Pak.j.stat.oper.res. Vol.20 No.2 2024 233-260 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.4458

Pakistan Journal of Statistics and Operation Research

A New Family of Heavy-Tailed Generalized Topp-Leone-G Distributions with Applications

Thatayaone Moakofi[1*], Broderick Oluyede[2], Bakang Tlhaloganyang[3] and Agolame Puoetsile[4]

*Corresponding author

1. Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology, Botswana, thatayaone.moakofi@gmail.com
2. Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology, Botswana, oluyedeo@biust.ac.bw
3. Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology, Botswana, tlhaloganyangs@gmail.com
4. Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology, Botswana, agolame.puoetsile@studentmail.biust.ac.bw

Abstract

In this article, we introduce a robust generalization of the generalized Topp-Leone-G (GEN-TL-G) family of distributions via the heavy-tailed technique, namely, heavy-tailed generalized Topp-Leone-G (HT-GEN-TL-G) family of distributions. Statistical properties of the HT-GEN-TL-G family of distributions including reliability functions, quantile function, density expansion, moments, moment generating function, incomplete moments, Rényi entropy, distribution of order statistics are derived. Different estimation methods including Maximum Likelihood, Anderson-Darling, Ordinary Least Squares, Weighted Least Squares, Cramér-von Mises and Maximum Product of Spacing are utilized to estimate the unknown parameters of the new distribution, and a simulation study is used to compare the results of the estimation methods. Risk measures for this distribution were also developed and finally the effectiveness of this new family of distributions was demonstrated using applications to two real data sets. Graphical and application results in this manuscript were obtained using R programming language.

Key Words: Heavy-tailed distribution; Topp-Leone-G distribution; Estimation; Applicability.

Mathematical Subject Classification: 60E05, 62E15.

1. Introduction

Recently, many classical or standard distributions fail when it comes to modeling real world events.This is due to the fact that they are not flexible enough to capture the characteristics of newly encountered data. To deal with this problem, researchers are still extending distributions by using different transformation techniques. Some of the distributions developed include among others the new odd log-logistic generalised half-normal distribution by Afshar et al. (2019), the Topp-Leone-Marshall-Olkin-G family of distributions by Chipepa et al. (2020), the Marshall-Olkin-Topp-Leone flexible Weibull distribution by Mohammad (2020), generalised odd Frechét family of distributions by Marganpoor et al. (2020), Topp–Leone modified Weibull distribution by Alyami et al. (2022), the odd Weibull inverse Topp–Leone distribution by Almetwally (2022), the gamma-Topp-Leone-type II-exponentiated half logistic-G family of distributions by Oluyede and Moakofi (2023), Gompertz Topp–Leone inverse Weibull distribution by Khaleel and Hammed (2023), generalised exponential-Gaussian distribution by Marmolejo-Ramos et al. (2023), and the Harris-Topp-Leone-G family of distributions by Moakofi and Oluyede (2023).

Despite their numerous advantages, many new proposed distributions are not flexible enough to provide an adequate fit to datasets that are positive, skewed, unimodal with heavy tails. Thus, reseachers are still on the lookout for heavy-tailed distributions in order to model heavy-tailed data. Some of the recently developed heavy-tailed distributions include LogPH distribution by Ahn et al. (2012), the exponent power-Weibull distribution by Ahmad et al. (2020), the logit slash distribution by Korkmaz (2020), heavy-tailed beta-power transformed Weibull distribution by Zhao et al. (2021), a new heavy-tailed Weibull distribution by Ahmad et al. (2022), the generalized geometric Rayleigh reciprocal Weibull distribution by Yousof et al. (2023) and Kavya-Manoharan power Lomax distribution by Riad et al. (2023).

The cumulative distribution function (cdf) and probability density function (pdf) of the type I heavy-tailed (TI-HT) family of distributions introduced by Zhao et al. (2020) are given by

$$F_{TI-HT-G}(x;\theta,\varphi) = 1 - \left(\frac{1-G(x;\varphi)}{1-(1-\theta)G(x;\varphi)}\right)^{\theta} \tag{1}$$

and

$$f_{TI-HT-G}(x;\theta,\varphi) = \frac{\theta^2 g(x;\varphi)\,(1-G(x;\varphi))^{\theta-1}}{(1-(1-\theta)G(x;\varphi))^{\theta+1}}, \tag{2}$$

respectively, for $\theta > 0$, $x \in \mathbb{R}$ and parameter vector φ, where $G(x;\varphi)$ is the cdf of the baseline distribution.

The cdf and pdf of the new GEN-TL-G family of distributions are given as

$$F_{GEN-TL-G}(x;b,\beta,\varphi) = 1 - \left[1 - \left(1 - \overline{G}^2(x;\varphi)\right)^b\right]^{\beta} \tag{3}$$

and

$$\begin{aligned} f_{GEN-TL-G}(x;b,\beta,\varphi) &= 2\beta b g(x;\varphi)\left(1-\overline{G}^2(x;\varphi)\right)^{b-1}\overline{G}(x;\varphi) \\ &\times \left(\left[1-\left(1-\overline{G}^2(x;\varphi)\right)^b\right]\right)^{\beta-1}, \end{aligned} \tag{4}$$

respectively, for $b, \beta > 0$ and baseline vector of parameters φ, where $G(x;\varphi)$ is the cdf of the baseline distribution.

Our basic motivations lies in the flexibility of the new family of distributions to model both monotonic and non-monotonic hazard rate functions by capturing different shapes; the ability of the new model in providing better fits than the baseline and several extended distributions available in the literature; the applicability of the special cases of the new family of distributions in real life scenarios. Another interesting part is the role played by the extra shape parameter(s) by introducing skewness and modulating the weight of the tails of any baseline distribution.

The main objective of this work is to introduce a new flexible family of distributions that can characterize several available or emerging data sets. The distribution is named HT-GEN-TL-G family of distributions.

This work is structured as follows: In Section 2, we develop the new family of distributions, namely the HT-GEN-TL-G distribution and some of its functions. We also present the series expansion of the density function. Statistical properties of the new family are presented in Section 3. Some special cases are offered in Section 4. Estimation of model parameters are presented in Section 5. In Section 6, Monte Carlo simulation study is conducted to examine the bias and mean square error of the maximum likelihood estimators for each parameter. Actuarial measures and their simulation study are given in Section 7. Applications of the proposed model to two real data are given in Section 8, followed by concluding remarks in Section 9.

2. The New Family of Distributions

In this section, we define the pdf, cdf, survival function (sf), hazard rate fuction (hrf) and density expansion of the new heavy-tailed generalized-Topp-Leone-G (HT-GEN-TL-G) family of distributions. By substituting the cdf and pdf of the new GEN-TL-G family of distributions defined in Eqns. (3) and (4) into Eqns. (1) and (2), we obtain the cdf and pdf of the HT-GEN-TL-G family of distributions given by

$$F_{HT-GEN-TL-G}(x;\theta,b,\beta,\varphi)=1-\left[\frac{\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}}{1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}\right)}\right]^{\theta} \tag{5}$$

and

$$\begin{aligned} f_{HT-GEN-TL-G}(x;\theta,b,\beta,\varphi) &= 2\beta b\theta^{2}g(x;\varphi)\left(1-\overline{G}^{2}(x;\varphi)\right)^{b-1}\overline{G}(x;\varphi) \\ &\times \left(\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]\right)^{\beta\theta-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}\right)\right]^{-(\theta+1)}, \end{aligned} \tag{6}$$

respectively, for $\theta,b,\beta>0$ and baseline vector of parameters φ.

The sf and hrf are given by

$$S(x;\theta,b,\beta,\varphi) = \left[\frac{\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}}{1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}\right)}\right]^{\theta},$$

and

$$\begin{aligned} h_{F}(x;\theta,b,\beta,\varphi) &= 2\beta b\theta^{2}g(x;\varphi)\left(1-\overline{G}^{2}(x;\varphi)\right)^{b-1}\overline{G}(x;\varphi) \\ &\times \left(\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]\right)^{-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}\right)\right]^{-1}. \end{aligned} \tag{7}$$

2.1. Sub-families

Several sub-families of the HT-GEN-TL-G family of distributions are presented in this sub-section.

- When $\theta=1$, we obtain the new generalized-Topp-Leone-G (GEN-TL-G) family of distributions with the cdf

$$F(x;b,\beta,\varphi)=1-\left[1-\left(1-\overline{G}^{2}(x;\varphi)\right)^{b}\right]^{\beta}$$

for $b,\beta>0$ and parameter vector φ.

- When $b = 1$, we obtain the new heavy-tailed family of distributions with the cdf

$$F(x;\theta,\beta,\varphi) = 1 - \left[\frac{\left[1-\left(1-\overline{G}^2(x;\varphi)\right)\right]^{\beta}}{1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^2(x;\varphi)\right)\right]^{\beta}\right)} \right]^{\theta} \tag{8}$$

for $\theta, \beta > 0$ and baseline vector of parameters φ.

- When $\beta = 1$, we obtain the heavy-tailed Topp-Leone-G (HT-TL-G) family of distributions with the cdf

$$F(x;\theta,b,\varphi) = 1 - \left[\frac{\left[1-\left(1-\overline{G}^2(x;\varphi)\right)^{b}\right]}{1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^2(x;\varphi)\right)^{b}\right]\right)} \right]^{\theta} \tag{9}$$

for $\theta, b > 0$ and baseline vector of parameters φ. This is a new family of distributions.

- When $\theta = b = 1$, we obtain the new family of distributions with the cdf

$$F(x;\beta,\varphi) = 1 - \overline{G}^{2\beta}(x;\varphi)$$

for $\beta > 0$ and parameter vector φ.

- When $\theta = \beta = 1$, we obtain the Topp-Leone-G (TL-G) family of distributions with the cdf

$$F(x;b,\varphi) = \left(1-\overline{G}^2(x;\varphi)\right)^{b}$$

for $b > 0$ and parameter vector φ. (See Al-Shomrani et al. (2016)).

- When $b = \beta = 1$, we obtain the new heavy-tailed family of distributions with the cdf

$$F(x;\theta,\varphi) = 1 - \left[\frac{\left[1-\left(1-\overline{G}^2(x;\varphi)\right)\right]}{1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^2(x;\varphi)\right)\right]\right)} \right]^{\theta} \tag{10}$$

for $\theta > 0$ and baseline vector of parameters φ.

- When $\theta = b = \beta = 1$, we obtain the new family of distributions with the cdf

$$F(x;\varphi) = 1 - \overline{G}^2(x;\varphi)$$

for parameter vector φ.

2.2. Quantile Function

The quantile function of the HT-GEN-TL-G family of distributions is given in this sub-section. Assuming a random variable X follows the HT-GEN-TL-G family of distributions, and the random variable U follows the uniform distribution, the quantile function of X can be obtained as follows:

$$1 - \left[\frac{\left[1-\left(1-\overline{G}^2(x;\varphi)\right)^{b}\right]^{\beta}}{1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^2(x;\varphi)\right)^{b}\right]^{\beta}\right)} \right]^{\theta} = u, \tag{11}$$

$0 \leq u \leq 1$, that is,

$$Q(u) = G^{-1}\left[1-\left(1-\left[1-\left(\theta\left[(1-u)^{\frac{-1}{\theta}}-(1-\theta)\right]^{-1}\right)^{\frac{1}{\beta}}\right]^{\frac{1}{b}}\right)^{\frac{1}{2}}\right]. \tag{12}$$

Therefore, random numbers from the HT-GEN-TL-G family of distributions can be obtained using Eqn. (12), when the baseline G is specified.

2.3. Density Expansion

In this sub-section, we express the pdf of the HT-GEN-TL-G family of distributions as an infinite linear combination of exponentiated-G (Exp-G) densities. Utilizing the generalized binomial series expansions, we can express the pdf of the HT-GEN-TL-G family of distributions as

$$f_{HT-GEN-TL-G}(x;\theta,b,\beta,\varphi) = \sum_{k=0}^{\infty} d_{k+1}^{*} g_{k+1}^{*}(x;\varphi), \tag{13}$$

where

$$\begin{aligned} d_{k+1}^{*} &= 2\beta b\theta^{2}\sum_{q,s,j,t=0}^{\infty}\frac{(1-\theta)^{q}(-1)^{q+s+j+t+k}}{(k+1)}\binom{-(\theta+1)}{q}\binom{q}{s}\binom{\beta s+\beta\theta-1}{j} \\ &\times \binom{\alpha(j+1)-1}{t}\binom{2t+1}{k}, \end{aligned} \tag{14}$$

and $g_{k+1}^{*}(x;\varphi) = (k+1)g(x;\varphi)G^{k}(x;\varphi)$ is the pdf of Exp-G distribution with power parameter $(k+1)$. **See appendix for details.**

Consequently, we can use the tractability property to obtain the statistical properties of the HT-GEN-TL-G family of distributions from those of the Exp-G family of distributions.

3. Statistical Properties

In the following sub-sections, we study some statistical properties of the HT-GEN-TL-G family of distributions including moments, moments generating function, incomplete moments, distribution of order statistics, Rényi entropy and parameter estimation.

3.1. Moments, Generating Function and Incomplete Moments

Let Y_{k+1} denote the Exp-G random variable with power parameter $(k+1)$. Then, the i^{th} raw moment of X, say μ_i', can be obtained directly from Eqn. (13) as

$$\mu_i' = E(X^i) = \sum_{k=0}^{\infty} d_{k+1}^{*} E(Y_{k+1}^{i}), \tag{15}$$

where $E(Y_{k+1}^{i})$ is the i^{th} raw moment of Exp-G distribution. The moment generating function (mgf) $M_X(t) = E(e^{tX})$ of the HT-GEN-TL-G family of distributions can be derived from Eqn. (13) as

$$M_X(t) = \sum_{k=0}^{\infty} d_{k+1}^{*} M_{k+1}(t),$$

where d_{k+1}^{*} is as given in Eqn. (14) and $M_{k+1}(t)$ is the mgf of Y_{k+1}.

The i^{th} incomplete moment of the HT-GEN-TL-G family of distributions is given by

$$m_i(t) = \int_{-\infty}^{t} x^i f(x;\theta,\alpha,\beta,\varphi)dx = \sum_{k=0}^{\infty} d_{k+1}^{*} \int_{-\infty}^{t} x^i g_{k+1}^{*}(x;\varphi)dx, \tag{16}$$

where d_{k+1}^{*} is as given in Eqn. (14) and $g_{k+1}^{*}(x;\varphi) = (k+1)g(x;\varphi)G^k(x;\varphi)$ is the pdf of Exp-G distribution with power parameter $(k+1)$. A general equation for $m_1(t)$ can be obtained from Eqn. (16) as

$$m_1(t) = \sum_{k=0}^{\infty} d_{k+1}^{*} D_{k+1}(t),$$

where $D_{k+1}(t) = \int_{-\infty}^{t} x g_{k+1}^{*}(x;\varphi)dx$ is the first incomplete moment of the Exp-G distribution.

