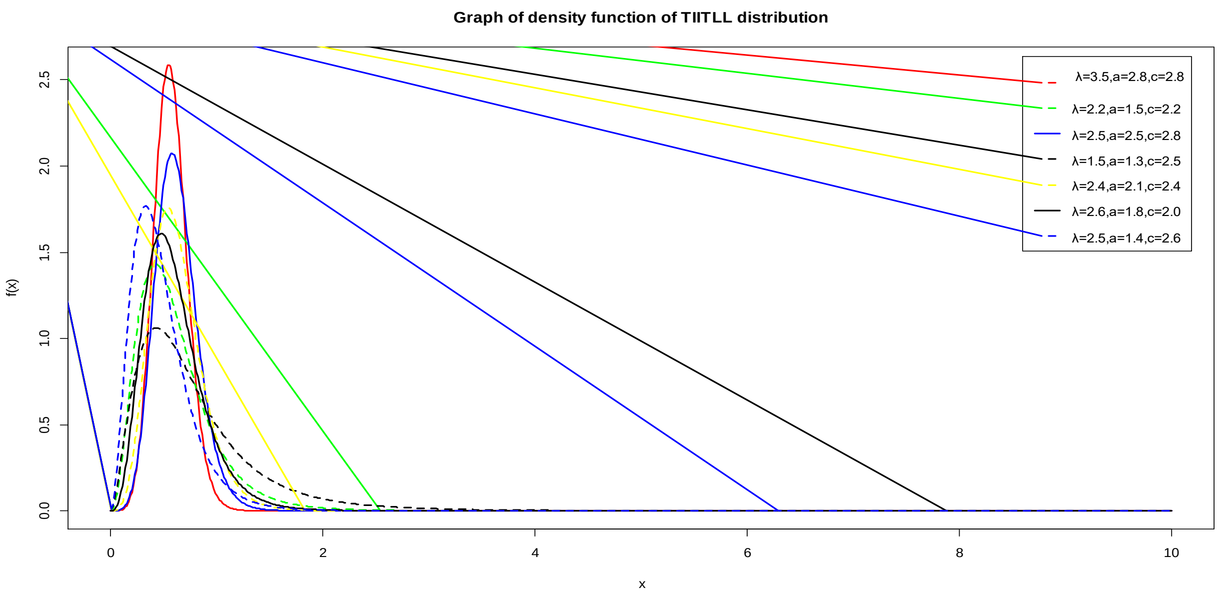


Figure 1: The graph of density function TIITLL distribution

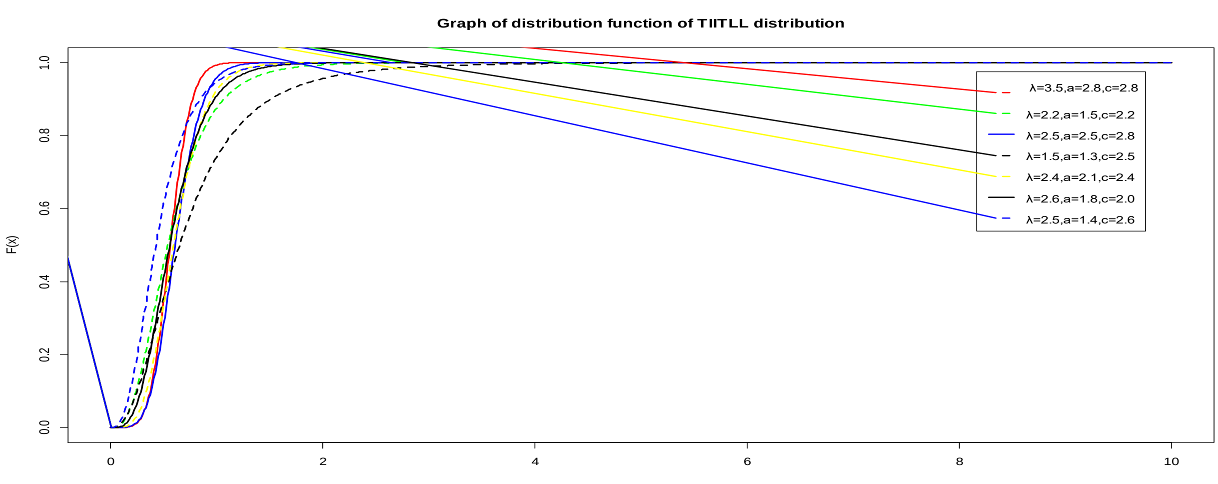


Figure 2: The graph of distribution function TIITLL distribution

2.1 Survival, Hazard, Reversed Hazard and the cumulative hazard function

The survival function of a random variable X that follows $TIITLBXII$ distribution is given by

$$S(x) = 1 - F(x) \tag{7}$$

$$S(x; a, c, \lambda) = \left[1 - \left(1 - \left(1 + \frac{x}{a}\right)^{-c}\right)^{2}\right]^{\lambda}, \quad x \in \mathbb{R}, \tag{8}$$

The graph of the survival function of $TIITLBXII$ distribution is given in figure 3 of various values of the parameters.

