**Marquette University**

**e-Publications@Marquette**

***Mathematics and Statistical Faculty Research and Publications/College of Arts and Sciences***

***This paper is NOT THE PUBLISHED VERSION;* but the author’s final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation below.

*Communications for Statistical Applications and Methods*, Vol. 26 (2019): 475-493. [DOI](https://doi.org/10.29220/CSAM.2019.26.5.473). This article is © Korean Statistical Society and permission has been granted for this version to appear in [e-Publications@Marquette](http://epublications.marquette.edu/). Korean Statistical Society does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Korean Statistical Society.

A New Extension of Lindley Distribution: Modified Validation Test, Characterizations and Different Methods of Estimation

Mohamed Ibrahim

Department of Applied Statistics and Insurance, Faculty of Commerce, Damietta University, Egypt

Abhimanyu Singh Yadav

Department of statistics, Central University of Rajasthan, India

Haitham M. Yousof

Department of Statistics, Mathematics and Insurance, Benha University, Egypt

Hafida Goual

Laboratory of Probability and Statistics University of Badji Mokhtar, Annaba, Algeria

G.G. Hamedani

Department of Mathematics, Statistics and Computer Science, Marquette University, USA

# Abstract

In this paper, a new extension of Lindley distribution has been introduced. Certain characterizations based on truncated moments, hazard and reverse hazard function, conditional expectation of the proposed distribution are presented. Besides, these characterizations, other statistical/mathematical properties of the proposed model are also discussed. The estimation of the parameters is performed through different classical methods of estimation. Bayes estimation is computed under gamma informative prior under the squared error loss function. The performances of all estimation methods are studied via Monte Carlo simulations in mean square error sense. The potential of the proposed model is analyzed through two data sets. A modified goodness-of-fit test using the Nikulin-Rao-Robson statistic test is investigated via two examples and is observed that the new extension might be used as an alternative lifetime model.

# Keywords

different characterizations, validation test, different method of estimation and applications

# 1. Introduction, motivation and physical interpretation

Statistical literature contains a large number of probability distributions for modeling real lifetime data, the most popular probability distributions are gamma (Ga), lognormal (Log-N), Weibull (W), and exponentiated exponential (Exp-E) distributions. However, these probability models suffer from some drawbacks. First, none of them exhibit bathtub shapes for their hazard rate functions (hrfs), the four models exhibit only monotonically decreasing, monotonically increasing or constant hrfs and this is a major weakness since most real life systems exhibit bathtub shapes for their hrfs. Second, at least three of these distributions exhibit constant hazard rates and this is an unrealistic feature since few real life systems have constant hrf. This work introduces a new three parameter lifetime distribution as an alternative to the Ga, Log-N, W, and the Exp-E probability models that does not have the above mentioned drawbacks. The new model, called the Topp Leone Generated Lindley (TLGLi), is constructed based on the Topp Leone Generated (TLG) family introduced by [Rezaei et al. (2017)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B22). The TLGLi can have unimodal or monotonically decreasing or monotonically increasing or constant or bathtub hrf ([Figure 2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F2)).

The cumulative distribution function (cdf) and the probability density function (pdf) of a random variable (r.v.) X with the Li distribution ([Lindley, 1958](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B16)) are given by

and

respectively. The scale parameter is a positive real number and can result in a unimodal or monotone decreasing density. The Li model has a thin tail since the distribution decreases exponentially for large values of . The Li distribution is one way to describe the lifetime of a process or device having increasing failure rate. It can be used in a wide variety of fields such as biology, engineering and medicine. The cdf of the TLGLi model can be expressed as

(1.1)

and the corresponding pdf is

(1.2)

[Equations (1.1)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD3) and ([1.2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)) are established based on the TLG distribution.

Suppose are independent r.v.’s distributed according to

and represent the failure times of the components of a series system, assumed to be independent. Then the probability that the system will fail before time is given by

So, [Equation (1.1)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD3) is the distribution of the failure of a series system with independent components if . We shall refer to the new distribution in ([1.1](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD3)) as the TLGLi model. For , TLGLi reduces to the TLLi model. The cdf ([1.1](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD3)) can be expressed as

(1.3)

where

and

is the cdf of the Exp-Li distribution with power parameter . The corresponding density function is obtained by differentiating ([1.3](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD7))

(1.4)

where

is the Exp-Li density with power parameter .Thus, several of its structural properties can be obtained from [Equation (1.4)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD10) and those of the Exp-Li distribution.

We provide some plots of the pdf and hrf of the TLGLi model to show its flexibility. [Figure 1](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F1) displays some plots of the TLGLi density for selected values of , , and . These plots reveal that the new density can be right-skewed with different flexible shapes. The hrf plots of the TLGLi distribution given in [Figure 2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F2) can be unimodal, decreasing, bathtub, increasing and constant shapes.

