REAL DATA APPLICATIONS OF TOPP-LEONE GENERATEDFAMILY OF DISTRIBUTIONS

Ruby Chanchal[1], Vaishali Gupta [1,*], Amit Kumar Misra [1]
[1]Department of Statistics
School of Physical & Decision Sciences
BabasahebBhimraoAmbedkar University, Lucknow - 226 025
*Corresponding Author: vaishali.gupta3091@gmail.com

Abstract
In this article, we compare the fits of three members of Topp-Leone generated family of distributions, namely, the Topp-Leone generated log-logistic distribution, the Topp-Leone generated Lomax distribution, and the Topp-Leonegenerated Weibull distribution with the help of two real data sets. All the computational studies are performed using R-software.
Keywords: Topp-Leone generated family; AIC; BIC; AICC; KS statistic.

1. Introduction

The Topp-Leone generated ($\text{TL-}G$) family of distributions was proposed by [1]. This distribution has the property to model bathtub shaped hazard rate functions for all $\nu \in (0,1)$. It is also used for modeling the lifetime phenomena. Let usconsider the probability density function (p.d.f.) of the $\text{TL-}G$ family of distributionswith parameters ν and θ given by

$$f(u;\nu,\theta,\varepsilon) = 2\nu\theta z(u;\varepsilon) Z(u;\varepsilon)^{\theta\nu-1}\left(1-Z(u;\varepsilon)^{\theta}\right)\left(2-Z(u;\varepsilon)^{\theta}\right)^{\nu-1}, \quad u \in \mathbb{R};\ \theta,\nu>0, \tag{1.1}$$

and the cumulative distribution function (c.d.f.) given by

$$F(u;\nu,\theta,\varepsilon) = \left(Z(u;\varepsilon)^{\theta}\left(2-Z(u;\varepsilon)^{\theta}\right)\right)^{\nu}, \quad u \in \mathbb{R};\ \theta,\nu>0.$$

Where $Z(u;\varepsilon)$ and $z(u;\varepsilon)$ denote the c.d.f. and the p.d.f. of the base-line distribution,respectively. Also, ε includes the parameters which specify the base-line distribution.

For convenience, a random variable U with p.d.f. given in Equation (1.1), we use $U \sim \text{TL-}G(\nu,\theta,\varepsilon)$. In particular, various cases of $\text{TL-}G(\nu,\theta,\varepsilon)$ can be obtained by using $Z(u;\varepsilon)$ as a parent distribution function such as exponential, normal, gamma, and weget TL-exponential, TL-normal, and TL-gamma distributions, respectively. In recentyears, some authors introduced and studied several such distributions by considering different $Z(u;\varepsilon)$, see, for example, [2], [3], and [4].

In this article, we choose log-logistic and Lomax (see, [5]) distributionsas base-line distributions with $Z(u;\beta,\lambda) = \frac{(\lambda u)^{\beta}}{1+(\lambda u)^{\beta}}$ and $Z^{*}(u;\beta,\lambda) = 1-(1+\lambda u)^{-\beta}$,respectively, and get Topp-Leone generated log-logistic (TL-log logistic) and Topp-Leone generated Lomax (TLGLo) distributions, respectively. The p.d.f. of the TL-loglogistic distribution and the TLGLo distribution are given by

$$f_U(u) = 2\nu\theta\beta\lambda^{\beta\theta\nu}\frac{u^{\beta\theta\nu-1}}{\left[1+(\lambda u)^{\beta}\right]^{1+\theta\nu}}\left[1-\left(\frac{(\lambda u)^{\beta}}{1+(\lambda u)^{\beta}}\right)^{\theta}\right]\left[2-\left(\frac{(\lambda u)^{\beta}}{1+(\lambda u)^{\beta}}\right)^{\theta}\right]^{\nu-1},$$

$$u>0, \nu>0, \theta>0, \beta>0, \lambda>0,$$

and

$$f_U^{*}(u)=2\nu\theta\beta\lambda(1+\lambda u)^{-(\beta+1)}\left[1-(1+\lambda u)^{-\beta}\right]^{\theta\nu-1}\left[1-\left(1-(1+\lambda u)^{-\beta}\right)^{\theta}\right]\left[2-\left(1-(1+\lambda u)^{-\beta}\right)^{\theta}\right]^{\nu-1},$$
$$u>0,\nu>0,\theta>0,\beta>0,\lambda>0,$$

respectively, and their corresponding c.d.f.'s are

$$F_U(u)=\left[\frac{(\lambda u)^{\beta}}{1+(\lambda u)^{\beta}}\right]^{\theta\nu}\left[2-\left(\frac{(\lambda u)^{\beta}}{1+(\lambda u)^{\beta}}\right)^{\theta}\right]^{\nu},\quad u>0,\nu>0,\theta>0,\beta>0,\lambda>0,$$

and

$$F_U^{**}(u)=\left(1-e^{-(\lambda u)^{\beta}}\right)^{\nu\theta}\left[2-\left(1-e^{-(\lambda u)^{\beta}}\right)^{\theta}\right]^{\nu},\quad u>0,\nu>0,\theta>0,\beta>0,\lambda>0.$$

To discuss the importance of the TL-G family of distributions, we compare the fitting of the above mentioned distributions under some criteria. For this comparison,we consider two real data sets defined in Section 2. Also, we estimate the unknownparameters of each model by using the method of maximum likelihood (ML) andevaluate the log-likelihood function for each model.