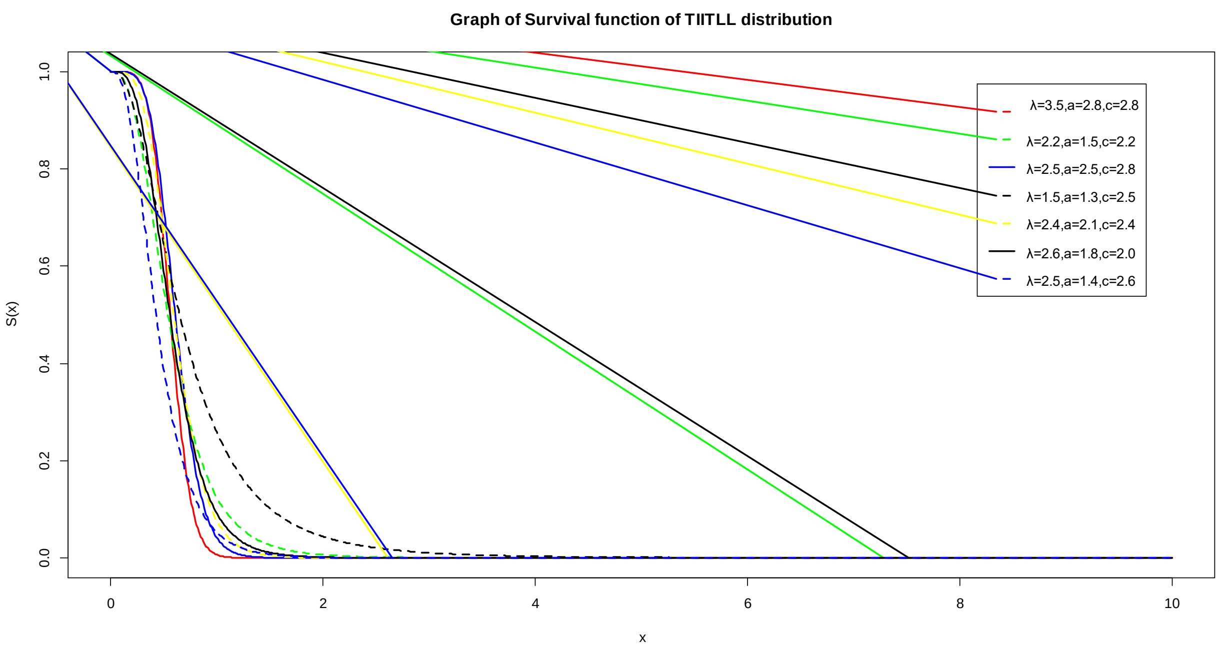


Figure 3: The graph of survival function TIITLL distribution

The hazard function is given by

$$h(x) = \frac{f(x)}{S(x)} \tag{9}$$

$$= \frac{\frac{2\lambda}{a} c \left(1 + \frac{x}{a}\right)^{-c-1} \left[1 - \left(1 + \frac{x}{a}\right)^{-c}\right] \left[1 - \left(1 - \left(1 + \frac{x}{a}\right)^{-c}\right)^{2}\right]^{\lambda-1}}{\left[1 - \left(1 - \left(1 + \frac{x}{a}\right)^{-c}\right)^{2}\right]^{\lambda}} \tag{10}$$

The graph of the hazard function of $TIITLL$ distribution is presented in Figure 4 for different values of the parameters.

