Characterizing log-logistic (L_L) distributions through methods of percentiles and L-moments

Article *in* Applied Mathematical Sciences · January 2017

DOI: 10.12988/ams.2017.612283

1 author:

Mohan Dev Pant
Eastern Virginia Medical School
36 PUBLICATIONS **150** CITATIONS

SEE PROFILE

Applied Mathematical Sciences, Vol. 11, 2017, no. 7, 311 - 329
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ams.2017.612283

Characterizing Log-Logistic (L_L) Distributions through Methods of Percentiles and L-Moments

Mohan D. Pant

Department of Curriculum and Instruction, 320-B Science Hall
University of Texas at Arlington, Arlington, TX-76019, USA

Abstract

The main purpose of this paper is to characterize the log-logistic (L_L) distributions through the methods of percentiles and L-moments and contrast with the method of (product) moments. The method of (product) moments (MoM) has certain limitations when compared with method of percentiles (MoP) and method of L-moments (MoLM) in the context of fitting empirical and theoretical distributions and estimation of parameters, especially when distributions with greater departure from normality are involved. Systems of equations based on MoP and MoLM are derived. A methodology to simulate univariate L_L distributions based on each of the two methods (MoP and MoLM) is developed and contrasted with MoM in terms of fitting distributions and estimation of parameters. Monte Carlo simulation results indicate that the MoP- and MoLM-based L_L distributions are superior to their MoM based counterparts in the context of fitting distributions and estimation of parameters.

Mathematics Subject Classification: 62G30, 62H12, 62H20, 65C05, 65C10, 65C60, 78M05

Keywords: Monte Carlo, Simulation, Product Moments, L-Moments

1 Introduction

The two-parameter log-logistic (L_L) distribution considered herein was derived by Tadikamalla and Johnson [1] by transforming Johnson's [2] S_L system

through a logistic variable. The L_L distribution is a continuous distribution with probability density function (pdf) and cumulative distribution function (cdf) expressed, respectively, as:

$$f(x) = \delta e^{-\gamma} x^{-(\delta+1)}(1 + e^{-\gamma}x^{-\delta})^{-2} \tag{1}$$

$$F(x) = (1 + e^{-\gamma}x^{-\delta})^{-1} \tag{2}$$

where $x \geq 0$ and $\delta > 0$. The pdf in (1) has a single mode, which is at $x = 0$ for $0 < \delta \leq 1$, and at $x = e^{-\gamma/\delta}((\delta - 1)/(\delta + 1))^{1/\delta}$ for $\delta > 1$. When $0 < \delta \leq 1$, the pdf in (1) has a shape of reverse J. For the pdf in (1), the rth moment exists only if $\delta > r$.

A variant of log-logistic distribution has received a wider application in a variety of research contexts such as hydrology [3], estimation of scale parameter [4], MCMC simulation for survival analysis [5], and Bayesian analysis [6]. The quantile function of L_L distribution with cdf in (2) is given as:

$$q(u) = e^{-\gamma/\delta}(u/(1 - u))^{1/\delta} \tag{3}$$

where $u \sim$ uniform(0, 1) is substitued for the cdf in (2).

The method of (product) moments (MoM)-based procedure used in fitting theoretical and empirical distributions involves matching of MoM-based indices (e.g., skew and kurtosis) computed from empirical and theoretical distributions [7]. In the context of L_L distributions, the MoM-based procedure has certain limitations. One of the limitations is that the parameters of skew and kurtosis are defined for L_L distributions only if $\delta > 3$ and $\delta > 4$, respectively. This limitation implies that the MoM-based procedure involving skew and kurtosis cannot be applied for L_L distributions with $\delta \leq 3$.

Another limitation associated with MoM-based application of L_L distributions is that the estimators of skew (α_3) and kurtosis (α_4) computed from sample data are algebraically bounded by the sample size (n) as $|\hat{\alpha_3}| \leq \sqrt{n}$ and $\alpha_4 \leq n$ [8]. This limitation implies that for simulating L_L distributions with kurtosis (α_4) = 48.6541 (as given in Figure 3C in Section 3.2) from samples of size (n) = 25, the largest possible value of the computed sample estimator ($\hat{\alpha}_4$) of kurtosis (α_4) is only 25, which is 51.38 % of the parameter value.

Another limitation associated with MoM-based application of non-normal distributions (e.g., the L_L distributions) is that the estimators of skew (α_3) and kurtosis (α_4) can be substantially biased, highly dispersed, or substantially influenced by outliers when the distributions with greater departure from normality are involved [8-13].

In order to obviate these limitations, this study proposes to characterize the L_L distributions through the methods of percentiles and L-moments. The method of percentiles (MoP) introduced by Karian and Dudewicz [14] and the method of L-moments (MoLM) introduced by Hosking [9] are attractive alternatives to the traditional method of (product) moments (MoM) in the context

of fitting theoretical and empirical distributions and in estimating parameters. In particular, the advantages of MoP-based procedure over the MoM-based procedure are that (a) MoP-based procedure can estimate parameters and obtain fits even when the MoM-based parameters do not exist, (b) the MoP-based estimators have relatively smaller variability than those obtained using MoM-based eprocedure, (c) the solving of MoP-based system of equations is far more efficient than that associated with the MoM-based system of equations [14-17]. Likewise, some of the advantages that MoLM-based estimators of L-skew and L-kurtosis have over MoM-based estimators of skew and kurtosis are that they (a) exist whenever the mean of the distribution exists, (b) are nearly unbiased for all sample sizes and distributions, and (c) are more robust in the presence of outliers [8-13, 18-22].

The rest of the paper is organized as follows. In Section 2, definitions of method of percentiles (MoP) and method of L-moments (MoLM) are provided and systems of equations associated with MoP- and MoLM-based procedures are derived. Also provided in Section 2 are the boundary graphs associated with these procedures. Further, provided in Section 2 are the steps for implimenting the MoP, MoLM, and MoM-based procedures for fitting L_L distributions to empirical and theoretical distributions. In Section 3, a comparison among the MoP-, MoLM-, and MoM-based procedures is provided in the context of fitting L_L distributions to empirical and theoretical distributions and in the context of estimating parameters using a Monte Carlo simulation example. In Section 4, the results are discussed and concluding remarks are provided.

2 Methodology

2.1 Method of Percentiles

Let X be a continuous random variable with quantile function $q(u)$ as in (3), then the method of percentiles (MoP) based analogs of location, scale, skew function, and kurtosis function associated with X are respectively defined by median (ρ_1), inter-decile range (ρ_2), left-right tail-weight ratio (ρ_3, a skew function), and tail-weight factor (ρ_4, a kurtosis function) and given as [14, pp. 154-155]

$$\rho_1 = q(u)_{u=0.50}, \tag{4}$$

$$\rho_2 = q(u)_{u=0.90} - q(u)_{u=0.10}, \tag{5}$$

$$\rho_3 = \frac{q(u)_{u=0.50} - q(u)_{u=0.10}}{q(u)_{u=0.90} - q(u)_{u=0.50}}, \tag{6}$$

$$\rho_4 = \frac{q(u)_{u=0.75} - q(u)_{u=0.25}}{q(u)_{u=0.90} - q(u)_{u=0.10}} \tag{7}$$

where $q(u)_{u=p}$ in (4)-(7) is the (100×p)th percentile with $p \in (0, 1)$.