In the literature, certain generalizations on the Li distribution are proposed and studied ([Ghitany et al., 2008a](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473" \l "B6), [b](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B7); [Deniz and Ojeda, 2011](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B5); [Ghitany et al., 2011](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B8); [Nadarajah et al., 2011](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B18); [Alizadeh et al., 2016](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B2), [2017](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B1)). This article shows how different estimators of the new distribution perform for different sample sizes and different parameter values and to promote a guideline for choosing the best estimation method for the new model, which we think would be of interest to applied statisticians. The unknown parameters of the new distribution are estimated using maximum likelihood, least squares, weighted least squares, Cramer-Von-Mises and Bayesian methods. The obtained estimators are compared using Monte Carlo simulations; in addition, it is observed that Bayesian estimators are more efficient compared to other estimators. The TLGLi distribution shows its suitability in modeling strength and relief times data sets.

The rest of the paper is outlined as follows: Certain characterizations of the TLGLi model are proposed in Section 2. In Section 3, we discuss some properties of this distribution. In Section 4, we describe four methods of estimation. In Section 5, the usefulness of the new distribution is illustrated by means of two real data sets. A modified goodness-of-fit test using a Nikulin-Rao-Robson (NRR) statistic test is presented in Section 6. In Section 7, a simulation study is carried out to compare the performance of the four methods of estimation. Section 8 offers some concluding remarks.

# 2. Characterizations

In this Section, we present a number of characterizations of the TLGLi distribution in the following directions: (i) in terms of the ratio of two truncated moments; (ii) in terms of the hazard function; (iii) based on the reverse hazard function; and (iv) in terms of the conditional expectation of certain function of the random variable. These characterizations are presented in four subsections.

## 2.1. Characterizations in terms of two truncated moments

Certain characterizations of TLGLi distribution based on a simple relationship between two truncated moments are presented. The first characterization employs a theorem due to [Glänzel (1987)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B9), see [Hamedani et al. (2018a](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B14); [2018b](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B15), Theorem 1) and in the [Appendix](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#APP1). However, the results holds also when the interval H is not closed since the condition of the Theorem is on the interior of H.

### **Proposition 1**

Let be a continuous random variable and let

and

The random variable has pdf ([1.2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)) only if the function defined in Theorem 1 is of the form

### **Proof**

Suppose the random variable X has pdf ([1.2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)), then

and

Further,

Conversely, if has the above form, then

and

Now, according to Theorem 1, has pdf ([1.2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)).

### **Corollary 1**

*Let be a continuous random variable and let be as given in Proposition 1. The random variable has pdf (*[*1.2*](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)*) only if there exist functions and defined in Theorem 1 satisfying the following differential equation*

### **Corollary 2**

*The general solution of the differential equation in Corollary 1 is*

*where D is a constant. One set of functions satisfying the above differential equation is given in Proposition 1 with D = 1/2. Of course, there are other triplets of function (q1, q2, ξ) which satisfy conditions of Theorem 1.*

## 2.2. Characterization based on hazard function

The hazard function, , of a twice differentiable distribution function, , satisfies the following differential equation

For many univariate continuous distributions, the above equation is the only differential equation available in terms of the hazard function. In this subsection we present a non-trivial characterization of TLGLi distribution, for , based on the hazard function.

### **Proposition 2**

*Let be a continuous random variable. The random variable has pdf (*[*1.2*](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)*), for , only if its hazard function satisfies the following first order differential equation*

### **Proof**

Is straightforward and hence omitted.

## 2.3. Characterization based on reverse hazard function

The reverse hazard function, , of a twice differentiable distribution function, , is defined as

In this subsection we present a characterization of TLGLi distribution, for , based on the reverse hazard function.

### **Proposition 3**

*Let be a continuous random variable. The random variable has pdf (*[*1.2*](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)*) for only if its reverse hazard function satisfies the following first order differential equation*

### **Proof**

Is straightforward and hence omitted.

## 2.4. Characterization based on the conditional expectation of certain function of the random variable

Here, we employ a function of and present characterize the distribution of in terms of the truncated moment of . The following proposition appeared in Hamedani’s previous work ([Hamedani, 2013](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B13)), so we just state it here and use it to characterize TLGLi distribution, for .

### **Proposition 4**

*Let be a continuous random variable with cdf F. Let be a differentiable function on with . Then for ,*

only if

### **Remark 1**

For , and , Proposition 4 provides a characterization of TLGLi distribution for λ = θ = β = 1.

# 3. Mathematical properties

The rth ordinary moment of is given by

(3.1)

where

and

The first four moments of are

The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships.

The , variance, skewness and kurtosis of the TLGLi distribution are computed numerically for selected values of , , and using the Mathcad program Version 15.0. The numerical values displayed in [Table 1](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T1) indicate that the skewness of the new model is always positive and can range in the interval (0.68, 3.18). The spread for its kurtosis is larger ranging from 3.77 to 16.92.

The moment generating function (mgf) of can be derived using [Equation (3.1)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD28) as

The characteristic function (cf) of , , and the cumulative generating function (cgf) of , re given by

and

respectively, where .

