Composition a Continuous Probability Distribution Resulting from Mixing Two Probability Distributions

Hassaneen Abd Aljabar Hashim 1, Kareema Abed Al-kadim Al-Khafaji 2

1 Department of Mathematics, College of Education for Pure Sciences, University of Babylon, Babylon, Iraq.

2 Department of Mathematics, College of Education for Pure Sciences, University of Babylon, Babylon, Iraq.

h_z_s82@yahoo.com

Abstract. In this paper, which includes three parameters of the new model resulting from the Topp-Leone Generated distribution mixture with the exponential distribution which we denoted by the symbol TLG-E distribution, finding a new probability density function and achieving its mathematical properties, as well as finding the moment, general moment, reliability and hazard functions, thereafter we clarify estimation by using maximum likelihood method as well as testing the fitting with some known distributions. In addition, we state the of statistical parameters like: scale and location.

Keywords: TLG-E, CDF, PDF, Moment function, Reliability, MLE.

1. Introduction

The statisticians are interested in expanding classical distributions to use them with a greater area of life applications. In this paper, our proposal was to find a new model used in several areas and obtain good results using known estimation methods to be the results in analysing data close to reality. A model was created by mixing one of the known distributions with a model that was recently rediscovered and they are exponential distribution with Topp-Leone distribution. Since the exponential distribution is known to most statisticians, we will give a brief summary of the Topp-Leone distribution. Topp-Leone distribution is a distribution of life mentioned by the two researchers Chester W. Topp and Fred C. Leone [9] where their research was published by Journal of the American Statistical Association in 1955. it was rediscovered by Nadarajah and Kotz in 2003 [5]. There are researchers who have worked to mixture this distribution with other distributions and from these [1], [4] and others.

2. Derive a New Model

In this section, we conclude how to find a Cumulative Distribution Function (CDF) and Probability Density Function (PDF) after using the general formula for Topp-Leone distribution which is presented below

\[
F(\rho) = \{J(\rho)^\alpha [2 - J(\rho)^\alpha]\}^B \quad \rho \geq 0
\]
Using the exponential distribution, which is considered as a random variable to generalize the Topp-Leone distribution so that

\[ j(\rho) = 1 - e^{-\lambda \rho}, \quad j(\rho) = \lambda e^{-\lambda \rho} \]

The CDF will be in the following format

\[ F(\rho) = \left\{ \frac{(1 - e^{-\lambda \rho})^\alpha}{2 - (1 - e^{-\lambda \rho})^\alpha} \right\}^\beta \]

By differentiating the above formula, we get the PDF

\[ f(\rho) = 2\beta \alpha \lambda e^{-\lambda \rho} \left(1 - e^{-\lambda \rho}\right)^{\beta\alpha - 1} \left[2 - (1 - e^{-\lambda \rho})^\alpha\right]^{\beta - 1} \left(1 - (1 - e^{-\lambda \rho})^\alpha\right) \]

**Figure 1.** The curves represent the CDF to TLG-E distribution, where different values of the parameters have been chosen in the function, and there are also, infinity of positive real values for the parameters. The CDF is achieved within its domain that belongs to the interval \([0, 1]\).

**Figure 2.** The curves in the figure represent a PDF to TLG-E distribution where are \(\alpha\) and \(\beta\) values fixed and the \(\lambda\) values are Changed.
Figure 3. The curves in the figure represent a PDF to TLG-E distribution where are $\alpha$ and $\lambda$ values fixed and the $\beta$ values are Changed.

Figure 4. The curves in the figure represent a PDF to TLG-E distribution where are $\beta$ and $\lambda$ values fixed and the $\alpha$ values are Changed.

Note that through Figures (2), (3) and (4) how the curves differ when the parameters change, so it is observe that the parameter $\lambda$ is scale and the two parameters $\beta$ and $\alpha$ are location.

The functions CDF and PDF can be writing by are new formulas they are obtained from distributive mixture The Topp-Leone Generated with exponential functions.