Substituting appropriate value of u into the quantile (percentile) function $q(u)$ in (3) and simplifying (4)-(7) yields the following MoP-based system of equations associated with L_L distributions:

$$\rho_1 = e^{-\gamma/\delta} \tag{8}$$

$$\rho_2 = 3^{-2/\delta} e^{-\gamma/\delta}(3^{4/\delta} - 1), \tag{9}$$

$$\rho_3 = 3^{-2/\delta}, \tag{10}$$

$$\rho_4 = \frac{3^{1/\delta}}{1 + 3^{2/\delta}}. \tag{11}$$

The parameters of median (ρ_1), inter-decile range (ρ_2), left-right tail-weight ratio (ρ_3), and tail-weight factor (ρ_4) for the L_L distribution are bounded as:

$$0 < \rho_1 < \infty, \rho_2 \geq 0, 0 \leq \rho_3 \leq 1, 0 \leq \rho_4 \leq 1/2, \tag{12}$$

where $\rho_3 = 1$ and $\rho_4 = 1/2$ are the limiting values when $\delta \to \infty$.

For a sample $(X_1, X_2, \cdots, X_n)$ of size n, let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(i)} \leq Y_{(i+1)} \cdots \leq Y_{(n)}$ denote the order statistics. Let $\widehat{q(u)}_{u=p}$ be the (100 × p)th percentile from this sample, where $p \in (0, 1)$. Then, $\widehat{q(u)}_{u=p}$ can be expressed as [14, p. 154]

$$\widehat{q(u)}_{u=p} = Y_{(i)} + (a/b)(Y_{(i+1)} - Y_{(i)}) \tag{13}$$

where i is a positive integer and a/b is a proper fraction such that $(n + 1)p = i + (a/b)$.

For a sample of data with size n, the MoP-based estimators $\hat{\rho}_1$-$\hat{\rho}_4$ of ρ_1-ρ_4 can be obtained in two steps as: (a) Use (13) to compute the values of the 10th, 25th, 50th, 75th, and 90th percentiles and (b) substitute these percentiles into (4)-(7) to obtain the sample estimators $\hat{\rho}_1$-$\hat{\rho}_4$ of ρ_1-ρ_4. See Section 3 for an example to demonstrate this methodology. Figure 1 (panel A) displays region for possible combinations of ρ_3 and ρ_4 for the MoP-based L_L distributions.

2.2 Preliminaries for L-moments

Let $X_1, X_2, \cdots, X_n$ be *i.i.d.* random variables each with pdf $f(x)$ and cdf $F(x)$, respectively. Then, the first four L-moments $(\lambda_1, \cdots \lambda_4)$ associated with each random variable X are expressed as linear combinations of probability weighted moments (PWMs), $\beta_{r=0,1,2,3}$, as [10, pp. 20-22]

$$\lambda_1 = \beta_0 \tag{14}$$

$$\lambda_2 = 2\beta_1 - \beta_0 \tag{15}$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{16}$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \tag{17}$$

where the $\beta_{r=0,1,2,3}$ in (14)-(17) are computed using

$$\beta_r = \int_0^\infty x\{F(x)\}^r f(x)dx \tag{18}$$

The coefficients associated with $\beta_{r=0,1,2,3}$ in (14)-(17) are based on shifted orthogonal Legendre polynomials and are computed as in [10, p. 20]. The first two L-moments, λ_1 and λ_2 in (14) and (15) are measures of location and scale, which are the arithmetic mean and one-half the coefficient of mean difference, respectively. The third- and fourth-order L-moments (λ_3 and λ_4) are transformed into dimensionless L-moment ratios defined as $\tau_3 = \lambda_3/\lambda_2$ and $\tau_4 = \lambda_4/\lambda_2$, which are referred to as the indices of L-skew and L-kurtosis, respectively. In general, L-moment ratios are bounded in the interval of $-1 < \tau_r < 1$, for $r \geq 3$, where a symmetric distribution ($\tau_3 = 0$) implies that all L-moment ratios with odd subscripts are zero. For further information on L-moments, see [9, 10].

2.3 Method of L-moments

In the context of L_L distributions, the derivation of the L-moment based system of equations begins with the derivation of PWMs. The PWMs are derived by substituting $f(x)$ and $F(x)$ from (1) and (2), respectively, into (18) and integrating the simplified integral for $r = 0, 1, 2, 3$ as:

$$\beta_0 = \{\pi e^{-\gamma/\delta} \csc(\pi/\delta)\}/\delta \tag{19}$$

$$\beta_1 = \{\pi e^{-\gamma/\delta}(1+\delta) \csc(\pi/\delta)\}/(2\delta^2) \tag{20}$$

$$\beta_2 = \{\pi e^{-\gamma/\delta}(1+\delta)(1+2\delta) \csc(\pi/\delta)\}/(6\delta^3) \tag{21}$$

$$\beta_3 = \{\pi e^{-\gamma/\delta}(1+\delta)(1+2\delta)(1+3\delta) \csc(\pi/\delta)\}/(24\delta^4) \tag{22}$$

where the PWMs in (19)-(22) take definite values only if $\delta > 1$.

Substituting β_0, β_1, β_2, β_3 from (19)-(22) into (14)-(17) and simplifying yields the first four L-moments. The terms for λ_3 and λ_4 are converted into the dimensionless L-moment ratios of L-skew and L-kurtosis and are presented in their simplest forms, preceded by the first two L-moments, as follows

$$\lambda_1 = \{\pi e^{-\gamma/\delta} \csc(\pi/\delta)\}/\delta \tag{23}$$

$$\lambda_2 = \{\pi e^{-\gamma/\delta} \csc(\pi/\delta)\}/\delta^2 \tag{24}$$

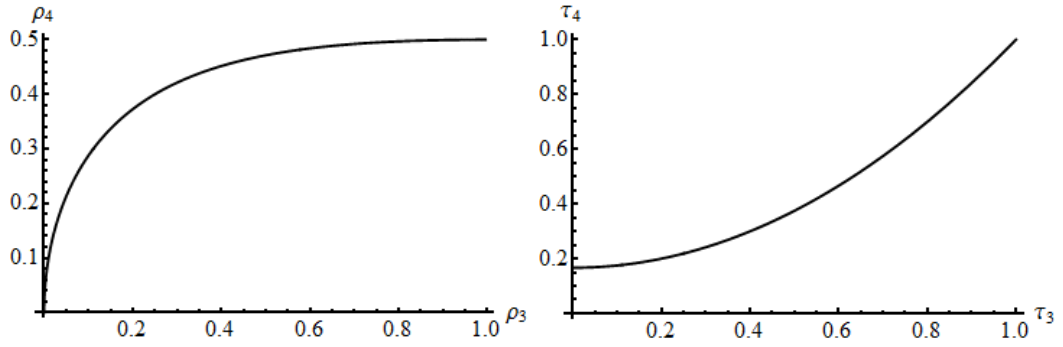
$$\tau_3 = \delta^{-1} \tag{25}$$

$$\tau_4 = \frac{1}{6} + \frac{5}{6\delta^2}. \tag{26}$$

For a sample $(X_1, X_2, \cdots, X_n)$ of size n, let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(i)} \leq Y_{(i+1)} \cdots \leq Y_{(n)}$ denote the sample order statistics, then the unbiased estimators $\hat{\beta}_0$-$\hat{\beta}_3$ of PWMs β_0-β_3 are given by [9, pp. 113-114]