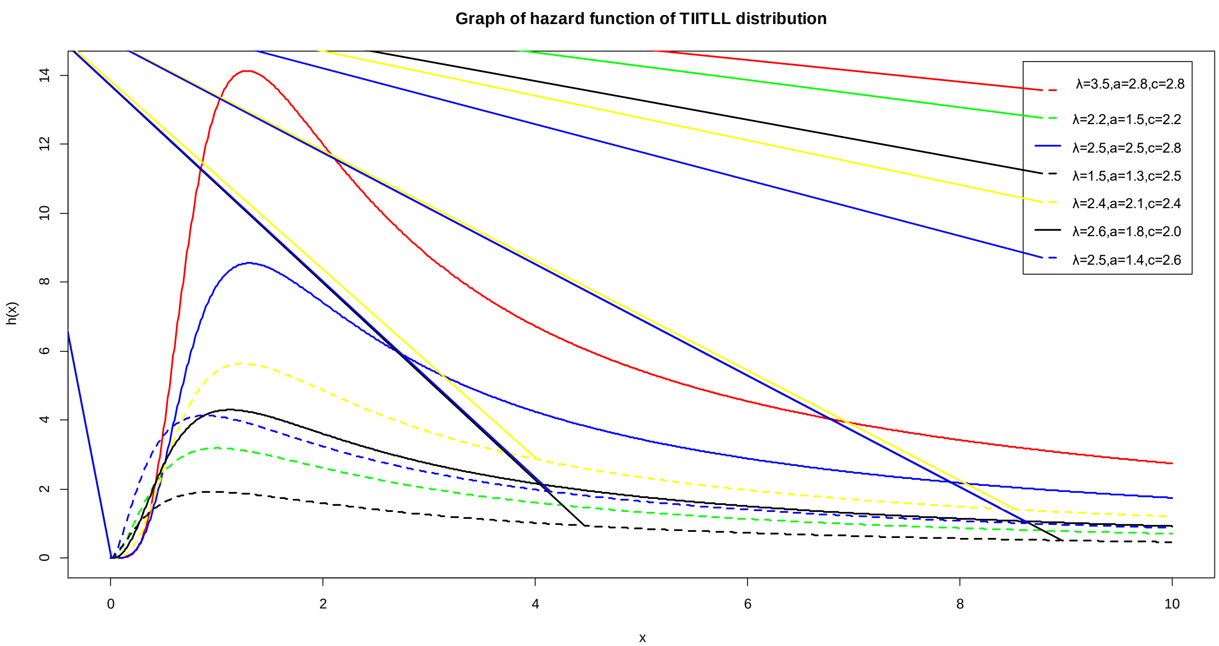


Figure 4: The graph of hazard function TIITLL distribution

The graph of the hazard function in figure 4 for various parameter values for $TIITLL$ distribution indicates that the new model exhibits monotonically increasing, decreasing, and upside-down bathtub shapes. This property clearly explains the flexibility of $TIITLl$ distribution and its applicability in modeling non-monotonic and monotonic hazard behavior which are often encountered in real-life data and

$$\begin{aligned} H(x) &= -log\left(1 - F(x)\right) \\ &= -log\left(1 - \left[1 - \left(1 - \left(1 + \frac{x}{a}\right)^{-c}\right)^{2}\right]^{\lambda}\right) \end{aligned} \tag{11}$$

for $x > 0, a > 0,$ and $c > 0$

2.2 Quantiles of the TIITLL distribution

The p^{th} quantile (x_p) of the TIITLL distribution is obtained by solving the equation.

