New Topp Leone-G Family with Mathematical Properties and Applications

A S Hassan¹, E A El-Sherpieny¹, S A El-Taweel²

¹Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt
²Faculty of Commerce, Al-Azhar University, Dakahlia, Egypt

h_s_elsafty@hotmail.com

Abstract. In the last few years, the Topp-Leone-G family has been regarded with various statistical applications. In this paper, efforts have been made to introduce a new family of distribution as an alternative to Topp-Leone-G distribution through a useful transformation with flexible hazard rate and greater reliability which we call new Topp-Leone-G (NTL-G) distribution. Some new special models and some statistical properties such as, expansions, probability weighted moments, moments and Rényi entropy are provided. The maximum likelihood method is used for estimating the model parameters. For the applied side, new models can be created in view of fitting data set. This last point is illustrated by consideration of the exponential distribution as baseline. A real data set is employed which represent the observations of ordered failure of some components. The obtained results show that, the NTL-G family is very competitive in comparison to other well implanted general families.

1. Introduction

Topp Leone (TL) distribution [1] is an attractive model for life testing and reliability studies, it is available to models bathtub shaped hazard rates and it is considering an alternative to beta distribution. The TL distribution is an appropriate for modeling lifetime of distributions with finite support and it has great importance in applications in many fields such as biological science, engineering, medicine, and others.

The cumulative distribution function (cdf) of TL distribution with one shape parameter α is given by

\[ G(t;\alpha) = (t(2-t))^{\alpha}; 0 < \alpha < 1, 0 \leq t \leq 1. \] (1)

In correspondence to (1), the associated probability density function (pdf) is as follows

\[ g(t;\alpha) = 2\alpha(1-t)(t(2-t))^{\alpha-1}; 0 < \alpha < 1, 0 \leq t \leq 1. \] (2)

Based on the TL distribution, several new generated families have been provided by several authors. Our interest here, with bounded support families, some of the notable families are; beta-G (B-G) [2], Kumaraswamy-G [3], generalized B-G [4], TL-G [5], truncated Fréchet-G [6], truncated Weibull [7], truncated inverted Kumaraswamy [8], truncated Lomax-G [9], truncated power Lomax-G [10]. The B-G family has the following cdf:

\[ F(x) = \frac{1}{B(\alpha,\beta)} \int_0^{G(x;\alpha,\beta)} t^{\alpha-1}(1-t)^{\beta-1}dt; \alpha, \beta > 0, \] (3)

and the pdf corresponding to (3) can be written in the form
where $B(a^*, b^*)$ is the beta function and $g(x; \xi) = \partial G(x; \xi) / \partial x$ is the pdf of the baseline distribution.

Recently, a new transformation [11] given by

$$H(x; \xi) = \frac{1}{1 - e^{-H(x; \xi)}},$$

is the odds ratio.

This transformation is more flexible to generate new families of probability distributions. Due to the importance of the transformation $1 - e^{-H(x; \xi)} \in [0, 1]$ and applicability of TL distribution, we introduce a more flexible family, called NTL-G family as an alternative to TL-G family. The main motivation of the NTL-G family is improving better fit than the families of distributions which having the same or high number of parameters.

2. NTL-G Family

In this section, we define the cdf, pdf, hazard rate function (hrf) and quantile function for the NTL-G family with shape parameter $\alpha$. Consider the pdf (2) as a generator and we take the upper limit to be $1 - e^{-H(x; \xi)}$, then:

$$F(x; a, \xi) = \int_0^1 2(1-t)(1-(1-t))^a dt = \left[1 - e^{-2H(x; \xi)}\right]^a, x \in R. \quad (4)$$

The pdf of the NTL-G family is

$$f(x; a, \xi) = \frac{2g(x; \xi)e^{-2H(x; \xi)}\left[1 - e^{-2H(x; \xi)}\right]^{a-1}}{\left[1-G(x; \xi)^a\right]^2 \left[1-(1-e^{-2H(x; \xi)})\right]^a}, x \in R. \quad (5)$$