The sth incomplete moment, say , of can be expressed as

where

and is the incomplete gamma function

where is a confluent hypergeometric function. If is an integer, then can be simplified to

The nth moment of the residual life, say

The nth moment of the residual life of X is given by

Therefore

The nth moment of the reversed residual life can be expressed as

or

Then, the nth moment of the reversed residual life of becomes

The reliability, say , of the system is the probability that the system is strong enough to overcome the stress imposed on it. Let and be two independent r.v.’s with TLGLi and TLGLi distributions. Thus can be expressed as

where

Let be a random sample (r.s.) from the TLGLi distribution and let be the corresponding order statistics. The pdf of ith order statistic, , can be written as

(3.2)

where is the beta function. Substituting ([1.1](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD3)) and ([1.2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD4)) in [Equation (3.2)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD47), the pdf of can be expressed as

where

and can be obtained recursively from

where . Then, the qth moment of can be expressed as

# 4. Different methods of estimation

## 4.1. Maximum likelihood method

Let be a r.s. from the TLGLi distribution with observed values , and be the vector of the model parameters. The log likelihood function for Ψ may be expressed as

(4.1)

where

The maximum likelihood estimators (MLEs) of , and are obtained from differentiating [Equation (4.1)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD52) with respect to , , and

where

Setting them equal to zero and solving the system simultaneously yields the MLE of . The above equations cannot be solved analytically. So, statistical software can be used to solve them numerically using the iterative methods such as the Newton-Raphson type algorithms (for more details see [Casella and Berger (2002)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B4)).

## 4.2. Method of ordinary least square and weighted least square estimation

The theory of ordinary least square (OLS) estimation and weighted least square (WLS) estimation was originally proposed by [Swain et al. (1988)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B23) to estimate the parameters of the Beta distribution. This theory is based on minimizing the sum of the square of differences of theoretical cdf and empirical cdf.

Suppose that denotes the distribution function of TLGLi distribution and if be the ordered random sample. The OLS estimators (OLSE) of the parameters are obtained by minimizing

then

The OLSE of the parameters are obtained by solving the following non-linear equations

where , and are the values of the first derivatives of the cdf with respect to (w.r.t.) parameters of TLGLi model.

The OLSE of the parameters are obtained via solving the above simultaneous equations by using any numerical approximation techniques. The WLS estimates (WLSE) are obtained by minimizing the given form of equation w.r.t. the parameters

The WLSE of the parameters are obtained by solving the following non-linear equations

where , , and are the values of first derivatives of the cdf of TLGLi distribution and

## 4.3. Method of Cramer-Von-Mises estimation

The Cramer-Von- Mises estimation method of the parameters is based on the theory of minimum distance estimation ([MacDonald, 1971](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B17)). The Crammer-Von Mises estimates (CVME) of the parameter , , and are obtained by minimizing the following expression w.r.t. the parameters respectively.

and

The CVME of the parameters are obtained by solving the following non-linear equations

and

where , , and are the values of the first derivatives of the cdf of TLGLi distribution w.r.t. , , respectively.

## 4.4. Bayes estimation

Using the squared error loss function (SELF), the Bayes estimators are computed under informative gamma priors for all , , and . The joint prior is given by

where the hyperparameters , , , , , and assume to be known and positive and

and is the estimated value of . The posterior distribution is needed to derive the Bayes estimators. The posterior distribution of the TLGLi distribution can be written as

Under this loss function, posterior mean is the Bayes estimate of the respective parameter. Thus, the Bayes estimators under SELF are obtained as

where

respectively. The above equations cannot be solved analytically, thus we use Markov chain Monte Carlo (MCMC) technique to generate the posterior sample from the full conditional posterior distribution. The full conditional posterior distributions are given by

The following steps are used to extract the posterior samples from full-conditional posterior density

* Starts with and set initial values of , , say , , ;
* Generate posterior samples from full conditional distribution using normal distribution as a proposal density;
* Repeat the above step for and simulate , ;
* Under SELF, the Bayes estimates of , , are given by
* where, is the burn in period.

# 5. Data analysis

In this section, real data sets are used to demonstrate the real life applicability of the proposed model using MLE, LS, WLS, and CVM. We consider the Cramér-Von Mises (W\*), the Anderson-Darling (A\*), and the Kolmogorov-Smirnov (KS) statistic. The W\* and A\* statistics are given by

and

where

where and the ’s values are the ordered observations.

## 5.1. Application 1

The first data with size 63 shows the strength measured in GPa for single carbon fibers and impregnated at gaugelengths of 20mm. The data are:

1.9010, 2.132, 2.203, 2.2280, 2.257, 2.350, 2.361, 2.3960, 2.397, 2.445, 2.4540, 2.474, 2.518, 2.5220, 2.525, 2.5320, 2.575, 2.614, 2.616, 2.6180, 2.624, 2.659, 2.6750, 2.738, 2.740, 2.8560, 2.917, 2.928, 2.9370, 2.937, 2.9770, 2.9960, 3.030, 3.125, 3.139, 3.1450, 3.220, 3.223, 3.2350, 3.243, 3.264, 3.272, 3.2940, 3.332, 3.346, 3.377, 3.408, 3.4350, 3.493, 3.501, 3.537, 3.554, 3.5620, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

In general, the smaller value of KS, the better fit to the data. [Table 2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T2) gives the values of estimators of , , and , the KS test statistics and its p-value for TLGLi for the four different estimation methods.