$$F(x_p) = \text{p}.$$

Hence solving equation (5) we get

$$x_p = a\left[\left\{1 - \left[1 - (1-p)^{1/\lambda}\right]^{1/2}\right\}^{-1/v} - 1\right] \tag{12}$$

the p^{th} quantile for $p \in (0,1)$ for $p = 0.25, 0.5, 0.75$, we have the lower quartile, middle quartile (median) and the upper quartile of the $TIITLL$ distribution respectively, given by

$$x_{0.25} = a\left[\left\{1 - \left[1 - 0.75^{1/\lambda}\right]^{1/2}\right\}^{-1/v} - 1\right], \tag{13}$$

$$x_{0.5} = a\left[\left\{1 - \left[1 - 0.5^{1/\lambda}\right]^{1/2}\right\}^{-1/v} - 1\right], \tag{14}$$

and

$$x_c = a\left[\left\{1 - \left[1 - 0.25^{1/\lambda}\right]^{1/2}\right\}^{-1/v} - 1\right] \tag{15}$$

2.3 Expansion for the density function

Consider the power series given by

$$(1 - J)^{m-1} = \sum_{i=0}^{\infty} (-1)^i \binom{m-1}{i} J^i \tag{16}$$

Which holds for $|J| < 1$ and $m > 0$ is a real non-integer. Using the power series in equation (16) and after some elementary algebra the pdf of $TIITLL$ model in equation (6) can be expressed as

$$f(x) = \frac{2c\lambda}{a} \sum_{i=j=0}^{\infty} (-1)^{i+j} \binom{\lambda - 1}{i} \binom{2i + 1}{j} \left(1 + \frac{x}{a}\right)^{-[c(j+1)+1]} \tag{17}$$

3 Statistical Properties of the $TIITLBXII$ distribution

3.1 The ordinary and incomplete moments of $TIITLBXII$ distribution

Moments are defined as the expected value of a certain function of a random variable. The moments of different types of $TIITL$L distribution can be obtained by direct calculations due to their mathematical tractability. Thus, the s^{th} ordinary moment of a distribution is given by

$$\mu_s' = E(x)^s$$

Thus the s^{th} moment of $TIITLL$ distribution is given by

$$\mu_s' = \int_0^{\infty} x^s f(x) dx. \tag{18}$$

Putting equation (17) in (18), we have

$$E(X^s) = \frac{2c\lambda}{a} \sum_{i=j=0}^{\infty} (-1)^{i+j} \binom{\lambda - 1}{i} \binom{2i + 1}{j} \int_0^{\infty} x^s \left(1 + \frac{x}{a}\right)^{-[c(j+1)+1]} dx. \tag{19}$$

Now by applying the change of variables, $w = {}^{x}/_{a}, x = aw, dx = adw,$ then we have

$$E(X^s) = 2c\lambda \sum_{i=j=0}^{\infty} (-1)^{i+j} \binom{\lambda-1}{i} \binom{2i+1}{j} a^s \int_{-\infty}^{\infty} w^s (1+w)^{-[c(j+1)+1]} \, dw. \tag{20}$$

Also, taking $z = u(1-u)^{-1}$, $dz = (1-u)^{-2}du$, then from equation (20), we obtain

$E(X^s) = 22c\lambda \sum_{i=j=0}^{\infty} (-1)^{i+j} \binom{\lambda-1}{i} \binom{2i+1}{j} a^s \int_0^{\infty} u^s (1-u)^{-s-2-[c(j+1)+1]} du.$

Finally, we have,

$$E(X^s) = \mu_s' = 2c\lambda \sum_{i=j=0}^{\infty} (-1)^{i+j} \binom{\lambda-1}{i} \binom{2i+1}{j} a^s B[(s+1), -(s+c(j+1)] \tag{21}$$

$where\ B(v,n) = \int_0^1 z^{v-1}(1-z)^{n-1}dz,$ is the standard beta function with $v > 0$ and $n > 0$.

It should be noted that equation (21) can serve as a numerical value of μ_s' and all other associated measures. The chief among them are the mean, the variance, and the coefficient of variation (CV) of X defined by $\mu_1' = \mu$, the variance given by $\mu_2 = \mu_2' - (\mu_1')^2$ and the $CV = {\mu_2}^{1/2}/\mu$.