From now on, a random variable $X$ has pdf (5) will be denoted by $X \sim \text{NTL-G}(a, \xi)$. The hrf for the NTL-G family is given by:

$$h(x; a, \xi) = \frac{2g(x; \xi)e^{-2H(x; \xi)}\left[1 - e^{-2H(x; \xi)}\right]^{a-1}}{\left[1-G(x; \xi)^a\right]^2 \left[1-(1-e^{-2H(x; \xi)})\right]^a}.$$  

Furthermore, the quantile function of the NTL-G family, say $Q(u) = F^{-1}(u)$ is given by

$$Q(u) = G^{-1}\left[\ln(1-u^{1/\alpha})^{-1/2}/(1-\ln(1-u^{1/\alpha})^{1/2}\right], u \in [0,1]. \quad (6)$$

To obtain the first, the second, and the third quartiles, $u = 0.25, 0.5$ and 0.75 must be put in (6) respectively.

3. Special Models

Two special distributions of the NTL-G family are defined and described, namely, NTL-Kumaraswamy (NTL-Kw), and NTL-Exponential (NTL-E) in this section.

3.1. NTL-Kw Distribution

Let us consider the Kumaraswamy (Kw) distribution with the pdf and the cdf given, respectively by

$$g(x; a_1, b_1) = a_1b_1x^{a_1-1}(1-x^{b_1-1}); 0 < x < 1, a_1, b_1 > 0,$$

$$G(x; a_1, b_1) = 1 - (1-x^{b_1})^{a_1}.$$  

The NTL-Kw distribution with three parameters $a, a_1$ and $b_1$, denoted by NTL-Kw$(a, a_1, b_1)$. The cdf and pdf of the NTL-Kw are obtained, respectively, by the following

$$F(x; a, a_1, b_1) = \left[1 - e^{-\left(\xi\left(1-x^{b_1-1}\right)\right)}\right]^a; 0 < x < 1, a, a_1, b_1 > 0. \quad (7)$$

The pdf corresponding to (7) is as follows
where \( c_2(x) = 1 - x^a \). Moreover, the hrf for the NTL-Kw distribution is as follows

\[
h(x; a, a, b) = \frac{2a b x^{a-1} e^{-2((c_2(x))^{b-1})} \left(1 - e^{-2((c_2(x))^{b-1}) a^2 - 1}\right)}{(c_2(x))^{b+1}}.
\]

Figure 1 shows plots of the pdf for the NTL-Kw distribution are unimodal, reversed J-shaped, and right skewed. Also, the hrf can be U-shaped and J-shaped.

3.2. NTL-E Distribution

Suppose that the exponential (E) distribution has cdf given by \( G(x; \alpha) = 1 - e^{-\alpha x}; \alpha > 0, x > 0 \), and pdf given by \( g(x; \alpha) = \alpha e^{-\alpha x}; \alpha > 0, x > 0 \), the cdf and pdf of NTL-E distribution with the shape parameter \( a \) and scale parameter \( \alpha \), denoted by NTL-E \((a, \alpha)\), are obtained respectively as following

\[
F(x; a, \alpha) = (1 - e^{-2(\alpha x - \alpha - 1)^a})^a; x > 0, a, \alpha > 0,
\]

and the pdf corresponds (8) is

\[
f(x; a, \alpha) = 2a \alpha e^{\alpha x - 2(\alpha x - \alpha - 1)^a}(1 - e^{-2(\alpha x - \alpha - 1)^a})^{a-1}; x > 0, a, \alpha > 0.
\]

The hrf for the NTL-E distribution will be as follows:

\[
f(x; a, \alpha) = \frac{2a \alpha e^{2(\alpha x - \alpha - 1)^a}(1 - e^{-2(\alpha x - \alpha - 1)^a})^{a-1}}{1 - (1 - e^{-2(\alpha x - \alpha - 1)^a})^a}; x > 0, a, \alpha > 0.
\]

Figure 2 shows that the plots of the pdf for the NTL-E distribution are decreasing and unimodal. Also, the hrf can be decreasing, increasing and decreasing and then increasing.
4. Some General Statistical Properties