It is possible to simplify the formulation of both PDF and CDF

$$
\begin{align*}
\eta_1 &= \left[2 - (1 - e^{-\lambda \rho})^\alpha\right]^\beta = \sum_{k=0}^{\infty} \binom{\beta}{k} 2^{\beta-k} (1 - e^{-\lambda \rho})^{\alpha k} \\
\eta_2 &= (1 - e^{-\lambda \rho})^{\alpha \beta}, \quad \eta_k = \binom{\beta}{k} 2^{\beta-k}
\end{align*}
$$

Thus, the CDF will be

$$F(\rho) = \sum_{k=0}^{\infty} \binom{\beta}{k} 2^{\beta-k} (1 - e^{-\lambda \rho})^{\alpha (\beta+k)}$$

And, respectively, the PDF format will be

$$f(\rho) = \sum_{k=0}^{\infty} \binom{\beta}{k} 2^{\beta-k} \alpha (\beta+k) \lambda e^{-\lambda \rho} (1 - e^{-\lambda \rho})^{\alpha (\beta+k)-1}$$
**Definition 1:** Let $u$ a random variable of continuous distribution on interval $[0, \infty)$ with PDF $j(u)$ and let $\rho$ random variable of continuous distribution with CDF $G(\rho)$ such that $\rho \in [0, \infty)$ the distribution arising from its combination can be represented by the following formula

$$F(\rho) = \int_0^{G(\rho)} j(u) \, du = \int \left[ j(u) \right]_0^{G(\rho)} \, d\rho \ .$$

hence the new pdf can be writing $f(\rho) = j(G(\rho)) \frac{d}{d\rho} G(\rho)$.

**Note:** The above definition can be used and applied to mixing two distributions.

### 2. Mathematical properties of CDF

1.1. Let $\alpha \geq 0$, $\beta \geq 0$, let $\gamma > 0$, arising from its combination can be represented by the following formula

$$\int_0^\infty \left[ \alpha \gamma \beta \right] \left( 1 - e^{-\gamma} \right)^\alpha \left( 1 - e^{-\beta} \right)^\beta \, d\gamma = \frac{\beta^\alpha}{\beta^\alpha} \ .$$

By using Laplace transform we get on $\alpha \gamma \beta > 0$.

$$\int_0^\infty \left[ \alpha \gamma \beta \right] \left( 1 - e^{-\gamma} \right)^\alpha \left( 1 - e^{-\beta} \right)^\beta \, d\gamma = \frac{\beta^\alpha}{\beta^\alpha} \ .$$

### 3. Moment, Central Moment and Moment Generating Functions

3.1. **Moment Function**

**Definition 1:**

$$\mu' \ = \int_0^\infty F_{\rho \in [0, \infty)} (\rho) \, d\rho$$

3.2. **Central Moment**

$$\mu \ = \ E(\rho - \mu)^2 = \sum_{r=0}^{\infty} \left( \frac{j'}{r} \right) (1) \left( \mu_{r+1} / \mu_1 \right)^r$$

Note that From the formula (7) we getting of the following:

1. $\mu_0 = 1 \ , \ \mu_1 = 0$
2. $\mu_2 = \text{Variance}(\sigma_2^2) = \mu_2 - \mu_1^2 = \mu_2 - 3 \mu_1 \mu_1 + 2 \mu_1^3$
3. $\mu_3 = \text{Skewness}(Sk) = \mu_3 - 3 \mu_2 \mu_1 + 2 \mu_1^3$
4. $\mu_4 = \text{Kurtosis}(Ku) = 6 \mu_4 \mu_2 - 4 \mu_1 \mu_3 + \mu_4 - 3 \mu_1^4$
\[
= \left( \sum_{k=0}^{\infty} \binom{\beta}{k} \right) 2^{\beta-k} \sum_{h=0}^{\infty} \left( \frac{\alpha(\beta+k)}{h+1} \right) (-1)^h \frac{\alpha(\beta+k) \Gamma(5)}{\lambda^h (h+1)^4} - 
\]

