



A new class of distributions based on the zero truncated Poisson distribution with properties and applications



T. H. M. Abouelmagd^{a,*}, Mohammed S. Hamed^a, Laba Handique^b, Hafida Goual^c, M. Masoom Ali^d, Haitham M. Yousof^e, Mustafa Ç. Korkmaz^f

^aManagement Information System Department, Taibah University, Saudi Arabia.

^bDepartment of Statistics, Dibrugarh University, Assam, India.

^cLaboratory of Probability and Statistics, University of Badji Mokhtar, Annaba, Algeria.

^dDepartment of Mathematical Sciences, Ball State University, Muncie, USA.

^eDepartment of Statistics, Mathematics and Insurance, Benha University, Egypt.

^fDepartment of Measurement and Evaluation, ArtvinÇoruh University, Artvin, Turkey.

Abstract

We study a new family of distributions defined by the minimum of the Poisson random number of independent and identically distributed random variables having the Topp Leone-G distribution. Some mathematical properties of the new family are derived. Maximum likelihood estimation of the model parameters is investigated. Two special models of the new family are discussed. We perform three applications to real data sets to show the potentiality of the proposed family. In order to test the validity of the new family, a modified Chi-squared goodness-of-fit test based on Nikulin-Rao-Robson statistics is proposed theoretically.

Keywords: Topp Leone-G family, order statistics, maximum likelihood estimation, quantile function, generating function, moments.

2010 MSC: 60E05, 62F10.

©2019 All rights reserved.

1. Introduction

In the statistical literature, several G classes of distributions have been constructed using zero truncated Poisson (ZTP) distribution such as the complementary generalized transmuted Poisson G family by Alizadeh et al. [1], the exponentiated generalized G Poisson family of distributions by Aryal and Yousof [3], the odd log-logistic Poisson G family of distributions by Alizadeh et al. [2], among others. On this

*Corresponding author

Email addresses: tabouelmagd@taibahu.edu.sa (T. H. M. Abouelmagd), moswilem@gmail.com (Mohammed S. Hamed), handique1aba@gmail.com (Laba Handique), goual.hafida@gmail.com (Hafida Goual), mali@bsu.edu (M. Masoom Ali), haitham.yousof@fcom.bu.edu.eg (Haitham M. Yousof), mcagatay@artvin.edu.tr (Mustafa Ç. Korkmaz)

doi: [10.22436/jnsa.012.03.03](https://doi.org/10.22436/jnsa.012.03.03)

Received: 2018-08-29 Revised: 2018-09-05 Accepted: 2018-09-19

article we introduce a new flexible version of the Topp Leone G family of distributions via the ZTP distribution in Section 2. Section 3 provides a useful representation. Some special models are presented in Section 4. In Section 5, we derive some of the mathematical properties of the new family. Maximum likelihood estimation for the model parameters under uncensored data is addressed in Section 6. In Section 7, the potentiality of the proposed class is illustrated by means of three real data sets. A goodness-of-fit test for TL-GP family in complete data case is presented in Section 8. Finally, Section 9 provides some concluding remarks.

2. Genesis and motivation of the new family

Suppose that a system has N subsystems functioning independently at a given time where N has ZTP distribution with parameter λ . It is the conditional probability distribution of a Poisson-distributed random variable (r.v.), given that the value of the random variable is not zero. The probability mass function (pmf) of N is given by

$$P(N = n) = \frac{\lambda^n e^{-\lambda}}{n!(1 - e^{-\lambda})}, \quad n = 1, 2, \dots$$

We note that for ZTP variable the expected value and variance are respectively given by

$$\mu(N) = E(N) = \lambda (1 - e^{-\lambda})^{-1} \quad \text{and} \quad V(N) = \frac{\lambda + \lambda^2}{1 - e^{-\lambda}} - \frac{\lambda^2}{(1 - e^{-\lambda})^2}.$$

Suppose that the failure time of each subsystem has the Topp Leone-G (TLG(α) for short) distribution [9] with CDF

$$H_{\text{TLG}}(x; \alpha, \psi) = \{G(x; \psi) [2 - G(x; \psi)]\}^\alpha = G(x; \psi)^\alpha [2 - G(x; \psi)]^\alpha, \quad x \in \mathbb{R}, \quad \alpha > 0,$$

where α is positive shape parameter, $G(x; \psi) = G(x)$ is the CDF of the baseline distribution and $\psi = (\psi_1, \psi_2, \dots, \psi_n)$ is the vector of the parameters. The corresponding PDF of the TLG family is given by

$$h_{\text{TLG}}(x; \alpha, \psi) = 2\alpha g(x; \psi) G(x; \psi)^{\alpha-1} [1 - G(x; \psi)] [2 - G(x; \psi)]^{\alpha-1}, \quad x \in \mathbb{R},$$

where $g(x; \psi) = g(x)$ is the PDF of the baseline distribution. Let Y_i denote the failure time of the i^{th} subsystem and X denote the time to failure of the first out of the N functioning subsystems. We can write

$$X = \min\{Y_1, \dots, Y_N\},$$

so the conditional CDF of X given N is

$$F(x | N) = 1 - P_r(X > x | N) = 1 - P_r(Y_1 > x)^N = 1 - (1 - H_{\text{TLG}}(x; \alpha, \psi))^N,$$

and the marginal CDF of X is

$$F(x) = F_{\text{TL-GP}}(x) = \left[\frac{1 - e^{-\lambda G(x; \psi)}}{1 - e^{-\lambda}} \right]^\alpha \left[2 - \frac{1 - e^{-\lambda G(x; \psi)}}{1 - e^{-\lambda}} \right]^\alpha, \quad x \in \mathbb{R}. \quad (2.1)$$

The CDF in (2.1) is called CDF of the TL-GP family. The corresponding PDF is

$$f(x) = f_{\text{TL-GP}}(x) = \frac{2\alpha\lambda}{(1 - e^{-\lambda})^\alpha} g(x; \psi) e^{-\lambda G(x; \psi)} \left[1 - e^{-\lambda G(x; \psi)} \right]^{\alpha-1} \\ \times \left\{ 1 - \frac{[1 - e^{-\lambda G(x; \psi)}]}{(1 - e^{-\lambda})} \right\} \left\{ 2 - \frac{[1 - e^{-\lambda G(x; \psi)}]}{(1 - e^{-\lambda})} \right\}^{\alpha-1}, \quad x \in \mathbb{R}. \quad (2.2)$$

The hazard rate function (HRF) can be calculated from $f(x)/[1 - F(x)]$.