4.1. Explicit Expansions

The generalized binomial theorem, for $a > 0$ is a real non-integer and $|z| < 1$, is written as

$$[1-z]^a = \sum_{i=0}^{\infty} \binom{a}{i} z^i.$$  \hspace{1cm} (10)

By applying (10) for $(1-e^{-2aG(x;\xi)})^{-1}$ in (5), gives

$$f(x;a,\xi) = \frac{2ag(x;\xi)}{(1-G(x;\xi))} \sum_{i=0}^{\infty} \binom{a-1}{i} (e^{-2aG(x;\xi)})^i.$$ \hspace{1cm} (11)

Using the power series;

$$e^{-2aG(x;\xi)} = \sum_{j=0}^{\infty} \binom{2j}{j} \frac{2j+1}{j!} (H(x;\xi))^j.$$ \hspace{1cm} (12)

where $H(x;\xi) = \frac{G(x;\xi)}{1-G(x;\xi)}$. Inserting (12) in (11) and using the generalized binomial theorem for $(1-G(x;\xi))^{-j-2}$, then $f(x;a,\xi)$ can be expressed as follows

$$f(x;a,\xi) = \sum_{j,k=0}^{\infty} \binom{2j+1}{j} \binom{a}{i} \frac{a(k+1)}{j!} g(x;\xi) G(x;\xi)^k.$$ \hspace{1cm} (13)

The pdf of the NTL-G in (13), can be expressed as exponentiated-$G$’s infinite linear combination density function, that is,

$$f(x;a,\xi) = \sum_{j,k=0}^{\infty} \omega_{j,k} g(x;\xi) (G(x;\xi))^k.$$ \hspace{1cm} (14)

where, $\omega_{j,k} = \binom{a}{j} \frac{a(k+1)}{j!}$. Furthermore, an expansion of the $(F(x;a,\xi))^m$ where $m$ is an integer, yields

$$(F(x;a,\xi))^m = \sum_{i=0}^{\infty} \sum_{a=0}^{\infty} \sum_{p=0}^{\infty} S_{i,p} (G(x;\xi))^e.$$ \hspace{1cm} (15)

$$S_{i,p} = \binom{am}{l} \binom{m}{p}.$$
4.2. The Probability
Replacing \( m \) with \( v \) in (15), the PWMs of NTL-G family is obtained as follows

\[
O_{r,v} = \sum_{i,j,k=0}^{\infty} \sum_{p=0}^{\infty} S_{i,j,p} \Phi_{i,j,k} \int_{-\infty}^{\infty} x^r g(x;\xi)(G(x;\xi))^{j+k+p} \, dx.
\]

Then,

\[
O_{r,v} = \sum_{i,j,k=0}^{\infty} \sum_{p=0}^{\infty} S_{i,j,p} \Phi_{i,j,k} O_{r,j+k+p},
\]

where \( O_{r,j+k+p} = \int_{-\infty}^{\infty} x^r g(x;\xi)(G(x;\xi))^{j+k+p} \, dx \) is the PWMs of the baseline distribution.

4.3. Moments
If \( X \) has the pdf (14), the \( r^{th} \) non-central moment of NTL-G is obtained as follows,

\[
\mu_r = \sum_{i,j,k=0}^{\infty} \Phi_{i,j,k} \int_{-\infty}^{\infty} x^r g(x;\xi)(G(x;\xi))^{j+k+p} \, dx = \sum_{i,j,k=0}^{\infty} \Phi_{i,j,k} O_{r,j+k+p},
\]

where \( O_{r,j+k+p} = \int_{-\infty}^{\infty} x^r g(x;\xi)(G(x;\xi))^{j+k+p} \, dx \) is the PWMs of the baseline distribution,

\[
\phi_r(t) = \sum_{i,j,k=0}^{\infty} \Phi_{i,j,k} \int_{-\infty}^{\infty} x^r g(x;\xi)(G(x;\xi))^{j+k+p} \, dx.
\]