$$\hat{\beta}_r = \frac{1}{n} \sum_{i=r+1}^{n} \frac{(i-1)(i-2)\cdots(i-r)}{(n-1)(n-2)\cdots(n-r)} Y_{(i)} \tag{27}$$

where $r = 0, 1, 2, 3$ and where $\hat{\beta}_0$ is the sample mean. The sample estimators $\hat{\lambda}_1$-$\hat{\lambda}_4$ of λ_1-λ_4 are obtained by substituting $\hat{\beta}_r$ from (27) into (14)-(17). The sample estimators of L-skew and L-kurtosis are denoted by $\hat{\tau}_3$ and $\hat{\tau}_4$, where $\hat{\tau}_3 = \hat{\lambda}_3/\hat{\lambda}_2$ and $\hat{\tau}_4 = \hat{\lambda}_4/\hat{\lambda}_2$. See Section 3 for an example to demonstrate this methodology. Figure 1 (panel B) displays region for possible combinations of τ_3 and τ_4 for the MoLM-based L_L distributions.



A B

Figure 1: Boundary graphs of MoP-based left-right tail-weight ratio (ρ_3) and tail-weight factor (ρ_4) (panel A) and MoLM-based L-skew (τ_3) and L-kurtosis (τ_4) (panel B) associated with the L_L distributions.

3 Comparison of MoP- with MoLM- and MoM-based Procedures

3.1 Fitting Empirical Distributions

Provided in Figure 2 and Table 1 is an example to demonstrate the advantages of MoP-based fit of L_L distributions over the MoLM- and MoM-based fits in the context of fitting empirical distributions (i.e., real-world data). Specifically, Fig. 2 displays the MoP-, MoLM- and MoM-based pdfs of L_L distributions superimposed on the histogram of total hospital charges (in US

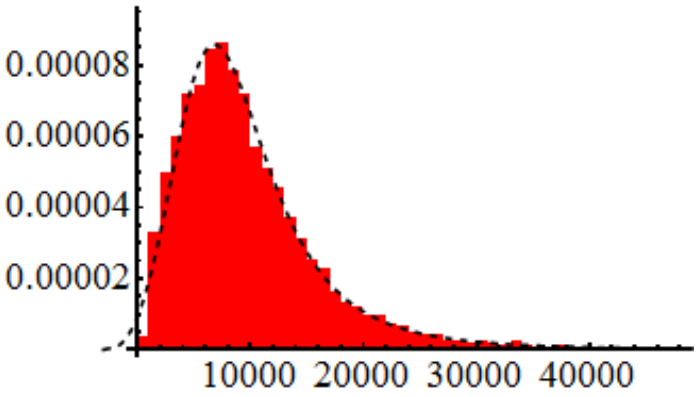
dollars) data of 12,145 heart attack patients discharged from all hospitals in the state of New York in 1993. These data were also used in [17] and can be accessed from the website `http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data_AMI_NY_1993_HeartAttacks`.

The estimates $(\hat{\rho}_1 - \hat{\rho}_4)$ of median, inter-decile range, left-right tail-weight ratio, and tail-weight factor $(\rho_1 - \rho_4)$ were computed from total hospital charges data in two steps as: (a) Obtain the values of the 10th, 25th, 50th, 75th, and 90th percentiles using (13) and (b) substitute these values of percentiles into (4)-(7) to compute the estimates $\hat{\rho}_1 - \hat{\rho}_4$. The parameter values of γ and δ associated with the MoP-based L_L distribution were determined by solving (9) and (10) after substituting the estimates of $\hat{\rho}_2$ and $\hat{\rho}_3$ into the right-hand sides of (9) and (10). The solved values of γ and δ can be used in (8) and (11), respectively, to compute the parameter values of median (ρ_1) and tail-weight factor (ρ_4). The MoP-based fit was obtained by using a linear transformation in the form $x = q(u) + (\hat{\rho}_1 - \rho_1)$.

The estimates $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\tau}_3$ and $\hat{\tau}_4)$ of L-mean, L-scale, L-skew, and L-kurtosis $(\lambda_1, \lambda_2, \tau_3$, and $\tau_4)$ were computed from total hospital charges data in three steps as: (a) Obtain the estimates $\hat{\beta}_0$-$\hat{\beta}_3$ of PWMs using (27), (b) substitute these estimates of PWMs into (14)-(17) to compute the estimates $\hat{\lambda}_1 - \hat{\lambda}_4$, and (c) compute the estimates of L-skew and L-kurtosis as $\hat{\tau}_3 = \hat{\lambda}_3/\hat{\lambda}_2$ and $\hat{\tau}_4 = \hat{\lambda}_4/\hat{\lambda}_2$. The parameter values of γ and δ associated with the MoLM-based L_L distribution were determined by solving (24) and (25) after substituting the estimates of $\hat{\lambda}_2$ and $\hat{\tau}_3$ into the right-hand sides of (24) and (25). The solved values of γ and δ can be used in (23) and (26), respectively, to compute the parameter values of L-mean (λ_1) and L-kurtosis (τ_4). The MoLM-based fit was obtained by using a linear transformation in the form $x = q(u) + (\hat{\lambda}_1 - \lambda_1)$.

The estimates $(\hat{\mu}, \hat{\sigma}, \hat{\alpha}_3$, and $\hat{\alpha}_4)$ of mean, standard deviation, skew, and kurtosis $(\mu, \sigma, \alpha_3$, and $\alpha_4)$ were computed from total hospital charges data using Fisher's k-statistics formulae [23, pp. 299-300]. The parameter values of γ and δ associated with the MoM-based L_L distribution were determined by solving (33) and (34) from Appendix A after substituting the estimates of $\hat{\sigma}$ and $\hat{\alpha}_3$ into the right-hand sides of (33) and (34). The solved values of γ and δ can be used in (32) and (35), respectively, to compute the parameter values of mean (μ) and kurtosis (α_4). The MoM-based fit was obtained by using a linear transformation in the form $x = q(u) + (\hat{\mu} - \mu)$.

The MoP-, MoLM-, and MoM-based estimates along with their corresponding solved values of γ and δ are shown in the top-right, middle-right, and bottom-right panels of Fig. 2. The goodness of MoP-based fit can be contrasted with that associted with MoLM- and MoM-based fits by inspecting the values of Euclidian distances $d = \sqrt{\sum (\text{OP} - \text{EP})^2}$, where OP = observed proportions and EP = expected proportions in each interval of percentiles, from Table 1.