3.2. Distribution of Order Statistics

Order statistics play an important role in probability and statistics. They find applications in reliability theory and quality control testing in engineering. By considering independent and identically distributed random variables $X_1, X_2, ..., X_n$ from the HT-GEN-TL-G family of distributions, the pdf of the i^{th} order statistic from the HT-GEN-TL-G family of distributions can be written as

$$f_{i:n}(x;\theta,\alpha,\beta,\varphi) = \sum_{k=0}^{\infty} p_{k+1}^{*} g_{k+1}^{*}(x;\varphi), \tag{17}$$

where

$$\begin{aligned} p_{k+1}^{*} &= 2\beta b\theta^2 \sum_{l=0}^{n-i} \sum_{e,q,s,j,t=0}^{\infty} \frac{n!(1-\theta)^q(-1)^{q+s+j+t+k+l+e}}{(n-i)!(i-1)!(k+1)} \binom{n-i}{l} \binom{i+l-1}{e} \\ &\times \binom{-(\theta(e+1)+1)}{q} \binom{q}{s} \binom{\beta s + \beta\theta(e+1)-1}{j} \binom{\alpha(j+1)-1}{t} \binom{2t+1}{k}, \end{aligned} \tag{18}$$

and $g_{k+1}^{*}(x;\varphi) = (k+1)g(x;\varphi)G^k(x;\varphi)$ is the pdf of Exp-G distribution with power parameter $(k+1)$. **See appendix for details.**

3.3. Rényi Entropy

Rényi entropy is a measure of randomness or uncertainty in the system. It is mostly used in information theory. Rényi entropy is defined to be

$$I_R(v) = (1-v)^{-1} log \left[\int_{-\infty}^{\infty} f^v(x)dx \right],$$

where $v > 0$ and $v \neq 1$. Rényi entropy for the HT-GEN-TL-G family of distributions can be obtained as follows:

$$\begin{aligned} I_R(v) &= (1-v)^{-1} \log \Bigg[(2b\beta\theta^2)^v \sum_{q,s,j,t,k=0}^{\infty} \frac{(1-\theta)^q(-1)^{q+s+j+t+k}}{(k+1)} \binom{-v(\theta+1)}{q} \binom{q}{s} \\ &\times \binom{\beta s + v(\beta\theta-1)}{j} \binom{\alpha(j+v)-v}{t} \binom{2t+v}{k} \frac{1}{\left[1+\frac{k}{v}\right]^v} \\ &\times \int_{-\infty}^{\infty} \left(\left[1+\frac{k}{v}\right] (G(x;\varphi))^{\frac{k}{v}} g(x;\varphi) \right)^v \Bigg] \\ &= \frac{1}{1-v} log \left[\sum_{k=0}^{\infty} r_k^{*} e^{(1-v)I_{REG}} \right], \end{aligned} \tag{19}$$

for $v>0$, $v\neq 1$, where $I_{REG}=\frac{1}{1-v}\log\left[\int_0^\infty\left(\left[1+\frac{k}{v}\right](G(x;\varphi))^{\frac{k}{v}}(g(x;\varphi))\right)^v dx\right]$ is the Rényi entropy of Exp-G distribution with power parameter $(\frac{k}{v}+1)$, and

$$\begin{aligned} r_k^* &= (2\alpha\beta\theta^2)^v\sum_{q,s,j,t=0}^{\infty}\frac{(1-\theta)^q(-1)^{q+s+j+t+k}}{k+1}\binom{-v(\theta+1)}{q}\binom{q}{s}\binom{\beta s+v(\beta\theta-1)}{j} \\ &\times \binom{-(\alpha(j+v)+v)}{t}\binom{2t+v}{k}\frac{1}{\left[1+\frac{k}{v}\right]^v}. \end{aligned} \tag{20}$$

See appendix for details.

4. Some Special Cases

In this section, we introduce three special cases of the HT-GEN-TL-G family of distributions by generalizing the classical distributions, namely, log-logistic, Weibull, and standard half-logistic distributions.

4.1. Heavy-Tailed Generalized Topp-Leone-Log Logistic (HT-GEN-TL-LLoG) Distribution

By considering the log-logistic distribution with the cdf and pdf given by $G(x;c)=1-(1+x^c)^{-1}$ and $g(x;c)=cx^{c-1}(1+x^c)^{-2}$ for $c>0$ and $x>0$, then the new HT-GEN-TL-LLoG distribution has cdf and pdf given by

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

and

$$\begin{aligned} f(x;\theta,b,\beta,c) &= 2b\beta\theta^2cx^{c-1}(1+x^c)^{-2}\left(1-(1+x^c)^{-2}\right)^{b-1}(1+x^c)^{-1} \\ &\times \left(\left[1-\left(1-(1+x^c)^{-2}\right)^b\right]\right)^{\beta\theta-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-(1+x^c)^{-2}\right)^b\right]^\beta\right)\right]^{-(\theta+1)}, \end{aligned}$$

respectively, for $\theta,b,\beta,c>0$. The hrf is given by

$$\begin{aligned} h(x;\theta,b,\beta,c) &= 2b\beta\theta^2cx^{c-1}(1+x^c)^{-2}\left(1-(1+x^c)^{-2}\right)^{b-1}(1+x^c)^{-1} \\ &\times \left(\left[1-\left(1-(1+x^c)^{-2}\right)^b\right]\right)^{-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-(1+x^c)^{-2}\right)^b\right]^\beta\right)\right]^{-1}, \end{aligned}$$

for $\theta,b,\beta,c>0$.

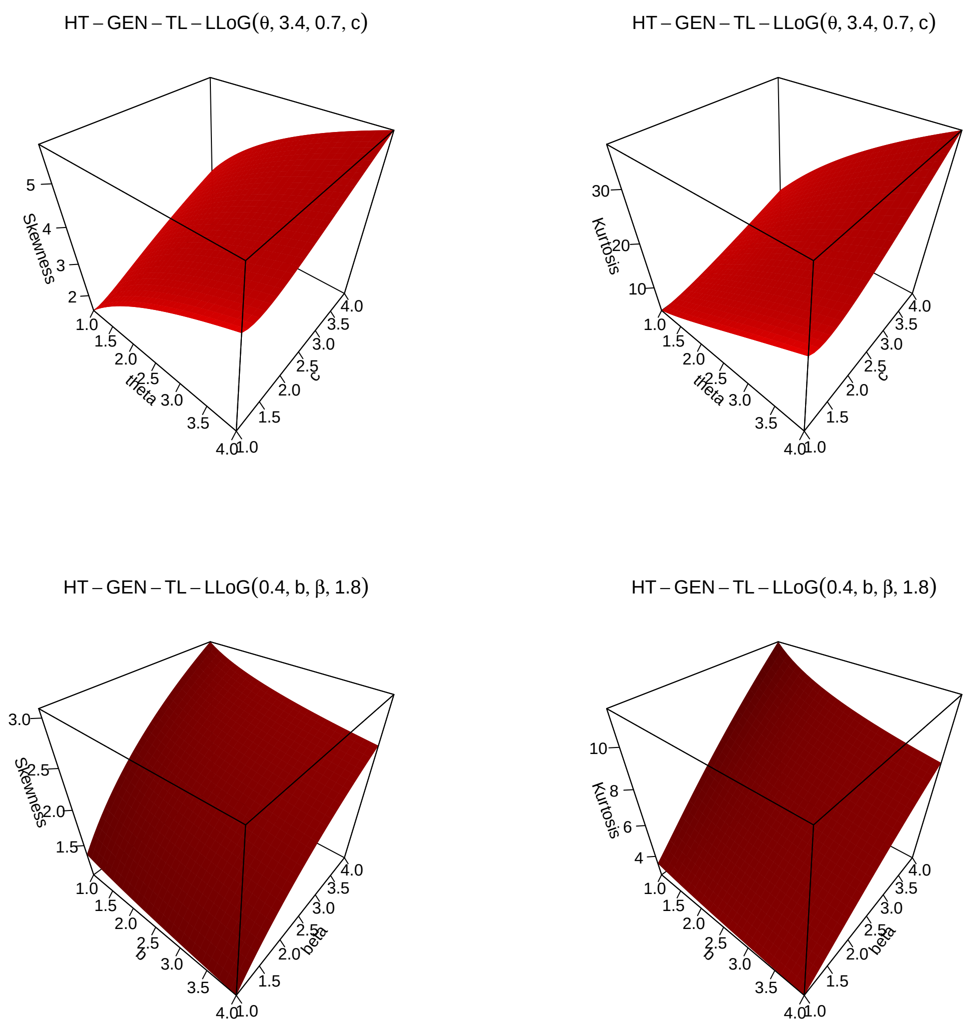

Figure 1: 3D-Plots of the skewness and kurtosis for HT-GEN-TL-LLoG distribution

Figure 1 shows plots of skewness and kurtosis for the HT-GEN-TL-LLoG distribution. We can see that for fixed values of b and β, skewness and kurtosis increase when θ and c increases. On another note, when we fix θ and c, there is an increase in skewness and kurtosis when b and β increases.

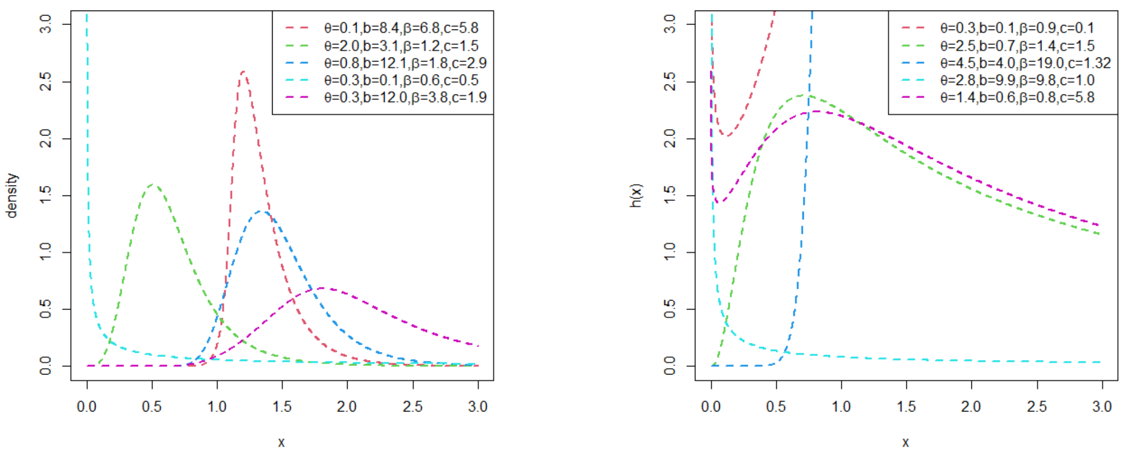

Figure 2: Plots of the pdf and hrf of the HT-GEN-TL-LLoG distribution

Figure 2 shows several plots of the pdf and the hrf of the HT-GEN-TL-LLoG distribution. The pdf can take several shapes including right-skewed, left-skewed, unimodal, and reverse-J shapes. It can be noted that the hrf of the HT-GEN-TL-LLoG distribution can be increasing, decreasing, bathtub, upside-down bathtub and bathtub followed by upside-down bathtub shapes.

4.2. Heavy-Tailed Generalized Topp-Leone-Weibull (HT-GEN-TL-W) Distribution

If we consider the Weibull distribution with cdf and pdf given by $G(x;\lambda) = 1-\exp(-x^{\lambda})$ and $g(x;\lambda) = \lambda x^{\lambda-1}\exp(-x^{\lambda})$, respectively, for $\lambda > 0$ and $x > 0$, as the baseline distribution, then we obtain the HT-GEN-TL-W distribution with cdf and pdf given by

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

and

$$\begin{aligned} f(x;\theta,b,\beta,\lambda) &= 2b\beta\theta^{2}\lambda x^{\lambda-1}\exp(-x^{\lambda})\left(1-\exp(-2x^{\lambda})\right)^{b-1}\exp(-x^{\lambda}) \\ &\times \left(\left[1-\left(1-\exp(-2x^{\lambda})\right)^{b}\right]\right)^{\beta\theta-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-\exp(-2x^{\lambda})\right)^{b}\right]^{\beta}\right)\right]^{-(\theta+1)}, \end{aligned}$$

respectively, for $\theta, b, \beta, \lambda > 0$. The hrf is given by

$$\begin{aligned} h(x;\theta,b,\beta,\lambda) &= 2b\beta\theta^{2}\lambda x^{\lambda-1}\exp(-x^{\lambda})\left(1-\exp(-2x^{\lambda})\right)^{b-1}\exp(-x^{\lambda}) \\ &\times \left(\left[1-\left(1-\exp(-2x^{\lambda})\right)^{b}\right]\right)^{-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-\exp(-2x^{\lambda})\right)^{b}\right]^{\beta}\right)\right]^{-1}, \end{aligned}$$

for $\theta, b, \beta, \lambda > 0$.

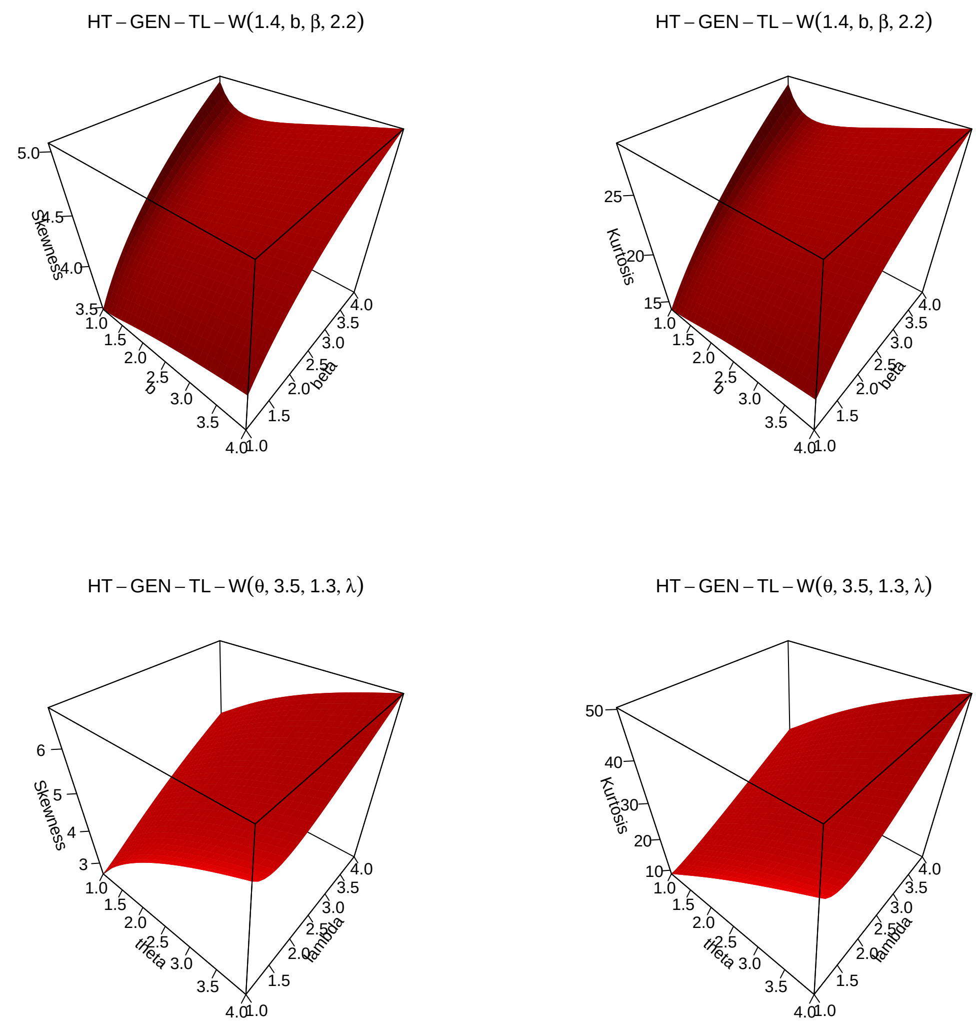

Figure 3: 3D-Plots of the skewness and kurtosis for HT-GEN-TL-W distribution

Figure 3 shows plots of skewness and kurtosis for the HT-GEN-TL-W distribution. We can see that for fixed values of θ and λ, skewness and kurtosis increase when b and β increases. On another note, when we fix b and β, there is an increase in skewness and kurtosis when θ and λ increases.