The justification for the practicality of new family is based on the wide use of the TL-G family. It can be viewed as a suitable model for fitting the symmetric, unimodal, and right skewed data. The proposed Topp Leone-exponential Poisson lifetime model is much better than the exponential, Moment exponential, Marshall-Olkin exponential, Generalized Marshall-Olkin exponential, Kumaraswamy exponential, Beta exponential, Marshall-Olkin Kumaraswamy exponential, Kumaraswamy Marshall-Olkin exponential models so the TL-exponential Poisson lifetime model is a good alternative to these models in modeling relief times and survival times data sets (see applications 1 and 2) and the proposed Topp Leone-Fréchet Poisson lifetime model is much better than the Fréchet, Marshall-Olkin-Fréchet, Generalized Marshall-Olkin-Fréchet, Kumaraswamy-Fréchet, Beta-Fréchet, Kumaraswamy Marshall-Olkin-Fréchet and Marshall-Olkin Kumaraswamy-Fréchet in modeling gauge lengths data (see application 3).

3. Useful representation

Now we can provide a useful linear representation for the TL-GP density function in (2.2). Consider the power series

$$(1 - z)^{b-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{j! \Gamma(b-j)} z^j = \sum_{j=0}^{\infty} (-1)^j \binom{b-1}{j} z^j, \tag{3.1}$$

which holds for $|z| < 1$ and $b > 0$ real non-integer. Using the power series in (3.1) and after some algebra the PDF of the TL-GP class in (3.1) can be expressed as

$$f_{TL-GP}(x) = \sum_{i=0}^{\infty} (-1)^i \binom{\alpha-1}{i} 2^{\alpha-i} \alpha \lambda g(x) e^{-\lambda G(x)} \overbrace{\frac{[1 - e^{-\lambda G(x)}]^{i+\alpha-1}}{(1 - e^{-\lambda})^{i+\alpha}}}}^A - \sum_{i=0}^{\infty} (-1)^i \binom{\alpha-1}{i} 2^{\alpha-i} \alpha \lambda g(x) e^{-\lambda G(x)} \overbrace{\frac{[1 - e^{-\lambda G(x)}]^{i+\alpha}}{(1 - e^{-\lambda})^{i+\alpha+1}}}}^B.$$

Applying (3.1) to the quantities A and B we get

$$f_{TL-GP}(x) = \sum_{i,j=0}^{\infty} (-1)^{i+j} \binom{\alpha-1}{i} \binom{i+\alpha-1}{j} 2^{\alpha-i} \alpha \lambda g(x) \overbrace{\frac{e^{-\lambda(j+1)G(x)}}{(1 - e^{-\lambda})^{i+\alpha}}}}^C - \sum_{i,j=0}^{\infty} (-1)^{i+j} \binom{\alpha-1}{i} \binom{i+\alpha}{j} 2^{\alpha-i} \alpha \lambda g(x) \overbrace{\frac{e^{-\lambda(j+1)G(x)}}{(1 - e^{-\lambda})^{i+\alpha+1}}}}^D.$$

Expanding the quantities C and D in power series and after some algebra the PDF of the TL-GP class can be expressed as

$$f_{TL-GP}(x) = \sum_{k=0}^{\infty} \sum_{i,j=0}^{\infty} \frac{\alpha (-1)^{i+j+k} \binom{\alpha-1}{i} \left[\binom{i+\alpha-1}{j} - \binom{i+\alpha}{j} (1 - e^{-\lambda})^{-1} \right]}{2^{-(\alpha-i)\lambda - (k+1)} (j+1)^{-k} k! (k+1) (1 - e^{-\lambda})^{i+\alpha}} \pi_{k+1}(x), \tag{3.2}$$

$$f(x) = \sum_{k=0}^{\infty} t_k \pi_{k+1}(x),$$

where

$$t_k = \sum_{i,j=0}^{\infty} \frac{\alpha(-1)^{i+j+k} \binom{\alpha-1}{i} \left[\binom{i+\alpha-1}{j} - \binom{i+\alpha}{j} (1-e^{-\lambda})^{-1} \right]}{2^{-(\alpha-i)} \lambda^{-(k+1)} (j+1)^{-k} k! (k+1) (1-e^{-\lambda})^{i+\alpha}},$$

and $\pi_{\gamma}(x) = \gamma g(x) G(x)^{\gamma-1}$. Equation (3.2) reveals that the PDF of X can be expressed as a simple linear representation of exp-G densities. So, several mathematical properties of the new generator can be obtained by knowing those of the exp-G density. The CDF of the TL-GP family can also be expressed as a simple mixture of exp-G densities. By integrating (3.2), we obtain the same mixture representation

$$F_{TL-GP}(x) = \sum_{k=0}^{\infty} t_k \Pi_{k+1}(x),$$

where $\Pi_{\gamma}(x)$ is the CDF of the exp-G family with power parameter (γ).

4. Special models

4.1. TLE-P model

The PDF and CDF of the Exponential (E) distribution are given by (for $x \geq 0$)

$$g(x; \beta) = \beta e^{-\beta x} \quad \text{and} \quad G(x; \beta) = (1 - e^{-\beta x}),$$

respectively, where $\beta > 0$ is a shape parameter. By $g(x; \beta)$ inserting and $G(x, \beta)$ in (2.2), the PDF of the TLE-P model can be written as

$$f_{TL-EP}(x) = \frac{2\alpha\lambda\beta}{(1-e^{-\lambda})^{\alpha}} e^{-\beta x - \lambda(1-e^{-\beta x})} \left[1 - e^{-\lambda(1-e^{-\beta x})} \right]^{\alpha-1} \times \left\{ 1 - \frac{[1 - e^{-\lambda(1-e^{-\beta x})}]}{(1-e^{-\lambda})} \right\} \left\{ 2 - \frac{[1 - e^{-\lambda(1-e^{-\beta x})}]}{(1-e^{-\lambda})} \right\}^{\alpha-1}.$$

Plots PDF and the HRF of the TLE-P model are given on Figure 1.