4.4. Rényi Entropy
A measure of uncertainty or variation can be defined as the entropy of a random variable \( X \) in many fields such as economics, engineering, and physics. The Rényi entropy [13] is defined by

\[
R_\rho(X) = \frac{1}{1-\rho} \log \left[ \int_{-\infty}^{\infty} f(x)^\rho \, dx \right], \quad \rho > 0 \text{ and } \rho \neq 1.
\]

Therefore, the Rényi entropy of NTL-G family of distributions is given by

\[
R_\rho(X) = (1-\rho)^{-1} \log \left[ \sum_{i,j,k=0}^{\infty} \Phi_{i,j,k} \int_{-\infty}^{\infty} g(x;\xi)^\rho (G(x;\xi))^{j+k+p} \, dx \right],
\]

where \( \Phi_{i,j,k} = \rho^p (1-\rho)^{j+k+1} \).

5. Maximum Likelihood Method
The log-likelihood can be expressed from (5) as follows:

\[
\ln L(\Theta) = n \ln(2\alpha) + \sum_{i,j} \ln(g(x_i;\xi)) - 2\sum_{i,j} \ln(1-G(x_i;\xi)) - 2\sum_{i,j} H(x_i;\xi) + (\alpha - 1) \sum_{i,j} \ln(1-e^{-2H(x_i;\xi)}).
\]

Differentiating log-likelihood with respect to the vector of parameter, the following gives the elements of the score function \( U(\Theta) = (U_a, U_s) \)

\[
U_a = \frac{n}{\alpha} + \sum_{i,j} \ln(1-e^{-2H(x_i;\xi)}),
\]

and

\[
U_s = \sum_{i,j} \frac{\partial \phi(x_i;\xi)}{\partial g(x_i;\xi)} + \sum_{i,j} \frac{\partial G(x_i;\xi)}{\partial g(x_i;\xi)} - \sum_{i,j} \frac{\partial H(x_i;\xi)}{\partial g(x_i;\xi)} + 2(\alpha-1) \sum_{i,j} \frac{(\partial H(x_i;\xi))}{1-e^{-2H(x_i;\xi)}} \frac{(\partial \phi(x_i;\xi))}{1-e^{-2H(x_i;\xi)}},
\]

where \( \partial H(x_i;\xi)/\partial g(x_i;\xi) = \frac{\partial g(x_i;\xi)}{G(x_i;\xi)} \cdot \frac{\partial \phi(x_i;\xi)}{G(x_i;\xi)} \).
6. Simulation Study \( n = 10, 30, 50, 100 \)

| Groups | N  | Initial values | MLEs   | Bias  | Variance | MSE   |
|--------|----|----------------|--------|-------|----------|-------|
|        | 10 | \( a = 1 \)    | 1.325  | 0.325 | 0.058    | 0.164 |
|        |    | \( \alpha = 1 \) | 1.176  | 0.176 | 0.013    | 0.044 |
|        | 30 | \( a = 1 \)    | 1.08   | 0.08  | 0.002    | 0.009 |
|        |    | \( \alpha = 1 \) | 1.051  | 0.051 | 9.728e-04 | 0.004 |
|        | 50 | \( a = 1 \)    | 1.051  | 0.051 | 6.862e-04 | 0.003 |
|        |    | \( \alpha = 1 \) | 1.032  | 0.032 | 3.269e-04 | 0.001 |
|        | 100| \( a = 1 \)    | 1.203  | 0.023 | 1.53e-04  | 7.003e-04 |
|        |    | \( \alpha = 1 \) | 1.015  | 0.015 | 7.456e-05  | 2.899e-04 |
|        | 10 | \( a = 0.2 \)  | 0.509  | 0.31  | 0.005    | 0.101 |
|        |    | \( \alpha = 0.5 \) | 1.348  | 0.866 | 0.086    | 0.848 |
|        | 30 | \( a = 0.2 \)  | 0.429  | 0.229 | 2.422e-04 | 0.053 |
|        |    | \( \alpha = 0.5 \) | 0.976  | 0.476 | 0.004    | 0.23  |
|        | 50 | \( a = 0.2 \)  | 0.413  | 0.213 | 7.804e-05 | 0.045 |
|        |    | \( \alpha = 0.5 \) | 0.905  | 0.405 | 0.001    | 0.165 |
|        | 100| \( a = 0.2 \)  | 0.403  | 0.203 | 1.821e-05 | 0.041 |
|        |    | \( \alpha = 0.5 \) | 0.86   | 0.36  | 2.611e-04 | 0.13  |
|        | 10 | \( a = 0.5 \)  | 0.638  | 0.138 | 0.009    | 0.028 |
|        |    | \( \alpha = 1 \) | 1.317  | 0.317 | 0.044    | 0.144 |
|        | 30 | \( a = 0.5 \)  | 0.536  | 0.036 | 4.458e-04 | 0.002 |
|        |    | \( \alpha = 1 \) | 1.086  | 0.086 | 0.002    | 0.009 |
|        | 50 | \( a = 0.5 \)  | 0.52   | 0.02  | 1.446e-04 | 5.629e-04 |
|        |    | \( \alpha = 1 \) | 1.049  | 0.049 | 5.94e-04 | 0.003 |
|        | 100| \( a = 0.5 \)  | 0.511  | 0.011 | 3.399e-05 | 1.506e-04 |
|        |    | \( \alpha = 1 \) | 1.026  | 0.026 | 1.361e-04 | 7.906e-04 |