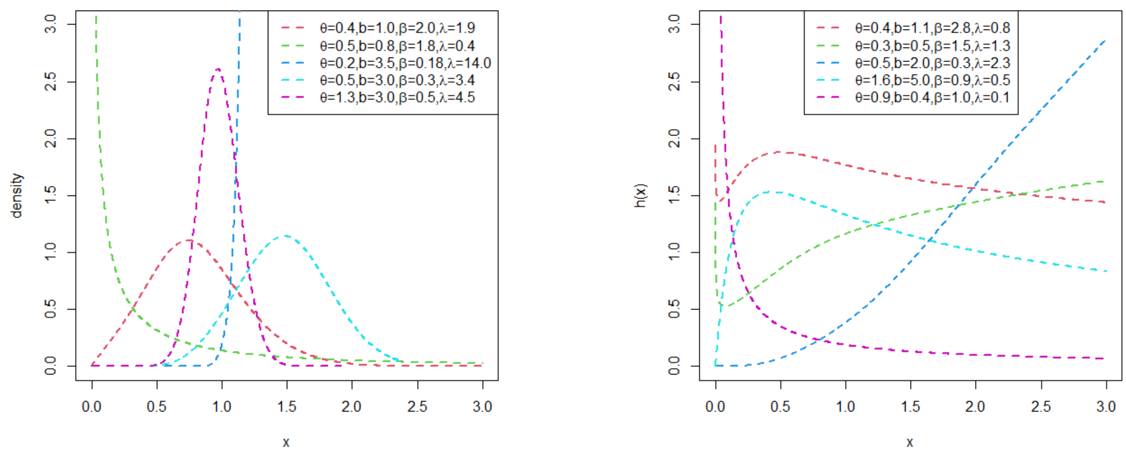

Figure 4: Plots of the pdf and hrf of the HT-GEN-TL-W distribution

Figure 4 shows the plots of the pdf and the hrf of the HT-GEN-TL-W distribution. The pdf of the HT-GEN-TL-W distribution can take several shapes including right-skewed, left-skewed, almost symmetric, J and reverse-J shapes. The hrf of the HT-GEN-TL-W distribution display increasing, decreasing, bathtub, upside-down bathtub, and bathtub followed by upside-down bathtub shapes.

4.3. Heavy-Tailed Generalized Topp-Leone-Standard Half Logistic (HT-GEN-TL-SHL) Distribution

Let the baseline distribution be standard half logistic distribution with pdf and cdf given by $g(x) = \frac{2\exp(-x)}{(1+\exp(-x))^2}$ and $G(x) = \frac{1-\exp(-x)}{1+\exp(-x)}$, for $x > 0$, respectively. Then, the cdf and pdf of HT-GEN-TL-SHL distribution are given by

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

and

$$\begin{aligned} f(x;\theta,b,\beta) &= 2b\beta\theta^2 \frac{2\exp(-x)}{(1+\exp(-x))^2}\left(1-\left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)^2\right)^{b-1} \\ &\times \left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)\left(\left[1-\left(1-\left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)^2\right)^b\right]\right)^{\beta\theta-1} \\ &\times \left[1-(1-\theta)\left(1-\left[1-\left(1-\left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)^2\right)^b\right]^\beta\right)\right]^{-(\theta+1)}, \end{aligned}$$

respectively, for $\theta, b, \beta > 0$. The hrf is given by

 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.4458

$$
\begin{aligned}
h(x;\theta,b,\beta) &= 2b\beta\theta^2\frac{2\exp(-x)}{(1+\exp(-x))^2}\left(1-\left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)^2\right)^{b-1} \\
&\times \left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)\left(\left[1-\left(1-\left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)^2\right)^b\right]\right)^{-1} \\
&\times \left[1-(1-\theta)\left(1-\left[1-\left(1-\left(1-\frac{1-\exp(-x)}{1+\exp(-x)}\right)^2\right)^b\right]^\beta\right)\right]^{-1},
\end{aligned}
$$

for $\theta,b,\beta>0$.

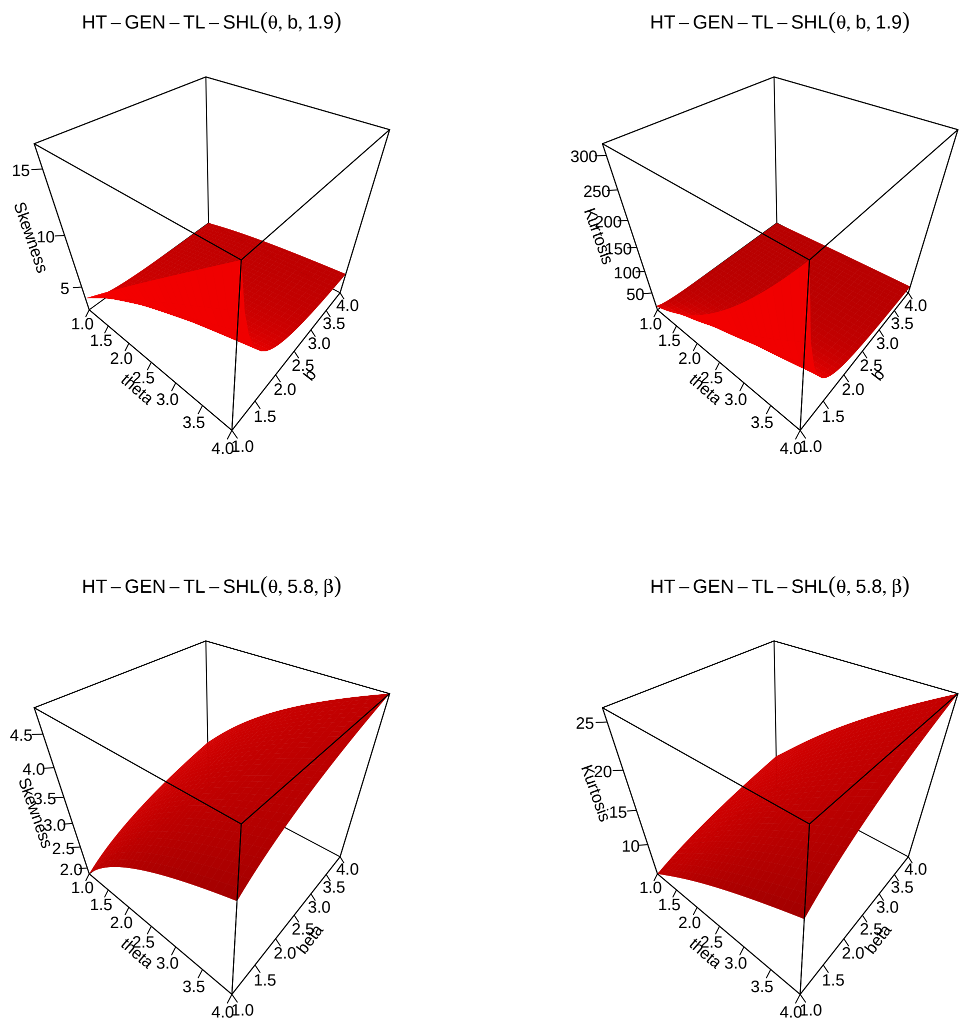

Figure 5: 3D-Plots of the skewness and kurtosis for HT-GEN-TL-SHL distribution

Figure 5 shows plots of skewness and kurtosis for the HT-GEN-TL-SHL distribution. We can see that for fixed value of β, skewness is positive (right skewed) and kurtosis is leptokurtic when θ and b increases. On another note, when we fix b, there is an increase in skewness and kurtosis when θ and β increases.

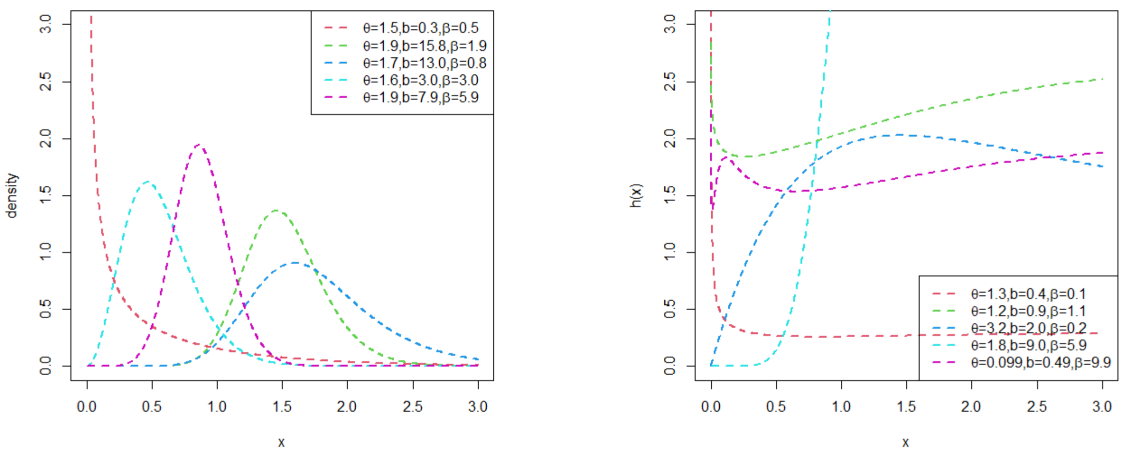

Figure 6: Plots of the pdf and hrf of the HT-GEN-TL-SHL distribution

Figure 6 shows the plots of the pdf and the hrf of the HT-GEN-TL-SHL distribution for different parameter values. The pdf can take several shapes including almost symmetric, right-skewed, left-skewed, unimodal, and reverse-J shapes. The hrf of the HT-GEN-TL-SHL distribution displays increasing, decreasing, bathtub, upside-down bathtub, and upside-down bathtub followed by bathtub shapes.

5. Different Methods of Estimation

In this section, different estimation methods including Maximum Likelihood, Anderson-Darling, Ordinary Least Squares, Weighted Least Squares, Cramér-von Mises and Maximum Product of Spacing are utilized to estimate the unknown parameters of the new family of distributions.

5.1. Maximum Likelihood Estimation

In this sub-section, we employ the maximum likelihood estimation technique to estimate the unknown parameters of the HT-GEN-TL-G family of distributions. By assuming $x_1, x_2, ..., x_n$ to be the sample of size n, obtained from the HT-GEN-TL-G family of distributions, the log-likelihood function $\ell_n = \ell_n(\Delta)$ for $\Delta = (\theta, b, \beta, \varphi)$ has the form:

$$\begin{aligned}
\ell_n(\Delta) &= (n)\ln(2b\beta\theta^2) + \sum_{i=1}^{n}\ln(g(x_i;\varphi)) + (b-1)\sum_{i=1}^{n}\ln\left(1-\overline{G}^2(x_i;\varphi)\right) \\
&+ (\beta\theta-1)\sum_{i=1}^{n}\ln\left[1-\left(1-\overline{G}^2(x_i;\varphi)\right)^b\right] \\
&- (\theta+1)\sum_{i=1}^{n}\ln\left[1-(1-\theta)\left(1-\left[1-\left(1-\overline{G}^2(x_i;\varphi)\right)^b\right]^\beta\right)\right].
\end{aligned}$$

In order to obtain the maximum likelihood (ML) estimates of the parameters $\theta, b, \beta, \varphi$, denoted by $\hat{\theta}, \hat{b}, \hat{\beta}, \hat{\varphi}$, we set the nonlinear system of equations
$(\frac{\partial \ell_n}{\partial \theta}, \frac{\partial \ell_n}{\partial b}, \frac{\partial \ell_n}{\partial \beta}, \frac{\partial \ell_n}{\partial \varphi_k})^T = \mathbf{0}$, and solve them simultaneously. It is clear that these equations are not in closed form and hence cannot be solved analytically. Hence, the ML estimates can be found by maximizing $\ell_n(\Delta)$ numerically with respect to the parameters, using a numerical method such as Newton-Raphson procedure. The partial derivatives of the log-likelihood function with respect to each component of the parameter vector are given in the appendix.

5.2. Anderson-Darling Estimation (ADE)

Suppose $x_{(1)}, x_{(2)}, ..., x_{(n)}$ are the order statistics of a random sample of size n from the HT-GEN-TL-G family of distributions. Then, the Anderson-Darling estimates (ADEs) of the HT-GEN-TL-G family of distributions are obtained by minimizing the following function

$$A(\theta, b, \beta, \varphi) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\log\left(F(x_{(i)};\theta,b,\beta,\varphi)\right) + \log\left(S(x_{(i)};\theta,b,\beta,\varphi)\right)\right],$$

where $F(x_{(i)};\theta,b,\beta,\varphi)$ and $S(x_{(i)};\theta,b,\beta,\varphi)$ be the cdf and sf of the i^{th} order statistic from the HT-GEN-TL-G family of distributions.
The ADEs can also be derived by solving the non-linear equations:

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right)}{F(x_{(i)};\theta,b,\beta,\varphi)} - \frac{\vartheta_z\left(x_{(n+1-i)};\theta,b,\beta,\varphi\right)}{S\left(x_{(n+1-i)};\theta,b,\beta,\varphi\right)}\right] = 0, z = 1,2,3,4, \tag{21}$$

where

$$\begin{aligned}
\vartheta_1\left(x_{(i)};\theta,b,\beta,\varphi\right) &= \frac{\partial F\left(x_{(i)};\theta,b,\beta,\varphi\right)}{\partial\theta}, \\
\vartheta_2\left(x_{(i)};\theta,b,\beta,\varphi\right) &= \frac{\partial F\left(x_{(i)};\theta,b,\beta,\varphi\right)}{\partial b}, \\
\vartheta_3\left(x_{(i)};\theta,b,\beta,\varphi\right) &= \frac{\partial F\left(x_{(i)};\theta,b,\beta,\varphi\right)}{\partial\beta},
\end{aligned}$$

and

$$\vartheta_4\left(x_{(i)};\theta,b,\beta,\varphi\right) = \frac{\partial F\left(x_{(i)};\theta,b,\beta,\varphi\right)}{\partial\varphi_k}, \tag{22}$$

for $k = 1,2,....,t$ and t is number of components of the baseline parameter vector φ.

5.3. Ordinary Least Squares (OLS)

The OLS estimates (OLSEs) of the parameters of the HT-GEN-TL-G family of distributions are obtained by minimizing the function:

$$V(\theta,b,\beta,\varphi) = \sum_{i=1}^{n}\left[F\left(x_{(i)};\theta,b,\beta,\varphi\right) - \frac{i}{n+1}\right]^2.$$

The OLSEs can be obtained by solving the non-linear equations

$$\sum_{i=1}^{n}\left[F\left(x_{(i)};\theta,b,\beta,\varphi\right) - \frac{i}{n+1}\right]\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right) = 0, z = 1,2,3,4,$$

where $\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right)$ are defined in Eqn. (22).
These non-linear equations can be solved using a numerical method such as Newton-Raphson procedure.

