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

Mohan D. Pant, University of Texas at Arlington

Available at: https://works.bepress.com/mohan_pant/19/
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:

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

$$F(x) = (1 + e^{-\gamma} x^{-\delta})^{-1}$$  \hspace{1cm} (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 $r$th 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:

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

where $u \sim \text{uniform}(0, 1)$ is substituted 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 ($\alpha_3$) and kurtosis ($\alpha_4$) 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 ($\alpha_4$) = 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 ($\alpha_4$) 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 ($\alpha_3$) and kurtosis ($\alpha_4$) 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 procedure, (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 implementing 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 ($\rho_1$), inter-decile range ($\rho_2$), left-right tail-weight ratio ($\rho_3$, a skew function), and tail-weight factor ($\rho_4$, a kurtosis function) and given as [14, pp. 154-155]

\[ \rho_1 = q(u)_{u=0.50}, \]
\[ \rho_2 = q(u)_{u=0.90} - q(u)_{u=0.10}, \]
\[ \rho_3 = \frac{q(u)_{u=0.50} - q(u)_{u=0.10}}{q(u)_{u=0.90} - q(u)_{u=0.50}}, \]
\[ \rho_4 = \frac{q(u)_{u=0.75} - q(u)_{u=0.25}}{q(u)_{u=0.90} - q(u)_{u=0.10}}. \]
where \( q(u)_{u=p} \) in (4)-(7) is the \((100 \times p)\)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:

\[
\begin{align*}
\rho_1 &= e^{-\gamma/\delta} \\
\rho_2 &= 3^{-2/\delta}e^{-\gamma/\delta}(3^{4/\delta} - 1), \\
\rho_3 &= 3^{-2/\delta}, \\
\rho_4 &= \frac{3^{1/\delta}}{1 + 3^{2/\delta}}.
\end{align*}
\]

The parameters of median (\( \rho_1 \)), inter-decile range (\( \rho_2 \)), left-right tail-weight ratio (\( \rho_3 \)), and tail-weight factor (\( \rho_4 \)) for the \( L_L \) distribution are bounded as:

\[
0 < \rho_1 < \infty, \rho_2 \geq 0, 0 \leq \rho_3 \leq 1, 0 \leq \rho_4 \leq 1/2,
\]

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 \( \hat{q}(u)_{u=p} \) be the \((100 \times p)\)th percentile from this sample, where \( p \in (0, 1) \). Then, \( \hat{q}(u)_{u=p} \) can be expressed as [14, p. 154]

\[
\hat{q}(u)_{u=p} = Y_{(i)} + \left(\frac{a}{b}\right)(Y_{(i+1)} - Y_{(i)})
\]

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 \( \rho_1-\rho_4 \) 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 \( \rho_1-\rho_4 \). See Section 3 for an example to demonstrate this methodology. Figure 1 (panel A) displays region for possible combinations of \( \rho_3 \) and \( \rho_4 \) 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]

\[
\begin{align*}
\lambda_1 &= \beta_0 \\
\lambda_2 &= 2\beta_1 - \beta_0 \\
\lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0
\end{align*}
\]
Characterizing log-logistic (\(L_L\)) distributions

\[
\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0
\]  
(17)

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

\[
\beta_r = \int_0^\infty x\{F(x)\}^rf(x)dx  
\]  
(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, \(\lambda_1\) and \(\lambda_2\) 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 (\(\lambda_3\) and \(\lambda_4\)) 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:

\[
\beta_0 = \left\{\pi e^{-\gamma/\delta} \csc (\pi/\delta)\right\}/\delta  
\]  
(19)

\[
\beta_1 = \left\{\pi e^{-\gamma/\delta}(1 + \delta) \csc (\pi/\delta)\right\}/(2\delta^2)  
\]  
(20)

\[
\beta_2 = \left\{\pi e^{-\gamma/\delta}(1 + \delta)(1 + 2\delta) \csc (\pi/\delta)\right\}/(6\delta^3)  
\]  
(21)

\[
\beta_3 = \left\{\pi e^{-\gamma/\delta}(1 + \delta)(1 + 2\delta)(1 + 3\delta) \csc (\pi/\delta)\right\}/(24\delta^4)  
\]  
(22)