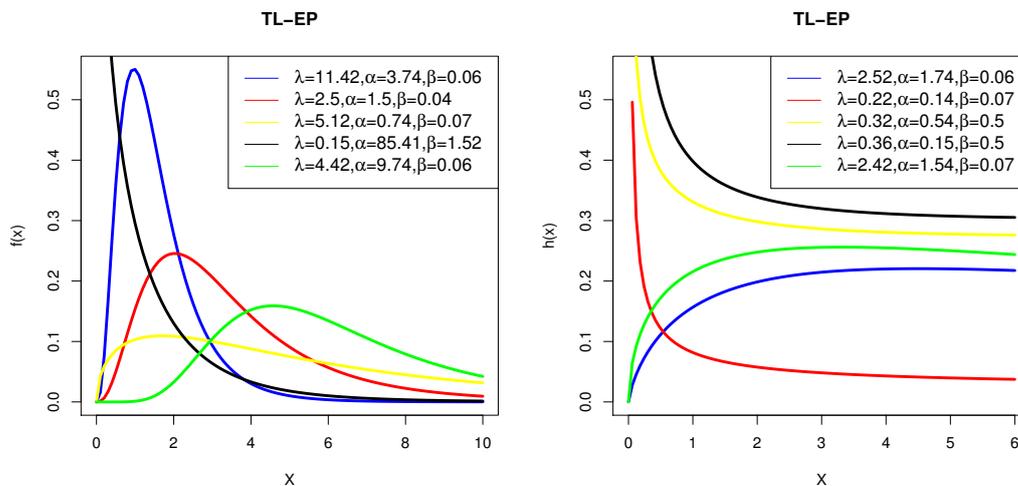


Figure 1: PDF and HRF plots for the TL-EP model.

4.2. TL Fréchet-P model

The PDF and CDF of the Fréchet (Fr) distribution are given by (for $x \geq 0$)

$$g(x; \beta, \delta) = \beta \delta^\beta x^{-(\beta+1)} e^{-\left(\frac{\delta}{x}\right)^\beta} \quad \text{and} \quad G(x; \beta, \delta) = e^{-\left(\frac{\delta}{x}\right)^\beta},$$

respectively, where $\delta > 0$ is a scale parameter and $\beta > 0$ is a shape parameter, respectively. By $g(x; \beta, \delta)$ inserting and $G(x, \beta, \delta)$ in (2.2), the PDF of the TLFr-P model can be written as

$$f_{\text{TL-FrP}}(x) = \frac{2\alpha\lambda\beta\delta^\beta}{(1-e^{-\lambda})^\alpha} x^{-(\beta+1)} e^{-\left(\frac{\delta}{x}\right)^\beta - \lambda e^{-\left(\frac{\delta}{x}\right)^\beta}} \left[1 - e^{-\lambda e^{-\left(\frac{\delta}{x}\right)^\beta}} \right]^{\alpha-1} \\ \times \left\{ 1 - \frac{\left[1 - e^{-\lambda e^{-\left(\frac{\delta}{x}\right)^\beta}} \right]}{(1-e^{-\lambda})} \right\} \left\{ 2 - \frac{\left[1 - e^{-\lambda e^{-\left(\frac{\delta}{x}\right)^\beta}} \right]}{(1-e^{-\lambda})} \right\}^{\alpha-1}.$$

Plots PDF and the HRF of the TLFr-P model are given on Figure 2.

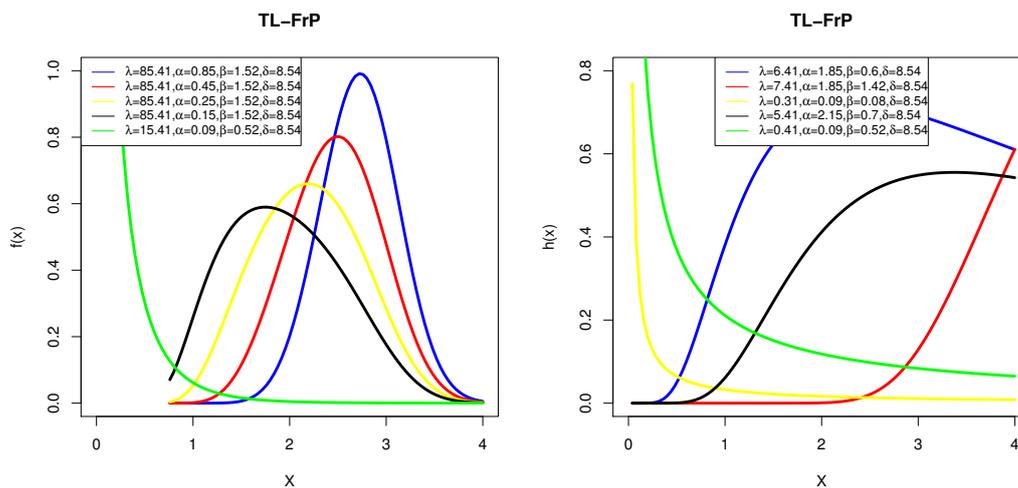


Figure 2: PDF and HRF plots for the TL-EFr model.