3.2 Renyi entropy

Renyi (1961), gave a useful mathematical expression that can be used to measure the entropy of a TIITLL distribution given by

$$I_R^{(v)} = \frac{1}{1-\Upsilon} log \left[\int_0^{\infty} f_{\text{TIITLL}}(x;\zeta)^v \right], \quad v > 0, v \neq 1. \tag{22}$$

Putting equation (6) in (22), we have

$$I_R^{(v)} = \frac{1}{1-v} log \left[\int_0^{\infty} \left(\frac{2\lambda}{a} c \left(1+\frac{x}{a}\right)^{-c-1} \left[1-\left(1+\frac{x}{a}\right)^{-c}\right] \left[1-\left(1-\left(1+\frac{x}{a}\right)^{-c}\right)^2\right]^{\lambda-1} \right)^v \right].$$

Using the binomial expression in (16), we have

$$I_R^{(v)} = \frac{1}{1-v} log \left[J^* \int_0^{\infty} x^0 \left(1+\frac{x}{a}\right)^{-cq} dx \right] \tag{23}$$

where

$$J^{*}=\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\binom{\left(v(\lambda-1)\right)}{p}\binom{2i+\lambda}{q}(-1)^{p+q}$$

It should be noted that for the above equation to have a feasible solution the sum of p and q be even. Then taking

$$Z=\int_{0}^{\infty}x^{0}\left(1+\frac{x}{a}\right)^{-cq}dx \tag{24}$$

In equation (24), let $y=\frac{x}{a}$, then $ady=dx$, then

$$Z=\int_{0}^{\infty}y^{0}(1+y)^{-cq}ady \tag{25}$$

Also taking $y=p(1-p)^{-1}$, $dy=(1-p)^{-2}dp$, and putting in (25), finally we have

$$I_{R}^{(v)}=\frac{1}{1-v}log\left[(c\lambda)^{v}a^{1-v}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\binom{\left(v(\lambda-1)\right)}{p}\binom{2i+\lambda}{q}(-1)^{p+q}B(1,1+\lambda-C(q+\lambda))\right]$$

(26)

$$where\ B(m,n)=\int_{0}^{\infty}z^{m-1}(1-z)^{n-1}dz\,,is\ the\ beta\ function$$

4 Maximum likelihood estimates of the parameters

The maximum likelihood approach is used to estimates the unknown parameters of the distribution. Let $\underline{x}=x_{1}\ldots,x_{n}$ represent a random sample obtained from the $TIITL$L distribution. The likelihood function $L(x;\ \zeta)$ and the log-likelihood function $logL(x;\ \zeta)=l(x;\ \zeta)$ corresponding to (43) are respectively given as

$$L(x;\zeta)=\frac{2\lambda}{a}c\prod_{i=1}^{n}\left(1+\frac{x}{a}\right)^{-c-1}\left[1-\left(1+\frac{x}{a}\right)^{-c}\right]\left[1-\left(1-\left(1+\frac{x}{a}\right)^{-c}\right)^{2}\right]^{\lambda-1} \tag{27}$$

and

$$\begin{aligned}l=\log\left(\frac{2\lambda}{a}c\right)-(c+1)\sum_{i=1}^{n}log\left(1+\frac{x}{a}\right)+\sum_{i=1}^{n}log\left[1-\left(1+\frac{x}{a}\right)^{-c}\right]\\+(\lambda-1)\sum_{i=1}^{n}log\left[1-\left(1-\left(1+\frac{x}{a}\right)^{-c}\right)^{2}\right]\end{aligned} \tag{28}$$

The element of the score vector is given as

$$\frac{\partial l}{\partial a} = -\frac{n}{a} - (c+1)\sum_{i=1}^{n}\frac{a^{-2}x}{\left(1+\frac{x}{a}\right)} - \sum_{i=1}^{n}\frac{cx\left(1+\frac{x}{a}\right)^{-c}}{a^{2}\left[1-\left(1+\frac{x}{a}\right)^{-c}\right]}$$
$$+(\lambda-1)\sum_{i=1}^{n} log\frac{\partial}{\partial a}\left[1-\left(1-\left(1+\frac{x}{a}\right)^{-c}\right)^{2}\right] \tag{29}$$

$$\frac{\partial l}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n} log\left(1+\frac{x}{a}\right) + \sum_{i=1}^{n}\frac{\left(1+\frac{x}{a}\right)^{-c} log\left(1+\frac{x}{a}\right)}{\left[1-\left(1+\frac{x}{a}\right)^{-c}\right]} \tag{30}$$