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

Substituting \(\beta_0, \beta_1, \beta_2, \beta_3\) from (19)-(22) into (14)-(17) and simplifying yields the first four \(L\)-moments. The terms for \(\lambda_3\) and \(\lambda_4\) 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

\[
\lambda_1 = \left\{\pi e^{-\gamma/\delta} \csc (\pi/\delta)\right\}/\delta  
\]  
(23)

\[
\lambda_2 = \left\{\pi e^{-\gamma/\delta} \csc (\pi/\delta)\right\}/\delta^2  
\]  
(24)

\[
\tau_3 = \delta^{-1}  
\]  
(25)

\[
\tau_4 = \frac{1}{6} + \frac{5}{6\delta^2}.  
\]  
(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 \(\beta_0 - \beta_3\) are given by [9, pp. 113-114]

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

(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 \(\lambda_1 - \lambda_4\) 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 \(\tau_3\) and \(\tau_4\) for the MoLM-based \(L_L\) distributions.

Figure 1: Boundary graphs of MoP-based left-right tail-weight ratio \((\rho_3)\) and tail-weight factor \((\rho_4)\) (panel A) and MoLM-based \(L\)-skew \((\tau_3)\) and \(L\)-kurtosis \((\tau_4)\) (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 $\gamma$ and $\delta$ 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 $\gamma$ and $\delta$ can be used in (8) and (11), respectively, to compute the parameter values of median ($\rho_1$) and tail-weight factor ($\rho_4$). The MoP-based fit was obtained by using a linear transformation in the form $x = q(u) + (\hat{\rho}_1 - \hat{\rho}_4)$.

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 $\hat{\beta}_3$ 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 $\gamma$ and $\delta$ associated with the MoLM-based $L_L$ distribution were determined by solving (24) and (25) after substituting the estimates of $\hat{\lambda}_2$ and $\hat{\lambda}_3$ into the right-hand sides of (24) and (25). The solved values of $\gamma$ and $\delta$ can be used in (23) and (26), respectively, to compute the parameter values of $L$-mean ($\lambda_1$) and $L$-kurtosis ($\tau_4$). 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 $\gamma$ and $\delta$ 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 $\gamma$ and $\delta$ can be used in (32) and (35), respectively, to compute the parameter values of mean ($\mu$) and kurtosis ($\alpha_4$). 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 $\gamma$ and $\delta$ 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 associated with MoLM- and MoM-based fits by inspecting the values of Euclidian distances $d = \sqrt{\sum (OP - EP)^2}$, where OP = observed proportions and EP = expected proportions in each interval of percentiles, from Table 1.
\( \hat{\rho}_1 = 8445.000 \)
\( \hat{\rho}_2 = 14684.800 \)
\( \hat{\rho}_3 = 0.5477 \)
\( \hat{\rho}_4 = 0.4868 \)
\( \gamma = -34.1207 \)
\( \delta = 3.6496 \)

\( \hat{\lambda}_1 = 9879.1009 \)
\( \hat{\lambda}_2 = 3380.6927 \)
\( \hat{\tau}_3 = 0.2498 \)
\( \hat{\tau}_4 = 0.1839 \)
\( \gamma = -37.6603 \)
\( \delta = 4.0030 \)

\( \hat{\mu} = 9879.1009 \)
\( \hat{\sigma} = 6558.3993 \)
\( \hat{\alpha}_3 = 1.7035 \)
\( \hat{\alpha}_4 = 4.3400 \)
\( \gamma = -62.2764 \)
\( \delta = 6.2693 \)

Figure 2: The pdfs of (A) MoP-, (B) MoLM-, and (C) MoM-based \( L_\lambda \) 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 \( \gamma \) and \( \delta \) associated with the MoP-based \( L_L \) distribution were determined by solving (9) and (10) after substituting the values of \( \rho_2 \) and \( \rho_3 \) of Dagum distribution into the right-hand sides of (9) and (10). These values of \( \gamma \) and \( \delta \) can be used in (8) and (11), respectively, to compute the parameter values of \( \rho_1 \) and \( \rho_4 \) 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 \( \lambda_1, \lambda_2, \tau_3, \tau_4 \) associated with the Dagum distribution were computed using (18) and (14)-(17) and using the formulae for \( \tau_3 \) and \( \tau_4 \) from Section 2.2. The parameter values of \( \gamma \) and \( \delta \) associated with the MoLM-based \( L_L \) distribution were determined by solving (24) and (25) after substituting the values of \( \lambda_2 \) and \( \tau_3 \) of Dagum distribution into the right-hand sides of (24) and (25). These values of \( \gamma \) and \( \delta \) can be used in (23) and (26), respectively, to compute the parameter values of \( \lambda_1 \) and \( \tau_4 \) 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 \( \mu, \sigma, \alpha_3 \) and \( \alpha_4 \) 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 \( \gamma \) and \( \delta \) associated with the MoM-based \( L_L \) distribution were determined by solving (33) and (34) from Appendix A after substituting the values of \( \sigma \) and \( \alpha_3 \) of Dagum distribution into the right-hand sides of (33) and (34). These values of \( \gamma \) and \( \delta \) can be used in (32) and (35), respectively, to compute the parameter values of \( \mu \) and \( \alpha_4 \) 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 \( \gamma \) and \( \delta \) 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 \( d = \sqrt{\sum (\text{OP} - \text{EP})^2} \), where \( \text{OP} = \text{observed percentile} \) and \( \text{EP} = \text{expected percentiles} \), from Table 2.
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 $b = 4$. 