From [Table 2](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T2) we conclude that the LS method is the best method for modelling the carbon fibers with KS = 0.06427, p-value = 0.9571, W\* = 0.0628, and A\* = 0.3331. However all other methods performed well. Finally, we can say that the CVM is better that the WLS method in modelling the carbon fibers. The fitted density, the relative histogram with the fitted density of the proposed model for various methods and fitted survival function for data-I, are also piloted and given in [Figure 3](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F3).

## 5.2. Application 2

The second data set represents on the relief times of twenty patients receiving an analgesic. The data are:

1.10, 1.4, 1.30, 1.7, 1.90, 1.8, 1.6, 2.20, 1.7, 2.70, 4.1, 1.80, 1.5, 1.20, 1.4, 30, 1.7, 2.30, 1.6, 2

([Gross and Clark, 1975](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B12)).

From [Table 3](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T3) we conclude that the CVM method is the best method for modelling the relief times data with KS = 0.09269, p-value = 0.9954, W\* = 0.0525, and A\* = 0.3078. However all other methods performed well.

The fitted density plot, the relative histogram plot with the fitted density of the proposed model and fitted survival function for data-II are given in [Figure 3](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F3). Afterall, from [Figures 3](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F3) and [4](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#F4), it has been noticed that the proposed model fitted well to the considered data set.

# 6. Goodness-of-fit test for TLGLi model in complete data case

## 6.1. Nikulin-Rao-Robson statistic test for complete data

The NRR statistic given by [Nikulin (1973)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B19) and [Rao-Robson (1974)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B21) is a famous modified chi-squared goodness-of-fit test ([Pearson, 1900](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B20)). This statistical test is based on the differences between theoretical and empirical probabilities that fall into grouping cells. It uses the maximum likelihood estimation on the initial data (for more details see Voinov et al. (2013)). Let be a random sample of independent and identically distributed r.v.s. We want to test the following null composite hypothesis

and is the vector of frequencies (where ) obtained by grouping into r intervals :

The boundaries of the intervals are:

The NRR statistic, with as a maximum likelihood estimator of the parameter vector , is defined by

The Pearson’s statistic , where

where vector of probabilities , with

The is the Fisher information matrix, and

represents the Fisher information matrix of any distribution with parameters , where

Under the null hypothesis , the NRR statistic follow a chi-square distribution with degrees of freedom. For any fixed , we have:

## 6.2. Validity of TLGLi model

We test the following null hypothesis . The distribution of the sample is given by

where is given by ([1.1](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#FD3)). To reveal the usefulness of the modified chi-square NRR statistic test and the pertinence of the TLGLi distribution, we consider two real demonstrative data sets.

### **Example 1**

Data was reported by [Badar and Priest (1982)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B3) with size 63 shows the strength measured in GPa for single carbon fibers and impregnated at gaugelengths of 20mm. We choose r = 5 intervals, Using the BB package of R language, the MLE’s parameters of the TLGLi distribution are the following:

We can aver that

since

### thus, the strength measured in GPa for single carbon fibers follows a TLGLi distribution with **parameters.**

### **Example 2**

The data constitutes twenty observations of the lifetime data relating to relief times (in minutes) of patients receiving an analgesic. This data was reported by [Gross and Clark (1975)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B12). We take intervals; then the values of the estimated parameters of our TLGLi are:

The NRR statistic tests . It is clear that

We conclude that we have a concordance of the analgesic failure time data and our TLGLi model.

# 7. Simulation study

A MCMC simulation study is conducted in this section, to compare the performance of the different estimators of the unknown parameters of the TLGLi distribution. This performance is evaluated regarding their mean squared errors (MSEs). Computations in this section are done by ‘Mathcad program Version 15.0’. We generate 1,000 samples of the TLGLi distribution, where n = (20, 50, 100, 200) and choosing , , and . The average values of estimates and MSEs of MLEs, LSEs, WLSEs, CVMEs, and Bayesian estimators are obtained and reported in [Tables 4](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T4)–[7](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T7). The Bayesian estimators of the parameters are evaluated with flexible gamma prior under the SELF by using the MCMC technique. The values of the hyperparameters are assumed known and chosen in such a way that the prior mean is equal to the true value, and prior variance is unity. From [Tables 4](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T4)–[7](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#T7), we observe that all the estimates show the property of consistency, i.e., the MSEs decrease as sample size increase. The MSEs of the Bayesian estimators are also less when compared to the other estimators; in addition, sometimes the MSEs of the Bayes and MLEs are very close to each other.

# 8. Conclusion

This paper introduces a new extension of the Lindley model. The estimation of the parameters is carried out via different methods. Bayes estimation is computed under gamma informative prior under the squared error loss function. The performances of the proposed estimation methods are studied through Monte Carlo simulations. The potentiality of the proposed model is analyzed through two data sets. A modified goodness-of-fit test using the NRR statistic test is investigated via two examples. Certain characterizations of the proposed distribution are presented. A modified goodness-of-fit test for the new model in complete data case is investigated via two examples. We propose the construction of a modified chi-squared goodness of fit statistic test for the new TLGLi model in complete data case. The new test is based on the NRR statistic separately proposed by [Nikulin (1973)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B19) and [Rao and Robson (1974)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B21). As a second step, an application to real data has been proposed to show the applicability of the proposed test and the new TLGLi model for modeling different data sets.