5. Mathematical properties

5.1. Moments and generating function

The r^{th} ordinary moment of X is given by $\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx$. Then we obtain

$$\mu'_r = \sum_{k=0}^{\infty} t_k E(Y_{k+1}^r). \tag{5.1}$$

Henceforth, Y_γ denotes the exp-G distribution with power parameter (γ). Setting $r = 1$ in (5.1), we have the mean of X . The s^{th} incomplete moment, say $\varphi_s(t)$, of X can be expressed from (3.2) as $\varphi_s(t) = \int_{-\infty}^t x^s f(x) dx$. Then

$$\varphi_s(t) = \sum_{k=0}^{\infty} t_k \int_{-\infty}^t x^s \pi_{k+1}(x) dx.$$

The moment generating function (mgf) $M_X(t) = E(e^{tX})$ of X can be derived from equation (3.2) as

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

where $M_Y(t)$ is the mgf of Y_γ . Hence, $M_X(t)$ can be determined from the exp-G generating function

5.2. Probability weighted moments (PWMs)

The $(s, r)^{\text{th}}$ PWM of X following the TL-GP family, say $\rho_{s,r}$, is formally defined by

$$\rho_{s,r} = E\{X^s F(X)^r\} = \int_{-\infty}^{\infty} x^s F(x)^r f(x) dx.$$

Using equations (2.1) and (2.2), we can write

$$f(x) F(x)^r = \sum_{k=0}^{\infty} p_k \pi_{k+1}(x),$$

where

$$p_k = \sum_{i,j=0}^{\infty} \frac{\alpha (-1)^{i+j+k} (j+1)^k \binom{\alpha r + \alpha - 1}{i} \left[\binom{\alpha r + \alpha + i - 1}{j} - \binom{\alpha r + \alpha + i}{j} (1 - e^{-\lambda})^{-1} \right]}{k! (k+1) 2^{-(\alpha r + \alpha - i)} \lambda^{-(k+1)} (1 - e^{-\lambda})^{\alpha r + \alpha + i}}.$$

Then, the $(s, r)^{\text{th}}$ PWM of X can be expressed as

$$\rho_{s,r} = \sum_{i,j=0}^{\infty} p_k E(Y_{k+1}^s) dx.$$

5.3. Moments of order statistics

Let X_1, \dots, X_n be a random sample from the TL-GP family of distributions and let $X_{(1)}, \dots, X_{(n)}$ be the corresponding order statistics. The PDF of i^{th} order statistic, say $X_{i:n}$, can be written as

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F^{j+i-1}(x), \tag{5.2}$$

where $B(\cdot, \cdot)$ is the beta function. Substituting (2.1) and (2.2) in equation (5.2) we get

$$f(x) F(x)^{(n-i+1)} = \sum_{k=0}^{\infty} d_k \pi_{k+1}(x).$$

The PDF of $X_{i:n}$ can be expressed as

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j}}{B(i, n-i+1)} \sum_{k=0}^{\infty} d_k \pi_{k+1}(x),$$

where

$$d_k = \sum_{w,m=0}^{\infty} \frac{\alpha (-1)^{w+m+k} \binom{\alpha(n-i+1)+\alpha-1}{w} \left[\binom{\alpha(n-i+1)+\alpha+w-1}{j} - \binom{\alpha(n-i+1)+\alpha+w}{m} (1 - e^{-\lambda})^{-1} \right]}{k! (k+1) (m+1)^{-k} 2^{-(\alpha(n-i+1)+\alpha-w)} \lambda^{-(k+1)} (1 - e^{-\lambda})^{\alpha(n-i+1)+\alpha+w}}.$$

Then, the density function of the TL-GP order statistics is a mixture of exp-G densities. Based on the last equation, we note that the properties of $X_{i:n}$ follow from those properties of Y_{k+1} . For example, the moments of $X_{i:n}$ can be expressed as

$$E(X_{i:n}^q) = \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j}}{B(i, n-i+1)} \sum_{k=0}^{\infty} d_k E(Y_{k+1}^q).$$

6. Estimation

Let x_1, \dots, x_n be a random sample from the TL-GP distribution with parameters δ, α , and ψ . Let $\Theta = (\alpha, \lambda, \psi^\top)^\top$ be the $p \times 1$ parameter vector. For determining the The maximum likelihood estimators (MLEs) of Θ , we have the log-likelihood function

$$\begin{aligned} \ell = \ell(\Theta) = & n \log(2) + n \log \alpha + n \log \lambda - n(\alpha - 1) \log(1 - e^{-\lambda}) + \sum_{i=1}^n \log g(x_i; \psi) - \lambda \sum_{i=1}^n G(x_i; \psi) \\ & + (\alpha - 1) \sum_{i=1}^n \log z_i + \sum_{i=1}^n \log \left[1 - \frac{z_i}{(1 - e^{-\lambda})} \right] + (\alpha - 1) \sum_{i=1}^n \log \left[2 - \frac{z_i}{(1 - e^{-\lambda})} \right], \end{aligned}$$

where

$$z_i = 1 - e^{-\lambda G(x_i; \psi)}.$$

The components of the score vector, $\mathbf{U}(\Theta) = \frac{\partial \ell}{\partial \Theta} = \left(\frac{\partial \ell}{\partial \alpha}, \frac{\partial \ell}{\partial \lambda}, \frac{\partial \ell}{\partial \psi} \right)^\top$, are

$$\begin{aligned} U_\alpha &= \frac{n}{\alpha} - n \log(1 - e^{-\lambda}) + \sum_{i=1}^n \log z_i + \sum_{i=1}^n \log \left(2 - \frac{z_i}{(1 - e^{-\lambda})} \right), \\ U_\lambda &= \frac{n}{\lambda} - \frac{ne^{-\lambda}(\alpha - 1)}{(1 - e^{-\lambda})} - \sum_{i=1}^n G(x_i; \psi) + (\alpha - 1) \sum_{i=1}^n \frac{w_i}{z_i} \\ &\quad - \sum_{i=1}^n \frac{(1 - e^{-\lambda})^{-1} w_i}{1 - (1 - e^{-\lambda})^{-1} z_i} - (\alpha - 1) \sum_{i=1}^n \frac{(1 - e^{-\lambda})^{-1} w_i}{2 - (1 - e^{-\lambda})^{-1} z_i}, \end{aligned}$$

and

$$\begin{aligned} U_\psi &= + \sum_{i=1}^n \frac{g'(x_i; \psi)}{g(x_i; \psi)} - \lambda \sum_{i=1}^n G'(x_i; \psi) + (\alpha - 1) \sum_{i=1}^n \frac{m_i}{z_i} \\ &\quad - \sum_{i=1}^n \frac{(1 - e^{-\lambda})^{-1} m_i}{1 - (1 - e^{-\lambda})^{-1} z_i} - (\alpha - 1) \sum_{i=1}^n \frac{(1 - e^{-\lambda})^{-1} m_i}{2 - (1 - e^{-\lambda})^{-1} z_i}, \end{aligned}$$