\[
4 \left( \sum_{k=0}^{\infty} \binom{\beta}{k} \right) 2^{\beta-k} \sum_{h=0}^{\infty} \left( \frac{\alpha(\beta+k)}{h+1} \right) (-1)^h \frac{\alpha(\beta+k) \Gamma(5)}{\lambda^h (h+1)^4} + 
\]

\[
\frac{\alpha(\beta+k) \Gamma(2)}{\lambda (h+1)} \left( \sum_{k=0}^{\infty} \binom{\beta}{k} \right) 2^{\beta-k} \sum_{h=0}^{\infty} \left( \frac{\alpha(\beta+k)}{h+1} \right) (-1)^h \frac{\alpha(\beta+k) \Gamma(4)}{\lambda^h (h+1)^3} 
\]

\[
6 \left( \sum_{k=0}^{\infty} \binom{\beta}{k} \right) 2^{\beta-k} \sum_{h=0}^{\infty} \left( \frac{\alpha(\beta+k)}{h+1} \right) (-1)^h \frac{\alpha(\beta+k) \Gamma(3)}{\lambda^h (h+1)^2} 
\]

\[
3 \left( \sum_{k=0}^{\infty} \binom{\beta}{k} \right) 2^{\beta-k} \sum_{h=0}^{\infty} \left( \frac{\alpha(\beta+k)}{h+1} \right) (-1)^h \frac{\alpha(\beta+k) \Gamma(2)}{\lambda^h (h+1)} 
\]

3.3. Moment Generating Function

The moment generating function (m.g.f) of random variable \( \rho \) for distribution function TLG-E

\[
M_{(\tau)}(x) = \sum_{r=0}^{\infty} \tau^r \mu_{(r)} 
\]

4. Reliability

We mentioned at the beginning that this model is from the distributions of life and what are the possibilities of its applications. Such a distribution can find a survival function or (Reliability) for it.

\[
R(\rho) = 1 - \left( (1 - e^{-\lambda \rho})^\alpha \left[ 2 - (1 - e^{-\lambda \rho})^\beta \right] \right) 
\]

Figure 5. The curves represent the reliability function to TLG-E distribution, where different values of the parameters \( \beta, \alpha, \lambda \).

It is noticed by drawing the reliability function when increasing the value of the location parameter, we get a good characteristic of the reliability function as it continues for a period of time at one or near to it, and then it begins to gradually decrease

4.1. Properties of Reliability

1) \( \lim_{\rho \to 0} R(\rho) = \lim_{\rho \to 0} (1 - \{(1 - e^{-\lambda \rho})^\alpha \left[ 2 - (1 - e^{-\lambda \rho})^\beta \right])^\beta = 1 \)

2) \( \lim_{\rho \to \infty} R(\rho) = \lim_{\rho \to \infty} (1 - \{(1 - e^{-\lambda \rho})^\alpha \left[ 2 - (1 - e^{-\lambda \rho})^\beta \right])^\beta = 0 \)
4.2. Hazard function

Reliability is one of the important branches of statistics in which specialists are interested in many areas that depend on the time period. One of these areas is the manufacture of electronic and electrical devices and the study of quality for them, as well as in the field of medicine, engineering and biological. In the reliability test, specialists calculate the failure rate of the manufactured device, and in the field of biology, the death rate of bacteria or virus.

The failure rate is found by calculating the rate $f(t)$ into $R(t)$.

$$H(t) = \frac{f(t)}{R(t)}$$

(11)

$$H(t) = \frac{2\beta \alpha \lambda e^{-\lambda t}(1-e^{-\lambda t})^\beta - 1}{-\left[(1-e^{-\lambda t})^\alpha \left[2 - (1-e^{-\lambda t})^\alpha \right]\right]^\beta}$$

Figure 6. The curves represent the hazard function to TLG-E distribution, where different values of the parameters $\beta$, $\alpha$, $\lambda$.