From the table 1, it is observed that, for the three groups I, II and III, the variance and MSE of the estimated parameters converge to zero as \( n \) increases.

7. Application
This section provides the application to real data set to illustrate the applicability of the NTL-E distribution as a particular case from the suggested family. The NTL-E distribution with pdf (9), is compared with Topp Leone- Exponential (TL-E) [5], new Kumaraswamy-Exponential (NKw-E) [11], Kumaraswamy-Exponential (Kw-E) [14], Type II generalized Topp Leone- Exponential (TII GTL-E) [15], generalized-exponential (G-E) [16] and exponential distributions.

The "best" distribution corresponds to the largest value of p-value criteria and the smallest value of K-S and smallest value (positive or negative signe) of -2lnL, AIC, BIC, CAIC and HQIC.

Data set: The data proposed that 20 items are put on test simultaneously and their ordered failure times are noted [17].
Table 2 gives the MLEs and standard error (S.E) of the model parameters for the dataset. While Table 3 gives the values of measurements for the data set.

### Table 2. The data set’s MLEs and S.E for parameters the model

| Model     | MLEs and (S.E) |  |  |  |
|-----------|----------------|---|---|---|
|           | \(a\)          | \(b\) | \(\alpha\) |
| NTL-E     | 0.674 (0.17)   | -  | 1.758 (0.385) |
| TL-E      | 0.793 (0.221)  | -  | 2.664 (0.824)  |
| NKw-E     | 0.677 (0.247)  | 1.479 (1.709) | 2.063 (1.666) |
| Kw-E      | 0.632 (0.411)  | 0.212 (0.462) | 26.826 (60.634) |
| TIIGTL-E  | 0.326 (0.202)  | 0.25 (0.553)  | 22.586 (51.884) |
| G-E       | 0.793 (0.221)  | -  | 5.329 (1.648)  |
| E         | 6.201 (1.387)  | -  | -               |

### Table 3. The data set’s Model Selection Criteria

| Model     | -2lnL | AIC  | CAIC | BIC  | p-value | K-S  | HQIC |
|-----------|-------|------|------|------|---------|------|------|
| NTL-E     | -34.541 | -30.541 | -29.836 | -28.55 | 0.998  | 0.088 | -30.153 |
| TL-E      | -33.722 | -29.722 | -29.016 | -27.73 | 0.925  | 0.122 | -29.333 |
| NKw-E     | -34.543 | -28.543 | -27.043 | -25.556 | 0.996  | 0.091 | -27.96 |
| Kw-E      | -34.036 | -28.036 | -26.536 | -25.049 | 0.967  | 0.111 | -27.453 |
| TIIGTL-E  | -34.017 | -28.017 | -26.517 | -25.03 | 0.962  | 0.113 | -27.434 |
| G-E       | -33.722 | -29.722 | -29.016 | -27.73 | 0.925  | 0.122 | -29.333 |
| E         | -32.989 | -28.989 | -28.284 | -26.998 | 0.987  | 0.101 | -28.601 |
The values illustrated in Table 3 show the NTL-E model has the lowest values of $-2\ln L$, CAIC, BIC, AIC, HQIC and K-S, and highest p-values among the real data set’s all fitted models. The NTL-E model can be chosen as the best model according to this. Figure 3 clears out that the best fits to real data set is provided by NTL-E distribution. Figure 3 confirm the result of Table 3.