where

$$w_i = G(x_i; \psi) e^{-\lambda G(x_i; \psi)}, g'(x_i; \psi) = \partial g(x_i; \psi) / \partial \psi, m_i = G'(x_i; \psi) e^{-\lambda G(x_i; \psi)},$$

and

$$G'(x_i; \psi) = \partial G(x_i; \psi) / \partial \psi.$$

Setting the nonlinear system of equations $U_\delta = U_\alpha = 0$ and $U_\psi = \mathbf{0}$ and solving them simultaneously yields the MLE $\hat{\Theta} = (\hat{\alpha}, \hat{\lambda}, \hat{\psi}^\top)^\top$. To solve these equations, it is usually more convenient to use nonlinear optimization methods such as the quasi-Newton algorithm to numerically maximize ℓ .

7. Applications

The first data set called the failure Time data: The data represents the lifetime data relating to relief times (in minutes) of patients receiving an analgesic. The data was reported by Gross and Clark [5] and it has twenty (20) observations. The second data set called the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal [4], also used by Shibu and Irshad [10]. The third data set (gauge lengths of 20 mm), which is introduced by [6], consists of 74 observations. We shall compare the fits of the TL-EP distribution with those of other competitive models, namely: exponential (E), Moment exponential (Mom-E), Marshall-Olkin exponential (MO-E), Generalized Marshall-Olkin exponential (GMO-E), Kumaraswamy exponential (Kw-E), Beta exponential

(B-E), Marshall-Olkin Kumaraswamy exponential (MOKw-E), Kumaraswamy Marshall-Olkin exponential (KwMO-E). We shall also compare the fits of the TL-FrP distribution with those of other competitive models, namely: Fr, MO-Fr, generalized MO-Fr(GMO-Fr), Kw-Fr, B-Fr, KwMO-Fr, and MOKw-Fr.

Tables 1, 3 and 5 give the MLEs, (SEs) and [C.I.s] values for the three data sets respectively. Figures 3, 4 and 5 gives the estimated pdfs and estimated cdfs for the three data sets respectively.

Table 1: Table 1: MLEs, (SEs) and [C.I.s] values for the relief times data.

Models	Estimates
E(β)	0.526 (0.117) [0.29, 0.75]
Mom-E(β)	0.950 (0.150) [0.66, 1.24]
MO-E(α, β)	54.474, 2.316 (35.582), (0.374) [0, 124.21], [1.58, 3.04]
GMO-E(λ, α, β)	0.519, 89.462, 3.169 (0.256), (66.278), (0.772) [0.02, 1.02], [0, 219.37], [1.66, 4.68]
KwE(a, b, β)	83.756, 0.568, 3.330 (42.361), (0.326), (1.188) [0.73, 166.78], [0, 1.21], [1.00, 5.66]
B-E(a, b, β)	81.633, 0.542, 3.514 (120.41), (0.327), (1.410) [0, 317.63], [0, 1.18], [0.75, 6.28]
MOKw-E(α, a, b, β)	0.133, 33.232, 0.571, 1.669 (0.332), (57.837), (0.721), (1.814) [0, 0.78], [0, 146.59], [0, 1.98], [0, 5.22]
KwMO-E(α, a, b, β)	8.868, 34.826, 0.299, 4.899 (9.146), (22.312), (0.239), (3.176) [10.94, 46.79], [0, 78.56], [0, 0.76], [0, 11.12]
TL-EP(λ, α, β)	66.194, 8.824, 0.169 (64.951), (4.739), (0.127) [0, 193.49], [0, 18.11], [0, 0.42]

Table 2: AIC, BIC, CAIC, HQIC, A^* , W^* , K.S., and (p-value) for the relief times data.

Models	AIC, BIC, CAIC, HQIC,	A^* , W^* , K.S., and (p-value)
E(β)	67.67, 68.67, 67.89, 67.87	4.60, 0.96, 0.44, (0.004)
Mom-E(β)	54.32, 55.31, 54.54, 54.50	2.76, 0.53, 0.32, (0.07)
MO-E(α, β)	43.51, 45.51, 44.22, 43.90	0.8, 0.14, 0.18, (0.55)
GMO-E(λ, α, β)	42.75, 45.74, 44.25, 43.34	0.51, 0.08, 0.15, (0.78)
Kw-E(a, b, β)	41.78, 44.75, 43.28, 42.32	0.45, 0.07, 0.14, (0.86)
B-E(a, b, β)	43.48, 46.45, 44.98, 44.02	0.70, 0.12, 0.16, (0.80)
MOKw-E(α, a, b, β)	41.58, 45.54, 44.25, 42.30	0.60, 0.11, 0.14, (0.87)
KwMO-E(α, a, b, β)	42.8, 46.84, 45.55, 43.60	1.08, 0.19, 0.15, (0.86)
TL-EP(λ, α, β)	37.56, 40.55, 39.06, 38.15	0.23, 0.04, 0.10, (0.99)

From tables 2 and 4, the proposed TL-EP lifetime model is much better than the Exp, M-E, MO-E, GMO-E, Kw-E, BE, MOK-E, and KMO-E models so the TL-EP lifetime model is a good alternative to these models in modeling relief times and survival times data sets as well as the proposed TL-Fr P lifetime model is much better than the Fr, MO-Fr, GMO-Fr, K-Fr, B-Fr, KMO-Fr, and MOKw-Fr in modeling gauge lengths data, see Table 6.

Table 3: MLEs, SEs, C.I.s (in parentheses) values for the survival times data.