# Appendix

## Theorem 1

Let be a given probability space and let be an interval for some . Let be a continuous random variable with the distribution function and let and be two real functions defined on such that

is defined with some real function . Assume that , and is twice continuously differentiable and strictly monotone function on the set . Finally, assume that the equation has no real solution in the interior of . Then F is uniquely determined by the functions , , and , particularly

where the function s is a solution of the differential equation and is the normalization constant, such that .

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence ([Glänzel, 1990](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473" \l "B10)), in particular, let us assume that there is a sequence of random variables with distribution functions such that the functions , , and satisfy the conditions of Theorem 1 and let , for some continuously differentiable real functions and . Let, finally, be a random variable with distribution . Under the condition that and are uniformly integrable and the family is relatively compact, the sequence converges to in distribution only if converges to η, where

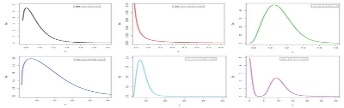
This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions , , and , respectively. It guarantees the ‘convergence’ of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if , as was pointed out in [Glänzel and Hamedani (2001)](http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.5.473#B11).

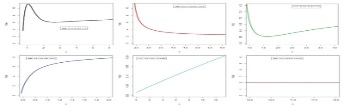
A further consequence of the stability property of Theorem 1 is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions , , and should be as simple as possible. Since the function triplet is not uniquely determined it is often possible to choose η as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics.

In some cases, one can take , as we did in Proposition 3, which reduces the condition of Theorem 1 to

We, however, believe that employing three functions , , and will enhance the domain of applicability of Theorem 1.

# Figures

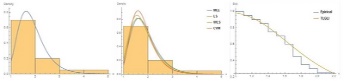


[Fig. 1.](http://www.csam.or.kr/journal/popup_file.html?uid=1865&file=CSAM026-05-04_Fig_01.jpg&md=fig&idx=1) Plots of the probability density function of the Topp Leone Generated Lindley for selected parameter values.  


[Fig. 2.](http://www.csam.or.kr/journal/popup_file.html?uid=1865&file=CSAM026-05-04_Fig_02.jpg&md=fig&idx=2) Plots of the hazard rate function of the Topp Leone Generated Lindley for selected parameter values.



[Fig. 3.](http://www.csam.or.kr/journal/popup_file.html?uid=1865&file=CSAM026-05-04_Fig_03.jpg&md=fig&idx=3) Fitted density, the histograms with the fitted density of the TLGLi distribution for various methods and fitted survival function for data I. TLGLi = Topp Leone Generated Lindley; MLE = maximum likelihood estimator; LS = Least square; WLS = weighted least square; CVM = Crammer-Von Mises.



[Fig. 4.](http://www.csam.or.kr/journal/popup_file.html?uid=1865&file=CSAM026-05-04_Fig_04.jpg&md=fig&idx=4) Fitted density, the histograms with the fitted density of the TLGLi distribution for various methods and fitted survival function for data II. TLGLi = Topp Leone Generated Lindley. TLGLi = Topp Leone Generated Lindley; MLE = maximum likelihood estimator; LS = least square; WLS = weighted least square; CVM = Crammer-Von Mises.

# TABLES

**Table 1** Mean, variance, skewness and kurtosis of the Topp Leone Generated Lindley distribution with different values of parameters

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **λ** | **β** | **θ** | ***μ*′1** | **Variance** | **Skewness** | **Kurtosis** |
| 0.5 | 0.5 | 0.5 | 0.5283 | 0.9050 | 3.1750 | 16.9179 |
|  |  | 1.5 | 1.7810 | 2.6030 | 1.4835 | 5.9315 |
|  |  | 3.0 | 3.0484 | 3.5869 | 1.0063 | 4.3890 |
|  |  | 5.0 | 4.1345 | 4.0290 | 0.8156 | 3.9783 |
|  |  | 10 | 5.7107 | 4.2990 | 0.6821 | 3.7671 |
| 1.0 | 1.5 | 0.5 | 0.5165 | 0.3515 | 2.1602 | 9.6757 |
|  |  | 1.5 | 1.4047 | 0.6927 | 1.1384 | 4.9824 |
|  |  | 3.0 | 2.1364 | 0.8069 | 0.8925 | 4.3499 |
|  |  | 5.0 | 2.7128 | 0.8435 | 0.8038 | 4.1739 |
|  |  | 10 | 3.5140 | 0.8562 | 0.7464 | 4.0793 |
| 2.0 | 2.5 | 0.5 | 0.3065 | 0.0882 | 1.9651 | 8.7477 |
|  |  | 1.5 | 0.7665 | 0.1568 | 1.1332 | 5.0869 |
|  |  | 3.0 | 1.1286 | 0.1785 | 0.9343 | 4.5435 |
|  |  | 5.0 | 1.4109 | 0.1858 | 0.8612 | 4.3793 |
|  |  | 10 | 1.8025 | 0.1887 | 0.8127 | 4.2841 |
| 4.0 | 5.0 | 0.5 | 0.2006 | 0.0229 | 1.6979 | 7.5324 |
|  |  | 1.5 | 0.4420 | 0.0351 | 1.1204 | 5.1709 |
|  |  | 3.0 | 0.6206 | 0.0386 | 0.9823 | 4.7674 |
|  |  | 5.0 | 0.7580 | 0.0399 | 0.9301 | 4.6324 |
|  |  | 10 | 0.9478 | 0.0406 | 0.8939 | 4.5461 |
| 10 | 10 | 0.5 | 0.1006 | 0.0036 | 1.5163 | 6.8066 |
|  |  | 1.5 | 0.1982 | 0.0049 | 1.1258 | 5.2763 |
|  |  | 3.0 | 0.2674 | 0.0053 | 1.0308 | 4.9781 |
|  |  | 5.0 | 0.3201 | 0.0055 | 0.9936 | 4.8694 |
|  |  | 10 | 0.3927 | 0.0056 | 0.9663 | 4.7934 |