\[\hat{\rho}_1 = 4.7711\]
\[\hat{\rho}_2 = 3.7403\]
\[\hat{\rho}_3 = 0.5601\]
\[\hat{\rho}_4 = 0.4839\]
\[\gamma = -4.2295\]
\[\delta = 3.7903\]

\[\hat{\lambda}_1 = 5.1310\]
\[\hat{\lambda}_2 = 0.8894\]
\[\hat{\tau}_3 = 0.2455\]
\[\hat{\tau}_4 = 0.2094\]
\[\gamma = -4.8296\]
\[\delta = 4.0727\]

\[\hat{\mu} = 5.1310\]
\[\hat{\sigma} = 1.8083\]
\[\hat{\alpha}_3 = 2.9648\]
\[\hat{\alpha}_4 = 34.4831\]
\[\gamma = -6.1462\]
\[\delta = 4.5926\]
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 ($\rho_2$, $\rho_3$, and $\rho_4$), the MoLM-based estimators ($\hat{\lambda}_2$, $\hat{\tau}_3$, and $\hat{\tau}_4$) of $L$-scale, $L$-skew, and $L$-kurtosis ($\lambda_2$, $\tau_3$, and $\tau_4$) and the MoM-based estimators ($\hat{\sigma}$, $\hat{\alpha}_3$, and $\hat{\alpha}_4$) of standard deviation, skew, and kurtosis ($\sigma$, $\alpha_3$, and $\alpha_4$) were computed for each of the $(2 \times 25,000)$ samples based on the solved values of $\gamma$ and $\delta$ 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 re-samples 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: $\text{RB}\% = \{(E - P)/P\} \times 100$. 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-
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.

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>EP</th>
<th>MoP OP</th>
<th>MoLM OP</th>
<th>MoM OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.05</td>
<td>0.0410</td>
<td>0.0462</td>
<td>0.0085</td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>0.0590</td>
<td>0.0631</td>
<td>0.0724</td>
</tr>
<tr>
<td>25</td>
<td>0.15</td>
<td>0.1534</td>
<td>0.1552</td>
<td>0.1881</td>
</tr>
<tr>
<td>50</td>
<td>0.25</td>
<td>0.2467</td>
<td>0.2432</td>
<td>0.2730</td>
</tr>
<tr>
<td>75</td>
<td>0.25</td>
<td>0.2455</td>
<td>0.2335</td>
<td>0.2235</td>
</tr>
<tr>
<td>90</td>
<td>0.15</td>
<td>0.1545</td>
<td>0.1510</td>
<td>0.1305</td>
</tr>
<tr>
<td>95</td>
<td>0.05</td>
<td>0.0513</td>
<td>0.0520</td>
<td>0.0445</td>
</tr>
<tr>
<td>100</td>
<td>0.05</td>
<td>0.0487</td>
<td>0.0558</td>
<td>0.0596</td>
</tr>
</tbody>
</table>

$d = 0.0151$ $d = 0.0239$ $d = 0.0735$

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.

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>EP</th>
<th>MoP OP</th>
<th>MoLM OP</th>
<th>MoM OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3.1185</td>
<td>3.1224</td>
<td>3.0975</td>
<td>3.0121</td>
</tr>
<tr>
<td>10</td>
<td>3.4283</td>
<td>3.4283</td>
<td>3.4174</td>
<td>3.3669</td>
</tr>
<tr>
<td>25</td>
<td>4.0000</td>
<td>4.0031</td>
<td>4.0084</td>
<td>4.0055</td>
</tr>
<tr>
<td>50</td>
<td>4.7711</td>
<td>4.7711</td>
<td>4.7823</td>
<td>4.8166</td>
</tr>
<tr>
<td>75</td>
<td>5.8098</td>
<td>5.7973</td>
<td>5.7959</td>
<td>5.8469</td>
</tr>
<tr>
<td>90</td>
<td>7.1686</td>
<td>7.1686</td>
<td>7.1234</td>
<td>7.1557</td>
</tr>
<tr>
<td>95</td>
<td>8.3011</td>
<td>8.3561</td>
<td>8.2540</td>
<td>8.2426</td>
</tr>
</tbody>
</table>