5. Maximum Likelihood Estimators

Let $L(\sigma) = L(\sigma; \rho_1, \ldots, \rho_n)$

Consider are $\xi_1, \xi_2, \ldots, \xi_n$ is a random sample of size $n$ from a population with density function $f(\rho; \sigma)$ where $\sigma$ is unknown parameter, the estimate of $\sigma$ is obtained by Maximum Likelihood

$$L(\sigma) = f(\rho_1; \sigma)f(\rho_2; \sigma) \ldots f(\rho_n; \sigma)$$

$$L(\sigma_1, \sigma_2, \ldots, \sigma_k) = \prod_{i=1}^{n} f(\rho_i; \sigma_1, \sigma_2, \ldots, \sigma_k)$$

Then

$$\frac{\partial L(\sigma_1, \sigma_2, \ldots, \sigma_k)}{\partial \theta_1} = 0$$

$$\frac{\partial L(\sigma_1, \sigma_2, \ldots, \sigma_k)}{\partial \theta_2} = 0$$

$$\frac{\partial L(\sigma_1, \sigma_2, \ldots, \sigma_k)}{\partial \sigma_k} = 0$$

Now, estimation of PDF an eq(3)

$$f(\rho) = 2\beta \alpha \lambda e^{-\lambda \rho}(1-e^{-\lambda \rho})^{\beta - 1}\left[2 - (1-e^{-\lambda \rho})^\alpha \right]^{-1}(1-e^{-\lambda \rho})^\alpha$$

\[ L(\sigma_i, \rho) = \prod_{i=1}^{n} f(\rho_i ; \sigma) \]
\[ = \prod_{i=1}^{n} 2\beta a \lambda e^{-\lambda \rho_i} (1 - e^{-\lambda \rho_i})^{\beta \alpha - 1} \left[ 2 - (1 - e^{-\lambda \rho_i})^{\alpha} \right]^{\beta - 1} \left( 1 - (1 - e^{-\lambda \rho_i})^{\alpha} \right) \]
\[ = (2\beta a \lambda)^n e^{-\lambda \sum_{i=1}^{n} \rho_i} \sum_{i=1}^{n} (1 - e^{-\lambda \rho_i})^{\beta \alpha - 1} \left[ 2 - (1 - e^{-\lambda \rho_i})^{\alpha} \right]^{\beta - 1} \left( 1 - (1 - e^{-\lambda \rho_i})^{\alpha} \right) \]

take the Ln to both sides
\[ (12) \]
\[ \ell = \ln L(\sigma_i, \rho) = n \ln 2 + n \ln \alpha + n \ln \beta + n \ln \lambda - \lambda \sum_{i=1}^{n} \rho_i + (\beta \alpha - 1) \sum_{i=1}^{n} \ln (1 - e^{-\lambda \rho_i}) + (\beta - 1) \sum_{i=1}^{n} \ln (2 - (1 - e^{-\lambda \rho_i})^{\alpha}) + \sum_{i=1}^{n} \ln (1 - (1 - e^{-\lambda \rho_i})^{\alpha}) \]

find the partial derivation with respect to \( \alpha, \beta, \lambda \)

\[ \frac{\partial \ell}{\partial \alpha} = n + \beta \sum_{i=1}^{n} \ln (1 - e^{-\lambda \rho_i}) + (\beta - 1) \cdot \sum_{i=1}^{n} \frac{-(1-e^{-\lambda \rho_i})^{\alpha} \ln (1-e^{-\lambda \rho_i})}{2(1-e^{-\lambda \rho_i})^\alpha} + \sum_{i=1}^{n} \ln (1 - (1 - e^{-\lambda \rho_i})^\alpha) = 0 \]  ... ... (I)