**Table 2** The values of estimators, KS, p-values, W\*, and A\*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** |  |  |  | **KS** | **p-value** | **W\*** | **A\*** |
| MLE | 19.04243 | 3.126965 | 1.249401 | 0.08472 | 0.7564 | 0.0658 | 0.3444 |
| LS | 18.06670 | 2.310553 | 1.128924 | 0.06427 | 0.9571 | 0.0628 | 0.3331 |
| WLS | 21.02794 | 2.369555 | 1.168441 | 0.07204 | 0.8993 | 0.0640 | 0.3372 |
| CVM | 21.01143 | 2.227783 | 1.145041 | 0.06687 | 0.9408 | 0.0635 | 0.3353 |

KS = Kolmogorov-Smirnov; W\* = Cramér-Von Mises; A\* = Anderson-Darling; MLE = maximum likelihood estimators; LS = least squares; WLS = weighted LS; CVM = Crammer-Von Mises.

**Table 3** The values of estimators, KS, p-values, W\*, and A\*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** |  |  |  | **KS** | **p-value** | **W\*** | **A\*** |
| MLE | 83.98383 | 0.450189 | 1.322298 | 0.13861 | 0.8369 | 0.0578 | 0.3399 |
| LS | 87.57996 | 0.45128 | 1.389133 | 0.11138 | 0.9651 | 0.0577 | 0.3396 |
| WLS | 89.28296 | 0.41184 | 1.316787 | 0.10808 | 0.9736 | 0.0591 | 0.3477 |
| CVM | 101.4514 | 0.52226 | 1.499851 | 0.09269 | 0.9954 | 0.0525 | 0.3078 |

KS = Kolmogorov-Smirnov; W\* = Cramér-Von Mises; A\* = Anderson-Darling; MLE = maximum likelihood estimator; LS = Least square; WLS = weighted LS; CVM = Crammer-Von Mises.

**Table 4** Average values of estimates and mean squared errors (in parentheses) for n = 20

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **MLE** | **LS** | **WLS** | **CVM** | **Bayesian** |
| β = 2.0 | 2.3532 (0.38887) | 2.1393 (2.08396) | 1.8615 (1.17179) | 1.8403 (0.78797) | 1.9825 (0.36291) |
| θ = 0.9 | 1.0345 (0.05332) | 1.0082 (0.28217) | 1.0695 (0.37182) | 0.9998 (0.27709) | 0.9725 (0.05102) |
| λ = 1.5 | 1.4189 (0.04491) | 1.6744 (0.49578) | 1.3668 (0.56632) | 1.8278 (0.75370) | 1.4062 (0.02231) |
| β = 1.2 | 1.3733 (0.14894) | 1.7486 (1.73004) | 1.6431 (1.47016) | 1.6111 (0.89955) | 1.0090 (0.11238) |
| θ = 1.25 | 1.3593 (0.07589) | 1.5477 (0.49841) | 1.4652 (0.53282) | 1.5500 (0.44179) | 1.1924 (0.06207) |
| λ = 0.5 | 0.4919 (0.00380) | 0.5382 (0.14299) | 0.6003 (0.31093) | 0.5259 (0.14304) | 0.4726 (0.00175) |
| β = 2.5 | 2.6570 (0.56914) | 2.4663 (1.45153) | 2.2088 (1.86197) | 2.4313 (1.59231) | 2.5121 (0.53294) |
| θ = 0.8 | 0.8241 (0.02121) | 0.9339 (0.10929) | 0.9314 (0.21013) | 0.9520 (0.14371) | 0.8321 (0.02019) |
| λ = 0.3 | 0.3012 (0.00128) | 0.3745 (1.19557) | 0.3345 (0.06211) | 0.5054 (3.00787) | 0.2752 (0.00089) |

MLE = maximum likelihood estimator; LS = least square; WLS = weighted least square; CVM = Crammer-Von Mises.