$d = 0.0566$ $d = 0.0722$ $d = 0.1489$

Theoretical 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
Characterizing log-logistic \((L_L)\) distributions

### Table 4: 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.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>Estimate</th>
<th>95%C.I.</th>
<th>SE</th>
<th>RB%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 2A</td>
<td>(\rho_2)</td>
<td>14684.8</td>
<td>17923</td>
<td>17853,17990</td>
<td>35.01</td>
</tr>
<tr>
<td></td>
<td>(\rho_3)</td>
<td>0.5477</td>
<td>0.5745</td>
<td>0.5710,0.5780</td>
<td>0.00178</td>
</tr>
<tr>
<td></td>
<td>(\rho_4)</td>
<td>0.4782</td>
<td>0.4317</td>
<td>0.4305,0.4333</td>
<td>0.00070</td>
</tr>
<tr>
<td>Fig. 3A</td>
<td>(\rho_2)</td>
<td>3.7403</td>
<td>4.3600</td>
<td>4.5204,4.5542</td>
<td>0.00853</td>
</tr>
<tr>
<td></td>
<td>(\rho_3)</td>
<td>0.5601</td>
<td>0.5885</td>
<td>0.5849,0.5920</td>
<td>0.00183</td>
</tr>
<tr>
<td></td>
<td>(\rho_4)</td>
<td>0.4797</td>
<td>0.4346</td>
<td>0.4332,0.4360</td>
<td>0.00071</td>
</tr>
</tbody>
</table>

### Table 3: MoL-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.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>Estimate</th>
<th>95%C.I.</th>
<th>SE</th>
<th>RB%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 2B</td>
<td>(\lambda_2)</td>
<td>3380.7</td>
<td>3382</td>
<td>3371.5,3393.8</td>
<td>5.686</td>
</tr>
<tr>
<td></td>
<td>(\tau_3)</td>
<td>0.2498</td>
<td>0.2271</td>
<td>0.2254,0.2287</td>
<td>0.00083</td>
</tr>
<tr>
<td></td>
<td>(\tau_4)</td>
<td>0.2187</td>
<td>0.2046</td>
<td>0.2032,0.2059</td>
<td>0.00071</td>
</tr>
<tr>
<td>Fig. 3B</td>
<td>(\lambda_2)</td>
<td>0.8894</td>
<td>0.8899</td>
<td>0.8871,0.8929</td>
<td>0.00149</td>
</tr>
<tr>
<td></td>
<td>(\tau_3)</td>
<td>0.2455</td>
<td>0.2234</td>
<td>0.2218,0.2250</td>
<td>0.00083</td>
</tr>
<tr>
<td></td>
<td>(\tau_4)</td>
<td>0.2169</td>
<td>0.2033</td>
<td>0.2019,0.2046</td>
<td>0.00070</td>
</tr>
</tbody>
</table>

### 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.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter</th>
<th>Estimate</th>
<th>95%C.I.</th>
<th>SE</th>
<th>RB%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 2B</td>
<td>(\lambda_2)</td>
<td>3380.7</td>
<td>3380</td>
<td>3377.5,3382.5</td>
<td>1.277</td>
</tr>
<tr>
<td></td>
<td>(\tau_3)</td>
<td>0.2498</td>
<td>0.2482</td>
<td>0.2478,0.2486</td>
<td>0.00021</td>
</tr>
<tr>
<td></td>
<td>(\tau_4)</td>
<td>0.2187</td>
<td>0.2175</td>
<td>0.2171,0.2178</td>
<td>0.00018</td>
</tr>
<tr>
<td>Fig. 3B</td>
<td>(\lambda_2)</td>
<td>0.8894</td>
<td>0.8891</td>
<td>0.8885,0.8898</td>
<td>0.00034</td>
</tr>
<tr>
<td></td>
<td>(\tau_3)</td>
<td>0.2455</td>
<td>0.2440</td>
<td>0.2436,0.2444</td>
<td>0.00021</td>
</tr>
<tr>
<td></td>
<td>(\tau_4)</td>
<td>0.2169</td>
<td>0.2158</td>
<td>0.2154,0.2161</td>
<td>0.00018</td>
</tr>
</tbody>
</table>
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 ($\rho_3$) and tail-weight factor ($\rho_4$) and MoLM-based estimators of $L$-skew ($\tau_3$) and $L$-kurtosis ($\tau_4$) have over their corresponding MoM-based estimators of skew ($\alpha_3$) and kurtosis ($\alpha_4$) is clearly obvious. For example, with samples of size $n = 25$ the estimates of $\alpha_3$ and $\alpha_4$ 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 $\rho_3$ and $\rho_4$ 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 $\rho_3$ and $\rho_4$ and MoLM-based estimators of $\tau_3$ and $\tau_4$ are more efficient estimators as their relative standard errors RSE = \{(SE/Estimate) × 100\} are considerably smaller than those associated with MoM-based estimators of $\alpha_3$ and $\alpha_4$. 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