\[ \frac{\partial \ell}{\partial \beta} = n + \alpha \sum_{i=1}^{n} \ln (1 - e^{-\lambda \rho_i}) + \sum_{i=1}^{n} \ln \left( 2 - (1 - e^{-\lambda \rho_i})^{\alpha} \right) = 0 \]  ... ... (II)

\[ \frac{\partial \ell}{\partial \lambda} = n \sum_{i=1}^{n} \rho_i + (\beta \alpha - 1) \sum_{i=1}^{n} \rho_i e^{-\lambda \rho_i} - \sum_{i=1}^{n} \frac{-\alpha (\beta - 1) \rho_i e^{-\lambda \rho_i} (1 - e^{-\lambda \rho_i})^{\alpha - 1}}{2(1-e^{-\lambda \rho_i})^\alpha} - \sum_{i=1}^{n} \frac{-\alpha (\beta - 1) \rho_i e^{-\lambda \rho_i} (1 - e^{-\lambda \rho_i})^{\alpha - 1}}{2(1-e^{-\lambda \rho_i})^\alpha} \]

\[ \sum_{i=1}^{n} \frac{-\alpha \rho_i e^{-\lambda \rho_i} (1 - e^{-\lambda \rho_i})^{\alpha - 1}}{(1-e^{-\lambda \rho_i})^\alpha} = 0 \]  ... ... (III)

Usually these equations are solved in one way by numerical analysis such as Newton-Raphson or other numerical methods.

6. Parameter Estimation

MLE is one of the most used estimation methods among researchers, which gives good and accurate results. We use this method with a data analysis program to find parameter values. Hence, comparing the model (TLG-E) with some distributions that have similar applications from the mentioned model, and from these distributions are Weibull(W), Exponentiated Weibull(EW) and Exponentiated Kumaraswamy(EK) where the formulas are as follows:

W: \( f(\rho) = ba^b \rho^{b-1} e^{-(a \rho)^b} \)

EW: \( f(\rho) = \theta ba^b \rho^{b-1} e^{-(a \rho)^b} \left[ 1 - e^{-(a \rho)^b} \right]^{\theta - 1} \)

EK: \( f(\rho) = \theta a \rho^{a-1} (1 - \rho^a)^{b-1} \left[ 1 - (1 - \rho^a)^b \right]^{\theta - 1} \)

And also our code in the table for a function MLE whereas

(MLE)_1: MLE of (TLG-E)

(MLE)_2: MLE of (W)

(MLE)_3: MLE of (EW)

(MLE)_4: MLE of (EK)

The process of calculating the parameters was performed by the MATLAB, where we mentioned the number of iterations that analyzed the system of equations, which was 122

Data:

This set of data represents a normal phenomenon and recorded waiting times (in seconds) resulting from successive explosions of Kiama Blowhole. These samples were recorded with the help of a digital clock on July 12, 1998 by Jim Irish and referred to recently has been referenced by Elgarhy [2]. The actual data are as follows [60, 28, 95, 29, 54, 91, 8, 17, 55, 10, 35, 47, 77, 36, 83, 51, 87, 60, 28, 95, 8, 27, 15, 10, 18, 16, 17, 21, 36, 18, 40, 10, 7, 34, 27, 28, 56, 8, 25, 68, 146, 89, 18, 73, 69, 9, 37, 10, 82,
29, 8, 60, 61, 18, 169, 25, 8, 26, 11, 83, 11, 42, 17, 14, 9, 12]. We presented the results of analyzing this data in the table below:

**Table 1.** Parameter estimation by using maximum likelihood estimator in MATLAB

| Iteration | Estimation | Parameters |
|-----------|------------|------------|
| index     | (MLE)_1    | X_1        |
| 1         | 4362.75    | 502.17     |
|           |            | 307.49     |
|           |            | 2935.4     |
|           |            | 0.031      |
| 8         | 3          | 5          |
| 5         |            | 5          |
| 2         | 1078.69    | 513.81     |
|           |            | 34.394     |
|           |            | 645.16     |
|           |            | 7.73       |
|           |            | 0.295      |
|           |            | 13.22      |
| 5         | 19.93      | 0.031      |
| 6         | 24.47      |            |
| 1         |            |            |