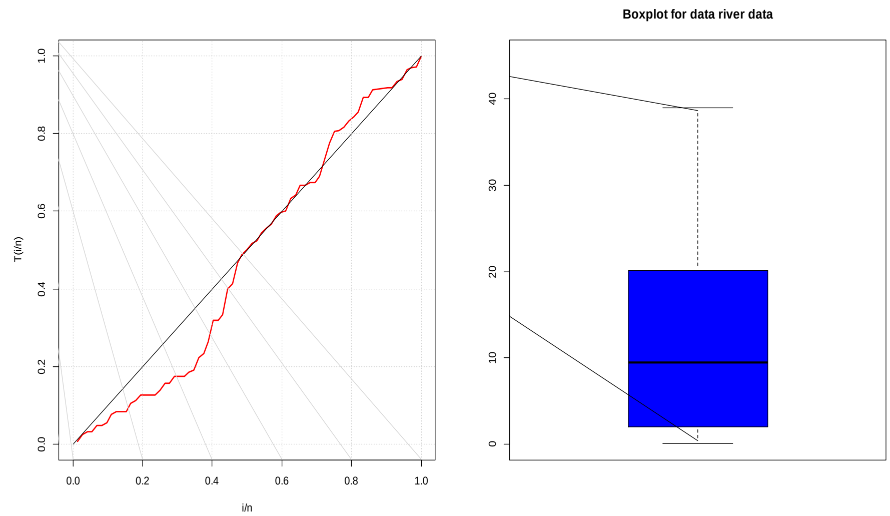
$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n} log\left[1-\left(1-\left(1+\frac{x}{a}\right)^{-c}\right)^{2}\right] \tag{31}$$

4.1 Application

In this section, a real data sets are provided to illustrate the importance of the $TIITLL$ distribution. We compare the $TIITLL$ model with other competitive models such as Type II half logistic Inverse Lomax and the Lomax (IL) distribution.

The data set consisting of exceedances (rounded to one decimal place) of flood peaks (in m3/s) of the Wheaton River, which is located in the Yukon Territory, Canada, for the years 1958-1984. The data set is the following: 1.7, 2.2, 14.4, 1.1, 0.4, 20.6, 5.3, 0.7, 1.9, 13.0, 12.0, 9.3, 1.4, 18.7, 8.5, 25.5, 11.6, 14.1, 22.1, 1.1, 2.5, 14.4, 1.7, 37.6, 0.6, 2.2, 39.0, 0.3, 15.0, 11.0, 7.3, 22.9, 1.7, 0.1, 1.1, 0.6, 9.0, 1.7, 7.0, 20.1, 0.4, 2.8, 14.1, 9.9, 10.4, 10.7, 30.0, 3.6, 5.6, 30.8, 13.3, 4.2, 25.5, 3.4, 11.9, 21.5, 27.6, 36.4, 2.7, 64.0, 1.5, 2.5, 27.4, 1.0, 27.1, 20.2, 16.8, 5.3, 9.7, 27.5, 2.5, 27.0.

The set data has been recently used by Mahmoud (2018). Some summary statistics of these data are: n = 72, $\bar{x}$= 12.20417, s = 12.29722, coefficient of skewness = 1.47251 and coefficient of kurtosis = 2.88955. The boxplot of these observations displayed in Figure 5 (b) indicates that the distribution is right-skewed. The TTT (total time on test) plot of these data is shown in Figure 5(a)



(a) TTT plot (b) Box plot

Figure 5. The graphs of TTT plot and the Box plots for the River data set.