Models	Estimates
E(β)	0.540 (0.063) [0.42, 0.66]
Mom-E(β)	0.925 (0.077) [0.62, 1.08]
MO-E(α, β)	8.778, 1.379 (3.555), (0.193) [1.81, 15.74], [1.00, 1.75]
GMO-E(λ, α, β)	0.179, 47.635, 4.465 (0.070), (44.901), (1.327) [0.04, 0.32], [0, 135.64], [1.86, 7.07]
KwE(a, b, β)	3.304, 1.100, 1.037 (1.106), (0.764), (0.614) [1.13, 5.47], [0, 2.59], [0, 2.24]
B-E(a, b, β)	0.807, 3.461, 1.331 (0.696), (1.003), (0.855) [0, 2.17], [1.49, 5.42], [0, 3.01]
MOKw-E(α, a, b, β)	0.008, 2.716, 1.986, 0.099 (0.002), 1.316), (0.784), (0.048) [0.004, 0.01], [0.14, 5.29], [0.449, 3.52], [0, 0.19]
KwMO-E(α, a, b, β)	0.373, 3.478, 3.306, 0.299 (0.136), (0.861), (0.779), (1.112) [0.11, 0.64], [1.79, 5.17], [1.78, 4.83], [0, 2.48]
TL-EP(λ, α, β)	3.642, 11.327, 0.053 (0.877), (14.012), (0.069) [1.92, 5.36], [0, 38.79], [0, 0.19]

Table 4: AIC, BIC, CAIC, HQIC, A^* , W^* , K.S., and (p-value) for survival times data.

Models	AIC, BIC, CAIC, HQIC	A^* , W^* , K.S., and (p-value)
E(β)	234.63, 236.91, 234.68, 235.54	6.53, 1.25, 0.27, (0.06)
Mom-E(β)	210.40, 212.68, 210.45, 211.30	1.52, 0.25, 0.14, (0.13)
MO-E(α, β)	210.36, 214.92, 210.53, 212.16	1.18, 0.17, 0.10, (0.43)
GMO-E(λ, α, β)	210.54, 217.38, 210.89, 213.24	1.02, 0.16, 0.09, (0.51)
KwE(a, b, β)	209.42, 216.24, 209.77, 212.12	0.74, 0.11, 0.08, (0.50)
B-E(a, b, β)	207.38, 214.22, 207.73, 210.08	0.98, 0.15, 0.11, (0.34)
MOKw-E(α, a, b, β)	209.44, 218.56, 210.04, 213.04	0.79, 0.12, 0.10, (0.44)
KwMO-E(α, a, b, β)	207.82, 216.94, 208.42, 211.42	0.61, 0.11, 0.08, (0.73)
TL-GP(λ, α, β)	205.23, 212.06, 205.58, 207.94	0.53, 0.08, 0.07, (0.83)

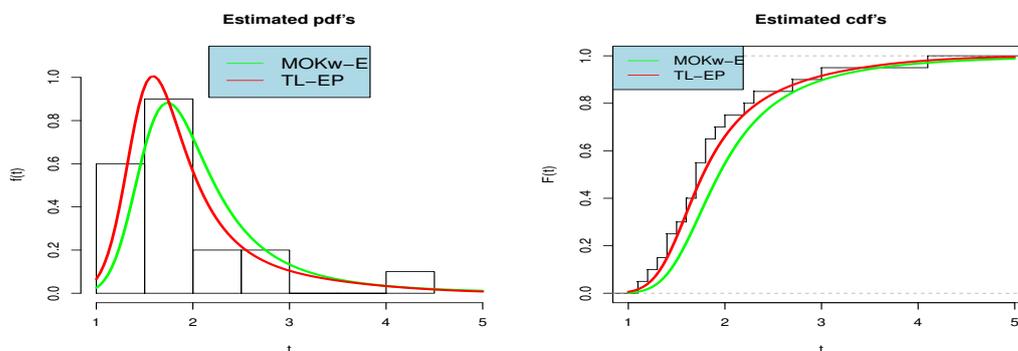


Figure 3: Estimated pdf's and estimated cdf's for failure time data.

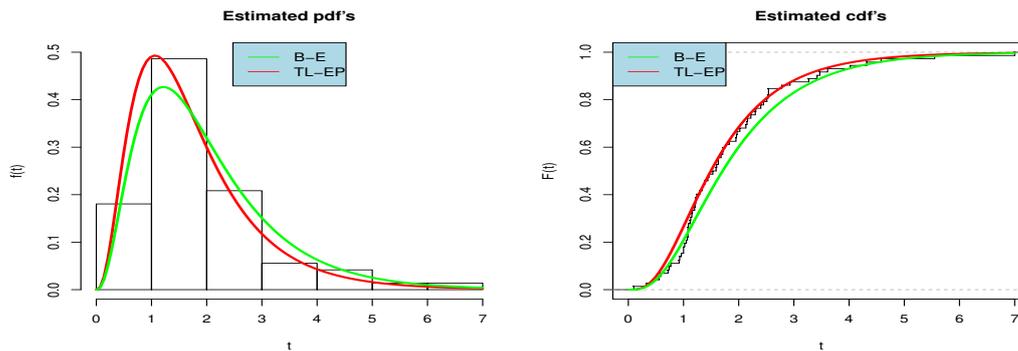


Figure 4: Estimated pdf's and estimated cdf's for the guinea pigs survival times data.

Table 5: MLEs, SEs, C.I.s (in parentheses) values for the gauge lengths data.