5.4. Weighted Least Squares (WLS)

The WLS estimates (WLSEs) of the parameters of the HT-GEN-TL-G family of distributions are obtained by minimizing the function

$$W(\theta,b,\beta,\varphi) = \sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-1+1)}\left[F\left(x_{(i)};\theta,b,\beta,\varphi\right) - \frac{i}{n+1}\right]^2,$$

with respect to θ, b, β, and parameter vector φ. The WLSEs can be obtained by solving the non-linear equations:

$$\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-1+1)} \left[F\left(x_{(i)};\theta,b,\beta,\varphi\right) - \frac{i}{n+1}\right] \vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right) = 0, z = 1,2,3,4,$$

where $\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right)$ are defined in equation (22).

5.5. Cramér-von Mises (CVM)

The CVM estimates (CVMEs) of the parameters of the HT-GEN-TL-G family of distributions are obtained through the minimization of the function:

$$C(\theta,b,\beta,\varphi) = -\frac{1}{12n}\sum_{i=1}^{n}\left[F\left(x_{(i)};\theta,b,\beta,\varphi\right) - \frac{2i-1}{2n}\right]^2,$$

with respect to θ, b, β, and parameter vector φ. The CVMEs can also be obtained by solving the non-linear equations

$$\sum_{i=1}^{n}\left[F\left(x_{(i)};\theta,b,\beta,\varphi\right) - \frac{2i-1}{2n}\right] \vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right) = 0, z = 1,2,3,4,$$

where $\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right)$ are defined in Eqn. (22) and $x_{(i)}$ is the i^{th} order statistic from the HT-GEN-TL-G family of distributions.

5.6. Maximum Product of Spacing (MPS)

The MPS method is used to estimate the parameters of a distribution as an alternative to the maximum likelihood method. Let $D_i\left(x_{(i)};\theta,b,\beta,\varphi\right) = F\left(x_{(i)};\theta,b,\beta,\varphi\right) - F\left(x_{(i-1)};\theta,b,\beta,\varphi\right)$, for $i = 1,2,..,n+1$, be the uniform spacing of a random sample from the HT-GEN-TL-G family of distributions, where $F\left(x_{(0)};\theta,b,\beta,\varphi\right) = 0$, $F\left(x_{(n+1)};\theta,b,\beta,\varphi\right) = 1$ and $\sum_{i=1}^{n+1} D_i\left(x_{(i)};\theta,b,\beta,\varphi\right) = 1$. The MPS estimates (MPSEs) for θ, b, β, and parameter vector φ can be obtained by maximizing

$$H(\theta,b,\beta,\varphi) = \frac{1}{n+1}\sum_{i=1}^{n+1}\log\left(D_i\left(x_{(i)};\theta,b,\beta,\varphi\right)\right).$$

The MPSEs of the HT-GEN-TL-G family of distributions can be obtained by solving the non-linear equations defined by

$$\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i\left(x_{(i)};\theta,b,\beta,\varphi\right)}\left[\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right) - \vartheta_z\left(x_{(i-1)};\theta,b,\beta,\varphi\right)\right] = 0, z = 1,2,3,4,$$

where $\vartheta_z\left(x_{(i)};\theta,b,\beta,\varphi\right)$ are defined in Eqn. (22).

6. Simulation

In this section, we carry out a Monte Carlo simulation study to assess the performance of parameter estimates of the HT-GEN-TL-LLoG distribution as a special case of the HT-GEN-TL-G family of distributions by utilizing the six estimation methods discussed in Section 5. A simulation study was carried out by generating $N = 3000$ random samples from the HT-GEN-TL-LLoG distribution for various sample sizes of $n = 25, 50, 100, 200, 400$, and 800.

To assess performance of the different estimation methods, we used the statistics: mean estimate (Mean), average bias (ABias) and root mean square error (RMSE). The ABias and RMSE for the estimated parameter, say, $\hat{\lambda}$, are given by:

$$ABias(\hat{\lambda}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\lambda}_i - \lambda), \quad \text{and} \quad RMSE(\hat{\lambda}) = \sqrt{\frac{\sum_{i=1}^{N}(\hat{\lambda}_i - \lambda)^2}{N}},$$

 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.4458

respectively.

Table 1: Simulation Results for Different Estimation Methods for $\theta = 1.3, b = 0.6, \beta = 0.2, c = 2.5$.

		MLE		LS		WLS		RADE		CVM		ADE	
n	Parameter	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE
25	θ	$1.8929^{\{1\}}$	$3.4209^{\{1\}}$	$5.8129^{\{4\}}$	$6.1402^{\{3\}}$	$4.7059^{\{2\}}$	$5.5121^{\{2\}}$	$5.6060^{\{3\}}$	$6.3409^{\{5\}}$	$5.8754^{\{5\}}$	$6.3236^{\{4\}}$	$6.0243^{\{6\}}$	$6.3741^{\{6\}}$
	b	$-0.0606^{\{1\}}$	$0.2404^{\{1\}}$	$-0.2822^{\{3\}}$	$0.8868^{\{4\}}$	$0.5666^{\{6\}}$	$1.4199^{\{6\}}$	$-0.5070^{\{5\}}$	$0.8607^{\{3\}}$	$-0.3127^{\{4\}}$	$0.9382^{\{5\}}$	$-0.2359^{\{2\}}$	$0.5698^{\{2\}}$
	β	$0.7444^{\{1\}}$	$2.0754^{\{1\}}$	$6.1285^{\{4\}}$	$7.006^{\{2\}}$	$5.7572^{\{2\}}$	$14.6241^{\{5\}}$	$10.1985^{\{6\}}$	$26.1081^{\{6\}}$	$6.1167^{\{3\}}$	$9.2257^{\{4\}}$	$8.3304^{\{5\}}$	$9.0433^{\{3\}}$
	c	$-0.5132^{\{1\}}$	$1.3313^{\{1\}}$	$-1.5440^{\{5\}}$	$2.0466^{\{4\}}$	$-0.7884^{\{2\}}$	$1.7141^{\{2\}}$	$-1.0162^{\{3\}}$	$2.1429^{\{5\}}$	$-1.4314^{\{4\}}$	$1.9827^{\{3\}}$	$-2.1333^{\{6\}}$	$2.3338^{\{6\}}$
	$\sum$ ranks	8		29		27		36		32		36	
50	θ	$1.2754^{\{1\}}$	$2.4925^{\{1\}}$	$5.7460^{\{5\}}$	$6.5845^{\{6\}}$	$3.6290^{\{2\}}$	$4.3465^{\{2\}}$	$5.5492^{\{4\}}$	$6.1934^{\{4\}}$	$5.3671^{\{3\}}$	$5.8272^{\{3\}}$	$5.8292^{\{6\}}$	$6.2139^{\{5\}}$
	b	$-0.0450^{\{1\}}$	$0.1984^{\{1\}}$	$-0.2264^{\{3\}}$	$0.9144^{\{4\}}$	$1.4081^{\{6\}}$	$1.8417^{\{6\}}$	$-0.4995^{\{5\}}$	$0.8580^{\{3\}}$	$-0.124^{\{2\}}$	$0.9184^{\{5\}}$	$-0.2551^{\{4\}}$	$0.5996^{\{2\}}$
	β	$0.1871^{\{1\}}$	$1.5415^{\{1\}}$	$5.5371^{\{4\}}$	$6.9009^{\{4\}}$	$3.9573^{\{2\}}$	$6.4258^{\{3\}}$	$7.4304^{\{6\}}$	$20.3083^{\{6\}}$	$5.3542^{\{3\}}$	$6.1818^{\{2\}}$	$6.154^{\{5\}}$	$7.9833^{\{5\}}$
	c	$-0.4507^{\{3\}}$	$1.0463^{\{1\}}$	$-1.3279^{\{5\}}$	$1.9070^{\{4\}}$	$-0.0447^{\{1\}}$	$1.4001^{\{2\}}$	$-0.1274^{\{2\}}$	$2.4341^{\{6\}}$	$-1.1964^{\{4\}}$	$1.8264^{\{3\}}$	$-1.988^{\{6\}}$	$2.2524^{\{5\}}$
	$\sum$ ranks	10		35		24		36		25		38	
100	θ	$0.1535^{\{1\}}$	$0.6467^{\{1\}}$	$5.0360^{\{4\}}$	$5.4824^{\{4\}}$	$2.8733^{\{2\}}$	$3.3958^{\{2\}}$	$5.7322^{\{6\}}$	$6.7767^{\{6\}}$	$4.9763^{\{3\}}$	$5.4662^{\{3\}}$	$5.2826^{\{5\}}$	$5.9648^{\{5\}}$
	b	$-0.0375^{\{2\}}$	$0.1598^{\{1\}}$	$0.0129^{\{1\}}$	$0.9737^{\{5\}}$	$1.9352^{\{6\}}$	$2.1994^{\{6\}}$	$-0.5431^{\{5\}}$	$0.8965^{\{3\}}$	$0.0696^{\{3\}}$	$0.9661^{\{4\}}$	$-0.2369^{\{4\}}$	$0.5457^{\{2\}}$
	β	$0.0495^{\{1\}}$	$0.1793^{\{1\}}$	$5.4199^{\{5\}}$	$7.9564^{\{5\}}$	$3.1016^{\{2\}}$	$3.3733^{\{2\}}$	$5.8465^{\{6\}}$	$20.6408^{\{6\}}$	$5.13163^{\{4\}}$	$6.1979^{\{4\}}$	$4.8514^{\{3\}}$	$5.7329^{\{3\}}$
	c	$-0.3723^{\{1\}}$	$0.8168^{\{1\}}$	$-0.9038^{\{5\}}$	$1.6286^{\{4\}}$	$0.5344^{\{2\}}$	$1.1230^{\{2\}}$	$0.8934^{\{4\}}$	$2.9671^{\{6\}}$	$-0.8579^{\{3\}}$	$1.5920^{\{3\}}$	$-1.7752^{\{6\}}$	$2.1206^{\{5\}}$
	$\sum$ ranks	9		33		24		42		27		33	
200	θ	$0.1282^{\{1\}}$	$0.5483^{\{1\}}$	$4.2532^{\{3\}}$	$4.8435^{\{3\}}$	$2.3675^{\{2\}}$	$2.5788^{\{2\}}$	$5.7459^{\{6\}}$	$7.0956^{\{6\}}$	$4.3123^{\{4\}}$	$4.8635^{\{4\}}$	$4.3176^{\{5\}}$	$5.3865^{\{5\}}$
	b	$-0.0324^{\{1\}}$	$0.1351^{\{1\}}$	$0.2082^{\{3\}}$	$0.9967^{\{3\}}$	$2.2779^{\{6\}}$	$2.4007^{\{6\}}$	$-0.6794^{\{5\}}$	$1.0031^{\{4\}}$	$0.1573^{\{2\}}$	$1.0262^{\{5\}}$	$-0.2233^{\{4\}}$	$0.5175^{\{2\}}$
	β	$0.0229^{\{1\}}$	$0.0955^{\{1\}}$	$4.6951^{\{5\}}$	$6.7472^{\{4\}}$	$2.9332^{\{2\}}$	$3.2191^{\{2\}}$	$4.5372^{\{4\}}$	$14.8686^{\{6\}}$	$5.0409^{\{6\}}$	$7.2590^{\{5\}}$	$3.9008^{\{3\}}$	$5.1276^{\{3\}}$
	c	$-0.2602^{\{1\}}$	$0.5907^{\{1\}}$	$-0.5596^{\{2\}}$	$1.3001^{\{3\}}$	$0.8176^{\{4\}}$	$0.9819^{\{2\}}$	$1.4661^{\{6\}}$	$3.2940^{\{6\}}$	$-0.5627^{\{3\}}$	$1.3085^{\{4\}}$	$-1.4369^{\{5\}}$	$1.9011^{\{5\}}$
	$\sum$ ranks	8		26		26		43		33		32	
400	θ	$0.0964^{\{1\}}$	$0.3873^{\{1\}}$	$4.0889^{\{5\}}$	$4.7368^{\{5\}}$	$2.2305^{\{2\}}$	$2.3290^{\{2\}}$	$6.2889^{\{6\}}$	$8.2140^{\{6\}}$	$3.9656^{\{4\}}$	$4.5828^{\{3\}}$	$3.0360^{\{3\}}$	$4.6208^{\{4\}}$
	b	$-0.0314^{\{1\}}$	$0.1090^{\{1\}}$	$0.0427^{\{3\}}$	$0.8905^{\{4\}}$	$2.3963^{\{6\}}$	$2.4463^{\{6\}}$	$-0.6281^{\{5\}}$	$0.9812^{\{5\}}$	$0.0315^{\{2\}}$	$0.12053^{\{2\}}$	$-0.2042^{\{4\}}$	$0.4647^{\{3\}}$
	β	$0.0106^{\{1\}}$	$0.0617^{\{1\}}$	$4.7787^{\{6\}}$	$5.5122^{\{5\}}$	$2.8312^{\{2\}}$	$2.8953^{\{2\}}$	$3.7496^{\{4\}}$	$6.6658^{\{6\}}$	$4.6406^{\{5\}}$	$5.2288^{\{4\}}$	$3.4285^{\{3\}}$	$4.510^{\{3\}}$
	c	$-0.1735^{\{1\}}$	$0.4213^{\{1\}}$	$-0.7597^{\{3\}}$	$1.3502^{\{4\}}$	$0.89915^{\{5\}}$	$0.9386^{\{2\}}$	$0.4399^{\{2\}}$	$3.0748^{\{6\}}$	$-0.7731^{\{4\}}$	$1.3424^{\{3\}}$	$-0.9750^{\{6\}}$	$1.5494^{\{5\}}$
	$\sum$ ranks	8		35		27		40		27		31	
800	θ	$0.0609^{\{1\}}$	$0.306^{\{1\}}$	$4.0305^{\{4\}}$	$4.6910^{\{5\}}$	$2.1874^{\{3\}}$	$2.2268^{\{2\}}$	$6.3672^{\{6\}}$	$7.4557^{\{6\}}$	$4.0504^{\{5\}}$	$4.6115^{\{4\}}$	$1.3464^{\{2\}}$	$3.0889^{\{3\}}$
	b	$-0.0267^{\{3\}}$	$0.0849^{\{1\}}$	$-0.0002^{\{1\}}$	$0.805^{\{4\}}$	$2.4303^{\{6\}}$	$2.4577^{\{6\}}$	$-0.6119^{\{5\}}$	$0.9876^{\{5\}}$	$0.0049^{\{2\}}$	$0.7810^{\{3\}}$	$-0.1391^{\{4\}}$	$0.3218^{\{2\}}$
	β	$0.0081^{\{1\}}$	$0.0455^{\{1\}}$	$4.6279^{\{5\}}$	$4.9633^{\{4\}}$	$2.7934^{\{3\}}$	$2.8268^{\{3\}}$	$3.5355^{\{4\}}$	$6.2253^{\{6\}}$	$4.6293^{\{6\}}$	$5.0128^{\{5\}}$	$1.0030^{\{2\}}$	$2.6572^{\{2\}}$
	c	$-0.0998^{\{1\}}$	$0.2756^{\{1\}}$	$-0.8610^{\{4\}}$	$1.3844^{\{5\}}$	$0.9167^{\{5\}}$	$0.9329^{\{2\}}$	$-1.21667^{\{6\}}$	$2.2528^{\{6\}}$	$-0.8579^{\{3\}}$	$1.3601^{\{4\}}$	$-0.4409^{\{2\}}$	$1.0190^{\{3\}}$
	$\sum$ ranks	10		32		30		44		32		20	

Table 2: Simulation Results for Different Estimation Methods for $\theta = 0.6, b = 1.3, \beta = 1.3, c = 1.3$.