$\hat{\rho}_1 = 8445.0000$

$\hat{\rho}_2 = 14684.8000$

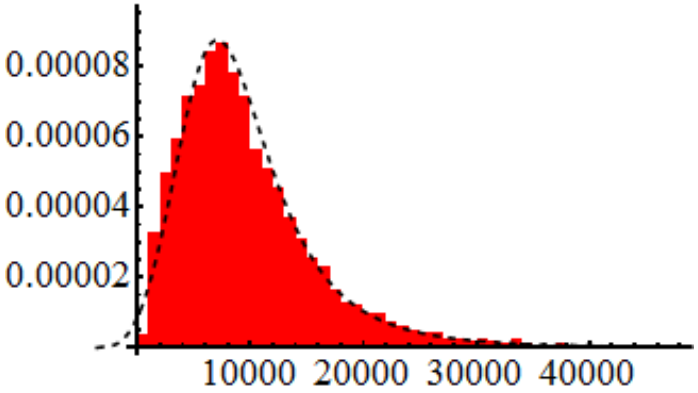
$\hat{\rho}_3 = 0.5477$

$\hat{\rho}_4 = 0.4868$

$\gamma = -34.1207$

$\delta = 3.6496$

(A)



$\hat{\lambda}_1 = 9879.1009$

$\hat{\lambda}_2 = 3380.6927$

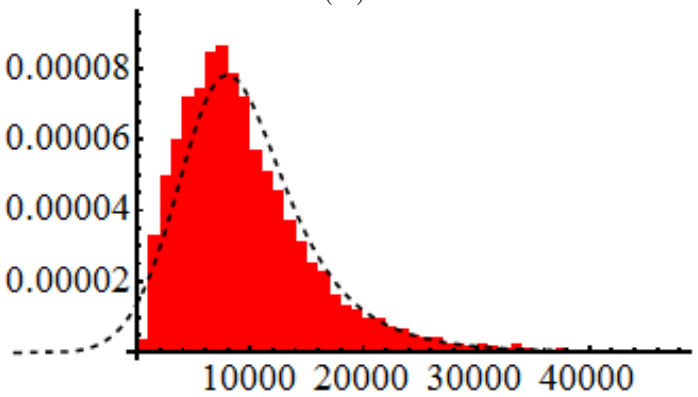
$\hat{\tau}_3 = 0.2498$

$\hat{\tau}_4 = 0.1839$

$\gamma = -37.6603$

$\delta = 4.0030$

(B)



$\hat{\mu} = 9879.1009$

$\hat{\sigma} = 6558.3993$

$\hat{\alpha}_3 = 1.7035$

$\hat{\alpha}_4 = 4.3400$

$\gamma = -62.2764$

$\delta = 6.2693$

(C)

Figure 2: The pdfs of (A) MoP-, (B) MoLM-, and (C) MoM-based L_L distributions superimposed on the histogram of total hospital charges (in US dollars) of 12,145 heart attack patients discharged from all hospitals in the State of New York in 1993.

3.2 Fitting Theoretical Distributions

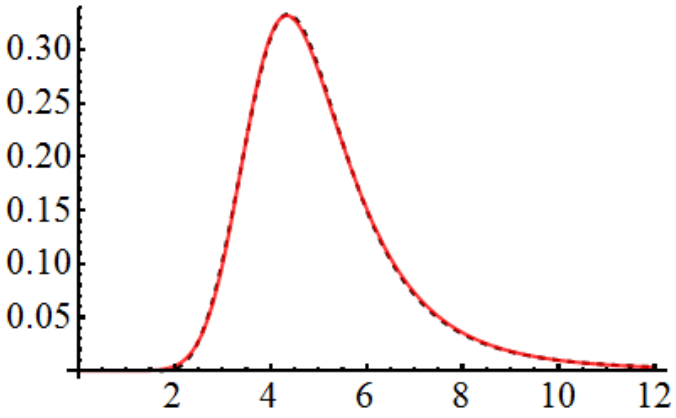
Provided in Figure 3 is an example to demonstrate the advantages of MoP-based fit of L_L distributions over the MoLM- and MoM-based fits in the context of fitting Dagum distribution with shape parameters: $p = 2$ and $a = 5$ and scale parameter $b = 4$. See [12] for a comparison of MoLM and MoM-based fits of Dagum distributions.

The values of $\rho_1 - \rho_4$ associated with the Dagum distribution were computed using (4)-(7), where the quantile function $q(u)$ of Dagum distribution was used. The parameter values of γ and δ associated with the MoP-based L_L distribution were determined by solving (9) and (10) after substituting the values of ρ_2 and ρ_3 of Dagum distribution into the right-hand sides of (9) and (10). These values of γ and δ can be used in (8) and (11), respectively, to compute the parameter values of ρ_1 and ρ_4 associated with the L_L distribution. The MoP-based fit was obtained by using a linear transformation $x = q(u) + (\hat{\rho}_1 - \rho_1)$, where $\hat{\rho}_1$ is the median of Dagum distribution.

The values of λ_1, λ_2, τ_3, τ_4 associated with the Dagum distribution were computed using (18) and (14)-(17) and using the formulae for τ_3 and τ_4 from Section 2.2. The parameter values of γ and δ associated with the MoLM-based L_L distribution were determined by solving (24) and (25) after substituting the values of λ_2 and τ_3 of Dagum distribution into the right-hand sides of (24) and (25). These values of γ and δ can be used in (23) and (26), respectively, to compute the parameter values of λ_1 and τ_4 associated with the L_L distribution. The MoLM-based fit was obtained by using a linear transformation $x = q(u) + (\hat{\lambda}_1 - \lambda_1)$, where $\hat{\lambda}_1$ is the L-mean of Dagum distribution.

The values of μ, σ, α_3 and α_4 associated with the Dagum distribution were computed using (28), formulae of mean and standard deviation and (30) and (31) from the Appendix. The parameter values of γ and δ associated with the MoM-based L_L distribution were determined by solving (33) and (34) from Appendix A after substituting the values of σ and α_3 of Dagum distribution into the right-hand sides of (33) and (34). These values of γ and δ can be used in (32) and (35), respectively, to compute the parameter values of μ and α_4 associated with the L_L distribution. The MoM-based fit was obtained by using a linear transformation $x = q(u) + (\hat{\mu} - \mu)$, where $\hat{\mu}$ is the mean of Dagum distribution.

The parameters (shown as estimates) of Dagum distribution based on the MoP, MoLM, and MoM procedures along with solved values of γ and δ associated with their respective L_L fits are shown, respectively, in the top-right, middle-right, and bottom-right panels of Fig. 3. The goodness of MoP-based fit can be contrasted with that associated with MoLM- and MoM-based fits by inspecting the values of Euclidian distances $d = \sqrt{\sum (\mathrm{OP} - \mathrm{EP})^2}$, where OP = observed percentile and EP = expected percentiles, from Table 2.