Models	Estimates
Fr(β, δ)	4.110, 2.169 (0.323) (0.065) [3.48, 4.74], [2.04, 2.29]
MO-Fr(α, β, δ)	80.338, 8.031 1.419 (62.007), (0.764), (0.109) [0, 201.87], [6.53, 9.53], [1.21, 1.63]
GMO-Fr($\lambda, \alpha, \beta, \delta$)	3.702, 63.707, 5.918, 1.577 (2.683), (38.655), (0.945), (0.139) [0, 8.96], [0, 139.47], [4.06, 7.77], [1.30, 1.85]
Kw-Fr(a, b, β, δ)	3.218, 217.031, 1.005, 4.384 (1.036), (268.565), (0.223), (1.012) [1.18, 5.25], [0, 743.42], [0.57, 1.44], [2.40, 6.37]
B-Fr(a, b, β, δ)	2.039, 5.857, 0.242, 37.179 (1.015), (1.813), (0.377), (33.510) [0.04, 4.02], [2.30, 9.41], [0, 0.98], [0, 102.85]
KwMO-Fr($\alpha, a, b, \beta, \delta$)	0.016, 0.827, 16.985, 0.894, 25.127 (0.023), (0.789), (24.975), (0.396), 19.688 [0, 0.06], [0, 2.37], [0, 65.93], [0.11, 1.67], [0, 63.72]
MOKw-Fr($\alpha, a, b, \beta, \delta$)	7.995, 2.933, 35.707, 1.221, 2.415 (13.063), (0.825), (41.500), (0.412), (1.032) [0, 33.59], [1.31, 4.55], [0, 117.05], [0.41, 2.03], [0.39, 4.44]
TL-FrP($\lambda, \alpha, \beta, \delta$)	71.411, 0.452, 1.524, 8.324 (82.496), (0.317), (0.588), (3.839) [0, 233.10], [0, 1.07], [0.37, 2.67], [0.79, 15.85]

Table 6: AIC, BIC, CAIC, HQIC, A^* , W^* , K.S., and (p-value) for gauge lengths data.

Models	AIC, BIC, CAIC, HQIC	A^* , W^* , K.S., and (p-value)
Fr(β, δ)	142.02, 146.63, 142.19, 143.86	2.93, 0.46, 0.15, (0.09)
MO-Fr(α, β, δ)	115.06, 121.96, 115.40, 117.81	0.78, 0.12, 0.07, (0.86)
GMO-Fr($\lambda, \alpha, \beta, \delta$)	112.80, 122.00, 113.37, 126.48	0.39, 0.07, 0.06, (0.94)
Kw-Fr(a, b, β, δ)	113.68, 122.82, 114.25, 117.36	0.37, 0.07, 0.05, (0.92)
B-Fr(a, b, β, δ)	112.63, 121.84, 113.21, 116.30,	0.44, 0.07, 0.06, (0.93)
KwMO-Fr($\alpha, a, b, \beta, \delta$)	113.30, 124.82, 114.18, 117.90	0.34, 0.05, 0.06, (0.95)
MOKw-Fr($\alpha, a, b, \beta, \delta$)	113.19, 124.68, 114.07, 117.78	0.35, 0.05, 0.05, (0.94)
TL-FrP($\lambda, \alpha, \beta, \delta$)	111.00, 120.22, 111.58, 114.68	0.30, 0.03, 0.06, (0.97)

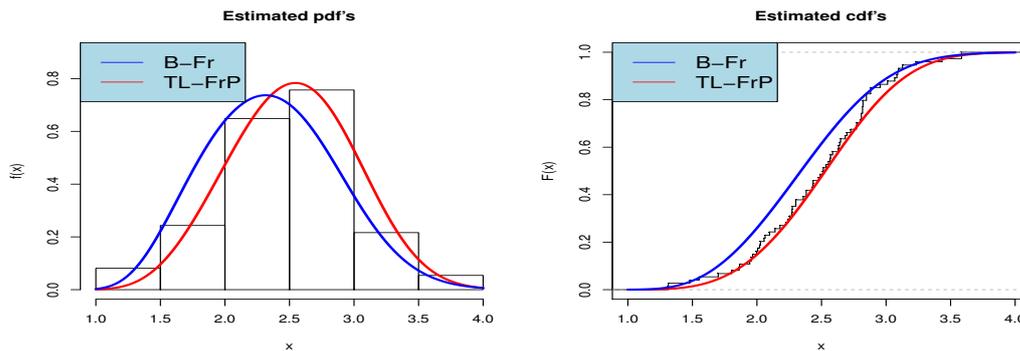


Figure 5: Estimated pdf's and estimated cdf's for the gauge lengths data.

8. Goodness-of- fit test for TL-GP family in complete data case

We present our approach to give explicitly a chi-squared test based on Nikulin-Rao-Robson (NRR) statistic for complete observations in sections, where we showed the practice of the proposed goodness of fit tests and the usefulness of the TLE-P family by applying them to two real data.

8.1. NRR statistic tes for complete data

The well-known NRR statistic ($Y_n^2(\hat{\Theta}_n)$) is one of famous modified classical chi-squared goodness-of-fit test (GOF) that was introduced by Nikulin [7] and Rao and Robson [8]. This statistic test is based on the differences between theoretical and empirical probabilities to fall into grouping cells. It use the maximum likelihood estimation on the initial data, follows a chi-square distribution (for more details see Voinov et al. [11]). Let $X = (X_1, X_2, \dots, X_n)^T$ be a random sample of n independent and identically distributed random variables. The problem is to test the null composite hypothesis H_0 according to which

$$H_0 : P_{\Theta}(X_i \leq x) = F(x, \Theta), \quad x \in \mathbb{R}, \quad \Theta = (\theta_1, \theta_2, \dots, \theta_s)^T,$$

and $v_j = (v_1, v_2, \dots, v_r)^T$ is the vector of frequencies (where $\sum_{j=1}^k v_j = n$) obtained by grouping X_i into r intervals I_j :

$$I_j =]a_{j-1}, a_j]; \quad -\infty < a_1 < \dots < a_{r-1} < a_r = +\infty.$$

The boundaries of intervals a_j are given by:

$$a_j = F^{-1} \left(\frac{j}{r} \right), \quad j = 1, \dots, r - 1.$$

The NRR statistic; with $\hat{\Theta}_n$ as a maximum likelihood estimator of the parameter vector Θ , is defined by

$$Y_n^2(\hat{\Theta}_n) = X_n^2(\hat{\Theta}_n) + \frac{1}{n} L^T(\hat{\Theta}_n) (I(\hat{\Theta}_n) - J(\hat{\Theta}_n))^{-1} L(\hat{\Theta}_n).$$

The Pearson's statistic is

$$X_n^2(\Theta) = X_n^T(\Theta) X_n(\Theta),$$

where

$$X_n(\Theta) = \left(\frac{v_1 - np_1(\Theta)}{\sqrt{np_1(\Theta)}}, \frac{v_2 - np_2(\Theta)}{\sqrt{np_2(\Theta)}}, \dots, \frac{v_r - np_r(\Theta)}{\sqrt{np_r(\Theta)}} \right)^T.$$