		MLE		LS		WLS		RADE		CVM		ADE	
n	Parameter	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE	ABIAS	RMSE
25	θ	$1.4167^{\{1\}}$	$2.9765^{\{1\}}$	$5.9377^{\{3\}}$	$5.987^{\{3\}}$	$5.3157^{\{2\}}$	$5.9089^{\{2\}}$	$7.9081^{\{6\}}$	$7.027^{\{6\}}$	$5.9850^{\{4\}}$	$6.0359^{\{4\}}$	$6.4660^{\{5\}}$	$6.7847^{\{5\}}$
	b	$0.2671^{\{1\}}$	$1.0407^{\{5\}}$	$0.3497^{\{3\}}$	$0.4896^{\{2\}}$	$0.6574^{\{6\}}$	$1.0529^{\{6\}}$	$-0.2754^{\{2\}}$	$0.9128^{\{4\}}$	$0.4602^{\{5\}}$	$0.4781^{\{1\}}$	$0.3836^{\{4\}}$	$0.5126^{\{3\}}$
	β	$-0.3112^{\{1\}}$	$0.5790^{\{1\}}$	$5.8895^{\{6\}}$	$6.0181^{\{4\}}$	$4.7895^{\{3\}}$	$7.7990^{\{6\}}$	$4.7220^{\{2\}}$	$5.0831^{\{2\}}$	$5.7593^{\{5\}}$	$5.8920^{\{3\}}$	$5.3219^{\{4\}}$	$6.5540^{\{5\}}$
	c	$-0.1169^{\{1\}}$	$0.5160^{\{1\}}$	$-1.1811^{\{4\}}$	$1.2975^{\{3\}}$	$-0.5997^{\{3\}}$	$1.1351^{\{2\}}$	$-0.3568^{\{2\}}$	$1.3733^{\{4\}}$	$-1.9272^{\{6\}}$	$2.3028^{\{6\}}$	$-1.4917^{\{5\}}$	$1.9030^{\{5\}}$
	$\sum$ ranks	12		28		30		28		34		36	
50	θ	$0.6528^{\{1\}}$	$1.590^{\{1\}}$	$4.3157^{\{3\}}$	$4.3552^{\{2\}}$	$3.5922^{\{2\}}$	$4.3948^{\{3\}}$	$7.1390^{\{6\}}$	$6.2256^{\{5\}}$	$4.3654^{\{4\}}$	$6.0293^{\{4\}}$	$6.3989^{\{5\}}$	$6.5249^{\{6\}}$
	b	$0.2539^{\{2\}}$	$0.7707^{\{4\}}$	$0.3393^{\{3\}}$	$0.4504^{\{2\}}$	$1.2447^{\{6\}}$	$1.5916^{\{6\}}$	$-0.2195^{\{1\}}$	$0.8394^{\{5\}}$	$0.3910^{\{5\}}$	$0.4331^{\{1\}}$	$0.3432^{\{4\}}$	$0.4910^{\{3\}}$
	β	$-0.2091^{\{1\}}$	$0.4422^{\{1\}}$	$5.1845^{\{6\}}$	$5.2915^{\{5\}}$	$4.0652^{\{2\}}$	$4.7618^{\{3\}}$	$4.2951^{\{3\}}$	$4.5401^{\{2\}}$	$5.1325^{\{5\}}$	$5.2517^{\{4\}}$	$4.9799^{\{4\}}$	$5.9989^{\{6\}}$
	c	$-0.0780^{\{2\}}$	$0.4166^{\{1\}}$	$-1.2632\{5\}$	$1.2961^{\{3\}}$	$0.03464^{\{1\}}$	$0.9633^{\{2\}}$	$-0.3028^{\{3\}}$	$1.3563^{\{4\}}$	$-1.6286^{\{6\}}$	$2.2916^{\{6\}}$	$-1.1335^{\{4\}}$	$1.4949^{\{5\}}$
	$\sum$ ranks	13		29		25		29		35		37	
100	θ	$0.2194^{\{1\}}$	$0.6549^{\{1\}}$	$3.5364^{\{4\}}$	$3.5790^{\{3\}}$	$2.4116^{\{2\}}$	$3.0530^{\{2\}}$	$3.8870^{\{5\}}$	$5.9852^{\{5\}}$	$3.5226^{\{3\}}$	$5.5620^{\{4\}}$	$6.1788^{\{6\}}$	$6.2767^{\{6\}}$
	b	$0.2071^{\{2\}}$	$0.5967^{\{4\}}$	$0.3067^{\{3\}}$	$0.4467^{\{2\}}$	$1.1747^{\{6\}}$	$1.0544^{\{6\}}$	$-0.1613^{\{1\}}$	$0.8353^{\{5\}}$	$0.3643^{\{5\}}$	$0.4250^{\{1\}}$	$0.3383^{\{4\}}$	$0.4749^{\{3\}}$
	β	$-0.1195^{\{1\}}$	$0.2886^{\{1\}}$	$4.8754^{\{5\}}$	$5.0599^{\{5\}}$	$3.3264^{\{2\}}$	$3.8826^{\{3\}}$	$3.5238^{\{3\}}$	$3.6984^{\{2\}}$	$4.9847^{\{6\}}$	$5.2014^{\{6\}}$	$4.5193^{\{4\}}$	$4.6872^{\{4\}}$
	c	$0.0537^{\{1\}}$	$0.2829^{\{1\}}$	$-1.1257^{\{5\}}$	$1.2959^{\{5\}}$	$0.5399^{\{3\}}$	$0.9543^{\{2\}}$	$0.2297^{\{2\}}$	$1.3873^{\{6\}}$	$-1.2237^{\{6\}}$	$1.2952^{\{4\}}$	$-1.1099^{\{4\}}$	$1.2878^{\{3\}}$
	$\sum$ ranks	12		32		26		29		35		34	
200	θ	$0.1028^{\{1\}}$	$0.2438^{\{1\}}$	$3.3721^{\{5\}}$	$3.3712^{\{4\}}$	$1.8596^{\{2\}}$	$2.3309^{\{2\}}$	$2.8141^{\{4\}}$	$2.9449^{\{3\}}$	$2.6376^{\{3\}}$	$4.6768^{\{5\}}$	$4.1169^{\{6\}}$	$5.9337^{\{6\}}$
	b	$0.1872^{\{2\}}$	$0.5199^{\{4\}}$	$0.3041^{\{3\}}$	$0.4452^{\{2\}}$	$1.1631^{\{6\}}$	$1.0414^{\{6\}}$	$-0.1446^{\{1\}}$	$0.6026^{\{5\}}$	$0.3607^{\{5\}}$	$0.4238^{\{1\}}$	$0.3240^{\{4\}}$	$0.4656^{\{3\}}$
	β	$-0.0831^{\{1\}}$	$0.1959^{\{1\}}$	$4.6955^{\{5\}}$	$4.8363^{\{6\}}$	$2.9015^{\{2\}}$	$3.1542^{\{3\}}$	$3.3958^{\{3\}}$	$3.1290^{\{2\}}$	$4.6997^{\{6\}}$	$4.8060^{\{5\}}$	$3.7971^{\{4\}}$	$3.8572^{\{4\}}$
	c	$-0.0549^{\{1\}}$	$0.2290^{\{1\}}$	$-1.043^{\{4\}}$	$1.2918^{\{4\}}$	$0.4124^{\{3\}}$	$0.9438^{\{2\}}$	$-0.1026^{\{2\}}$	$1.3660^{\{6\}}$	$-1.2129^{\{6\}}$	$1.2925^{\{5\}}$	$-1.0911^{\{5\}}$	$1.2378^{\{3\}}$
	$\sum$ ranks	12		33		26		26		36		35	
400	θ	$0.0582^{\{1\}}$	$0.1369^{\{1\}}$	$1.6972^{\{4\}}$	$1.7420^{\{3\}}$	$1.6250^{\{3\}}$	$1.8736^{\{4\}}$	$1.5937^{\{2\}}$	$1.7057^{\{2\}}$	$1.7126^{\{5\}}$	$2.752^{\{5\}}$	$2.7981^{\{6\}}$	$3.0728^{\{6\}}$
	b	$0.1103^{\{2\}}$	$0.3495^{\{1\}}$	$0.3014^{\{3\}}$	$0.4413^{\{5\}}$	$1.0776^{\{6\}}$	$1.0312^{\{6\}}$	$-0.0116^{\{1\}}$	$0.4169^{\{3\}}$	$0.3603^{\{5\}}$	$0.41443^{\{2\}}$	$0.3062^{\{4\}}$	$0.4334^{\{4\}}$
	β	$-0.0563^{\{1\}}$	$0.1365^{\{1\}}$	$4.4101^{\{5\}}$	$4.5286^{\{5\}}$	$2.7780^{\{2\}}$	$2.9747^{\{2\}}$	$3.0230^{\{3\}}$	$3.0911^{\{3\}}$	$4.5877^{\{6\}}$	$4.7856^{\{6\}}$	$3.1315^{\{4\}}$	$3.1722^{\{4\}}$
	c	$-0.0420^{\{1\}}$	$0.1750^{\{1\}}$	$-1.0085^{\{5\}}$	$1.2907^{\{4\}}$	$0.3619^{\{3\}}$	$0.9346^{\{2\}}$	$-0.0713^{\{2\}}$	$1.3184^{\{6\}}$	$-1.2051^{\{6\}}$	$1.2919^{\{5\}}$	$-0.9752^{\{4\}}$	$1.1589^{\{3\}}$
	$\sum$ ranks	9		34		28		22		40		35	
800	θ	$0.0361^{\{1\}}$	$0.0829^{\{1\}}$	$0.6845^{\{3\}}$	$0.7478^{\{4\}}$	$1.5391^{\{5\}}$	$1.5872^{\{5\}}$	$0.1738^{\{2\}}$	$0.2676^{\{2\}}$	$0.6959^{\{4\}}$	$0.7394^{\{3\}}$	$1.7727^{\{6\}}$	$2.1862^{\{6\}}$
	b	$0.0553^{\{1\}}$	$0.2097^{\{2\}}$	$0.2334^{\{3\}}$	$0.4151^{\{4\}}$	$1.0294^{\{6\}}$	$1.0173^{\{6\}}$	$-0.0680^{\{2\}}$	$0.1588^{\{1\}}$	$0.2601^{\{5\}}$	$0.4014^{\{3\}}$	$0.2429^{\{4\}}$	$0.4190^{\{5\}}$
	β	$-0.0418^{\{1\}}$	$0.0928^{\{1\}}$	$4.0893^{\{6\}}$	$3.9859^{\{6\}}$	$2.7236^{\{4\}}$	$2.7506^{\{3\}}$	$0.3821^{\{2\}}$	$0.5874^{\{2\}}$	$3.8550^{\{5\}}$	$3.9600^{\{5\}}$	$2.5860^{\{3\}}$	$2.9307^{\{4\}}$
	c	$-0.0254^{\{1\}}$	$0.1130^{\{1\}}$	$-0.807^{\{5\}}$	$0.9752^{\{4\}}$	$0.2997^{\{4\}}$	$0.9178^{\{3\}}$	$0.0711^{\{2\}}$	$1.3169^{\{6\}}$	$-1.0932^{\{6\}}$	$1.0978^{\{5\}}$	$-0.1388^{\{3\}}$	$0.2336^{\{2\}}$
	$\sum$ ranks	9		35		36		19		36		33	

Table 3: Partial and Overall Ranks of all Estimation Methods of HT-GEN-TL-LLoG Distribution by Various Model Parameter Values.

Parameters	n	MLE	LS	WLS	RADE	CVM	ADE
$\theta = 1.3, b = 0.6, \beta = 0.2, c = 2.5.$	25	1	3	2	5.5	4	5.5
	50	1	4	2	5	3	6
	100	1	4.5	2	6	3	4.5
	200	1	2.5	2.5	6	5	4
	400	1	5	2.5	6	2.5	4
	800	1	4.5	3	6	4.5	2
$\theta = 0.6, b = 1.3, \beta = 1.5, c = 1.3.$	25	1	2.5	4	2.5	5	6
	50	1	3.5	2	3.5	5	6
	100	1	4	2	3	6	5
	200	1	4	2.5	2.5	6	5
	400	1	4	3	2	6	5
	800	1	4	5.5	2	5.5	3
$\sum$ ranks		12.0	45.5	33	50	55.5	56
Overall rank		1	3	2	4	5	6

In Tables 1 and 2, the row indicating $\sum$ Ranks represents the partial sum of the ranks. Among all the estimators for a given metric, the superscript indicates their rank. Table 1 presents, for example, the ABIAS of $\hat{\theta}$ obtained via MLE method as $1.8929^{\{1\}}$ for $n = 25$. This indicates that the ABIAS of $\hat{\theta}$ obtained using the MLE method ranks first among all other estimators. Accordingly, $n = 25$, MLE offers the best ABIAS of $\hat{\theta}$ when compared with all other estimators. Table 3 shows the partial and overall ranks of all the estimation methods of HT-GEN-TL-LLoG distribution by various model parameter values. Based on the results in Tables 1 and 2, the HT-GEN-TL-LLoG distribution is stable, as the ABIAS and RMSE values for its four parameters are modest. With increasing sample size, the bias occasionally decreases while the RMSE decreases for all estimations as the sample size increases. In general, all estimation methods provide accurate bias and mean squared error estimates for large sample sizes. Table 3 shows that MLE method allows us to obtain better estimates of HT-GEN-TL-LLoG parameters, followed by WLS and then LS methods. According to the rankings, the ADE method performs the least well.

7. Risk Measures

In this section, we discuss risk measures including: value at risk (VaR), tail value at risk (TVaR), tail variance (TV), and tail variance premium (TVP) which are routinely employed by financial and actuarial professionals to assess the exposure to market risk in a portfolio of instruments.

7.1. Value at Risk

VaR is an actuarial measure that is often used to assess risk in the financial markets. It is referred to as the quantile risk measure or the quantile premium principle, and it is always provided with a stated degree of confidence, such as (90%, 95%, or 99%). The VaR of the HT-GEN-TL-G family of distributions is given by

$$VaR_q = G^{-1}\left[1-\left(1-\left[1-\left(\theta\left[(\bar{q})^{\frac{-1}{\theta}}-(1-\theta)\right]^{-1}\right)^{\frac{1}{\beta}}\right]^{\frac{-1}{b}}\right)^{\frac{1}{2}}\right], \tag{23}$$

where $q \in (0,1)$ is a specified level of significance and $\bar{q} = (1-q)$.

7.2. Tail Value at Risk

Another significant and extensively used risk measure is the TVaR, which is used to express the expected value of loss in the case that an event beyond the predetermined probability threshold has actually occurred. The conditional tail expectation (CTE) or tail conditional expectation (TCE) are some of its other names. The TVaR of the HT-GEN-EHL-G family of distributions is given as

$$\begin{aligned} TVaR_q &= E(X \mid X > x_q) = \frac{1}{1-q}\int_{VaR_q}^{\infty} x f(x)dx \\ &= \frac{1}{1-q}\sum_{k=0}^{\infty}\int_{VaR_q}^{\infty} x d_{k+1}^{*} g_{k+1}^{*}(x;\varphi)dx, \end{aligned} \tag{24}$$

where d_{k+1}^{*} is given by Eqn. (14) and $g_{k+1}^{*}(x;\varphi) = (k+1)[G(x;\varphi)]^k g(x;\varphi)$ is the pdf of Exp-G distribution with the power parameter $(k+1)$.