$\hat{\rho}_1 = 4.7711$

$\hat{\rho}_2 = 3.7403$

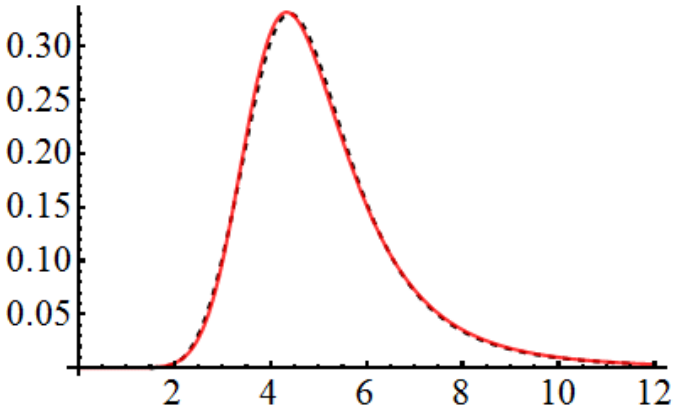
$\hat{\rho}_3 = 0.5601$

$\hat{\rho}_4 = 0.4839$

$\gamma = -4.2295$

$\delta = 3.7903$

(A)



$\hat{\lambda}_1 = 5.1310$

$\hat{\lambda}_2 = 0.8894$

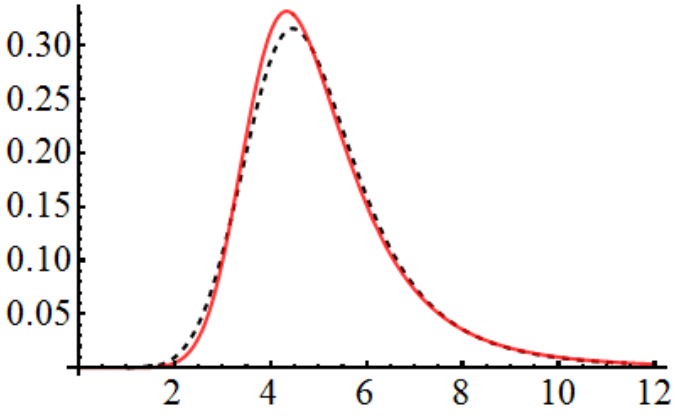
$\hat{\tau}_3 = 0.2455$

$\hat{\tau}_4 = 0.2094$

$\gamma = -4.8296$

$\delta = 4.0727$

(B)



$\hat{\mu} = 5.1310$

$\hat{\sigma} = 1.8083$

$\hat{\alpha}_3 = 2.9648$

$\hat{\alpha}_4 = 34.4831$

$\gamma = -6.1462$

$\delta = 4.5926$

(C)

Figure 3: The pdfs (dashed curves) of (A) MoP-, (B) MoLM-, and (C) MoM-based L_L distributions superimposed on the pdf of Dagum distribution with shape parameters: $p = 2$ and $a = 5$ and scale parameter $b = 4$.

3.3 Estimation of Parameters

In the context of L_L distributions, the MoP-based estimators have certain advantages over MoLM- and MoM-based estimators. For example, provided in Tables 3-5 are the MoP-, MoLM-, and MoM-based parameter values (of second- and higher-order) for the six L_L distributions (dashed curves) in Figures 2 and 3. Specifically, Tables 3-5 provide the parameter values and the results of a Monte Carlo simulation associated with their corresponding estimators along with indices of standard errors (SE) and relative bias (RB%).

For the Monte Carlo simulation, a Fortran algorithm was written to generate 25,000 independent samples of sizes $n = 25$ and $n = 500$ from the six distributions in Figures 2 and 3. The MoP-based estimators ($\hat{\rho}_2$, $\hat{\rho}_3$, and $\hat{\rho}_4$) of inter-decile range, left-right tail-weight ratio, and tail-weight factor (ρ_2, ρ_3, and ρ_4), the MoLM-based estimators ($\hat{\lambda}_2$, $\hat{\tau}_3$, and $\hat{\tau}_4$) of L-scale, L-skew, and L-kurtosis (λ_2, τ_3, and τ_4) and the MoM-based estimators ($\hat{\sigma}$, $\hat{\alpha}_3$, and $\hat{\alpha}_4$) of standard deviation, skew, and kurtosis (σ, α_3, and α_4) were computed for each of the ($2 \times 25{,}000$) samples based on the solved values of γ and δ parameters listed in the right panels of Figures 2 and 3. The MoP-based estimates ($\hat{\rho}_2$, $\hat{\rho}_3$, and $\hat{\rho}_4$) were computed in two steps as: (a) Obtain the 10th, 25th, 50th, 75th, and 90th percentiles using (13) and (b) substitute these percentiles into (5)-(7). The MoLM-based estimators ($\hat{\lambda}_2$, $\hat{\tau}_3$, and $\hat{\tau}_4$) were computed in three steps as: (a) Obtain the estimates of PWMs ($\hat{\beta}_{0,1,2,3}$) using (27), (b) substitute these values of $\hat{\beta}_{0,1,2,3}$ into (15)-(17) to obtain the estimates $\hat{\lambda}_{2,3,4}$, and (c) compute the estimates of L-skew and L-kurtosis as $\hat{\tau}_3 = \hat{\lambda}_3/\hat{\lambda}_2$ and $\hat{\tau}_4 = \hat{\lambda}_4/\hat{\lambda}_2$. The MoM-based estimators ($\hat{\sigma}$, $\hat{\alpha}_3$, and $\hat{\alpha}_4$) were based on Fisher's k-statistics formulae [23, pp. 299-300], which are used by most commercial software packages such as SAS, SPSS, Minitab, etc., for computing indices of standard deviation, skew and kurtosis. Bias-corrected accelerated bootstrapped average estimates, 95% bootstrap confidence intervals (95% C.I.), and associated standard errors (SE) were subsequently obtained for each type of estimates using 10,000 resamples via the commercial software package Spotfire S+ [24]. If a parameter (P) was outside its associated 95% C.I., then the percentage of relative bias (RB%) was computed for the estimate (E) as: RB% $= \{((E - P)/P) \times 100\}$. The results of the simulation are reported in Tables 3-5 and are discussed in the next section.

4 Discussion and Conclusion

One of the advantages of MoP- and MoLM-based procedures over the traditional MoM-based procedure is that the distributions characterized through the former procedures can provide better fits to real-world data and some the-

Percentiles	EP	MoP OP	MoLM OP	MoM OP
5	0.05	0.0410	0.0462	0.0085
10	0.05	0.0590	0.0631	0.0724
25	0.15	0.1534	0.1552	0.1881
50	0.25	0.2467	0.2432	0.2730
75	0.25	0.2455	0.2335	0.2235
90	0.15	0.1545	0.1510	0.1305
95	0.05	0.0513	0.0520	0.0445
100	0.05	0.0487	0.0558	0.0596
		$d = 0.0151$	$d = 0.0239$	$d = 0.0735$

Table 1: Percentiles, expected proportions (EP), the MoP-based observed proportions (OP) (MoP OP), MoLM-based OP (MoLM OP), and MoM-based OP (MoM OP) of L_L distribution fits to total hospital charges data in Fig. 2.