To check the validity of the fitted model, Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Hannan-Quinn criterion (HQIC), Bayesian information criterion (BIC), Crammer VAN Von-Misses (CV) goodness of fit test are obtained. In general, it considered that the smaller the values of AIC, BIC, CAIC, HQIC and CV statistics the better the fit of the model

Table 1: MLEs and Measures of goodness of fits for River data

$Model$	a	c	λ	$-l$	AIC	BIC	$CAIC$	$HQIC$	CV
$TLTIIL$	-0.279	0.551	2.608	250.315	506.631	513.461	506.984	509.350	0.099
$TIIHLIL$	0.402	9.804	2.977	259.258	524.516	531.346	524.869	527.236	0.312
L	0.279	1.832	–	240.418	544.835	549.389	545.009	546.648	0.503

4.2 Concluding remarks

In this work, we have discussed a new generalization of the Lomax model, called the Type-II Topp-Leone Lomax model. This generalization is a product of the transformation of the two-parameter Lomax model using Type II Topp-Leone generated family of distributions developed by Elgarhy et al. (2018). The model is more flexible than the Type Half Logic Lomax, and Lomax model. The newly developed model is useful in modeling lifetime data with different shapes of the hazard function. The properties of the proposed distribution are discussed.

We give explicit expressions for ordinary, order statistics, and entropy. Maximum likelihood estimation is discussed within the framework of asymptotic log-likelihood inferences. The three-parameter Type-II Topp-Leone Lomax model produced monotonically increasing, decreasing, and inverted bathtub hazard rates. In terms of the statistical significance of the model adequacy, the Type-

II Topp-Leone Lomax model gives a better fit than its sub-models and other competing models considered in this work.

References

[1] Abdul-Moniem, I. B. and Abdel-Hameed, H. F. (2012). On the exponentiated Lomax distribution. International Journal of Mathematical Achieve, 3: 2144-2150.

[2] Alzaghal, A., and Hamed, D. (2019). New families of Generalized Lomax distribution: Properties and Applications. International Journal of Statistics and Probability, 8: 51-68.

[3] Al. Zahrani, B., and Sagor, H. (2014). The Poisson-Lomax distribution. Rev. Colomb Estad., 37: 225-245.

[4] Bryson, M. C. (1974). Heavy-tailed distribution: properties and tests. Technometrics, 16: 61-68.

[5] Elgarhy, M.., Nasir, M.A., Jamal, F., and Ozel, G. (2018). The type II Topp-Leone generated family of distributions: Properties and Applications. Journal of Statistical Management System, 21: 1529–1551.

[6] Kilany, N. M. (2016). Weighted Lomax distribution. Sringer Plus, 5: 1862.

[7] Ghitany, M. E., Al-Awadhi, F. A., and Alkhalfan, L. (2007). Marshall–Olkin extended Lomax distribution and its application to censored data. Communications in Statistics-Theory and Methods, 36(10): 1855-1866.

[8] El-Bassiouny, A. H., Abdo, N. F., and Shahen, H. S. (2015). The Exponential Lomax distribution. International Journal of Computer Applications, 121: 24-29.

[9] Lomax, K. S. (1954). Business failures: another example of the analysis of failure data. Journal of the American Statistical Association, 49(268): 847-852.

[10] Cordeiro, G.M., Ortega, E.M., and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. Journal of Frankling Institute, 347: 1399–1429.

[11] Mahmoud M. Mansour, Enayat M. Abd Elrazik , Emrah Altun, Ahmed Z. Afify and Zafar Iqbal (2018). A new three-parameter frechet distribution: properties and applications. Pakistan Journal of Statistics, 4(6), 441-458.

[12] Ogunde, A. A., and Adeniji O. E. (2022). Type II Topp-Leone Bur XII distribution: Properties and applications to failure time data. Scientific African. 16: e01200.

[13] Rady, E. A., Hassaine, W. A., Elhaddad, T. A. (2016). The power Lomax distribution with an Application to bladder cancer data. SpringerPlus, 5: 1838.

[14] Renyi, A. (1961). On measures of entropy and information. Proceedings of the 4th Berkeley Symposiumon Mathematical Statistics and Probability, Vol. I, University of California Press, Berkeley. pp. 547-561.