7.3. Tail Variance

One of the most significant actuarial measures that examines variation outside of the VaR is the tail variance. The TV of the HT-GEN-TL-G family of distributions is given by

$$\begin{aligned} TV_q &= E(X^2 \mid X > x_q) - (TVaR_q)^2 \\ &= \frac{1}{1-q}\int_{VaR_q}^{\infty} x^2 f(x)dx - (TVaR_q)^2 \\ &= \frac{1}{1-q}\sum_{k=0}^{\infty} d_{k+1}^{*}\int_{VaR_q}^{\infty} x^2 g_{k+1}^{*}(x;\varphi)dx - (TVaR_q)^2, \end{aligned} \tag{25}$$

where $g_{k+1}^{*}(x;\varphi) = (k+1)[G(x;\varphi)]^k g(x;\varphi)$ is the pdf of Exp-G distribution with the power parameter $(k+1)$ and parameter vector φ, and d_{k+1}^{*} is given by Eqn. (14). Thus, TV of HT-GEN-TL-G family of distributions can be obtained from those of Exp-G distribution.

7.4. Tail Variance Premium

The TVP is a significant risk measure that is crucial to the study of insurance. The TVP of the HT-GEN-TL-G family distributions is given by

$$TVP_q = TVaR_q + \delta(TV_q), \tag{26}$$

where $0 < \delta < 1$. The TVP of the HT-GEN-TL-G family of distributions can be obtained by substituting Eqns. (24) and (25) into Eqn. (26).

7.5. Numerical Study for the Risk Measures

In this sub-section, we perform numerical simulation of the risk measures: VaR, TVaR, TV and TVP of the heavy-tailed generalized Topp-Leone-log logistic (HT-GEN-TL-LLoG) distribution, and compare to those of the sub-models HT-GEN-LLoG, HT-TL-LLoG, HT-LLoG, GEN-TL-LLoG, and the non-nested alpha power Weibull (APW) distribution by Nassar et al. (2017), alpha power Topp-Leone Weibull (APTLW) distribution by Benkhelifa (2022) and alpha power exponentiated log-logistic (APExLLD) distribution by Teamah et al. (2021).
Simulation results are obtained as follows:
1. Random samples of size $n = 100$ are generated from each one of the used distributions and parameters have been estimated via maximum likelihood method.
2. 1000 repetitions are made to calculate the VaR, TVaR, TV and TVP for these distributions.
Tables 4 shows the numerical findings of VaR, TVaR, TV and TVP for the six compared heavy-tailed distributions. A model with higher values of VaR, TVaR, TV and TVP is said to have a heavier tail. From the figures in Table 4, we conclude that the HT-GEN-TL-LLoG distribution have a heavier tail than the HT-GEN-LLoG, HT-TL-LLoG, HT-LLoG, GEN-TL-LLoG, APW, APTLW and APExLLD distributions, hence it is suitable for modelling heavy-tailed data.

Table 4: Simulation Results of VaR, TVaR, TV and TVP

Significance level		0.7	0.75	0.8	0.85	0.9	0.95
HT-GEN-TL-LLoG($\theta = 1.3, b = 0.8, \beta = 1.5, c = 0.7$)	VaR	0.9376	1.1814	1.5297	2.0738	3.0645	5.5948
	TVaR	4.0729	4.6769	5.5101	6.7538	8.8749	13.6675
	TV	48.6618	56.0125	66.2249	81.5165	107.4176	164.0391
	TVP	38.1362	46.6863	58.4901	76.0429	105.5508	169.5047
HT-GEN-LLoG($\theta = 1.3, \beta = 1.5, c = 0.7$)	VaR	0.6670	0.8568	1.1307	1.5624	2.3547	4.3897
	TVaR	3.2591	3.7596	4.4534	5.4951	7.2866	11.4005
	TV	37.8878	43.8301	52.1598	64.7870	86.5718	135.5825
	TVP	29.7807	36.6323	46.1813	60.5641	85.2013	140.2039
HT-TL-LLoG($\theta = 1.3, b = 0.8, c = 0.7$)	VaR	0.0229	0.0493	0.1044	0.2255	0.5300	1.6321
	TVaR	1.6317	1.9511	2.4206	3.1752	4.5867	8.2312
	TV	34.8068	41.1439	50.3063	64.7515	91.0271	154.8907
	TVP	25.9965	32.8091	42.6656	58.2140	86.5111	155.3774
HT-LLoG($\theta = 1.3, c = 1.7$)	VaR	0.4280	0.5142	0.6289	0.7927	1.0579	1.6204
	TVaR	1.1735	1.3144	1.5009	1.7660	2.1926	3.0907
	TV	1.9466	2.2156	2.5935	3.1729	4.2041	6.7485
	TVP	2.5362	2.9761	3.5757	4.4631	5.9764	9.5018
GEN-TL-LLoG($b = 0.8, \beta = 1.5, c = 0.7$)	VaR	0.5460	0.6818	0.8709	1.1560	1.6496	2.8074
	TVaR	2.0749	2.3678	2.7671	3.3553	4.3447	6.5640
	TV	12.9308	14.954	17.8163	22.2204	30.0370	48.8054
	TVP	11.1265	13.5838	17.0202	22.2426	31.3780	52.9292
APW($\alpha = 0.5, \beta = 0.9, \lambda = 1.3$)	VaR	0.7329	0.8659	1.0342	1.2592	1.5900	2.1866
	TVaR	1.5454	1.6951	1.8822	2.1294	2.4881	3.1250
	TV	0.81422	0.8385	0.8672	0.9024	0.9491	1.0232
	TVP	2.1153	2.3240	2.5760	2.8965	3.3423	4.0971
APTLW($\theta = 1.9, \alpha = 1.1, \beta = 0.7, \lambda = 0.99$)	VaR	0.5061	0.5527	0.6061	0.6703	0.7536	0.8815
	TVaR	1.3355	1.4019	1.4801	1.5679	1.6369	1.7046
	TV	0.0861	0.1774	0.3357	0.6580	1.4934	4.8542
	TVP	1.3958	1.5350	1.7487	2.1272	2.9810	6.0161
APExLLD($\alpha = 0.1, a = 0.8, b = 1.5, c = 0.2$)	VaR	0.0103	0.0219	0.0476	0.1100	0.2876	1.0724
	TVaR	1.2303	1.4733	1.8333	2.4198	3.5387	6.5115
	TV	31.9238	37.8542	46.4831	60.1944	85.3673	147.1997
	TVP	23.5770	29.8639	39.0198	53.5850	80.3693	146.3512

8. Applications

In this section, we assess the goodness-of-fit of the heavy-tailed generalized Topp-Leone-log logistic (HT-GEN-TL-LLoG) distribution. The fit of the HT-GEN-TL-LLoG distribution is compared to that of the alpha power exponentiated log-logistic distribution (APExLLD) by Teamah et al. (2021), alpha power Topp-Leone Weibull (APTLW) distribution by Benkhelifa (2022), the Marshall-Olkin odd Burr III log-logistic (MOO-BIII-LLoG) distribution by Afify et al. (2020), Topp-Leone odd Burr III log-logistic (TL-OBIII-LLoG) by Moakofi et al. (2022) and odd exponentiated half logistic-Burr XII (OEHL-BXII) distribution by Aldahlan et al. (2018). The pdf of these competing models are given in the appendix.

The goodness-of-fit is assessed using the following statistics: -2log-likelihood ($-2\ln(L)$), Akaike Information Criterion ($AIC = 2p - 2\ln(L)$), Consistent Akaike Information Criterion ($CAIC = AIC + 2\frac{p(p+1)}{n-p-1}$), Bayesian Information Criterion ($BIC = p\ln(n) - 2\ln(L)$), (n is the number of observations, and p is the number of estimated parameters), Cramér-von Mises statistic (W^*), Anderson-Darling statistic (A^*), Kolmogorov-Smirnov (K-S) statistic, and its $p-value$. The model with the smallest values of the goodness-of-fit statistics and a bigger p-value for the K-S statistic is regarded as the best model.

Plots of the fitted densities, the histogram of the data, probability plots, fitted Kaplan-Meier (K-M) survival curve, empirical cumulative distribution function (ECDF), the total time on test (TTT) scaled plot, and the fitted hazard rate function (hrf) are also presented under each example.

For the probability plot, we plotted $F(x_{(j)}) = F(x_{(j)};\hat{\theta},\hat{b},\hat{\beta},\hat{c})$ against $\dfrac{j-0.375}{n+0.25}$, $j = 1, 2, \cdots, n$, where $x_{(j)}$ are the ordered values of the observed data. The measures of closeness are given by the sum of squares (SS) as

$$SS = \sum_{j=1}^{n}\left[F(x_{(j)}) - \left(\frac{j-0.375}{n+0.25}\right)\right]^2.$$

All graphics and numerical results in this section were obtained using the R programming language.

8.1. Climate Data

The data were obtained from the National Climate Data Center (NCDC) in Asheville, USA and contain measurements of wind speeds in knots over a 30-day period. There were a total of 24 observations. The data was recently analyzed by Hasaballah et al. (2023). The data are given in the appendix.

The estimated variance-covariance matrix for HT-GEN-TL-LLoG model on climate data is given by

$$\begin{bmatrix} 0.0611 & -1.373 & 9.9239\times10^{-04} & -6.0529\times10^{-03} \\ -1.3733 & 30.8647 & -2.2303\times10^{-02} & 1.3603\times10^{-01} \\ 0.0009 & -0.0223 & 1.6116\times10^{-05} & -9.8301\times10^{-05} \\ -0.0060 & 0.1360 & -9.8301\times10^{-05} & 6.2172\times10^{-04} \end{bmatrix}$$

and the 95% confidence intervals for the model parameters are given by
$\theta \in [1.4685\times10^{02} \pm 0.4845]$, $b \in [9.2405\times10^{01} \pm 10.8889]$, $\beta \in [4.5634\times10^{03} \pm 0.0078]$ and $c \in [1.6946\times10^{-01} \pm 0.0488]$, respectively.

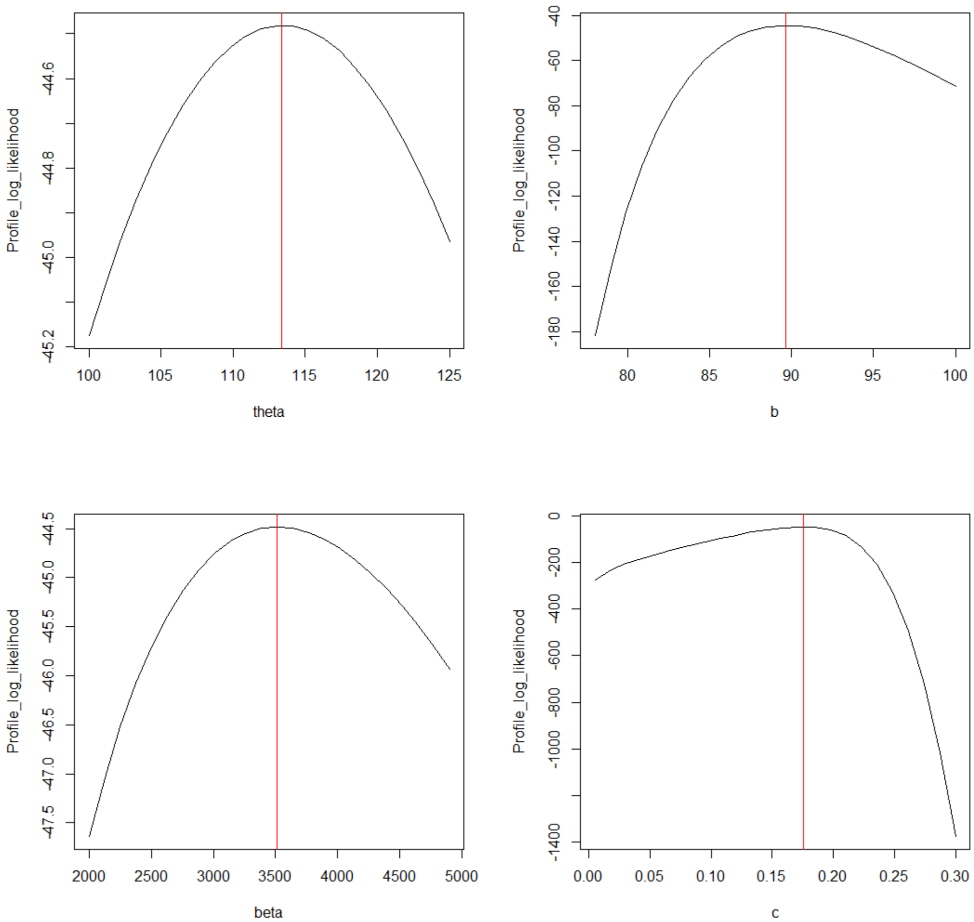

Figure 7: Profile log-likelihood function plots for parameters of HT-GEN-TL-LLoG distribution on the climate data

Figure 7 shows the profile log-likelihood plots for parameters of the HT-GEN-TL-LLoG distribution on the climate data. It can be seen that the MLEs obtained for the HT-GEN-TL-LLoG distribution are unique. This shows that the parameters of the HT-GEN-TL-LLoG distribution are identifiable.

Table 5: MLEs and Goodness-of-Fit Statistics for Climate Data

	Estimates (SE)				Statistics							
Model	θ	b	β	c	$-2\log L$	AIC	$AICC$	BIC	W^*	A^*	K-S	p-value
HT-GEN-TL-LLoG	1.4685×10^{02} (2.4720×10^{-01})	9.2405×10^{01} (5.5556)	4.5634×10^{03} (4.0145×10^{-03})	1.6946×10^{-01} (2.4934×10^{-02})	88.96236	96.96236	99.06763	101.6746	0.0323	0.2342	0.1088	0.9388
APExLLD	α 4.6888×10^{01} (6.6085×10^{-06})	a 9.6818×10^{01} (1.6290×10^{-04})	b 7.8865 (1.4318×10^{-01})	c 6.7968×10^{-03} (1.6723×10^{-03})	90.20456	98.20456	100.3098	102.9168	0.074	0.4377	0.1516	0.6389
APTLW	θ 1.0010×10^{03} (1.0528×10^{-04})	α 1.0000×10^{06} (1.0442×10^{-08})	β 3.3191×10^{-01} (4.9297×10^{-02})	λ 2.8809 (2.2317×10^{-01})	96.99553	104.9955	107.1008	109.7077	0.1198	0.8184	0.14705	0.677
MOO-BIII-LLoG	δ 0.9987 (1.6317)	α 14.6625 (0.0715)	β 60.0925 (47.3114)	λ 0.1893 (0.0825)	100.8747	108.8747	110.98	113.5869	0.1749	1.1308	0.1755	0.4500
TL-OBIII-LLoG	α 8.6415×10^{-02} (7.6491×10^{-03})	β 9.7640×10^{01} (6.1855×10^{-04})	b 4.4507×10^{-01} (1.6923×10^{-01})	λ 2.7125×10^{01} (2.4374×10^{-05})	97.14276	105.1428	107.248	109.855	0.1241	0.8445	0.1589	0.5795
OEHL-BXII	α 4.8208×10^{-01} (1.1898×10^{-01})	λ 7.7751×10^{-05} (6.0565×10^{-05})	a 2.4970 (7.6868×10^{-02})	b 2.0081 (9.9251×10^{-02})	98.35474	106.3549	108.4602	111.0671	0.0400	0.2326	0.1194	0.8836

Table 5 shows the the maximum likelihood estimates, their respective standard errors (in parenthesis) and goodness-of-fit measures: $-2log(L), AIC, AICC, BIC, W^*, A^*, K-S$ and $K-S$ p-value for the climate data. In comparison with other fitted distributions, the HT-GEN-TL-LLoG distribution provides a better fit for the climate data since it has the smallest values for all goodness-of-fit measures. It also has the highest p-value of the $K-S$ statistic. These results are further visually confirmed by the density plots and probability plots with the sum of square (SS) in Figure 9.