Percentiles	EP	MoP OP	MoLM OP	MoM OP
5	3.1185	3.1224	3.0975	3.0121
10	3.4283	3.4283	3.4174	3.3669
25	4.0000	4.0031	4.0084	4.0055
50	4.7711	4.7711	4.7823	4.8166
75	5.8098	5.7973	5.7959	5.8469
90	7.1686	7.1686	7.1234	7.1557
95	8.3011	8.3561	8.2540	8.2426
		$d = 0.0566$	$d = 0.0722$	$d = 0.1489$

Table 2: Percentiles, expected percentiles (EP), the MoP-based observed percentiles (OP) (MoP OP), MoLM-based OP (MoLM OP), and MoM-based OP (MoM OP) of L_L distribution fits to Dagum distribution in Fig. 3.

oretical distributions [8-17]. In case of L_L distributions, inspection of Figures 2 and 3 indicates that the MoP- and MoLM-based procedures provide better fits than the MoM-based procedure in the context of both fitting real-world data and theoretical distributions. Furthermore, the Euclidian distances related with MoP- and MoLM-based fits in Tables 1 and 2 are substantially smaller than those associated with the MoM-based fits. For example, inspection of Table 1 indicates that $d = 0.0151$ associated with MoP-based fit of L_L distribution is approximately one-fifth of $d = 0.0735$ associated with MoM-based fit of L_L distribution over total hospital charges data in Fig. 2. Similarly, $d = 0.0239$ associated with MoLM-based fit is approximately one-third of $d = 0.0735$ associated with MoM-based fit.

The MoP-based estimators can be far less biased and less dispersed than the MoM-based estimators when distributions with larger departure from normality are involved [14-17]. The MoLM-based estimators can also be far less

$n = 25$					
Distribution	Parameter	Estimate	95%C.I.	SE	RB%
Fig. 2A	$\rho_2 = 14684.8$	$\hat{\rho}_2 = 17923$	$17853, 17990$	35.01	22.05
	$\rho_3 = 0.5477$	$\hat{\rho}_3 = 0.5745$	$0.5710, 0.5780$	0.00178	4.89
	$\rho_4 = 0.4782$	$\hat{\rho}_4 = 0.4317$	$0.4305, 0.4333$	0.00070	-9.72
Fig. 3A	$\rho_2 = 3.7403$	$\hat{\rho}_2 = 4.5360$	$4.5204, 4.5542$	0.00853	21.27
	$\rho_3 = 0.5601$	$\hat{\rho}_3 = 0.5885$	$0.5849, 0.5920$	0.00183	5.07
	$\rho_4 = 0.4797$	$\hat{\rho}_4 = 0.4346$	$0.4332, 0.4360$	0.00071	-9.40
$n = 500$					
Distribution	Parameter	Estimate	95%C.I.	SE	RB%
Fig. 2A	$\rho_2 = 14684.8$	$\hat{\rho}_2 = 14730$	$14720, 14742$	5.563	0.31
	$\rho_3 = 0.5477$	$\hat{\rho}_3 = 0.5489$	$0.5481, 0.5495$	0.00036	0.22
	$\rho_4 = 0.4782$	$\hat{\rho}_4 = 0.4774$	$0.4770, 0.4777$	0.00018	-0.17
Fig. 3A	$\rho_2 = 3.7403$	$\hat{\rho}_2 = 3.7540$	$3.7517, 3.7572$	0.00139	0.37
	$\rho_3 = 0.5601$	$\hat{\rho}_3 = 0.5616$	$0.5609, 0.5623$	0.00037	0.27
	$\rho_4 = 0.4797$	$\hat{\rho}_4 = 0.4788$	$0.4785, 0.4791$	0.00017	-0.19

Table 3: MoP-based parameters and their bootstrap estimates along with 95% confidence intervals (95%C.I.), standard errors (SE), and indices of relative bias (RB%) for the L_L distributions in Figures 2A and 3A.

$n = 25$					
Distribution	Parameter	Estimate	95%C.I.	SE	RB%
Fig. 2B	$\lambda_2 = 3380.7$	$\hat{\lambda}_2 = 3382$	$3371.5, 3393.8$	5.686	—
	$\tau_3 = 0.2498$	$\hat{\tau}_3 = 0.2271$	$0.2254, 0.2287$	0.00083	-9.09
	$\tau_4 = 0.2187$	$\hat{\tau}_4 = 0.2046$	$0.2032, 0.2059$	0.00071	-6.45
Fig. 3B	$\lambda_2 = 0.8894$	$\hat{\lambda}_2 = 0.8899$	$0.8871, 0.8929$	0.00149	—
	$\tau_3 = 0.2455$	$\hat{\tau}_3 = 0.2234$	$0.2218, 0.2250$	0.00083	-9.00
	$\tau_4 = 0.2169$	$\hat{\tau}_4 = 0.2033$	$0.2019, 0.2046$	0.00070	-6.27
$n = 500$					
Distribution	Parameter	Estimate	95%C.I.	SE	RB%
Fig. 2B	$\lambda_2 = 3380.7$	$\hat{\lambda}_2 = 3380$	$3377.5, 3382.5$	1.277	—
	$\tau_3 = 0.2498$	$\hat{\tau}_3 = 0.2482$	$0.2478, 0.2486$	0.00021	-0.64
	$\tau_4 = 0.2187$	$\hat{\tau}_4 = 0.2175$	$0.2171, 0.2178$	0.00018	-0.55
Fig. 3B	$\lambda_2 = 0.8894$	$\hat{\lambda}_2 = 0.8891$	$0.8885, 0.8898$	0.00034	—
	$\tau_3 = 0.2455$	$\hat{\tau}_3 = 0.2440$	$0.2436, 0.2444$	0.00021	-0.61
	$\tau_4 = 0.2169$	$\hat{\tau}_4 = 0.2158$	$0.2154, 0.2161$	0.00018	-0.51

Table 4: MoLM-based parameters and their bootstrap estimates along with 95% confidence intervals (95%C.I.), standard errors (SE), and indices of relative bias (RB%) for the L_L distributions in Figures 2B and 3B.

$n = 25$

Distribution	Parameter	Estimate	95%C.I.	SE	RB%
Fig. 2C	$\sigma = 6558.4$	$\hat{\sigma} = 6202$	$6182.7, 6224.3$	10.58	-5.43
	$\alpha_3 = 1.7035$	$\hat{\alpha}_3 = 0.8131$	$0.8042, 0.8232$	0.00480	-52.27
	$\alpha_4 = 10.1527$	$\hat{\alpha}_4 = 1.1650$	$1.1344, 1.1956$	0.01580	-88.53
Fig. 3C	$\sigma = 1.8083$	$\hat{\sigma} = 1.6730$	$1.6666, 1.6812$	0.00367	-7.48
	$\alpha_3 = 2.9648$	$\hat{\alpha}_3 = 1.0860$	$1.0751, 1.0962$	0.00534	-63.37
	$\alpha_4 = 48.6541$	$\hat{\alpha}_4 = 1.7800$	$1.7399, 1.8174$	0.0198	-96.34