**Table 5** Average values of estimates and mean squared errors for n = 50

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **MLE** | **LS** | **WLS** | **CVM** | **Bayesian** |
| β = 2.0 | 2.2743 (0.15772) | 1.8969 (0.39803) | 1.8156 (0.93618) | 1.8654 (0.38163) | 1.9625 (0.13194) |
| θ = 0.9 | 1.0101 (0.02351) | 0.8766 (0.06023) | 1.0299 (0.32983) | 0.9116 (0.18209) | 0.8562 (0.02127) |
| λ = 1.5 | 1.4205 (0.02041) | 1.6965 (0.32712) | 1.4223 (0.52627) | 1.8211 (0.59441) | 1.2213 (0.01749) |
| β = 1.2 | 1.3246 (0.05230) | 1.6139 (0.54912) | 1.6080 (1.11703) | 1.6155 (0.65038) | 1.1230 (0.05067) |
| θ = 1.25 | 1.3331 (0.02870) | 1.4953 (0.18328) | 1.4126 (0.51303) | 1.5109 (0.29059) | 1.1960 (0.02513) |
| λ = 0.5 | 0.4908 (0.00147) | 0.4684 (0.01874) | 0.5942 (0.29852) | 0.4873 (0.09689) | 0.4238 (0.00071) |
| β = 2.5 | 2.5587 (0.17512) | 3.1008 (1.41519) | 2.3333 (1.73189) | 2.7694 (0.97663) | 2.5670 (0.16327) |
| θ = 0.8 | 0.8084 (0.00710) | 0.9115 (0.04340) | 0.9317 (0.14571) | 0.9154 (0.05181) | 0.7865 (0.00591) |
| λ = 0.3 | 0.3009 (0.00048) | 0.3104 (0.36087) | 0.3183 (0.06067) | 0.3284 (1.04234) | 0.3421 (0.00029) |

MLE = maximum likelihood estimator; LS = least square; WLS = weighted least square; CVM = Crammer-Von Mises.

**Table 6** Average values of estimates and mean squared errors for n = 100

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **MLE** | **LS** | **WLS** | **CVM** | **Bayesian** |
| β = 2.0 | 2.2659 (0.09170) | 1.7649 (0.17525) | 1.8554 (0.79055) | 1.8139 (0.17631) | 1.9920 (0.08923) |
| θ = 0.9 | 1.0084 (0.01832) | 0.8492 (0.05150) | 0.9972 (0.27431) | 0.8493 (0.05137) | 0.9321 (0.01652) |
| λ = 1.5 | 1.4113 (0.01541) | 1.6787 (0.17699) | 1.4944 (0.47237) | 1.6295 (0.19440) | 1.4621 (0.01469) |
| β = 1.2 | 1.3222 (0.02832) | 1.5396 (0.27328) | 1.6031 (0.93027) | 1.5404 (0.27639) | 1.0951 (0.02343) |
| θ = 1.25 | 1.3334 (0.01804) | 1.4637 (0.09777) | 1.5482 (0.48893) | 1.4641 (0.09843) | 1.1970 (0.01601) |
| λ = 0.5 | 0.4885 (0.00082) | 0.4587 (0.00348) | 0.5638 (0.27453) | 0.4631 (0.01789) | 0.4235 (0.00063) |
| β = 2.5 | 2.5538 (0.09009) | 2.9853 (0.67365) | 2.4582 (1.57146) | 2.9866 (0.68129) | 2.6120 (0.07321) |
| θ = 0.8 | 0.8083 (0.00367) | 0.8962 (0.02395) | 0.9102 (0.10320) | 0.8963 (0.02412) | 0.7126 (0.00296) |
| λ = 0.3 | 0.2997 (0.00024) | 0.2826 (0.00084) | 0.3148 (0.05267) | 0.2826 (0.00084) | 0.3712 (0.00018) |

MLE = maximum likelihood estimator; LS = least square; WLS = weighted least square; CVM = Crammer-Von Mises.

**Table 7** Average values of estimates and mean squared errors for n = 200

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameters** | **MLE** | **LS** | **WLS** | **CVM** | **Bayesian** |
| β = 2.0 | 2.0022 (0.02182) | 1.7879 (0.10016) | 2.1824 (0.61904) | 1.7845 (0.09545) | 1.8921 (0.00861) |
| θ = 0.9 | 1.0113 (0.01584) | 0.8332 (0.01154) | 0.9739 (0.07393) | 0.8316 (0.01086) | 0.7829 (0.01329) |
| λ = 1.5 | 1.4067 (0.01229) | 1.6824 (0.09175) | 1.6873 (0.18729) | 1.6813 (0.08527) | 1.5700 (0.01017) |
| β = 1.2 | 1.3189 (0.02400) | 1.4896 (0.13249) | 1.4998 (0.61342) | 1.4914 (0.13387) | 1.1921 (0.02091) |
| θ = 1.25 | 1.3331 (0.01279) | 1.4420 (0.05639) | 1.5302 (0.37676) | 1.4431 (0.05649) | 1.3624 (0.00972) |
| λ = 0.5 | 0.4875 (0.00051) | 0.4576 (0.00268) | 0.4958 (0.07906) | 0.4573 (0.00264) | 0.3905 (0.00047) |
| β = 2.5 | 2.5496 (0.04815) | 2.9140 (0.34917) | 2.4723 (1.41766) | 2.9138 (0.32176) | 2.6412 (0.02938) |
| θ = 0.8 | 0.8080 (0.00195) | 0.8860 (0.08602) | 0.9093 (0.08668) | 0.8863 (0.01300) | 0.9214 (0.00072) |
| λ = 0.3 | 0.2991 (0.00012) | 0.2819 (0.00059) | 0.2911 (0.01006) | 0.2819 (0.00057) | 0.2764 (0.00011) |

MLE = maximum likelihood estimator; LS = least square; WLS = weighted least square; CVM = Crammer-Von Mises.