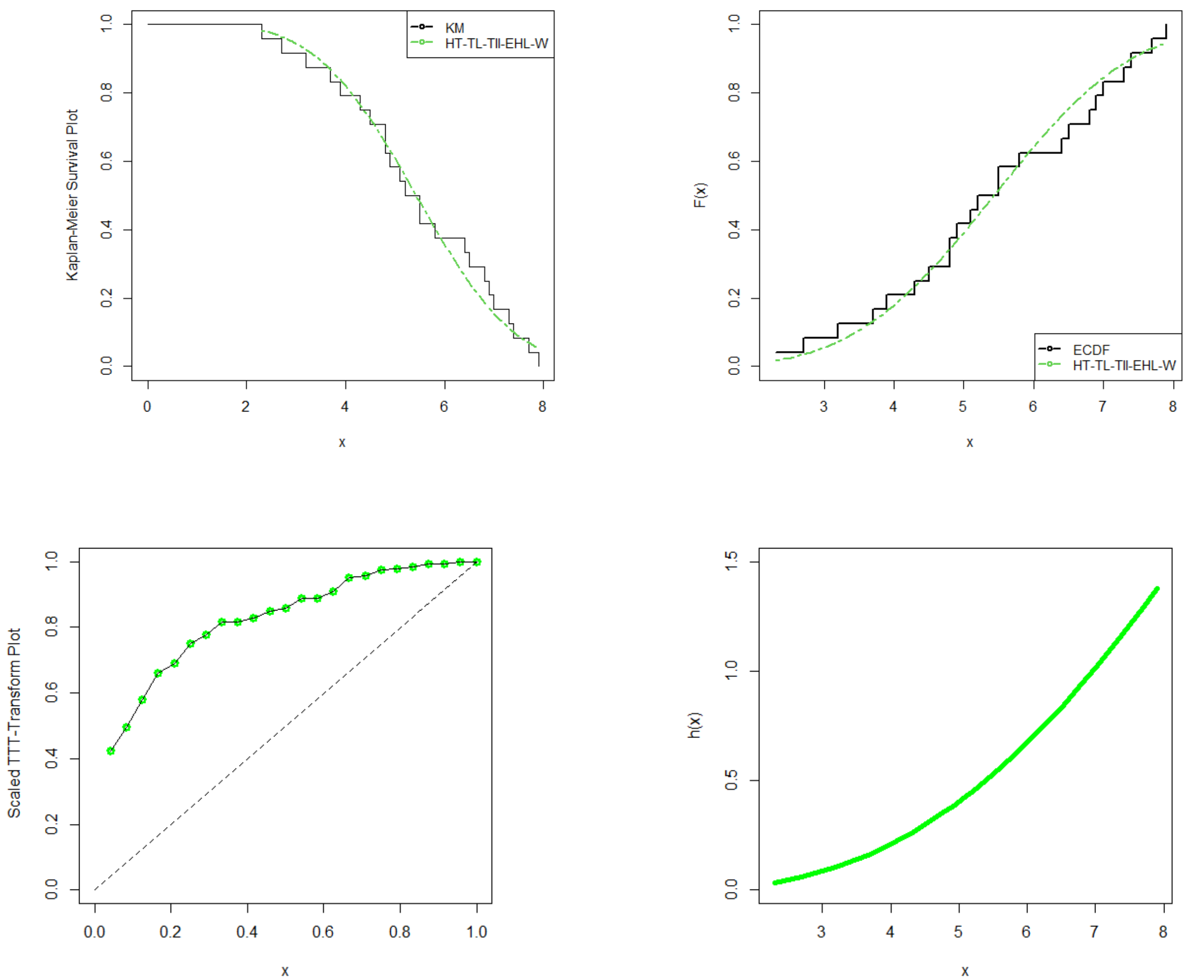

Figure 8: Fitted Kaplan-Meier survival curve, empirical cumulative distribution function, the total time on test scaled plot, and the fitted hazard rate function for the climate data

Figure 8 shows the observed and the fitted Kaplan-Meier (K-M) survival curves, theoretical and empirical cumulative distribution (ECDF), total test on time (TTT) scaled plot, and hazard rate function (hrf) plot. We can see that the HT-GEN-TL-LLoG distribution follows the empirical cdf, and Kaplan-Meier survival curves very closely. The TTT scaled plot indicates an increasing hrf because the plot exhibits a nearly concave shape and remains mostly above the 45-degree line. Furthermore, the estimated hrf in is agreement with the TTT scaled plot as it also displays an increasing shape for climate data.

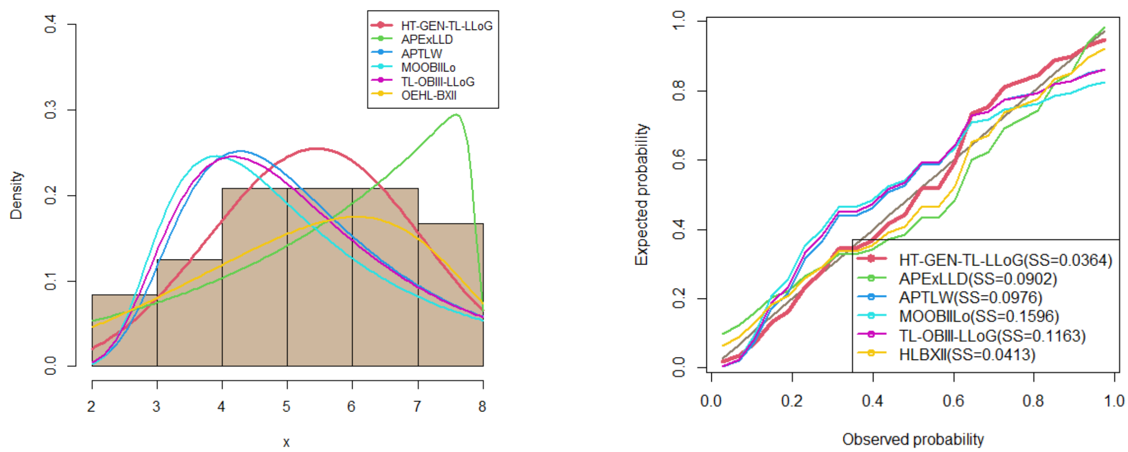

Figure 9: Histogram superposed by fitted density (left) and observed vs expected probability plots (right) for the climate data

The plots in Figure 9 depict that the HT-GEN-TL-LLoG distribution follows the fitted histogram closely and has the smallest sum of squares (SS) value from the probability plots. This supports the conclusion made from Table 5.

8.2. COVID-19 Data

The second data is concerned with the COVID-19 pandemic. These data are recorded between March 4, 2022, and July 20, 2020. The considered data set was analyzed by Almazah et al. (2022), has one hundred and eight observations and is given in the appendix.

The estimated variance-covariance matrix for HT-GEN-TL-LLoG model on COVID-19 data is given by

$$\begin{bmatrix} 0.4997 & 3.3576 & -0.2221 & -0.0238 \\ 3.3576 & 25.5058 & -1.6863 & -0.1297 \\ -0.2221 & -1.6863 & 0.1114 & 0.0085 \\ -0.0238 & -0.1297 & 0.0085 & 0.0015 \end{bmatrix}$$

and the 95% confidence intervals for the model parameters are given by
$\theta \in [1.1751 \pm 1.3855]$, $b \in [31.8877 \pm 9.8986]$, $\beta \in [54.7369 \pm 0.6544]$ and $c \in [0.3274 \pm 0.0768]$, respectively.

 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.4458

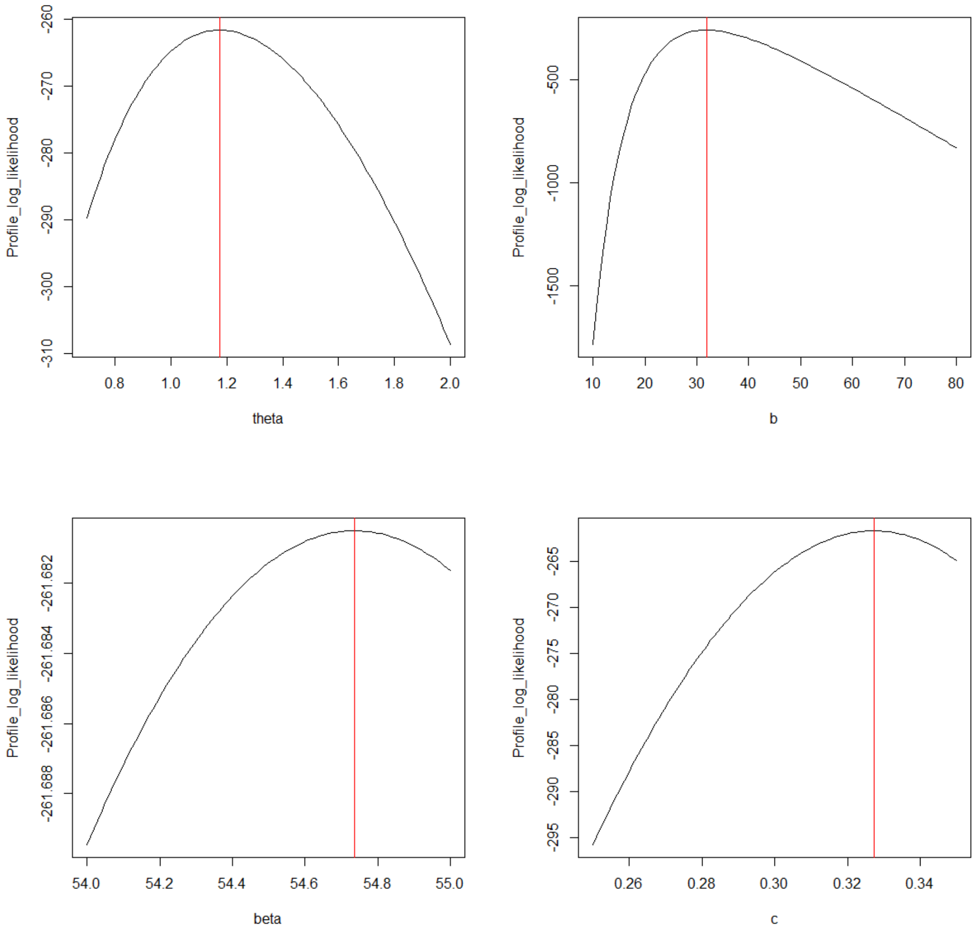

Figure 10: Profile log-likelihood function plots for parameters of HT-GEN-TL-LLoG distribution on the COVID-19 data

Figure 10 shows the profile log-likelihood plots for parameters of the HT-GEN-TL-LLoG distribution on the COVID-19 data. From the plots we can see that the estimates for the parameters of the HT-GEN-TL-LLoG distribution are achieved at single and different points.

Table 6: MLEs and Goodness-of-Fit Statistics for COVID-19 Data

	Estimates (SE)				**Statistics**							
Model	θ	b	β	c	$-2\log L$	AIC	$AICC$	BIC	W^*	A^*	K-S	p-value
HT-GEN-TL-LLoG	1.1751	31.8877	54.7369	0.3274	523.3610	531.361	531.757	542.0148	0.0529	0.2952	0.0681	0.7085
	(0.7069)	(5.0502)	(0.3339)	(0.0392)								
	α	a	b	c								
APExLLD	5.9519×10^{01}	2.0779	5.4001×10^{-02}	2.0001×10^{03}	536.3646	544.3646	544.7606	555.0183	0.1868	1.1491	0.0839	0.4435
	(4.1629×10^{-06})	(1.4757×10^{-01})	(1.6288×10^{-02})	(2.3632×10^{-07})								
	θ	α	β	λ								
APTLW	1.0011×10^{03}	1.0000×10^{06}	1.9823×10^{-01}	3.6248	533.7399	541.7400	542.1360	552.3937	0.1573	0.9735	0.0864	0.4066
	(3.9544×10^{-05})	(3.5162×10^{-09})	(1.4290×10^{-02})	(8.0466×10^{-02})								
	δ	α	β	λ								
MOO-BIII-LLoG	32.9837	5.1065	3.2202	0.5666	528.4689	536.4689	536.8649	547.1226	0.1029	0.5898	0.0688	0.6971
	(38.2888)	(0.0053)	(2.8670)	(0.0512)								
	α	β	b	λ								
TL-OBIII-LLoG	4.2909	28.5053	0.3680	0.3543	532.2005	540.2005	540.5965	550.8542	0.1349	0.8391	0.0840	0.4417
	(72.0651)	(28.0193)	(0.3576)	(5.9507)								
	α	λ	a	b								
OEHL-BXII	0.3348	7.7814×10^{-04}	58.0010	0.0557	568.9019	576.9019	577.2979	587.5557	0.2900	1.8394	0.1254	0.0712
	(0.0631)	(9.3956×10^{-04})	(1.3570×10^{-05})	(0.0077)								

The data analysis results given in Table 6 indicates that the HT-GEN-TL-LLoG distribution outperforms the other fitted distributions since it has the lowest values of the goodness-of-fit statistics: $-2\ln(L), AIC, CAIC, BIC, W^*, A^*, K-S$ and largest p-value of the $K-S$ statistic.