8. Conclusion

Present paper introduces New Topp Leone-G family as a new family of probability distributions. Family’s several structural properties, such as, expressions, probability weighted moments, moments and Rényi entropy are investigated. To estimate parameters, the method of maximum likelihood is employed. As shown by one of the applications of real data set, the new Topp Leone-Exponential distribution is a stronger competitor for the corresponding other distributions.

References

[1] Topp C W and Leone F C 1955 A family of J-shaped frequency functions Journal of American Statistical Association 50 pp 209-219
[2] Eugene N, Lee C and Famoye F 2002 Beta-normal distribution and its applications Communication in Statistics- Theory and Methods 31 (4) pp 497–512
[3] Cordeiro G M and de Castro M 2011 A new family of generalized distribution Journal of Statistical Computation and Simulation 81 (7) pp 883-898
[4] Alexander C, Cordiero G M, Ortega E M M and Sarabia J M 2012 Generalized beta-generated distributions Computational statistics and data analysis 56 (6) pp 1880-1896
[5] Al-Shomrani A, Arif O, Shawky K, Hanif S and Shhabaz M Q 2016 Topp–Leone family of distributions: some properties and application Pak.j.Stat.Oper.Res XII (3) pp 443-451
[6] Abid S H and Abdulrazazk R K 2017 [0,1] truncated Fréchet-G generator of distributions Applied Mathematics 7 pp 51-66
[7] Najarzadegan H, Alamatsaz M H and Hayati S 2017 Truncated Weibull-G more flexible and more reliable than beta-G distribution International Journal of Statistics and Probability 6 (5) pp 1-17
[8] Bantan R A, Jamal F, Chesneau C and Elgarhy M 2019 Truncated inverted Kumaraswamy generated family of distributions with applications Entropy 21 (11) pp 1-22
[9] Hassan A S, Sabry M A H and Elsehery A M 2020a A New Probability Distribution Family Arising from Truncated power Lomax Distribution with Application to Weibull Model Pakistan Journal of Statistics and Operation Research 16 (4) pp 661-674
[10] Hassan A S, Sabry M A H and Elsehery A M 2020b A new family of upper-truncated distributions: properties and estimation Thailand Statistician 18 (2) pp 196-214
[11] El-Taweel S A 2019 New Transformation for generating families of continuous distributions with Application to Weibull distribution *The 54-th, Annual Conference on Statistics & Computer Science and Operations Research, FGSSR- Cairo University, Egypt* pp 167-184

[12] Greenwood J A, Landwehr J M and N C Matalas 1979 Probability weighted moments: Definitions and relations of parameters of several distributions expressible in inverse form *Water Resources Research* **15** pp 1049-1054

[13] Renyi A 1961 On measures of entropy and information *Proceedings of the 4th Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley* pp 547-561

[14] Cordeiro G M, Ortega E M M and Nadarajah S 2010 The Kumaraswamy Weibull distribution with application to failure data *Journal of the Franklin Institute-Engineering and Applied Mathematics* **347** pp 1399-1429

[15] Hassan A S, Elgarhy M. and Ahmad Z 2019 Type II Generalized Topp-Leone family of distributions: Properties and applications *Journal of Data Science* **17** (4) pp 638-659

[16] Gupta R D and Kundu D 1999 Generalized exponential distribution. *Australian and New Zealand Journal of Statistics* **41** (2) pp 173-188

[17] Ngim A M, Al-Hussaini E K and Jaheen Z F 2003 Bayesian one-sample prediction of future observations under Pareto distribution *Statistics* **37** (6) pp 527-536