The vector of probabilities is

$$p(\Theta) = (p_1(\Theta), p_2(\Theta), \dots, p_r(\Theta))^T,$$

with

$$p_j(\Theta) = \int_{a_{j-1}}^{a_j} f(x, \Theta) dx, \quad j = 1, 2, \dots, r.$$

The $I(\Theta_n)$ represents the Fisher information matrix, and

$$I(\Theta) = (I_1(\Theta), \dots, I_s(\Theta))^T,$$

with

$$I_k(\Theta) = \sum_{i=1}^r \frac{\nu_i}{p_i} \frac{\partial p_i(\Theta)}{\partial \Theta_k}.$$

where $J(\Theta) = B(\Theta)^T B(\Theta)$ is called the Fisher information matrix of multinational distribution with parameters $p(\Theta)$, where

$$B(\Theta) = \left[\frac{1}{\sqrt{p_i}} \frac{\partial p_i(\Theta)}{\partial \mu} \right]_{r \times s}, \quad i = 1, 2, \dots, r \quad \forall k = 1, \dots, s.$$

Under the null hypothesis H_0 ; the NRR ($Y_n^2(\hat{\Theta}_n)$) statistic follow a chi-square distribution with $r - 1$ degrees of freedom. So, for any fixed $x > 0$, we have:

$$\lim_{n \rightarrow \infty} \mathbf{P}(Y_n^2(\hat{\Theta}_n) \geq x) = \mathbf{P}(\chi_{r-1}^2 \geq x).$$

8.2. Validity of TL-EP and TL-FrP models

We want to test the null hypothesis H_0 according to which the distribution of the sample $\mathbb{X} = (X_1, X_2, \dots, X_n)^T$ verifies

$$H_0 : P(X_i \leq x) = F_{TL-EP}(x; \Theta), \quad \Theta = (\alpha, \lambda, \beta)^T, \quad x \geq 0,$$

and

$$H_0 : P(X_i \leq x) = F_{TL-FrP}(x; \Theta), \quad \Theta = (\alpha, \lambda, \beta, \gamma)^T, \quad x \geq 0,$$

where $F_{TL-EP}(x; \Theta)$ and $F_{TL-FrP}(x; \Theta)$ are the CDF of the TL-EP and TL-FrP distributions respectively.

9. Conclusions

We study a new family of distributions defined by the minimum of the Poisson random number of independent identically distributed random variables having the Topp Leone-G distribution. Some mathematical properties of the new family including are derived. Maximum likelihood estimation of the model parameters is investigated. The new family can be viewed as a suitable model for fitting the symmetric, unimodal, and right skewed data sets. Two special models of the new family are discussed. We perform three applications to real data sets to show the potentiality of the proposed family. We conclude that the proposed Topp Leone-exponential Poisson lifetime model is much better than the exponential, Moment exponential, Marshall-Olkin exponential, Generalized Marshall-Olkin exponential, Kumaraswamy exponential, Beta exponential, Marshall-Olkin Kumaraswamy exponential, and Kumaraswamy Marshall-Olkin exponential models so the Topp Leone-exponential Poisson lifetime model is a good alternative to these models in modeling relief times and survival times data sets as well as the proposed Topp Leone-Fréchet Poisson lifetime model is much better than the Fréchet, Marshall-Olkin-Fréchet, Generalized Marshall-Olkin-Fréchet, Kumaraswamy-Fréchet, Beta-Fréchet, Kumaraswamy Marshall-Olkin-Fréchet, and Marshall-Olkin Kumaraswamy-Fréchet in modeling gauge lengths data. In order to test the validity of the new family, a modified Chi-squared goodness-of-fit test based on Nikulin-Rao-Robson statistics is proposed and developed for the new family with unknown parameters by using the maximum likelihood estimation. We also demonstrate the applications of this model by applying it to two real data.

References

- [1] M. Alizadeh, H. M. Yousof, A. Z. Afify, G. M. Cordeiro, M. Mansoor, *The complementary generalized transmuted Poisson-G family of distributions*, Austrian J. Stat., **47** (2018), 51–71. 1
- [2] M. Alizadeh, H. M. Yousof, M. Rasekhi, E. Altun, *The odd log-logistic Poisson-G family of distributions*, J. Math. Exten., **12** (2018). 1
- [3] G. R. Aryal, H. M. Yousof, *The exponentiated generalized-G Poisson family of distributions*, Stoch. Qual. Control, **32** (2017), 7–23. 1
- [4] T. Bjerkedal, *Acquisition of resistance in Guinea pigs infected with different doses of virulent tubercle bacilli*, Amer. J. Hygiene, **72** (1960), 130–148. 7
- [5] A. J. Gross, V. Clark, *Survival Distributions: Reliability Applications in the Biometrical Sciences*, John Wiley, New York, (1975). 7
- [6] D. Kundu, M. Z. Raqab, *Estimation of $R = P(Y < X)$ for three parameter Weibull distribution*, Statist. Probab. Lett., **79**, (2009), 1839–1846. 7
- [7] M. S. Nikulin, *Chi-square test for continuous distribution with shift and scale parameters*, Theory Probab. Appl., **19** (1973), 559–568. 8.1
- [8] K. C. Rao, D. S. Robson, *A chi-square statistic for goodness-of-fit for tests with in the exponential family*, Comm. Statist., **3** (1974), 1139–1153. 8.1
- [9] Y. Sangsanit, W. Bodhisuwan. *The Topp-Leone generator of distributions: properties and inferences*, Songklanakarin J. Sci. Technol., **38** (2016), 537–548. 2
- [10] D. S. Shibu, M. R. Irshad, *Extended new generalized Lindley distribution*, Statistica, **76** (2016), 41–56. 7
- [11] V. Voinov, M. Nikulin, N. Balakrishnan, *Chi-Squared Goodness of Fit Tests with Applications*, Academic Press, London, (2013). 8.1