$n = 500$

Distribution	Parameter	Estimate	95%C.I.	SE	RB%
Fig. 2C	$\sigma = 6558.4$	$\hat{\sigma} = 6541$	$6534.1, 6546.5$	3.1330	-0.27
	$\alpha_3 = 1.7035$	$\hat{\alpha}_3 = 1.5310$	$1.5206, 1.5397$	0.00490	-10.13
	$\alpha_4 = 10.1527$	$\hat{\alpha}_4 = 6.6640$	$6.5355, 6.7937$	0.06590	-34.36
Fig. 3C	$\sigma = 1.8083$	$\hat{\sigma} = 1.7980$	$1.7952, 1.8007$	0.00140	-0.57
	$\alpha_3 = 2.9648$	$\hat{\alpha}_3 = 2.2980$	$2.2822, 2.3150$	0.00835	-22.49
	$\alpha_4 = 48.6541$	$\hat{\alpha}_4 = 13.320$	$13.069, 13.607$	0.13780	-72.62

Table 5: MoM-based parameters and their bootstrap estimates along with 95% confidence intervals (95%C.I.), standard errors (SE), and indices of relative bias (RB%) for the L_L distributions in Figures 2C and 3C.

biased and less dispersed than the MoM-based estimators when sampling is from distributions with more severe departures from normality [8-13, 18-22]. Inspection of the simulation results in Tables 3-5 clearly indicates that in the context of L_L distributions, the MoP- and MoLM-based estimators are superior to their MoM-based counterparts for the estimators of third- and fourth-order parameters. That is, the superiority that MoP-based estimators of left-right tail-weight ratio (ρ_3) and tail-weight factor (ρ_4) and MoLM-based estimators of L-skew (τ_3) and L-kurtosis (τ_4) have over their corresponding MoM-based estimators of skew (α_3) and kurtosis (α_4) is clearly obvious. For example, with samples of size $n = 25$ the estimates of α_3 and α_4 for the L_L distribution in Fig. 3C were, on average, only 36.63% and 3.66% of their respective parameters, whereas the estimates of ρ_3 and ρ_4 for the L_L distribution in Fig. 3A were, on average, 105.07% and 90.60% of their respective parameters and the estimates of L-skew and L-kurtosis for the L_L distribution in Fig. 3B were, on average, 91% and 93.73% of their respective parameters.

From inspection of Tables 3-5, it is also evident that MoP-based estimators of ρ_3 and ρ_4 and MoLM-based estimators of τ_3 and τ_4 are more efficient estimators as their relative standard errors RSE = $\{(\text{SE/Estimate}) \times 100\}$ are considerably smaller than those associated with MoM-based estimators of α_3 and α_4. For example, inspection of Tables 3-5 for $n = 500$, indicates RSE measures of: RSE($\hat{\rho}_3$) = 0.07% and RSE($\hat{\rho}_4$) = 0.04% for the L_L distribution in Fig. 3A compared with RSE($\hat{\tau}_3$) = 0.09% and RSE($\hat{\tau}_4$) = 0.08% for the

L_L distribution in Fig. 3B and RSE($\hat{\alpha}_3$) = 0.36% and RSE($\hat{\alpha}_4$) = 1.03% for the L_L distribution in Fig. 3C. Thus, MoP-based estimators of ρ_3 and ρ_4 have about the same degree of precision compared to the MoLM-based estimators of τ_3 and τ_4, whereas both MoP- and MoLM-based estimators have substantially higher precision when compared to the MoM-based estimators of α_3 and α_4.

In conclusion, the proposed MoP- and MoLM-based procedures are more attractive alternatives to the traditional MoM-based procedure. In particular, the proposed MoP- and MoLM-based procedures have distinctive advantages over MoM-based procedure when distributions with large departures from normality are involved. Finally, Mathematica Version 9.0.0.0 [25] source code is available from the author for implementing all three procedures.

References

[1] P.R. Tadikamalla and N.L. Johnson, Systems of frequency curves generated by transformations of logistic variables, *Biometrika,* **69** (1982), 461 - 465. https://doi.org/10.1093/biomet/69.2.461

[2] N.L. Johnson, Systems of frequency curves generated by methods of translation, *Biometrika,* **36** (1949), 149 - 176. https://doi.org/10.2307/2332539

[3] F. Ashkar and S. Mahdi, Fitting the log-logistic distribution by generalized moments, *Journal of Hydrology,* **328** (2006), 694 - 703. https://doi.org/10.1016/j.jhydrol.2006.01.014

[4] G. Lesitha and P.Y. Thomas, Estimation of the scale parameter of a log-logistic distribution, *Metrika,* **76** (2013), 427 - 448. https://doi.org/10.1007/s00184-012-0397-5

[5] A.A. Al-Shomrani, A.I. Shawky, O.H. Arif and M. Aslam, Log-logistic distribution for survival data analysis using MCMC, *SpringerPlus,* **5** (2016), 1 - 16. https://doi.org/10.1186/s40064-016-3476-7

[6] K. Abbas and Y. Tang, Objective Bayesian analysis for log-logistic distribution, *Communications in Statistics - Simulation and Computation,* **45** (2016), 2782 - 2791. https://doi.org/10.1080/03610918.2014.925925

[7] K. Pearson, Method of moments and method of maximum likelihood, *Biometrika,* **28** (1936), 34 - 59. https://doi.org/10.2307/2334123

[8] T.C. Headrick, A characterization of power method transformations through L-moments, *Journal of Probability and Statistics,* **2011** (2011), Article ID 497463, 1 - 22. https://doi.org/10.1155/2011/497463

[9] J.R.M. Hosking, *L*-moments: Analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society, Series B,* **52** (1990), 105 - 124.

[10] J.R.M. Hosking and J.R. Wallis, *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, Cambridge, UK, 1997. https://doi.org/10.1017/cbo9780511529443

[11] M.D. Pant and T.C. Headrick, A method for simulating Burr Type III and Type XII distributions through *L*-moments and *L*-correlations, *ISRN Applied Mathematics,* **2013** (2013), Article ID 191604, 1 - 14. https://doi.org/10.1155/2013/191604

[12] M.D. Pant and T.C. Headrick, An *L*-moment based characterization of the family of Dagum distributions, *Journal of Statistical and Econometric Methods,* **2** (2013), 17 - 40.

[13] M.D. Pant and T.C. Headrick, Simulating Burr Type VII distributions through the method of *L*-moments and *L*-correlations, *Journal of Statistical and Econometric Methods,* **3** (2014), 23 - 63.