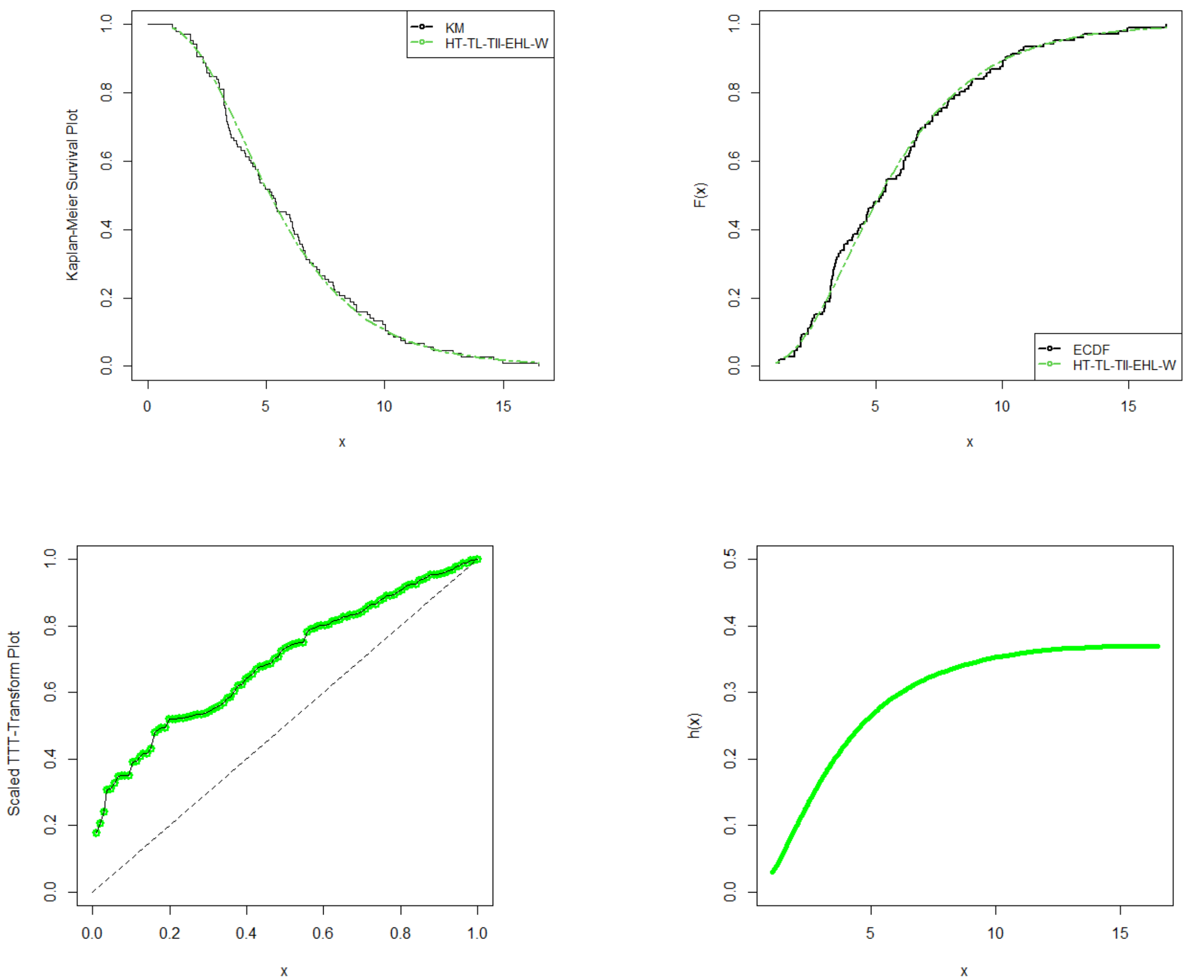

Figure 11: Fitted Kaplan-Meier survival curve, empirical cumulative distribution function, the total time on test scaled plot, and the fitted hazard rate function for the COVID-19 data

Figure 11 shows the observed and the fitted Kaplan-Meier survival curves, ECDF plots, TTT scaled plot and hrf plot. We can see that the HT-GEN-TL-LLoG distribution follows the empirical cdf, and Kaplan-Meier survival curves very closely. The TTT scaled plot and hrf plot shows that the hrf for the COVID-19 data is increasing. This is seen from the nearly concave shape of the plot above the 45-degree line. These characteristics suggest a rising hazard rate over time, indicating an escalating risk or likelihood of the event (in this case, COVID-19 cases or related events) occurring as time progresses.

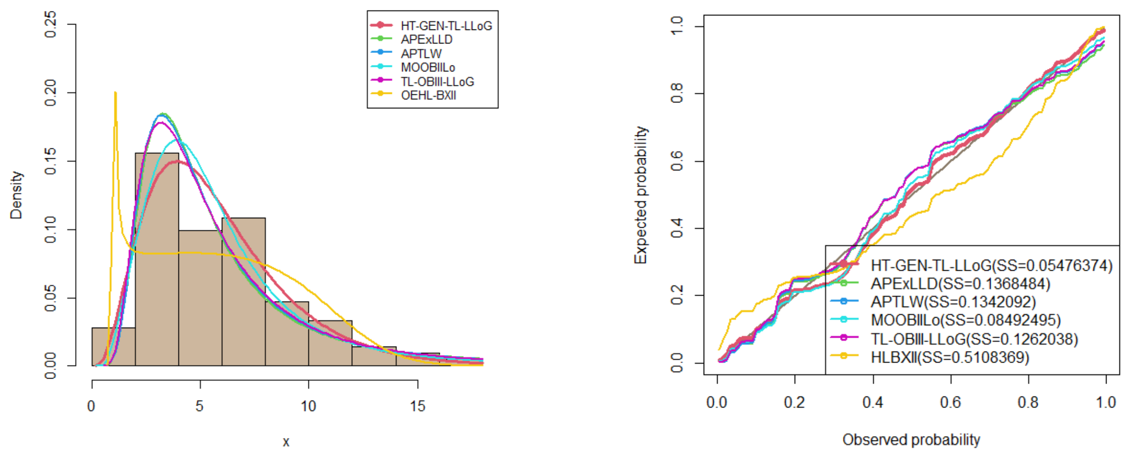

Figure 12: Histogram superposed by fitted density (left) and observed vs expected probability plots (right) for the COVID-19 data

To support the results in Table 6, a visual illustration is provided in Figure 12. These plots show that the HT-GEN-TL-LLoG capture the COVID-19 data better than other fitted distributions.

9. Conclusions

We have proposed a new family of distributions called heavy-tailed generalized Topp-Leone-G (HT-GEN-TL-G) family of distributions. Several statistical properties of the new distribution are derived and studied. The unknown parameters of the new distribution are estimated using different estimation methods, namely, maximum likelihood estimation, least-squares estimation, weighted least-squares estimation, maximum product spacing estimation, Cramér-von Mises, and Anderson-Darling estimation. Monte Carlo simulations were used to evaluate the consistency properties of the six estimation methods for the HT-GEN-TL-LLoG distribution. Risk measures for this distribution were also presented, and the results revealed that the HT-GEN-TL-LLoG distribution is heavy-tailed. Finally, the superiority and importance of a member of the HT-GEN-TL-G family of distributions was illustrated by using two real data sets.

In the future, the HT-GEN-TL-G family of distributions can be applied to different kinds of data including censored data, record values and extreme values. In this work, we only used classical methods to estimate the parameters of the new HT-GEN-TL-LLoG distribution. In the future, parameters of the HT-GEN-TL-G family of distributions can be estimated using Bayesian methods. Also, machine learning alogrithms, namely, support vector regression, random forest and neural network autoregression can be used to forecast the data set analyzed using the HT-GEN-TL-G family of distributions.

Appendix

Click on the link below for results in the appendix.
`https://drive.google.com/file/d/1knP4A9QD5rE5obkH8jEXRI8peOXDpU2_/view?usp=sharing`

References

Afify, A. Z., Cordeiro, G. M., Ibrahim, N. A., Jamal, F., Elgarhy, M., & Nasir, M. A. (2020). The Marshall-Olkin odd Burr III-G family: Theory, estimation, and engineering applications. IEEE Access, 9, 4376-4387.

Afshari, M. A., Abdi, M., Karamikabir, H., Mozafari, M., & Alizadeh, M. (2019). The new odd log-logistic generalised half-normal distribution: Mathematical properties and simulations. Pakistan Journal of Statistics and Operation Research, 15(2), 277.

Ahmad, Z., Mahmoudi, E., Hamedani, G. G., & Kharazmi, O. (2020). New methods to define heavy-tailed distributions with applications to insurance data. Journal of Taibah University for Science, 14(1), 359-382.

Ahmad, Z., Mahmoudi, E., & Dey, S. (2022). A new family of heavy tailed distributions with an application to the heavy tailed insurance loss data. Communications in Statistics-Simulation and Computation, 51(8), 4372-4395.

Ahn, S., Kim, J. H., & Ramaswami, V. (2012). A new class of models for heavy tailed distributions in finance and insurance risk. Insurance: Mathematics and Economics, 51(1), 43-52.

Aldahlan, M., & Afify, A. Z. (2018). The odd exponentiated half-logistic Burr XII distribution. Pakistan Journal of Statistics and Operation Research, 14(2), 305-317.

Aldahlan, M. A., Jamal, F., Chesneau, C., Elgarhy, M., & Elbatal, I. (2020). The truncated Cauchy power family of distributions with inference and applications. Entropy, 22(3), 346.

Alghamdi, S. M., Shrahili, M., Hassan, A. S., Mohamed, R. E., Elbatal, I., & Elgarhy, M. (2023). Analysis of milk production and failure data: Using unit exponentiated half logistic power series class of distributions. Symmetry, 15(3), 714.

Almazah, M., Ullah, K., Hussam, E., Hossain, M. M., Aldallal, R., & Riad, F. H. (2022). New statistical approaches for modeling the COVID-19 data set: A case study in the medical sector. Complexity, 2022.

Almetwally, E. M. (2022). The odd Weibull inverse Topp–Leone distribution with applications to COVID-19 data. Annals of Data Science, 9(1), 121-140.

Al-Moisheer, A. S., Elbatal, I., Almutiry, W., & Elgarhy, M. (2021). Odd inverse power generalized Weibull generated family of distributions: Properties and applications. Mathematical Problems in Engineering, 2021, 1-17.

Al-Shomrani, A., Arif, O., Shawky, A., Hanif, S., & Shahbaz, M. Q. (2016). Topp-Leone family of distributions: Some properties and application. Pakistan Journal of Statistics and Operation Research, 12(3), 443-451.

Alyami, S. A., Elbatal, I., Alotaibi, N., Almetwally, E. M., Okasha, H. M., & Elgarhy, M. (2022). Topp–Leone modified Weibull model: Theory and applications to medical and engineering data. Applied Sciences, 12(20), 10431.

Atchadé, M. N., N'bouké, M., Djibril, A. M., Shahzadi, S., Hussam, E., Aldallal, R., Alshanbari, H.M., Gemeay, A.M., & El-Bagoury, A. A. H. (2023). A new power Topp–Leone distribution with applications to engineering and industry data. PLoS one, 18(1), e0278225.

Badr, M., & Ijaz, M. (2021). The exponentiated exponential Burr XII distribution: Theory and application to lifetime and simulated data. Plos one, 16(3), e0248873.

Benkhelifa, L. (2022). Alpha power Topp-Leone Weibull distribution: Properties, characterizations, regression modeling and applications. Journal of Statistics and Management Systems, 25(8), 1945-1970.

Castillo, E. (2005). Extreme value and related models with applications in engineering and science. (No Title).

Chipepa, F., Oluyede, B., & Makubate, B. (2020). The Topp-Leone-Marshall-Olkin-G family of distributions with applications. International Journal of Statistics and Probability, 9(4), 15-32.

Fayomi, A., Khan, S., Tahir, M. H., Algarni, A., Jamal, F., & Abu-Shanab, R. (2022). A new extended gumbel distribution: Properties and application. Plos one, 17(5), e0267142.

Forbes, F., & Wraith, D. (2014). A new family of multivariate heavy-tailed distributions with variable marginal amounts of tail weight: Application to robust clustering. Statistics and computing, 24(6), 971-984.

Khaleel, M. A., & Hammed, S. S. (2023). Gompertz Topp–Leone invers Rayleigh distributions: Some properties and application. Tikrit Journal of Administration and Economics Sciences, 19(61, 2), 417-435.

Hasaballah, M. M., Al-Babtain, A. A., Hossain, M. M., & Bakr, M. E. (2023). Theoretical aspects for bayesian predictions based on three-parameter Burr-XII distribution and its applications in climatic Data. Symmetry, 15(8), 1552.

Kilai, M., Waititu, G. A., Kibira, W. A., Alshanbari, H. M., & El-Morshedy, M. (2022). A new generalization of gull alpha power family of distributions with application to modeling COVID-19 mortality rates. Results in Physics, 36, 105339.

Korkmaz, M. Ç. (2020). A new heavy-tailed distribution defined on the bounded interval: The logit slash distribution and its application. Journal of Applied Statistics, 47(12), 2097-2119.

Korkmaz, M. Ç., Altun, E., Chesneau, C., & Yousof, H. M. (2022). On the unit-Chen distribution with associated quantile regression and applications. Mathematica Slovaca, 72(3), 765-786.

Mahmoud, M. R., Ismail, A. E., & Ahmad, M. A. (2023). Weibull-inverse exponential [Loglogistic]: A new distribution. Asian Journal of Probability and Statistics, 21(3), 33-44.

Marganpoor, S., Ranjbar, V., Alizadeh, M., & Abdollahnezhad, K. (2020). Generalised odd Fréchet family of distributions: Properties and applications. Statistics in Transition new series, 21(3), 109-128.

Marmolejo-Ramos, F., Barrera-Causil, C., Kuang, S., Fazlali, Z., Wegener, D., Kneib, T., De Bastiani, F., & Martinez-Florez, G. (2023). Generalised exponential-Gaussian distribution: A method for neural reaction time analysis. Cognitive Neurodynamics, 17(1), 221-237.

Moakofi, T., & Oluyede, B. (2023). The Harris-Topp-Leone-G family of distributions: properties and applications. International Journal of Mathematics in Operational Research, 25(2), 208-241.

Moakofi, T., Oluyede, B., & Gabanakgosi, M. (2022). The Topp-Leone odd Burr III-G family of distributions: Model, properties and applications. Statistics, Optimization & Information Computing, 10(1), 236-262.

Mohammad, A.A. (2020). The Marshall-Olkin Topp Leone flexible Weibull distribution properties with application. Tikrit Journal of Administration and Economics Sciences, 16(51 part 2).

Nassar, M., Alzaatreh, A., Mead, M., & Abo-Kasem, O. (2017). Alpha power Weibull distribution: Properties and applications. Communications in Statistics-Theory and Methods, 46(20), 10236-10252.

Oluyede, B., & Moakofi, T. (2023). The gamma-Topp-Leone-type II-exponentiated half logistic-G family of distributions with applications. Stats, 6(2), 706-733.

Riad, F. H., Radwan, A., Almetwally, E. M., & Elgarhy, M. (2023). A new heavy tailed distribution with actuarial measures. Journal of Radiation Research and Applied Sciences, 16(2), 100562.

Shah, Z., Khan, D. M., Khan, Z., Faiz, N., Hussain, S., Anwar, A., Ahmad, T., & Kim, K. I. (2023). A new generalized logarithmic–X family of distributions with biomedical data analysis. Applied Sciences, 13(6), 3668.

Shafiq, A., Sindhu, T. N., Hussain, Z., Mazucheli, J., & Alves, B. (2023). A flexible probability model for proportion data: Unit Gumbel type-II distribution, development, properties, different method of estimations and applications. Austrian Journal of Statistics, 52(2), 116-140.

Shrahili, M., Muhammad, M., Elbatal, I., Muhammad, I., Bouchane, M., & Abba, B. (2023). Properties and applications of the type I half-logistic Nadarajah-Haghighi distribution. Austrian Journal of Statistics, 52(2), 1-21.

Tahir, M. H., Hussain, M. A., Cordeiro, G. M., El-Morshedy, M., & Eliwa, M. S. (2020). A new Kumaraswamy generalized family of distributions with properties, applications, and bivariate extension. Mathematics, 8(11), 1989.

Teamah, A. E. A., Elbanna, A. A., & Gemeay, A. M. (2021). Heavy-tailed log-logistic distribution: Properties, risk measures and applications. Statistics, Optimization & Information Computing, 9(4), 910-941.

Tomazella, V. L., Ramos, P. L., Ferreira, P. H., Mota, A. L., & Louzada, F. (2022). The Lehmann type II inverse Weibull distribution in the presence of censored data. Communications in Statistics-Simulation and Computation, 51(12), 7057-7073.

Yousof, H. M., Tashkandy, Y., Emam, W., Ali, M. M., & Ibrahim, M. (2023). A new reciprocal Weibull extension for modeling extreme values with risk analysis under insurance data. Mathematics, 11(4), 966.

Zhao, W., Khosa, S. K., Ahmad, Z., Aslam, M., & Afify, A. Z. (2020). Type-I heavy tailed family with applications in medicine, engineering and insurance. PloS one, 15(8), e0237462.

Zhao, J., Ahmad, Z., Mahmoudi, E., Hafez, E. H., & Mohie El-Din, M. M. (2021). A new class of heavy-tailed distributions: Modeling and simulating actuarial measures. Complexity, 2021, 1-18.