# References

1. Alizadeh M, Rasekhi M, Yousof HM, and Hamedani GG (2017). The transmuted Weibull G family of distributions, Hacettepe Journal of Mathematics and Statistics, 47, 1-20.
2. Alizadeh M, Yousof HM, Afify AZ, Cordeiro GM, and Mansoor M (2016). The complementary generalized transmuted Poisson-G family of distributions, Austrian Journal of Statistics, 47, 51-71.
3. Badar MG and Priest AM (1982). Statistical aspects of fiber and bundle strength in hybrid composites. In Hayashi T, Kawata K, and Umekawa S (Eds), Progress in Science and Engineering Composites (pp. 1129-1136), ICCM-IV, Tokyo.
4. Casella G and Berger RL (2002). Statistical Inference, Cengage Learning India Private Limited, Delhi.
5. Deniz E and Ojeda E (2011). The discrete Lindley distribution-properties and applications, Journal of Statistical Computation and Simulation, 81, 1405-1416.
6. Ghitany ME, Al-Mutairi DK, and Nadarajah S (2008a). Zero-truncated Poisson-Lindley distribution and its application, Mathematics and Computers in Simulation, 79, 279-287.
7. Ghitany ME, Atieh B, and Nadarajah S (2008b). Lindley distribution and its application, Mathematics and Computers in Simulation, 78, 493-506.
8. Ghitany ME, Alqallaf F, Al-Mutairi DK, and Husain HA (2011). A two-parameter weighted Lindley distribution and its applications to survival data, Mathematics and Computers in Simulation, 81, 1190-1201.
9. Glänzel W (1987). A characterization theorem based on truncated moments and its application to some distribution families. Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986) (B)(pp. 75-84), Reidel, Dordrecht.
10. Glänzel W (1990). Some consequences of a characterization theorem based on truncated moments, Statistics: A Journal of Theoretical and Applied Statistics, 21, 613-618.
11. GlänzelWand Hamedani GG (2001). Characterizations of univariate continuous distributions, Studia Scientiarum Mathematicarum Hungarica, 37, 83-118.
12. Gross J and Clark VA (1975). Survival Distributions: Reliability Applications in the Biometrical Sciences, John Wiley, New York.
13. Hamedani GG (2013). On certain generalized gamma convolution distributions II, MSCS, Marquette University,.
14. Hamedani GG, Altun E, Korkmaz MC, Yousof HM, and Butt NS (2018a). A new extended G family of continuous distributions with mathematical properties, characterizations and regression modeling, Pakistan Journal of Statistics and Operation Research, 14, 737-758.
15. Hamedani GG, Yousof HM, Rasekhi M, Alizadeh M, and Najibi SM (2018b). Type I general exponential class of distributions, Pakistan Journal of Statistics and Operation Research, 14, 39-55.
16. Lindley DV (1958). Fiducial distributions and Bayes’ theorem, Journal of the Royal Statistical Society Series B (Methodological), , 102-107.
17. MacDonald PDM (1971). Comment on an estimation procedure for mixtures of distributions by Choi and Bulgren, Journal of the Royal Statistical Society. Series B, 33, 326-329.
18. Nadarajah S, Bakouch HS, and Tahmasbi R (2011). A generalized Lindley distribution, Sankhya B, 73, 331-359.
19. Nikulin MS (1973). Chi-square test for continuous distribution with shift and scale parameters, Theory of Probability and its Applications, 19, 559-568.
20. Pearson KFRS (1900). On the criterion that a given system of deviation from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 50, 157-175.
21. Rao KC and Robson DS (1974). A chi-square statistic for goodness-of-fit for tests within the exponential family, Communications in Statistics, 3, 1139-1153.
22. Rezaei S, Sadr BB, Alizadeh M, and Nadarajah S (2017). Topp-Leone generated family of distributions: Properties and applications, Communications in Statistics-Theory and Methods, 46, 2893-2909.
23. Swain JJ, Venkatraman S, and andWilson JR (1988). Least squares estimation of distribution functions in Johnson’s translation system, Journal of Statistical Computation and Simulation, 29, 271-297.
24. Voinov V, Alloyarova R, and Pya N (2008). Recent achievements in modified chi-squared goodness-of-fit testing. In Vonta F, Nikulin M, Limnios N, and Huber C (Eds), Statistical Models and Methods for Biomedical and Technical Systems (pp. 245-262), Birkhauser, Boston.