In the table, parameters of TLG-E \( \alpha_1 = X_1 \), \( \lambda = X_2 \), \( \beta = X_3 \), as for the rest of the distributions, then \( \alpha = X_1 \), \( \beta = X_2 \), \( \theta = X_3 \). In the MLE we look for parameters that make the maximum likelihood function large possible this means choosing the values for the parameters that appeared in the first row because it made the value of the (MLE)_1 large possible, and therefore, \( \alpha = 19.935 \), \( \lambda = 0.031 \), \( \beta = 24.472 \). As for the rest of the functions (MLE)_3 and (MLE)_4 we also choose the parameters that make them the largest possible as noted in the fourth row of table (1). That is \( \alpha = 7.647 \), \( \beta = 0.457 \), \( \theta = 21.06 \).

### 7. Results

Parameter values were found by estimating using the maximum likelihood estimation with data (1) we will compare the models Weibull(W), Exponentiated Weibull(EW) and Exponentiated Kumaraswamy(EK) with (TLG-E) and find the fit using the scales which are (AIC), (CAIC (HQIC) and (BIC) also known as Schars Information Criterion (SIC).

**Table 2.** Statistical Criteria

| Model     | AIC   | BIC   | CAIC  | HQIC  |
|-----------|-------|-------|-------|-------|
| TLG-E     | 533.7223 | 523.0868 | 533.3223 | 531.1709 |
| \( \lambda, \alpha, \beta \) |       |       |       |       |
In table (2) it is clear that the model (TLG-E) contains the lowest values according to the criteria mentioned in the table, this results in the distribution getting the best fit. Thus, it can be chosen as the most suitable for data analysis.

8. Conclusion
In the TLG-E model, some statistical and mathematical properties were mentioned, as well as the suitability of the data for a natural phenomenon with some of the indicated distributions. This was appropriate through the use of some statistical measures. It concluded that the (TLG-E) model belongs to the family of natural distributions and is possible Selecting appropriate data as well as applying it in several fields, including economics, biology, health, and others.

References

[1] Chipepa, F., & Makubate, B. O. B. (2020). The Topp-Leone Marshall-Olkin-G Family of Distributions with Applications. International Journal of Statistics and Probability, 9(4).

[2] Elgarhy M, Nasir M. A, Jamal F. & Ozel G. (2018) The type II Topp-Leone generated family of distributions Properties and applications: Journal of Statistics and Management Systems, 21:8, 1529-1551, DOI: 10.1080/09720510.2018.1516725.

[3] M. Arshad & Q. Azhad Jamal (2018): Interval Estimation for Topp-Leone Generated Family of Distributions based on Dual Generalized Order Statistics, American Journal of Mathematical and Management Sciences, DOI: 10.1080/01966324.2018.1509033.

[4] Nanthaprut, P., Patummasut, M., & Bodhisuwan, W. (2018). Topp-Leone generalized Rayleigh distribution and its applications. Songklanakarin Journal of Science & Technology, 40(5).

[5] Nadarajah, S., & Kotz, S. (2003). Moments of some J-shaped distributions. Journal of Applied Statistics, 30(3), 311-317.

[6] Nick T. Thomopoulos (2018): Probability Distributions with Truncated, Log and Bivariate Extensions: Springer International Publishing AG, part of Springer Nature.

[7] O'Connor, A. N. (2011). Probability distributions used in reliability engineering. RIAC.

[8] Rezaei, S., Sadr, B. B., Alizadeh, M., & Nadarajah, S. (2017). Topp–Leone generated family of distributions: Properties and applications. Communications in Statistics-Theory and Methods, 46(6), 2893-2909.

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