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

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 ($\rho_3$) and tail-weight factor ($\rho_4$) and MoLM-based estimators of $L$-skew ($\tau_3$) and $L$-kurtosis ($\tau_4$) have over their corresponding MoM-based estimators of skew ($\alpha_3$) and kurtosis ($\alpha_4$) is clearly obvious. For example, with samples of size $n = 25$ the estimates of $\alpha_3$ and $\alpha_4$ 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 $\rho_3$ and $\rho_4$ 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 $\rho_3$ and $\rho_4$ and MoLM-based estimators of $\tau_3$ and $\tau_4$ are more efficient estimators as their relative standard errors RSE = \{(SE/Estimate) × 100\} are considerably smaller than those associated with MoM-based estimators of $\alpha_3$ and $\alpha_4$. 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 $\rho_3$ and $\rho_4$ have about the same degree of precision compared to the MoLM-based estimators of $\tau_3$ and $\tau_4$, whereas both MoP- and MoLM-based estimators have substantially higher precision when compared to the MoM-based estimators of $\alpha_3$ and $\alpha_4$.

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


Characterizing log-logistic ($L_L$) distributions


Appendices

A Method of (Product) Moments

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

$$
\mu_r = \int_0^\infty x^r f(x)dx
$$

(28)

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

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

(29)

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

Provided that the first four moments ($\mu_{r=1,2,3,4}$) exist, the MoM-based parameters of mean ($\mu$) and standard deviation ($\sigma$) are respectively given as
\[ \mu = \mu_1 \text{ and } \sigma = \sqrt{\mu_2 - \mu_1^2}, \text{ whereas the parameters of skew and kurtosis are obtained by substituting these four moments into the following formulae for skew (} \alpha_3 \text{) and kurtosis (} \alpha_4 \text{) from [23]:} \]

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

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

Extracting the first four moments \( (\mu_{r=1,2,3,4}) \) from (29) and substituting them into (30) and (31), the indices of skew \( (\alpha_3) \) and kurtosis \( (\alpha_4) \), preceded by mean and standard deviation of \( L_L \) distributions are expressed as:

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

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

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

\[ \alpha_4 = 2\{2\delta^3 \csc(4\pi/\delta) + 6\pi^2 \delta \csc(\pi/\delta)^3 \sec(\pi/\delta) \\
- 6\pi \delta^2 \csc(\pi/\delta) \csc(3\pi/\delta) - 6\pi \delta^2 \csc(2\pi/\delta)^2 \\
- 3\pi^3 \csc(\pi/\delta)^4\} / \{\pi(2\delta \csc(2\pi/\delta) - \pi \csc(\pi/\delta)^2)^2\}. \tag{35} \]

In the context of \( L_L \) distributions, the MoM-based procedure involves solving of (33) and (34) for the parameters of \( \gamma \) and \( \delta \) after given values (or, estimates) of standard deviation \( (\sigma) \) and skew \( (\alpha_3) \) are substituted into the right-hand sides of (33) and (34). The solved values of \( \gamma \) and \( \delta \) can be substituted into (32) and (35), respectively, for computing the values of mean and kurtosis.
 Boundary Graph of MoM-based Measures of Skew ($\alpha_3$) and Kurtosis ($\alpha_4$) for the $L_L$ Distributions.

Figure B: Boundary graph of MoM-based skew ($\alpha_3$) and kurtosis ($\alpha_4$) associated with the $L_L$ distributions. Fig. B can be used to find a possible combination of skew ($\alpha_3$) and kurtosis ($\alpha_4$) of a valid $L_L$ distribution.

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