2. Applications based on real data sets

In this section, we evaluate two real data sets to compare the fits of TL-log logistic, TLGLo, and TLGW distributions. We choose some criteria to decide best fitteddistribution whichare Akaike information criterion (AIC), Bayesian information criterion (BIC), Akaike information criterion corrected (AICC), Kolmogorov-Smirnov (KS)statistic with its p -value.

The first data set contains observed number of hours spent repairing an airbornecommunications transceiver by [6], which was previously provided by [7]. The data are described below:
0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0, 24.5.

The second data set contains 63 observations of the gauge lengths of each of the10 mm recorded by [8]. The data are described below:
1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

A few descriptive statistics for both data sets are shown in Table 1.

Table 1: Descriptive statistics for the Data Sets 1 and 2

Descriptive statistics	Data Set 1	Data Set 2
Minimum	0.200	1.901
Median	1.750	2.996
Mean	3.607	3.059
Maximum	24.500	5.020
Standard deviation	4.944195	0.6209216

Further, we have drawn the probability-probability plots (PP-plots) for both data setsin Figures 1 and 2, from which we can easily observe that the models TL-log logistic,TLGW, and TLGLo provide a suitable fit to the data sets.

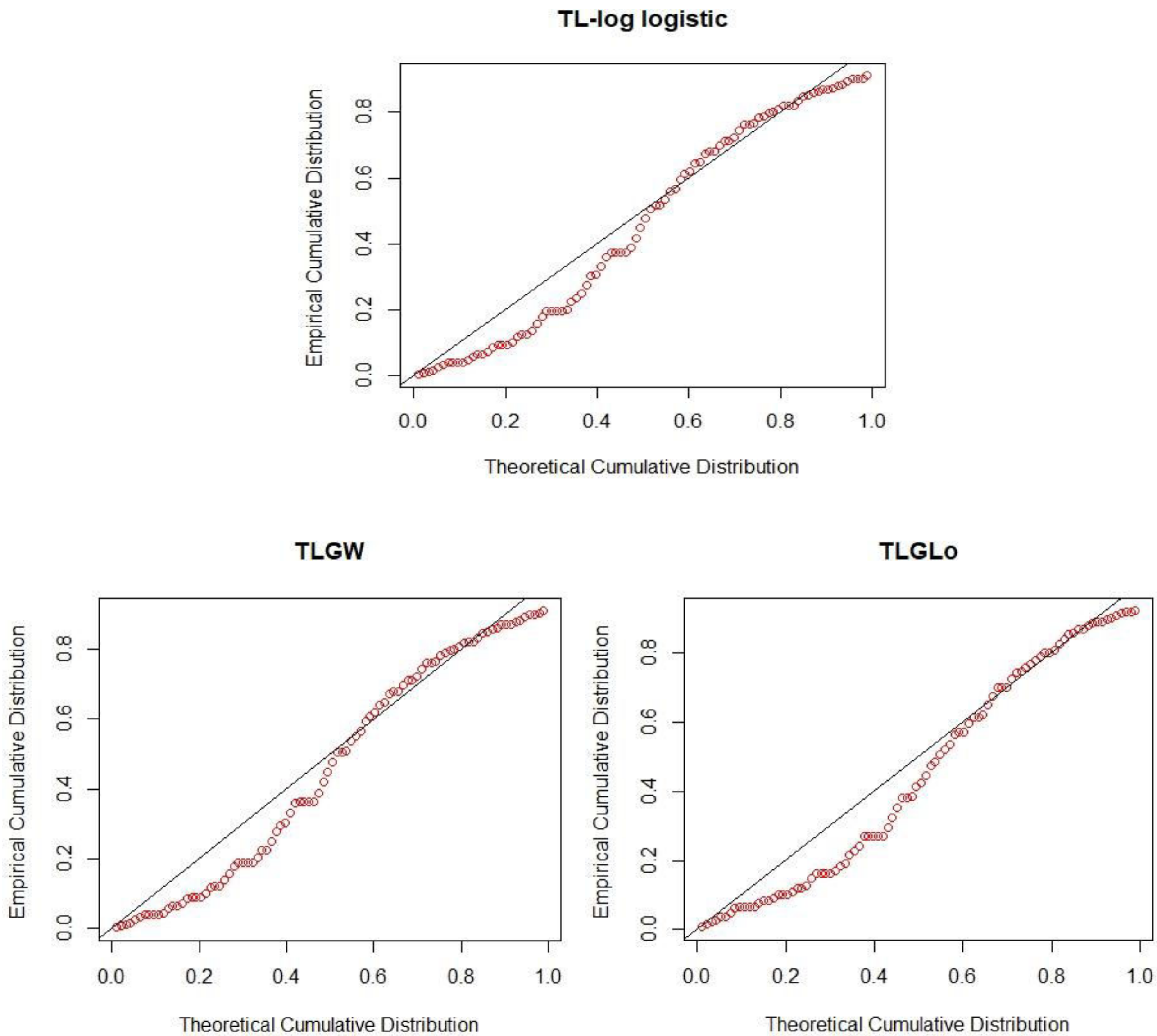

Figure 1:The PP-plots of the repairing time of airborne communications transceiver data for the TL-log logistic, TLGW, and TLGLo distributions

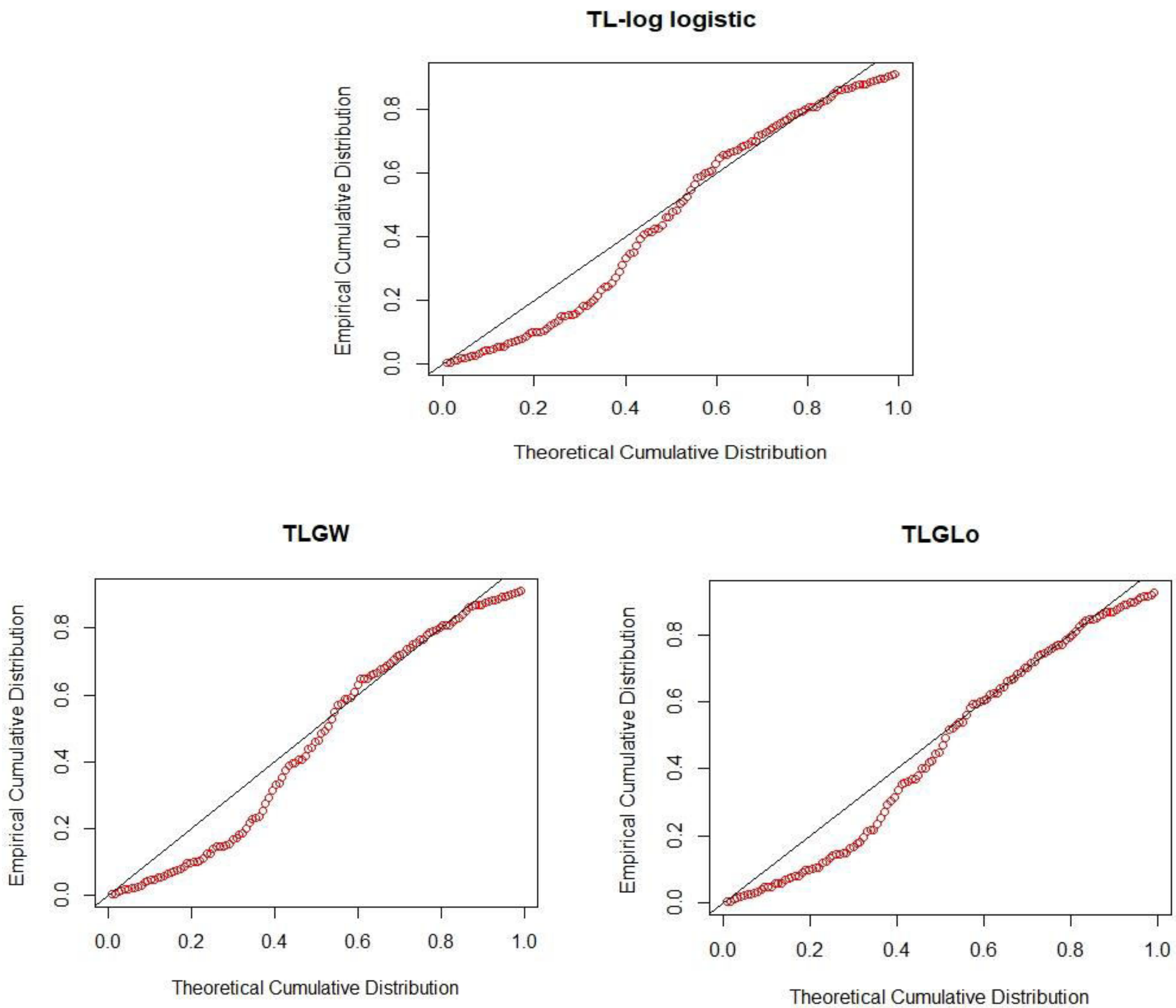

Figure 2: The PP-plots for the data of the gauge lengths for the TL-loglogistic, TLGW, and TLGLo distributions

In order to estimate the unknown parameters in each model, we use the method ofML estimation, and then to compare the results of the models, we use AIC, AICC, BIC,KS, and p-value. Also, we determine -log l (log-likelihood function) for each models (see, Tables 2 and 3). The strongest model meets the smallest AIC, BIC, AICC, andKS-statistic, and the highest p-value.