[14] Z.A. Karian and E.J. Dudewicz, *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Method*, Chapman and Hall/CRC, Boca Raton, Florida, USA, 2000. https://doi.org/10.1201/9781420038040

[15] J. Koran, T.C. Headrick and T.C. Kuo, Simulating univariate and multivariate nonnormal distributions through the method of percentiles, *Multivariate Behavioral Research,* **50** (2015), 216 - 232. https://doi.org/10.1080/00273171.2014.963194

[16] T.C. Kuo and T.C. Headrick, Simulating univariate and multivariate Tukey *g*-and-*h* distributions based on the method of percentiles, *ISRN Probability and Statistics,* **2014** (2014), Article ID 645823, 1 - 10. https://doi.org/10.1155/2014/645823

[17] M.D. Pant and T.C. Headrick, A characterization of the Burr Type III and Type XII distributions through the method of percentiles and the Spearman correlation, *Communications in Statistics-Simulation and Computation,* (2015), 1 - 17. https://doi.org/10.1080/03610918.2015.1048878

[18] T.C. Headrick and M.D. Pant, Simulating non-normal distributions with specified *L*-moments and *L*-correlations, *Statistica Neerlandica,* **66** (2012), 422 - 441. https://doi.org/10.1111/j.1467-9574.2012.00523.x

[19] T.C. Headrick and M.D. Pant, A method for simulating nonnormal distributions with specified L-skew, L-kurtosis, and L-correlation, *ISRN Applied Mathematics,* **2012** (2012), Article ID 980827, 1 - 23. https://doi.org/10.5402/2012/980827

[20] T.C. Headrick and M.D. Pant, A logistic L-moment-based analog for the Tukey g-h, g, h, and h-h system of distributions, *ISRN Probability and Statistics,* **2012** (2012), Article ID 245986, 1 - 23. https://doi.org/10.5402/2012/245986

[21] T.C. Headrick and M.D. Pant, Characterizing Tukey h and hh-distributions through L-moments and the L-correlations, *ISRN Applied Mathematics,* **2012** (2012), Article ID 980153, 1 - 20. https://doi.org/10.5402/2012/980153

[22] M.D. Pant and T.C. Headrick, A doubling technique for the power method transformations, *Applied Mathematical Sciences,* **6** (2012), 6437 - 6475.

[23] M. Kendall and A. Stuart, *The Advanced Theory of Statistics,* 4th edition, Macmillan, New York, USA, 1977.

[24] TIBCO, *Spotfire S+ 8.1 for Windows,* TIBCO Software, Palo Alto, CA, 2008.

[25] Wolfram Research Inc., *Mathematica,* Version 9.0.0.0, Wolfram Research Inc., Champaign, IL, 2012.

Appendices

A Method of (Product) Moments

Let $X > 0$ be a continuous random variable from a probability distribution with pdf $f(x)$, then the rth (product) moment of X is given as

$$\mu_r = \int_0^\infty x^r f(x) dx \tag{28}$$

Substituting $f(x)$ from (1) and integrating the simplified integral yields the rth (product) moment of L_L distribution as:

$$\mu_r = \frac{r\pi e^{-r\gamma/\delta} \csc(r\pi/\delta)}{\delta}. \tag{29}$$

where $\delta > r$ so that the rth moment exists.

Provided that the first four moments ($\mu_{r=1,2,3,4}$) exist, the MoM-based parameters of mean (μ) and standard deviation (σ) are respectively given as

$\mu = \mu_1$ and $\sigma = \sqrt{\mu_2 - \mu_1^2}$, whereas the parameters of skew and kurtosis are obtained by substituting these four moments into the following formulae for skew (α_3) and kurtosis (α_4) from [23]:

$$\alpha_3 = (\mu_3 - 3\mu_2\mu_1 + 2{\mu_1}^3)/(\mu_2 - {\mu_1}^2)^{3/2}, \tag{30}$$

$$\alpha_4 = (\mu_4 - 4\mu_3\mu_1 - 3{\mu_2}^2 + 12\mu_2{\mu_1}^2 - 6{\mu_1}^4)/(\mu_2 - {\mu_1}^2)^2. \tag{31}$$

Extracting the first four moments ($\mu_{r=1,2,3,4}$) from (29) and substituting them into (30) and (31), the indices of skew (α_3) and kurtosis (α_4), preceded by mean and standard deviation of L_L distributions are expressed as:

$$\mu = \frac{\pi e^{-\gamma/\delta}\csc(\pi/\delta)}{\delta}, \tag{32}$$

$$\sigma = \frac{\sqrt{\pi}e^{-\gamma/\delta}\sqrt{2\delta\csc(2\pi/\delta) - \pi\csc(\pi/\delta)^2}}{\delta}, \tag{33}$$

$$\alpha_3 = \frac{2\pi^2\csc(\pi/\delta)^3 - 6\pi\delta\csc(\pi/\delta)\csc(2\pi/\delta) + 3\delta^2\csc(3\pi/\delta)}{\sqrt{\pi}(2\delta\csc(2\pi/\delta) - \pi\csc(\pi/\delta)^2)^{3/2}}, \tag{34}$$

$$\begin{aligned}\alpha_4 \;=\; & 2\{2\delta^3\csc(4\pi/\delta) + 6\pi^2\delta\csc(\pi/\delta)^3\sec(\pi/\delta) \\ & -6\pi\delta^2\csc(\pi/\delta)\csc(3\pi/\delta) - 6\pi\delta^2\csc(2\pi/\delta)^2 \\ & -3\pi^3\csc(\pi/\delta)^4\}/\{\pi(2\delta\csc(2\pi/\delta) - \pi\csc(\pi/\delta)^2)^2\}.\end{aligned} \tag{35}$$

In the context of L_L distributions, the MoM-based procedure involves solving of (33) and (34) for the parameters of γ and δ after given values (or, estimates) of standard deviation (σ) and skew (α_3) are substituted into the right-hand sides of (33) and (34). The solved values of γ and δ can be substituted into (32) and (35), respectively, for computing the values of mean and kurtosis.

B Boundary Graph of MoM-based Measures of Skew (α_3) and Kurtosis (α_4) for the L_L Distributions.

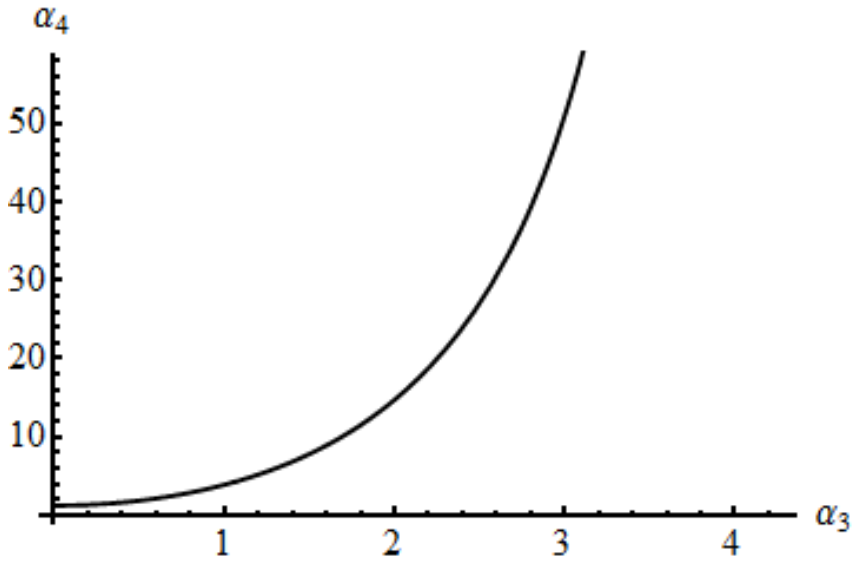


Figure B: Boundary graph of MoM-based skew (α_3) and kurtosis (α_4) associated with the L_L distributions. Fig. B can be used to find a possible combination of skew (α_3) and kurtosis (α_4) of a valid L_L distribution.

Received: January 2, 2017; Published: January 25, 2017