Table 2: ML-estimates and loglikelihood functions for the TL-log logistic, TLGW, and TLGLo models, and thestatistics AIC, AICC, BIC, and KS statistics with its p-values for repairing time of airborne communicationstransceiver

Models	ML-estimates	$-\log l$	AIC	BIC	AICC	KS statistic	p-values
TL-log logistic		–99.55465	**207.1093**	**214.4239**	**208.0849**	**0.086957**	**0.995**

	$\hat{\nu}=0.1048540$ $\hat{\theta}=92.3674740$ $\hat{\beta}$=0.9768692 $\hat{\lambda}$=7.4358103						
TLGW	$\hat{\nu}=3.013868$ $\hat{\theta}=20.536544$ $\hat{\beta}$=0.181533 $\hat{\lambda}$=571.376529	–99.81147	207.6229	214.9375	208.5985	0.1087	0.9487
TLGLo	$\hat{\nu}=3.539238e-04$ $\hat{\theta}=2.670313e+03$ $\hat{\beta}$=1.079512e+03 $\hat{\lambda}$=2.455555e−04	–104.972	217.944	225.2586	218.9196	0.15217	0.6612

Table 3: ML-estimates and log-likelihood functions for the TL-log logistic, TLGW, and TLGLo models, and the statistics AIC, AICC, BIC, and KS statistics with its p -values for the gauge lengths data

Models	**ML-estimates**	**$-\log l$**	**AIC**	**BIC**	**AICC**	**KS statistic**	**p - values**
TL-log logistic	$\hat{\nu}=0.1324324$ $\hat{\theta}=30.9651194$ $\hat{\beta}$=5.6684962 $\hat{\lambda}$=0.4442241	–56.44526	**120.8905**	**129.4631**	**121.5802**	**0.079365**	**0.9888**
TLGW	$\hat{\nu}=13.49923566$ $\hat{\theta}=0.02088991$ $\hat{\beta}$=28.27730026 $\hat{\lambda}$=0.20264199	–66.73314	141.4663	150.0388	142.1559	0.19048	0.2032

TLGLo	$\hat{\nu}=7.271080e-02$ $\hat{\theta}=1.395350e+03$ $\hat{\beta}=1.942276e+02$ $\hat{\lambda}=8.457492e-03$	-59.38903	126.7781	135.3506	127.4677	0.12698	0.69

In this analysis, it is easy to see that the TL-log logistic model is the best modelamong TLGW and TLGLo distributions for both data sets because TL-log logistic has the smallest values of AIC, AICC, BIC, KS statistic, and the highest p -value.

3. Conflicts of Interest

All the authors have equally participated towards the article and there is no potentialconflicts of interest.

References

1. Rezaei S, Sadr BB, Alizadeh M, Nadarajah S. (2017). Topp-Leone generated family of distributions: Properties and applications. Communications in Statistics - Theory and Methods. 2017; 46, 2893-2909.
2. Aryal GR, Ortega EM, Hamedani GG, Yousof, HM.The Topp-Leone generatedWeibull distribution: Regression model, characterizationsand applications. International Journal of Statistics and Probability. 2017; 6, 126-141.
3. Brito E, Cordeiro GM, Yousof, HM, Alizadeh M, Silva GO. The Topp-Leone odd log-logistic family of distributions. Journal of Statistical Computation and Simulation. 2017; 87, 3040-3058.
4. Korkmaz M, Yousof HM, Alizadeh M, Hamedani GG. (2019). The Topp-Leone generalized odd log-logistic family of distributions: Properties, characterizations and applications. Communications Faculty of Sciences University of Ankara Series A1- Mathematics and Statistics. 2019; 68, 1506-1527.
5. Lomax KS. Business Failures: Another example of the analysis of failure data. Journal of the American Statistical Association. 1954; 49, 847-852.
6. Balakrishanan N, Leiva V, Sanhuzea A, Cabrera E. Mixture inverse Gaussian distributions and its transformations, moments and applications. Statistics. 2009; 43, 91-104.
7. Chhikara RS, Folks JL. The inverse Gaussian distribution.Marcel Dekker, New York; 1989.
8. Kundu D, Gupta RD. Estimation of R = P(Y < X) for threeparameter Weibull distribution. Statistics and Probability Letters. 2009; 79, 1839-1846.